Graduate Texts in Mathematics

Serge Lang

Complex Analysis

Fourth Edition

Springer

Graduate Texts in Mathematics 103

Springer Science+Business Media, LLC

BOOKS OF RELATED INTEREST BY SERGE LANG

Short Calculus
2002, ISBN 0-387-95327-2

Calculus of Several Variables, Third Edition
1987, ISBN 0-387-96405-3

Undergraduate Analysis, Second Edition
1996, ISBN 0-387-94841-4

Introduction to Linear Algebra
1997, ISBN 0-387-96205-0

Math Talks for Undergraduates
1999, ISBN 0-387-98749-5

OTHER BOOKS BY LANG PUBLISHED BY SPRINGER-VERLAG

Math! Encounters with High School Students • The Beauty of Doing Mathematics • Geometry: A High School Course • Basic Mathematics • Short Calculus • A First Course in Calculus • Introduction to Linear Algebra • Calculus of Several Variables • Linear Algebra • Undergraduate Analysis • Undergraduate Algebra • Complex Analysis • Math Talks for Undergraduates • Algebra • Real and Functional Analysis • Introduction to Differentiable Manifolds • Fundamentals of Differential Geometry • Algebraic Number Theory • Cyclotomic Fields I and II • Introduction to Diophantine Approximations • $SL_2(\mathbf{R})$ • Spherical Inversion on $SL_n(\mathbf{R})$ (with Jay Jorgenson) • Elliptic Functions • Elliptic Curves: Diophantine Analysis • Introduction to Arakelov Theory • Riemann-Roch Algebra (with William Fulton) • Abelian Varieties • Introduction to Algebraic and Abelian Functions • Complex Multiplication • Introduction to Modular Forms • Modular Units (with Daniel Kubert) • Introduction to Complex Hyperbolic Spaces • Number Theory III • Survey on Diophantine Geometry

Collected Papers I-V, including the following: Introduction to Transcendental Numbers in volume I, Frobenius Distributions in GL2-Extensions (with Hale Trotter in volume II, Topics in Cohomology of Groups in volume IV, Basic Analysis of Regularized Series and Products (with Jay Jorgenson) in volume V and Explicit Formulas for Regularized Products and Series (with Jay Jorgenson) in volume V

THE FILE • CHALLENGES

Serge Lang

Complex Analysis

Fourth Edition

With 139 Illustrations

Springer

Serge Lang
Department of Mathematics
Yale University
New Haven, CT 06520
USA

Mathematics Subject Classification (2000): 30-01

Library of Congress Cataloging-in-Publication Data
Lang, Serge, 1927-
Complex analysis / Serge Lang. — 4th ed.
p. cm. — (Graduate texts in mathematics ; 103)
Includes bibliographical references and index.
ISBN 978-1-4419-3135-1 ISBN 978-1-4757-3083-8 (eBook)
DOI 10.1007/978-1-4757-3083-8

1. Functions of complex variables. 2. Mathematical analysis.
I. Title. II. Series
QA331.7 .L36 1999
515'.9—dc21 98-29992

Printed on acid-free paper.

Originally published by Springer-Verlag New York, Inc. in 1999
Softcover reprint of the hardcover 4th edition 1999

9 8 7 6 5 4

www.springer-ny.com

Foreword

The present book is meant as a text for a course on complex analysis at the advanced undergraduate level, or first-year graduate level. The first half, more or less, can be used for a one-semester course addressed to undergraduates. The second half can be used for a second semester, at either level. Somewhat more material has been included than can be covered at leisure in one or two terms, to give opportunities for the instructor to exercise individual taste, and to lead the course in whatever directions strikes the instructor's fancy at the time as well as extra reading material for students on their own. A large number of routine exercises are included for the more standard portions, and a few harder exercises of striking theoretical interest are also included, but may be omitted in courses addressed to less advanced students.

In some sense, I think the classical German prewar texts were the best (Hurwitz–Courant, Knopp, Bieberbach, etc.) and I would recommend to anyone to look through them. More recent texts have emphasized connections with real analysis, which is important, but at the cost of exhibiting succinctly and clearly what is peculiar about complex analysis: the power series expansion, the uniqueness of analytic continuation, and the calculus of residues. The systematic elementary development of formal and convergent power series was standard fare in the German texts, but only Cartan, in the more recent books, includes this material, which I think is quite essential, e.g., for differential equations. I have written a short text, exhibiting these features, making it applicable to a wide variety of tastes.

The book essentially decomposes into two parts.

The *first part*, Chapters I through VIII, includes the basic properties of analytic functions, essentially what cannot be left out of, say, a one-semester course.

I have no fixed idea about the manner in which Cauchy's theorem is to be treated. In less advanced classes, or if time is lacking, the usual hand waving about simple closed curves and interiors is not entirely inappropriate. Perhaps better would be to state precisely the homological version and omit the formal proof. For those who want a more thorough understanding, I include the relevant material.

Artin originally had the idea of basing the homology needed for complex variables on the winding number. I have included his proof for Cauchy's theorem, extracting, however, a purely topological lemma of independent interest, not made explicit in Artin's original *Notre Dame* notes [Ar 65] or in Ahlfors' book closely following Artin [Ah 66]. I have also included the more recent proof by Dixon, which uses the winding number, but replaces the topological lemma by greater use of elementary properties of analytic functions which can be derived directly from the local theorem. The two aspects, homotopy and homology, both enter in an essential fashion for different applications of analytic functions, and neither is slighted at the expense of the other.

Most expositions usually include some of the global geometric properties of analytic maps at an early stage. I chose to make the preliminaries on complex functions as short as possible to get quickly into the analytic part of complex function theory: power series expansions and Cauchy's theorem. The advantages of doing this, reaching the heart of the subject rapidly, are obvious. The cost is that certain elementary global geometric considerations are thus omitted from Chapter I, for instance, to reappear later in connection with analytic isomorphisms (Conformal Mappings, Chapter VII) and potential theory (Harmonic Functions, Chapter VIII). I think it is best for the coherence of the book to have covered in one sweep the basic analytic material before dealing with these more geometric global topics. Since the proof of the general Riemann mapping theorem is somewhat more difficult than the study of the specific cases considered in Chapter VII, it has been postponed to the second part.

The *second and third parts* of the book, Chapters IX through XVI, deal with further assorted analytic aspects of functions in many directions, which may lead to many other branches of analysis. I have emphasized the possibility of defining analytic functions by an integral involving a parameter and differentiating under the integral sign. Some classical functions are given to work out as exercises, but the gamma function is worked out in detail in the text, as a prototype.

The chapters in Part II allow considerable flexibility in the order they are covered. For instance, the chapter on analytic continuation, including the Schwarz reflection principle, and/or the proof of the Riemann mapping theorem could be done right after Chapter VII, and still achieve great coherence.

As most of this part is somewhat harder than the first part, it can easily be omitted from a one-term course addressed to undergraduates. In the same spirit, some of the harder exercises in the first part have been starred, to make their omission easy.

Comments on the Third and Fourth Editions

I have rewritten some sections and have added a number of exercises. I have added some material on harmonic functions and conformal maps, on the Borel theorem and Borel's proof of Picard's theorem, as well as D.J. Newman's short proof of the prime number theorem, which illustrates many aspects of complex analysis in a classical setting. I have made more complete the treatment of the gamma and zeta functions. I have also added an Appendix which covers some topics which I find sufficiently important to have in the book. The first part of the Appendix recalls summation by parts and its application to uniform convergence. The others cover material which is not usually included in standard texts on complex analysis: difference equations, analytic differential equations, fixed points of fractional linear maps (of importance in dynamical systems), Cauchy's formula for C^∞ functions, and Cauchy's theorem for locally integrable vector fields in the plane. This material gives additional insight on techniques and results applied to more standard topics in the text. Some of them may have been assigned as exercises, and I hope students will try to prove them before looking up the proofs in the Appendix.

I am very grateful to several people for pointing out the need for a number of corrections, especially Keith Conrad, Wolfgang Fluch, Alberto Grunbaum, Bert Hochwald, Michal Jastrzebski, José Carlos Santos, Ernest C. Schlesinger, A. Vijayakumar, Barnet Weinstock, and Sandy Zabell. Finally, I thank Rami Shakarchi for working out an answer book.

New Haven 1998 SERGE LANG

Prerequisites

We assume that the reader has had two years of calculus, and has some acquaintance with epsilon–delta techniques. For convenience, we have recalled all the necessary lemmas we need for continuous functions on compact sets in the plane. Section §1 in the Appendix also provides some background.

We use what is now standard terminology. A function

$$f: S \to T$$

is called **injective** if $x \neq y$ in S implies $f(x) \neq f(y)$. It is called **surjective** if for every z in T there exists $x \in S$ such that $f(x) = z$. If f is surjective, then we also say that f maps S **onto** T. If f is both injective and surjective then we say that f is **bijective**.

Given two functions f, g defined on a set of real numbers containing arbitrarily large numbers, and such that $g(x) \geqq 0$, we write

$$f \ll g \qquad \text{or} \qquad f(x) \ll g(x) \qquad \text{for} \quad x \to \infty$$

to mean that there exists a number $C > 0$ such that for all x sufficiently large, we have

$$|f(x)| \leqq Cg(x).$$

Similarly, if the functions are defined for x near 0, we use the same symbol $\ll$ for $x \to 0$ to mean that there exists $C > 0$ such that

$$|f(x)| \leqq Cg(x)$$

for all x sufficiently small (there exists $\delta > 0$ such that if $|x| < \delta$ then $|f(x)| \leqq Cg(x)$). Often this relation is also expressed by writing

$$f(x) = O(g(x)),$$

which is read: $f(x)$ is **big oh of** $g(x)$, for $x \to \infty$ or $x \to 0$ as the case may be.

We use $]a, b[$ to denote the **open** interval of numbers

$$a < x < b.$$

Similarly, $[a, b[$ denotes the half-open interval, etc.

Contents

PART ONE

Basic Theory

CHAPTER I

Complex Numbers and Functions

One of the advantages of dealing with the real numbers instead of the rational numbers is that certain equations which do not have any solutions in the rational numbers have a solution in real numbers. For instance, $x^2 = 2$ is such an equation. However, we also know some equations having no solution in real numbers, for instance $x^2 = -1$, or $x^2 = -2$. We define a new kind of number where such equations have solutions. The new kind of numbers will be called **complex** numbers.

I, §1. DEFINITION

The **complex numbers** are a set of objects which can be added and multiplied, the sum and product of two complex numbers being also a complex number, and satisfy the following conditions.

1. Every real number is a complex number, and if α, β are real numbers, then their sum and product as complex numbers are the same as their sum and product as real numbers.
2. There is a complex number denoted by i such that $i^2 = -1$.
3. Every complex number can be written uniquely in the form $a + bi$ where a, b are real numbers.
4. The ordinary laws of arithmetic concerning addition and multiplication are satisfied. We list these laws:

   If α, β, γ are complex numbers, then $(\alpha\beta)\gamma = \alpha(\beta\gamma)$, and

   $$(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma).$$

We have $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$, and $(\beta + \gamma)\alpha = \beta\alpha + \gamma\alpha$.

We have $\alpha\beta = \beta\alpha$, and $\alpha + \beta = \beta + \alpha$.

If 1 is the real number one, then $1\alpha = \alpha$.

If 0 is the real number zero, then $0\alpha = 0$.

We have $\alpha + (-1)\alpha = 0$.

We shall now draw consequences of these properties. With each complex number $a + bi$, we associate the point (a, b) in the plane. Let $\alpha = a_1 + a_2 i$ and $\beta = b_1 + b_2 i$ be two complex numbers. Then

$$\alpha + \beta = a_1 + b_1 + (a_2 + b_2)i.$$

Hence addition of complex numbers is carried out "componentwise". For example, $(2 + 3i) + (-1 + 5i) = 1 + 8i$.

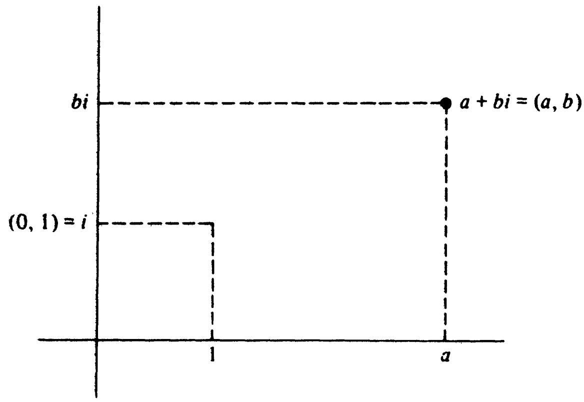


Figure 1

In multiplying complex numbers, we use the rule $i^2 = -1$ to simplify a product and to put it in the form $a + bi$. For instance, let $\alpha = 2 + 3i$ and $\beta = 1 - i$. Then

$$\begin{aligned}\alpha\beta = (2 + 3i)(1 - i) &= 2(1 - i) + 3i(1 - i)\\ &= 2 - 2i + 3i - 3i^2\\ &= 2 + i - 3(-1)\\ &= 2 + 3 + i\\ &= 5 + i.\end{aligned}$$

Let $\alpha = a + bi$ be a complex number. We define $\bar{\alpha}$ to be $a - bi$. Thus if $\alpha = 2 + 3i$, then $\bar{\alpha} = 2 - 3i$. The complex number $\bar{\alpha}$ is called the

conjugate of α. We see at once that

$$\alpha\bar{\alpha} = a^2 + b^2.$$

With the vector interpretation of complex numbers, we see that $\alpha\bar{\alpha}$ is the square of the distance of the point (a, b) from the origin.

We now have one more important property of complex numbers, which will allow us to divide by complex numbers other than 0.

If $\alpha = a + bi$ is a complex number $\neq 0$, and if we let

$$\lambda = \frac{\bar{\alpha}}{a^2 + b^2}$$

then $\alpha\lambda = \lambda\alpha = 1$.

The proof of this property is an immediate consequence of the law of multiplication of complex numbers, because

$$\alpha \frac{\bar{\alpha}}{a^2 + b^2} = \frac{\alpha\bar{\alpha}}{a^2 + b^2} = 1.$$

The number λ above is called the **inverse** of α, and is denoted by α^{-1} or $1/\alpha$. If α, β are complex numbers, we often write β/α instead of $\alpha^{-1}\beta$ (or $\beta\alpha^{-1}$), just as we did with real numbers. We see that we can divide by complex numbers $\neq 0$.

Example. To find the inverse of $(1 + i)$ we note that the conjugate of $1 + i$ is $1 - i$ and that $(1 + i)(1 - i) = 2$. Hence

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

Theorem 1.1. *Let α, β be complex numbers. Then*

$$\overline{\alpha\beta} = \bar{\alpha}\bar{\beta}, \qquad \overline{\alpha + \beta} = \bar{\alpha} + \bar{\beta}, \qquad \bar{\bar{\alpha}} = \alpha.$$

Proof. The proofs follow immediately from the definitions of addition, multiplication, and the complex conjugate. We leave them as exercises (Exercises 3 and 4).

Let $\alpha = a + bi$ be a complex number, where a, b are real. We shall call a the **real part** of α, and denote it by $\mathrm{Re}(\alpha)$. Thus

$$\alpha + \bar{\alpha} = 2a = 2\,\mathrm{Re}(\alpha).$$

The real number b is called the **imaginary part** of α, and denoted by $\mathrm{Im}(\alpha)$.

We define the **absolute value** of a complex number $\alpha = a_1 + ia_2$ (where a_1, a_2 are real) to be

$$|\alpha| = \sqrt{a_1^2 + a_2^2}.$$

If we think of α as a point in the plane (a_1, a_2), then $|\alpha|$ is the length of the line segment from the origin to α. In terms of the absolute value, we can write

$$\alpha^{-1} = \frac{\bar{\alpha}}{|\alpha|^2}$$

provided $\alpha \neq 0$. Indeed, we observe that $|\alpha|^2 = \alpha\bar{\alpha}$.

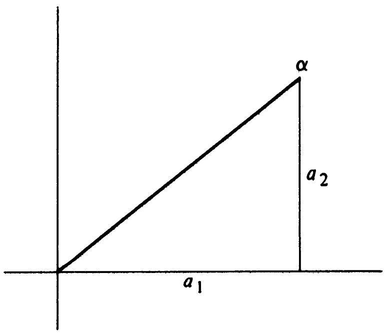


Figure 2

If $\alpha = a_1 + ia_2$, we note that

$$|\alpha| = |\bar{\alpha}|$$

because $(-a_2)^2 = a_2^2$, so $\sqrt{a_1^2 + a_2^2} = \sqrt{a_1^2 + (-a_2)^2}$.

Theorem 1.2. *The absolute value of a complex number satisfies the following properties. If α, β are complex numbers, then*

$$|\alpha\beta| = |\alpha||\beta|,$$
$$|\alpha + \beta| \leqq |\alpha| + |\beta|.$$

Proof. We have

$$|\alpha\beta|^2 = \alpha\beta\overline{\alpha\beta} = \alpha\bar{\alpha}\beta\bar{\beta} = |\alpha|^2|\beta|^2.$$

Taking the square root, we conclude that $|\alpha||\beta| = |\alpha\beta|$, thus proving the first assertion. As for the second, we have

$$\begin{aligned}|\alpha + \beta|^2 = (\alpha + \beta)(\overline{\alpha + \beta}) &= (\alpha + \beta)(\bar{\alpha} + \bar{\beta}) \\ &= \alpha\bar{\alpha} + \beta\bar{\alpha} + \alpha\bar{\beta} + \beta\bar{\beta} \\ &= |\alpha|^2 + 2\,\mathrm{Re}(\beta\bar{\alpha}) + |\beta|^2\end{aligned}$$

because $\alpha\bar{\beta} = \overline{\beta\bar{\alpha}}$. However, we have

$$2\,\mathrm{Re}(\beta\bar{\alpha}) \leqq 2|\beta\bar{\alpha}|$$

because the real part of a complex number is $\leqq$ its absolute value. Hence

$$\begin{aligned}|\alpha + \beta|^2 &\leqq |\alpha|^2 + 2|\beta\bar{\alpha}| + |\beta|^2 \\ &\leqq |\alpha|^2 + 2|\beta||\alpha| + |\beta|^2 \\ &= (|\alpha| + |\beta|)^2.\end{aligned}$$

Taking the square root yields the second assertion of the theorem.

The inequality

$$|\alpha + \beta| \leqq |\alpha| + |\beta|$$

is called the **triangle inequality**. It also applies to a sum of several terms. If $z_1, \ldots, z_n$ are complex numbers then we have

$$|z_1 + \cdots + z_n| \leqq |z_1| + \cdots + |z_n|.$$

Also observe that for any complex number z, we have

$$|-z| = |z|.$$

Proof?

I, §1. EXERCISES

1. Express the following complex numbers in the form $x + iy$, where x, y are real numbers.
   (a) $(-1 + 3i)^{-1}$ (b) $(1 + i)(1 - i)$
   (c) $(1 + i)i(2 - i)$ (d) $(i - 1)(2 - i)$
   (e) $(7 + \pi i)(\pi + i)$ (f) $(2i + 1)\pi i$
   (g) $(\sqrt{2}i)(\pi + 3i)$ (h) $(i + 1)(i - 2)(i + 3)$
2. Express the following complex numbers in the form $x + iy$, where x, y are real numbers.

(a) $(1+i)^{-1}$ (b) $\dfrac{1}{3+i}$ (c) $\dfrac{2+i}{2-i}$ (d) $\dfrac{1}{2-i}$

(e) $\dfrac{1+i}{i}$ (f) $\dfrac{i}{1+i}$ (g) $\dfrac{2i}{3-i}$ (h) $\dfrac{1}{-1+i}$

3. Let α be a complex number $\neq 0$. What is the absolute value of $\alpha/\bar{\alpha}$? What is $\bar{\bar{\alpha}}$?

4. Let α, β be two complex numbers. Show that $\overline{\alpha\beta} = \bar{\alpha}\bar{\beta}$ and that

$$\overline{\alpha + \beta} = \bar{\alpha} + \bar{\beta}.$$

5. Justify the assertion made in the proof of Theorem 1.2, that the real part of a complex number is $\leqq$ its absolute value.

6. If $\alpha = a + ib$ with a, b real, then b is called the **imaginary part** of α and we write $b = \mathrm{Im}(\alpha)$. Show that $\alpha - \bar{\alpha} = 2i\,\mathrm{Im}(\alpha)$. Show that

$$\mathrm{Im}(\alpha) \leqq |\mathrm{Im}(\alpha)| \leqq |\alpha|.$$

7. Find the real and imaginary parts of $(1+i)^{100}$.

8. Prove that for any two complex numbers z, w we have:
   (a) $|z| \leqq |z - w| + |w|$
   (b) $|z| - |w| \leqq |z - w|$
   (c) $|z| - |w| \leqq |z + w|$

9. Let $\alpha = a + ib$ and $z = x + iy$. Let c be real >0. Transform the condition

$$|z - \alpha| = c$$

into an equation involving only x, y, a, b, and c, and describe in a simple way what geometric figure is represented by this equation.

10. Describe geometrically the sets of points z satisfying the following conditions.
    (a) $|z - i + 3| = 5$ (b) $|z - i + 3| > 5$
    (c) $|z - i + 3| \leqq 5$ (d) $|z + 2i| \leqq 1$
    (e) $\mathrm{Im}\, z > 0$ (f) $\mathrm{Im}\, z \geqq 0$
    (g) $\mathrm{Re}\, z > 0$ (h) $\mathrm{Re}\, z \geqq 0$

I, §2. POLAR FORM

Let $(x, y) = x + iy$ be a complex number. We know that any point in the plane can be represented by polar coordinates (r, θ). We shall now see how to write our complex number in terms of such polar coordinates.

Let θ be a real number. We define the expression $e^{i\theta}$ to be

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

Thus $e^{i\theta}$ is a complex number.

For example, if $\theta = \pi$, then $e^{i\pi} = -1$. Also, $e^{2\pi i} = 1$, and $e^{i\pi/2} = i$. Furthermore, $e^{i(\theta + 2\pi)} = e^{i\theta}$ for any real θ.

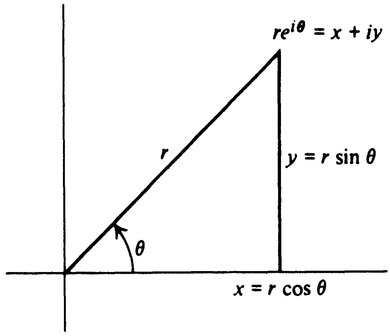


Figure 3

Let x, y be real numbers and $x + iy$ a complex number. Let

$$r = \sqrt{x^2 + y^2}.$$

If (r, θ) are the polar coordinates of the point (x, y) in the plane, then

$$x = r \cos \theta \qquad \text{and} \qquad y = r \sin \theta.$$

Hence

$$x + iy = r \cos \theta + ir \sin \theta = re^{i\theta}.$$

The expression $re^{i\theta}$ is called the **polar form** of the complex number $x + iy$. The number θ is sometimes called the **angle**, or **argument** of z, and we write

$$\theta = \arg z.$$

The most important property of this polar form is given in Theorem 2.1. It will allow us to have a very good geometric interpretation for the product of two complex numbers.

Theorem 2.1. *Let θ, φ be two real numbers. Then*

$$e^{i\theta + i\varphi} = e^{i\theta} e^{i\varphi}.$$

Proof. By definition, we have

$$e^{i\theta + i\varphi} = e^{i(\theta + \varphi)} = \cos(\theta + \varphi) + i \sin(\theta + \varphi).$$

Using the addition formulas for sine and cosine, we see that the preceding expression is equal to

$$\cos\theta\cos\varphi - \sin\theta\sin\varphi + i(\sin\theta\cos\varphi + \sin\varphi\cos\theta).$$

This is exactly the same expression as the one we obtain by multiplying out

$$(\cos\theta + i\sin\theta)(\cos\varphi + i\sin\varphi).$$

Our theorem is proved.

Theorem 2.1 justifies our notation, by showing that the exponential of complex numbers satisfies the same formal rule as the exponential of real numbers.

Let $\alpha = a_1 + ia_2$ be a complex number. We define e^α to be

$$e^{a_1}e^{ia_2}.$$

For instance, let $\alpha = 2 + 3i$. Then $e^\alpha = e^2e^{3i}$.

Theorem 2.2. *Let α, β be complex numbers. Then*

$$e^{\alpha+\beta} = e^\alpha e^\beta.$$

Proof. Let $\alpha = a_1 + ia_2$ and $\beta = b_1 + ib_2$. Then

$$\begin{aligned} e^{\alpha+\beta} &= e^{(a_1+b_1)+i(a_2+b_2)} = e^{a_1+b_1}e^{i(a_2+b_2)} \\ &= e^{a_1}e^{b_1}e^{ia_2+ib_2}. \end{aligned}$$

Using Theorem 2.1, we see that this last expression is equal to

$$e^{a_1}e^{b_1}e^{ia_2}e^{ib_2} = e^{a_1}e^{ia_2}e^{b_1}e^{ib_2}.$$

By definition, this is equal to $e^\alpha e^\beta$, thereby proving our theorem.

Theorem 2.2 is very useful in dealing with complex numbers. We shall now consider several examples to illustrate it.

Example 1. Find a complex number whose square is $4e^{i\pi/2}$.

Let $z = 2e^{i\pi/4}$. Using the rule for exponentials, we see that $z^2 = 4e^{i\pi/2}$.

Example 2. Let n be a positive integer. Find a complex number w such that $w^n = e^{i\pi/2}$.

It is clear that the complex number $w = e^{i\pi/2n}$ satisfies our requirement. In other words, we may express Theorem 2.2 as follows:

Let $z_1 = r_1 e^{i\theta_1}$ and $z_2 = r_2 e^{i\theta_2}$ be two complex numbers. To find the product $z_1 z_2$, we multiply the absolute values and add the angles. Thus

$$z_1 z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)}.$$

In many cases, this way of visualizing the product of complex numbers is more useful than that coming out of the definition.

Warning. We have not touched on the logarithm. As in calculus, we want to say that $e^z = w$ if and only if $z = \log w$. Since $e^{2\pi ik} = 1$ for all integers k, it follows that the inverse function $z = \log w$ is defined only up to the addition of an integer multiple of $2\pi i$. We shall study the logarithm more closely in Chapter II, §3, Chapter II, §5, and Chapter III, §6.

I, §2. EXERCISES

1. Put the following complex numbers in polar form.
   (a) $1 + i$ (b) $1 + i\sqrt{2}$ (c) -3 (d) $4i$
   (e) $1 - i\sqrt{2}$ (f) $-5i$ (g) -7 (h) $-1 - i$

2. Put the following complex numbers in the ordinary form $x + iy$.
   (a) $e^{3i\pi}$ (b) $e^{2i\pi/3}$ (c) $3e^{i\pi/4}$ (d) $\pi e^{-i\pi/3}$
   (e) $e^{2\pi i/6}$ (f) $e^{-i\pi/2}$ (g) $e^{-i\pi}$ (h) $e^{-5i\pi/4}$

3. Let α be a complex number $\neq 0$. Show that there are two distinct complex numbers whose square is α.

4. Let $a + bi$ be a complex number. Find real numbers x, y such that

$$(x + iy)^2 = a + bi,$$

   expressing x, y in terms of a and b.

5. Plot all the complex numbers z such that $z^n = 1$ on a sheet of graph paper, for $n = 2, 3, 4$, and 5.

6. Let α be a complex number $\neq 0$. Let n be a positive integer. Show that there are n distinct complex numbers z such that $z^n = \alpha$. Write these complex numbers in polar form.

7. Find the real and imaginary parts of $i^{1/4}$, taking the fourth root such that its angle lies between 0 and $\pi/2$.

8. (a) Describe all complex numbers z such that $e^z = 1$.
   (b) Let w be a complex number. Let α be a complex number such that $e^\alpha = w$. Describe all complex numbers z such that $e^z = w$.

9. If $e^z = e^w$, show that there is an integer k such that $z = w + 2\pi k i$.

10. (a) If θ is real, show that

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad \text{and} \quad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

(b) For arbitrary complex z, suppose we define $\cos z$ and $\sin z$ by replacing θ with z in the above formula. Show that the only values of z for which $\cos z = 0$ and $\sin z = 0$ are the usual real values from trigonometry.

11. Prove that for any complex number $z \neq 1$ we have

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

12. Using the preceding exercise, and taking real parts, prove:

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

for $0 < \theta < 2\pi$.

13. Let z, w be two complex numbers such that $\bar{z}w \neq 1$. Prove that

$$\left|\frac{z - w}{1 - \bar{z}w}\right| < 1 \quad \text{if } |z| < 1 \text{ and } |w| < 1,$$

$$\left|\frac{z - w}{1 - \bar{z}w}\right| = 1 \quad \text{if } |z| = 1 \text{ or } |w| = 1.$$

(There are many ways of doing this. One way is as follows. First check that you may assume that z is real, say $z = r$. For the first inequality you are reduced to proving

$$(r - w)(r - \bar{w}) < (1 - rw)(1 - r\bar{w}).$$

Expand both sides and make cancellations to simplify the problem.)

I, §3. COMPLEX VALUED FUNCTIONS

Let S be a set of complex numbers. An association which to each element of S associates a complex number is called a **complex valued function**, or a **function** for short. We denote such a function by symbols like

$$f\colon S \to \mathbf{C}.$$

If z is an element of S, we write the association of the **value** $f(z)$ to z by the special arrow

$$z \mapsto f(z).$$

We can write

$$f(z) = u(z) + iv(z),$$

where $u(z)$ and $v(z)$ are real numbers, and thus

$$z \mapsto u(z), \qquad z \mapsto v(z)$$

are real valued functions. We call u the **real part** of f, and v the **imaginary part** of f.

We shall usually write

$$z = x + iy,$$

where x, y are real. Then the values of the function f can be written in the form

$$f(z) = f(x + iy) = u(x, y) + iv(x, y),$$

viewing u, v as functions of the two real variables x and y.

Example. For the function

$$f(z) = x^3 y + i \sin(x + y),$$

we have the real part,

$$u(x, y) = x^3 y,$$

and the imaginary part,

$$v(x, y) = \sin(x + y).$$

Example. The most important examples of complex functions are the power functions. Let n be a positive integer. Let

$$f(z) = z^n.$$

Then in polar coordinates, we can write $z = re^{i\theta}$, and therefore

$$f(z) = r^n e^{in\theta} = r^n(\cos n\theta + i \sin n\theta).$$

For this function, the real part is $r^n \cos n\theta$, and the imaginary part is $r^n \sin n\theta$.

Let $\bar{D}$ be the closed disc of radius 1 centered at the origin in $\mathbf{C}$. In other words, $\bar{D}$ is the set of complex numbers z such that $|z| \leqq 1$. If z is an element of $\bar{D}$, then z^n is also an element of $\bar{D}$, and so $z \mapsto z^n$ maps $\bar{D}$ into itself. Let S be the sector of complex numbers $re^{i\theta}$ such that

$$0 \leqq \theta \leqq 2\pi/n,$$

as shown on Fig. 4.

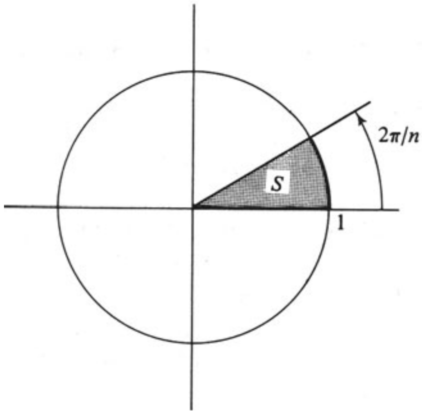


Figure 4

The function of a real variable

$$r \mapsto r^n$$

maps the unit interval [0, 1] onto itself. The function

$$\theta \mapsto n\theta$$

maps the interval

$$[0, 2\pi/n] \to [0, 2\pi].$$

In this way, we see that the function $f(z) = z^n$ maps the sector S onto the full disc of all numbers

$$w = te^{i\varphi},$$

with $0 \leqq t \leqq 1$ and $0 \leqq \varphi \leqq 2\pi$. We may say that the power function wraps the sector around the disc.

We could give a similar argument with other sectors of angle $2\pi/n$

as shown on Fig. 5. Thus we see that $z \mapsto z^n$ wraps the disc n times around.

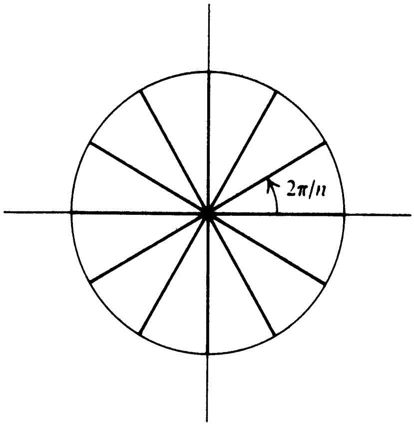


Figure 5

Given a complex number $z = re^{i\theta}$, you should have done Exercise 6 of the preceding section, or at least thought about it. For future reference, we now give the answer explicitly. We want to describe all complex numbers w such that $w^n = z$. Write

$$w = te^{i\varphi}.$$

Then

$$w^n = t^n e^{in\varphi}, \qquad 0 \leqq t.$$

If $w^n = z$, then $t^n = r$, and there is a unique real number $t \geqq 0$ such that $t^n = r$. On the other hand, we must also have

$$e^{in\varphi} = e^{i\theta},$$

which is equivalent with

$$in\varphi = i\theta + 2\pi ik,$$

where k is some integer. Thus we can solve for φ and get

$$\varphi = \frac{\theta}{n} + \frac{2\pi k}{n}.$$

The numbers

$$w_k = e^{i\theta/n}e^{2\pi ik/n}, \qquad k = 0, 1, \ldots, n - 1$$

are all distinct, and are drawn on Fig. 6. These numbers w_k may be described pictorially as those points on the circle which are the vertices of a regular polygon with n sides inscribed in the unit circle, with one vertex being at the point $e^{i\theta/n}$.

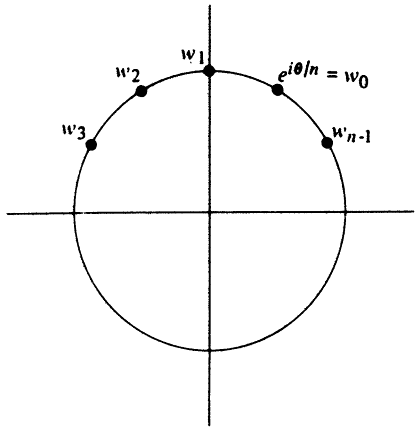


Figure 6

Each complex number

$$\zeta^k = e^{2\pi ik/n}$$

is called a **root of unity**, in fact, an n-th root of unity, because its n-th power is 1, namely

$$(\zeta^k)^n = e^{2\pi ikn/n} = e^{2\pi ik} = 1.$$

The points w_k are just the product of $e^{i\theta/n}$ with all the n-th roots of unity,

$$w_k = e^{i\theta/n}\zeta^k.$$

One of the major results of the theory of complex variables is to reduce the study of certain functions, including most of the common functions you can think of (like exponentials, logs, sine, cosine) to power series, which can be approximated by polynomials. Thus the power function is in some sense the unique basic function out of which the others are constructed. For this reason it was essential to get a good intuition of the power function. We postpone discussing the geometric aspects of the other functions to Chapters VII and VIII, except for some simple exercises.

I, §3. EXERCISES

1. Let $f(z) = 1/z$. Describe what f does to the inside and outside of the unit circle, and also what it does to points on the unit circle. This map is called **inversion** through the unit circle.

2. Let $f(z) = 1/\bar{z}$. Describe f in the same manner as in Exercise 1. This map is called **reflection** through the unit circle.

3. Let $f(z) = e^{2\pi i z}$. Describe the image under f of the set shaded in Fig. 7, consisting of those points $x + iy$ with $-\frac{1}{2} \leqq x \leqq \frac{1}{2}$ and $y \geqq B$.

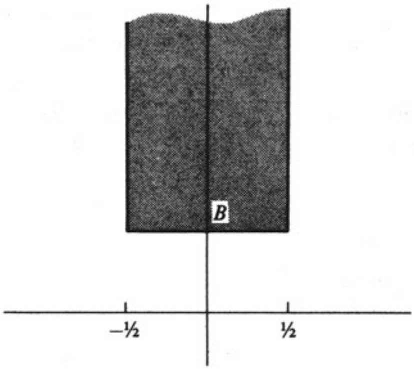


Figure 7

4. Let $f(z) = e^z$. Describe the image under f of the following sets:
   (a) The set of $z = x + iy$ such that $x \leqq 1$ and $0 \leqq y \leqq \pi$.
   (b) The set of $z = x + iy$ such that $0 \leqq y \leqq \pi$ (no condition on x).

I, §4. LIMITS AND COMPACT SETS

Let α be a complex number. By the **open disc** of radius $r > 0$ centered at α we mean the set of complex numbers z such that

$$|z - \alpha| < r.$$

For the **closed** disc, we use the condition $|z - \alpha| \leqq r$ instead. We shall deal only with the open disc unless otherwise specified, and thus speak simply of the **disc**, denoted by $D(\alpha, r)$. The closed disc is denoted by $\bar{D}(\alpha, r)$.

Let U be a subset of the complex plane. We say that U is **open** if for every point α in U there is a disc $D(\alpha, r)$ centered at α, and of some radius $r > 0$ such that this disc $D(\alpha, r)$ is contained in U. We have illustrated an open set in Fig. 8.

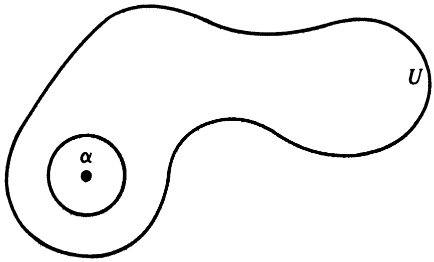


Figure 8

Note that the radius r of the disc depends on the point α. As α comes closer to the boundary of U, the radius of the disc will be smaller.

Examples of Open Sets. The first quadrant, consisting of all numbers $z = x + iy$ with $x > 0$ and $y > 0$ is open, and drawn on Fig. 9(a).

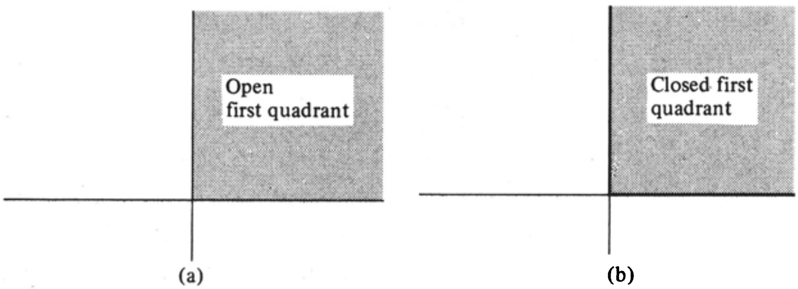


Figure 9

On the other hand, the set consisting of the first quadrant and the vertical and horizontal axes as on Fig. 9(b) is not open.

The **upper half plane** by definition is the set of complex numbers

$$z = x + iy$$

with $y > 0$. It is an open set.

Let S be a subset of the plane. A **boundary point** of S is a point α such that **every** disc $D(\alpha, r)$ centered at α and of radius $r > 0$ contains both points of S and points not in S. In the closed first quadrant of Fig. 9(b), the points on the x-axis with $x \geqq 0$ and on the y-axis with $y \geqq 0$ are boundary points of the quadrant.

A point α is said to be **adherent** to S if **every** disc $D(\alpha, r)$ with $r > 0$ contains **some** element of S. A point α is said to be an **interior** point of S if **there exists** a disc $D(\alpha, r)$ which is contained in S. Thus an adherent point can be a boundary point or an interior point of S. A set is called

closed if it contains all its boundary points. The complement of a closed set is then open.

The **closure** of a set S is defined to be the union of S and all its boundary points. We denote the closure by $\bar{S}$.

A set S is said to be **bounded** if there exists a number $C > 0$ such that

$$|z| \leqq C \qquad \text{for all } z \text{ in } S.$$

For instance, the set in Fig. 10 is bounded. The first quadrant is not bounded.

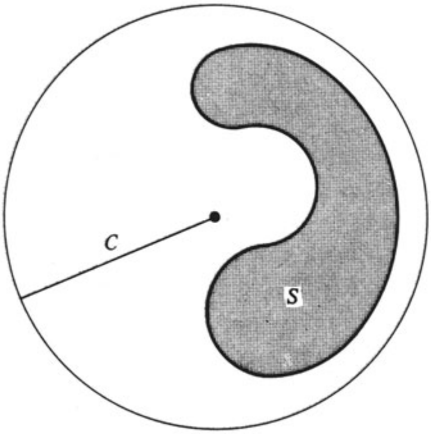


Figure 10

The upper half plane is not bounded. The condition for boundedness means that the set is contained in the closed disc of radius C, as shown on Fig. 10.

Let f be a function on S, and let α be an adherent point of S. Let w be a complex number. We say that

$$w = \lim_{\substack{z \to \alpha \\ z \in S}} f(z)$$

if the following condition is satisfied. Given $\epsilon > 0$ there exists $\delta > 0$ such that if $z \in S$ and $|z - \alpha| < \delta$, then

$$|f(z) - w| < \epsilon.$$

We usually omit the symbols $z \in S$ under the limit sign, and write merely

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

In some applications $\alpha \in S$ and in some applications, $\alpha \notin S$.

Let $\alpha \in S$. We say that f is **continuous** at α if

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

These definitions are completely analogous to those which you should have had in some analysis or advanced calculus course, so we don't spend much time on them. As usual, we have the rules for limits of sums, products, quotients as in calculus.

If $\{z_n\}$ $(n = 1, 2, \ldots)$ is a sequence of complex numbers, then we say that

$$w = \lim_{n \to \infty} z_n$$

if the following condition is satisfied:

Given $\epsilon > 0$ there exists an integer N such that if $n \geqq N$, then

$$|z_n - w| < \epsilon.$$

Let S be the set of fractions $1/n$, with $n = 1, 2, \ldots$. Let $f(1/n) = z_n$. Then

$$\lim_{n \to \infty} z_n = w \quad \text{if and only if} \quad \lim_{\substack{z \to 0 \\ z \in S}} f(z) = w.$$

Thus basic properties of limits for $n \to \infty$ are reduced to similar properties for functions. Note that in this case, the number 0 is not an element of S.

A sequence $\{z_n\}$ is said to be a **Cauchy sequence** if, given ϵ, there exists N such that if $m, n \geqq N$, then

$$|z_n - z_m| < \epsilon.$$

Write

$$z_n = x_n + iy_n.$$

Since

$$|z_n - z_m| = \sqrt{(x_n - x_m)^2 + (y_n - y_m)^2},$$

and

$$|x_n - x_m| \leqq |z_n - z_m|, \qquad |y_n - y_m| \leqq |z_n - z_m|,$$

we conclude that $\{z_n\}$ is Cauchy if and only if the sequences $\{x_n\}$ and $\{y_n\}$ of real and imaginary parts are also Cauchy. Since we know that real Cauchy sequences converge (i.e. have limits), we conclude that complex Cauchy sequences also converge.

We note that all the usual theorems about limits hold for complex numbers: Limits of sums, limits of products, limits of quotients, limits of composite functions. The proofs which you had in advanced calculus hold without change in the present context. It is then usually easy to compute limits.

Example. Find the limit

$$\lim_{n\to\infty} \frac{nz}{1+nz}$$

for any complex number z.

If $z = 0$, it is clear that the limit is 0. Suppose $z \neq 0$. Then the quotient whose limit we are supposed to find can be written

$$\frac{nz}{1+nz} = \frac{z}{\dfrac{1}{n} + z}.$$

But

$$\lim_{n\to\infty} \left(\frac{1}{n} + z\right) = z.$$

Hence the limit of the quotient is $z/z = 1$.

Compact Sets

We shall now go through the basic results concerning compact sets. Let S be a set of complex numbers. Let $\{z_n\}$ be a sequence in S. By a **point of accumulation of** $\{z_n\}$ we mean a complex number v such that given ϵ (always assumed > 0) there exist infinitely many integers n such that

$$|z_n - v| < \epsilon.$$

We may say that given an open set U containing v, there exist infinitely many n such that $z_n \in U$.

Similarly we define the notion of **point of accumulation** of an infinite set S. It is a complex number v such that given an open set U containing v, there exist infinitely many elements of S lying in U. In particular, a point of accumulation of S is adherent to S.

We assume that the reader is acquainted with the **Weierstrass–Bolzano theorem** about sets of real numbers: *If S is an infinite bounded set of real numbers, then S has a point of accumulation.*

We define a set of complex numbers S to be **compact** if every sequence of elements of S has a point of accumulation in S. This property is equivalent to the following properties, which could be taken as alternate definitions:

(a) Every infinite subset of S has a point of accumulation in S.

(b) Every sequence of elements of S has a convergent subsequence whose limit is in S.

We leave the proof of the equivalence between the three possible definitions to the reader.

Theorem 4.1. *A set of complex numbers is compact if and only if it is closed and bounded.*

Proof. Assume that S is compact. If S is not bounded, for each positive integer n there exists $z_n \in S$ such that

$$|z_n| > n.$$

Then the sequence $\{z_n\}$ does not have a point of accumulation. Indeed, if v is a point of accumulation, pick $m > 2|v|$, and note that $|v| > 0$. Then

$$|z_m - v| \geqq |z_m| - |v| \geqq m - |v| > |v|.$$

This contradicts the fact that for infinitely many m we must have z_m close to v. Hence S is bounded. To show S is closed, let v be in its closure. Given n, there exists $z_n \in S$ such that

$$|z_n - v| < 1/n.$$

The sequence $\{z_n\}$ converges to v, and has a subsequence converging to a limit in S because S is assumed compact. This limit must be v, whence $v \in S$ and S is closed.

Conversely, assume that S is closed and bounded, and let B be a bound, so $|z| \leqq B$ for all $z \in S$. If we write

$$z = x + iy,$$

then $|x| \leqq B$ and $|y| \leqq B$. Let $\{z_n\}$ be a sequence in S, and write

$$z_n = x_n + iy_n.$$

There is a subsequence $\{z_{n_1}\}$ such that $\{x_{n_1}\}$ converges to a real number a, and there is a sub-subsequence $\{z_{n_2}\}$ such that $(y_{n_2}\}$ converges to a real number b. Then

$$\{z_{n_2} = x_{n_2} + iy_{n_2}\}$$

converges to $a + ib$, and S is compact. This proves the theorem.

Theorem 4.2. *Let S be a compact set and let $S_1 \supset S_2 \supset \cdots$ be a sequence of non-empty closed subsets such that $S_n \supset S_{n+1}$. Then the intersection of all S_n for all $n = 1, 2, \ldots$ is not empty.*

Proof. Let $z_n \in S_n$. The sequence $\{z_n\}$ has a point of accumulation in S. Call it v. Then v is also a point of accumulation for each subsequence $\{z_k\}$ with $k \geqq n$, and hence lies in the closure of S_n for each n, But S_n is assumed closed, and hence $v \in S_n$ for all n. This proves the theorem.

Theorem 4.3. *Let S be a compact set of complex numbers, and let f be a continuous function on S. Then the image of f is compact.*

Proof. Let $\{w_n\}$ be a sequence in the image of f, so that

$$w_n = f(z_n) \qquad \text{for} \qquad z_n \in S.$$

The sequence $\{z_n\}$ has a convergent subsequence $\{z_{n_k}\}$, with a limit v in S. Since f is continuous, we have

$$\lim_{k\to\infty} w_{n_k} = \lim_{k\to\infty} f(z_{n_k}) = f(v).$$

Hence the given sequence $\{w_n\}$ has a subsequence which converges in $f(S)$. This proves that $f(S)$ is compact.

Theorem 4.4. *Let S be a compact set of complex numbers, and let*

$$f: S \to \mathbf{R}$$

be a continuous function. Then f has a maximum on S, that is, there exists $v \in S$ such that $f(z) \leqq f(v)$ for all $z \in S$.

Proof. By Theorem 4.3, we know that $f(S)$ is closed and bounded. Let b be its least upper bound. Then b is adherent to $f(S)$, whence in $f(S)$ because $f(S)$ is closed. So there is some $v \in S$ such that $f(v) = b$. This proves the theorem.

Remarks. In practice, one deals with a continuous function $f: S \to \mathbf{C}$ and one applies Theorem 4.4 to the absolute value of f, which is also continuous (composite of two continuous functions).

Theorem 4.5. *Let S be a compact set, and let f be a continuous function on S. Then f is uniformly continuous, i.e. given ϵ there exists δ such that whenever z, $w \in S$ and $|z - w| < \delta$, then $|f(z) - f(w)| < \epsilon$.*

Proof. Suppose the assertion of the theorem is false. Then there exists ϵ, and for each n there exists a pair of elements z_n, $w_n \in S$ such that

$$|z_n - w_n| < 1/n \qquad \text{but} \qquad |f(z_n) - f(w_n)| > \epsilon.$$

There is an infinite subset J_1 of positive integers and some $v \in S$ such that $z_n \to v$ for $n \to \infty$ and $n \in J_1$. There is an infinite subset J_2 of J_1 and $u \in S$ such that $w_n \to u$ for $n \to \infty$ and $n \in J_2$. Then, taking the limit for $n \to \infty$ and $n \in J_2$ we obtain $|u - v| = 0$ and $u = v$ because

$$|v - u| \leqq |v - z_n| + |z_n - w_n| + |w_n - u|.$$

Hence $f(v) - f(u) = 0$. Furthermore,

$$|f(z_n) - f(w_n)| \leqq |f(z_n) - f(v)| + |f(v) - f(u)| + |f(u) - f(w_n)|.$$

Again taking the limit as $n \to \infty$ and $n \in J_2$, we conclude that

$$f(z_n) - f(w_n)$$

approaches 0. This contradicts the assumption that

$$|f(z_n) - f(w_n)| > \epsilon,$$

and proves the theorem.

Let A, B be two sets of complex numbers. By the **distance** between them, denoted by $d(A, B)$, we mean

$$d(A, B) = \text{g.l.b.}|z - w|,$$

where the greatest lower bound g.l.b. is taken over all elements $z \in A$ and $w \in B$. If B consists of one point, we also write $d(A, w)$ instead of $d(A, B)$.

We shall leave the next two results as easy exercises.

Theorem 4.6. *Let S be a closed set of complex numbers, and let v be a complex number. There exists a point $w \in S$ such that*

$$d(S, v) = |w - v|.$$

[*Hint*: Let E be a closed disc of some suitable radius, centered at v, and consider the function $z \mapsto |z - v|$ for $z \in S \cap E$.]

Theorem 4.7. *Let K be a compact set of complex numbers, and let S be a closed set. There exist elements $z_0 \in K$ and $w_0 \in S$ such that*

$$d(K, S) = |z_0 - w_0|.$$

[*Hint*: Consider the function $z \mapsto d(S, z)$ for $z \in K$.]

Theorem 4.8. *Let S be compact. Let r be a real number > 0. There exists a finite number of open discs of radius r whose union contains S.*

Proof. Suppose this is false. Let $z_1 \in S$ and let D_1 be the open disc of radius r centered at z_1. Then D_1 does not contain S, and there is some $z_2 \in S$, $z_2 \neq z_1$. Proceeding inductively, suppose we have found open discs $D_1, \ldots, D_n$ of radius r centered at points $z_1, \ldots, z_n$, respectively, such that z_{k+1} does not lie in $D_1 \cup \cdots \cup D_k$. We can then find z_{n+1} which does not lie in $D_1 \cup \cdots \cup D_n$, and we let D_{n+1} be the disc of radius r centered at z_{n+1}. Let v be a point of accumulation of the sequence $\{z_n\}$. By definition, there exist positive integers m, k with $k > m$ such that

$$|z_k - v| < r/2 \qquad \text{and} \qquad |z_m - v| < r/2.$$

Then $|z_k - z_m| < r$ and this contradicts the property of our sequence $\{z_n\}$ because z_k lies in the disc D_m. This proves the theorem.

Let S be a set of complex numbers, and let I be some set. Suppose that for each $i \in I$ we are given an open set U_i. We denote this association by $\{U_i\}_{i \in I}$, and call it a **family of open sets**. The **union** of the family is the set U consisting of all z such that $z \in U_i$ for some $i \in I$. We say that the family **covers** S if S is contained in this union, that is, every $z \in S$ is contained in some U_i. We then say that the family $\{U_i\}_{i \in I}$ is an **open covering** of S. If J is a subset of I, we call the family $\{U_j\}_{j \in J}$ a **subfamily**, and if it covers S also, we call it a **subcovering** of S. In particular, if

$$U_{i_1}, \ldots, U_{i_n}$$

is a finite number of the open sets U_i, we say that it is a **finite subcovering** of S if S is contained in the finite union

$$U_{i_1} \cup \cdots \cup U_{i_n}.$$

Theorem 4.9. *Let S be a compact set, and let $\{U_i\}_{i \in I}$ be an open covering of S. Then there exists a finite subcovering, that is, a finite number of open sets $U_{i_1}, \ldots, U_{i_n}$ whose union covers S.*

Proof. By Theorem 4.8, for each n there exists a finite number of open discs of radius $1/n$ which cover S. Suppose that there is no finite subcovering of S by open sets U_i. Then for each n there exists one of the open discs D_n from the preceding finite number such that $D_n \cap S$ is not covered by any finite number of open sets U_i. Let $z_n \in D_n \cap S$, and let w be a point of accumulation of the sequence $\{z_n\}$. For some index i_0 we have $w \in U_{i_0}$. By definition, U_{i_0} contains an open disc D of radius $r > 0$ centered at w. Let N be so large that $2/N < r$. There exists $n > N$ such that

$$|z_n - w| \leqq 1/N.$$

Any point of D_n is then at a distance $\leqq 2/N$ from w, and hence D_n is contained in D, and thus contained in U_{i_0}. This contradicts the hypothesis made on D_n, and proves the theorem.

I, §4. EXERCISES

1. Let α be a complex number of absolute value <1. What is $\lim_{n\to\infty} \alpha^n$? Proof?
2. If $|\alpha| > 1$, does $\lim_{n\to\infty} \alpha^n$ exist? Why?
3. Show that for any complex number $z \neq 1$, we have

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

If $|z| < 1$, show that

$$\lim_{n\to\infty} (1 + z + \cdots + z^n) = \frac{1}{1 - z}.$$

4. Let f be the function defined by

$$f(z) = \lim_{n\to\infty} \frac{1}{1 + n^2 z}.$$

Show that f is the characteristic function of the set $\{0\}$, that is, $f(0) = 1$, and $f(z) = 0$ if $z \neq 0$.

5. For $|z| \neq 1$ show that the following limit exists:

$$f(z) = \lim_{n\to\infty} \left(\frac{z^n - 1}{z^n + 1}\right).$$

Is it possible to define $f(z)$ when $|z| = 1$ in such a way to make f continuous?

6. Let

$$f(z) = \lim_{n\to\infty} \frac{z^n}{1 + z^n}.$$

(a) What is the domain of definition of f, that is, for which complex numbers z does the limit exist?
(b) Give explicitly the values of $f(z)$ for the various z in the domain of f.

7. Show that the series

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

converges to $1/(1-z)^2$ for $|z| < 1$ and to $1/z(1-z)^2$ for $|z| > 1$. Prove that the convergence is uniform for $|z| \leqq c < 1$ in the first case, and $|z| \geqq b > 1$ in the second. [*Hint*: Multiply and divide each term by $1 - z$, and do a partial fraction decomposition, getting a telescoping effect.]

I, §5. COMPLEX DIFFERENTIABILITY

In studying differentiable functions of a real variable, we took such functions defined on intervals. For complex variables, we have to select domains of definition in an analogous manner.

Let U be an open set, and let z be a point of U. Let f be a function on U. We say that f is **complex differentiable** at z if the limit

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

exists. This limit is denoted by $f'(z)$ or df/dz.

In this section, differentiable will always mean complex differentiable.

The usual proofs of a first course in calculus concerning basic properties of differentiability are valid for complex differentiability. We shall run through them again.

We note that if f is differentiable at z then f is continuous at z because

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

and since the limit of a product is the product of the limits, the limit on the right-hand side is equal to 0.

We let f, g be functions defined on the open set U. We assume that f, g are differentiable at z.

Sum. *The sum $f + g$ is differentiable at z, and*

$$(f+g)'(z) = f'(z) + g'(z).$$

Proof. This is immediate from the theorem that the limit of a sum is the sum of the limits.

Product. *The product fg is differentiable at z, and*

$$(fg)'(z) = f'(z)g(z) + f(z)g'(z).$$

Proof. To determine the limit of the Newton quotient

$$\frac{f(z+h)g(z+h)-f(z)g(z)}{h}$$

we write the numerator in the form

$$f(z+h)g(z+h)-f(z)g(z+h)+f(z)g(z+h)-f(z)g(z).$$

Then the Newton quotient is equal to a sum

$$\frac{f(z+h)-f(z)}{h}g(z+h)+f(z)\frac{g(z+h)-g(z)}{h}.$$

Taking the limits yields the formula.

Quotient. *If* $g(z) \neq 0$, *then the quotient* f/g *is differentiable at* z, *and*

$$(f/g)'(z)=\frac{g(z)f'(z)-f(z)g'(z)}{g(z)^2}.$$

Proof. This is again proved as in ordinary calculus. We first prove the differentiability of the quotient function $1/g$. We have

$$\frac{\dfrac{1}{g(z+h)}-\dfrac{1}{g(z)}}{h}=-\frac{g(z+h)-g(z)}{h}\frac{1}{g(z+h)g(z)}$$

Taking the limit yields

$$-\frac{1}{g(z)^2}g'(z),$$

which is the usual value. The general formula for a quotient is obtained from this by writing

$$f/g=f\cdot 1/g,$$

and using the rules for the derivative of a product, and the derivative of $1/g$.

Examples. As in ordinary calculus, from the formula for a product and induction, we see that for any positive integer n,

$$\frac{dz^n}{dz}=nz^{n-1}.$$

The rule for a quotient also shows that this formula remains valid when n is a negative integer.

The derivative of $z^2/(2z-1)$ is

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

This formula is valid for any complex number z such that $2z - 1 \neq 0$. More generally, let

$$f(z) = P(z)/Q(z),$$

where P, Q are polynomials. Then f is differentiable at any point z where $Q(z) \neq 0$.

Last comes the chain rule. Let U, V be open sets in $\mathbf{C}$, and let

$$f\colon U \to V \qquad \text{and} \qquad g\colon V \to \mathbf{C}$$

be functions, such that the image of f is contained in V. Then we can form the composite function $g \circ f$ such that

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

Chain Rule. *Let $w = f(z)$. Assume that f is differentiable at z, and g is differentiable at w. Then $g \circ f$ is differentiable at z, and*

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

Proof. Again the proof is the same as in calculus, and depends on expressing differentiability by an equivalent property not involving denominators, as follows.

Suppose that f is differentiable at z, and let

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

Then

$$f(z+h) - f(z) = f'(z)h + h\varphi(h), \tag{1}$$

and

$$\lim_{h \to 0} \varphi(h) = 0. \tag{2}$$

Furthermore, even though φ is at first defined only for sufficiently small h and $h \neq 0$, we may also define $\varphi(0) = 0$, and formula (1) remains valid for $h = 0$.

Conversely, suppose that there exists a function φ defined for sufficiently small h and a number a such that

$$f(z+h) - f(z) = ah + h\varphi(h) \tag{1'}$$

and

$$\lim_{h \to 0} \varphi(h) = 0. \tag{2}$$

Dividing by h in formula (1′) and taking the limit as $h \to 0$, we see that the limit exists and is equal to a. Thus $f'(z)$ exists and is equal to a. Hence the existence of a function φ satisfying (1′), (2) is equivalent to differentiability.

We apply this to a proof of the chain rule. Let $w = f(z)$, and

$$k = f(z+h) - f(z),$$

so that

$$g\big(f(z+h)\big) - g\big(f(z)\big) = g(w+k) - g(w).$$

There exists a function $\psi(k)$ such that $\lim_{k \to 0} \psi(k) = 0$ and

$$\begin{aligned} g(w+k) - g(w) &= g'(w)k + k\psi(k) \\ &= g'(w)\big(f(z+h) - f(z)\big) + \big(f(z+h) - f(z)\big)\psi(k). \end{aligned}$$

Dividing by h yields

$$\frac{g \circ f(z+h) - g \circ f(z)}{h} = g'(w)\frac{f(z+h) - f(z)}{h} + \frac{f(z+h) - f(z)}{h}\psi(k).$$

As $h \to 0$, we note that $k \to 0$ also by the continuity of f, whence $\psi(k) \to 0$ by assumption. Taking the limit of this last expression as $h \to 0$ proves the chain rule.

A function f defined on an open set U is said to be **differentiable** if it is differentiable at every point. We then also say that f is **holomorphic** on U. The word holomorphic is usually used in order not to have to specify *complex* differentiability as distinguished from real differentiability.

In line with general terminology, a holomorphic function

$$f\colon U \to V$$

from an open set into another is called a **holomorphic isomorphism** if there exists a holomorphic function

$$g\colon V \to U$$

such that g is the inverse of f, that is,

$$g \circ f = \mathrm{id}_U \qquad \text{and} \qquad f \circ g = \mathrm{id}_V.$$

A holomorphic isomorphism of U with itself is called a holomorphic **automorphism**. In the next chapter we discuss this notion in connection with functions defined by power series.

I, §6. THE CAUCHY–RIEMANN EQUATIONS

Let f be a function on an open set U, and write f in terms of its real and imaginary parts,

$$f(x + iy) = u(x, y) + iv(x, y).$$

It is reasonable to ask what the condition of differentiability means in terms of u and v. We shall analyze this situation in detail in Chapter VIII, but both for the sake of tradition, and because there is some need psychologically to see right away what the answer is, we derive the equivalent conditions on u, v for f to be holomorphic.

At a fixed z, let $f'(z) = a + bi$. Let $w = h + ik$, with h, k real. Suppose that

$$f(z + w) - f(z) = f'(z)w + \sigma(w)w,$$

where

$$\lim_{w \to 0} \sigma(w) = 0.$$

Then

$$f'(z)w = (a + bi)(h + ki) = ah - bk + i(bh + ak).$$

On the other hand, let

$$F\colon U \to \mathbf{R}^2$$

be the map (often called vector field) such that

$$F(x, y) = \big(u(x, y), v(x, y)\big).$$

We call F the (real) **vector field associated with** f. Then

$$F(x + h, y + k) - F(x, y) = (ah - bk, bh + ak) + \sigma_1(h, k)h + \sigma_2(h, k)k,$$

where $\sigma_1(h, k)$, $\sigma_2(h, k)$ are functions tending to 0 as (h, k) tends to 0. Hence if we assume that f is holomorphic, we conclude that F is differentiable in the sense of real variables, and that its derivative is represented by the (Jacobian) matrix

$$J_F(x, y) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix}.$$

This shows that

$$\boxed{f'(z) = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}}$$

and

$$\boxed{\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.}$$

These are called the **Cauchy–Riemann equations.**

Conversely, let $u(x, y)$ and $v(x, y)$ be two functions satisfying the Cauchy–Riemann equations, and continuously differentiable in the sense of real functions. Define

$$f(z) = f(x + iy) = u(x, y) + iv(x, y).$$

Then it is immediately verified by reversing the above steps that f is complex-differentiable, i.e. holomorphic.

The Jacobian determinant Δ_F of the associated vector field F is

$$\Delta_F = a^2 + b^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2.$$

Hence $\Delta_F \geqq 0$, and is $\neq 0$ if and only if $f'(z) \neq 0$. We have

$$\boxed{\Delta_F(x, y) = |f'(z)|^2.}$$

We now drop these considerations until Chapter VIII.

The study of the real part of a holomorphic function and its relation with the function itself will be carried out more substantially in Chapter VIII. It is important, and much of that chapter depends only on elementary facts. However, the most important part of complex analysis at the present level lies in the power series aspects and the immediate applications of Cauchy's theorem. The real part plays no role in these matters. Thus we do not wish to interrupt the straightforward flow of the book now towards these topics.

However, the reader may read immediately the more elementary parts §1 and §2 of Chapter VIII, which can be understood already at this point.

I, §6. EXERCISE

1. Prove in detail that if u, v satisfy the Cauchy–Riemann equations, then the function
$$f(z) = f(x + iy) = u(x, y) + iv(x, y)$$
is holomorphic.

I, §7. ANGLES UNDER HOLOMORPHIC MAPS

In this section, we give a simple geometric property of holomorphic maps. Roughly speaking, they preserve angles. We make this precise as follows.

Let U be an open set in C and let

$$\gamma\colon [a, b] \to U$$

be a curve in U, so we write

$$\gamma(t) = x(t) + iy(t).$$

We assume that γ is differentiable, so its derivative is given by

$$\gamma'(t) = x'(t) + iy'(t).$$

Let $f: U \to \mathbf{C}$ be holomorphic. We let the reader verify the chain rule

$$\frac{d}{dt} f(\gamma(t)) = f'(\gamma(t))\gamma'(t).$$

We interpret $\gamma'(t)$ as a vector in the direction of a tangent vector at the point $\gamma(t)$. This derivative $\gamma'(t)$, if not 0, defines the direction of the curve at the point.

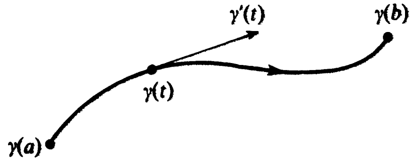


Figure 11

Consider two curves γ and η passing through the same point z_0. Say

$$z_0 = \gamma(t_0) = \eta(t_1).$$

Then the tangent vectors $\gamma'(t_0)$ and $\eta'(t_1)$ determine an angle θ which is defined to be the **angle between the curves.**

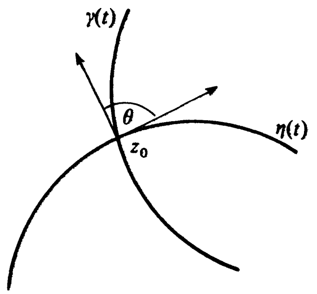


Figure 12

Applying f, the curves $f \circ \gamma$ and $f \circ \eta$ pass through the point $f(z_0)$, and by the chain rule, tangent vectors of these image curves are

$$f'(z_0)\gamma'(t_0) \qquad \text{and} \qquad f'(z_0)\eta'(t_1).$$

Theorem 7.1. *If $f'(z_0) \neq 0$ then the angle between the curves γ, η at z_0 is the same as the angle between the curves $f \circ \gamma$, $f \circ \eta$ at $f(z_0)$.*

Proof. Geometrically speaking, the tangent vectors under f are changed by multiplication with $f'(z_0)$, which can be represented in polar coordinates as a dilation and a rotation, so preserves the angles.

We shall now give a more formal argument, dealing with the cosine and sine of angles.

Let z, w be complex numbers,

$$z = a + bi \qquad \text{and} \qquad w = c + di,$$

where a, b, c, d are real. Then

$$z\bar{w} = ac + bd + i(bc - ad).$$

Define the **scalar product**

$$\langle z, w\rangle = \mathrm{Re}(z\bar{w}). \tag{1}$$

Then $\langle z, w\rangle$ is the ordinary scalar product of the vectors (a, b) and (c, d) in $\mathbf{R}^2$. Let $\theta(z, w)$ be the angle between z and w. Then

$$\cos\theta(z, w) = \frac{\langle z, w\rangle}{|z||w|}. \tag{2}$$

Since $\sin\theta = \cos\left(\theta - \frac{\pi}{2}\right)$, we can define

$$\sin\theta(z, w) = \frac{\langle z, -iw\rangle}{|z||w|}. \tag{3}$$

This gives us the desired precise formulas for the cosine and sine of an angle, which determine the angle.

Let $f'(z_0) = \alpha$. Then

$$\langle \alpha z, \alpha w\rangle = \mathrm{Re}(\alpha z\bar{\alpha}\bar{w}) = \alpha\bar{\alpha}\,\mathrm{Re}(z\bar{w}) = |\alpha|^2\langle z, w\rangle \tag{4}$$

because $\alpha\bar{\alpha} = |\alpha|^2$ is real. It follows immediately from the above formulas that

$$\cos\theta(\alpha z, \alpha w) = \cos\theta(z, w) \qquad \text{and} \qquad \sin\theta(\alpha z, \alpha w) = \sin\theta(z, w). \tag{5}$$

This proves the theorem.

A map which preserves angles is called **conformal**. Thus we can say that a holomorphic map with non-zero derivative is conformal. The complex conjugate of a holomorphic map also preserves angles, if we disregard the orientation of an angle.

In Chapter VII, we shall consider holomorphic maps which have inverse holomorphic maps, and therefore such that their derivatives are never equal to 0. The theorem proved in this section gives additional geometric information concerning the nature of such maps. But the emphasis of the theorem in this section is local, whereas the emphasis in Chapter VII will be global. The word "conformal", however, has become a code word for this kind of map, even in the global case, which explains the title of Chapter VII. The reader will notice that the local property of preserving angles is irrelevant for the global arguments given in Chapter VII, having to do with inverse mappings. Thus in Chapter VII, we shall use a terminology which emphasizes the invertibility, namely the terminology of isomorphisms and automorphisms.

In this terminology, we can say that a holomorphic isomorphism is conformal. The converse is false in general. For instance, let U be the open set obtained by deleting the origin from the complex numbers. The function

$$f\colon U \to U \qquad \text{given by} \qquad z \mapsto z^2$$

has everywhere non-zero derivative in U, but it does not admit an inverse function. This function f is definitely conformal. The invertibility is true locally, however. See Theorem 5.1 of Chapter II.

CHAPTER II

Power Series

So far, we have given only rational functions as examples of holomorphic functions. We shall study other ways of defining such functions. One of the principal ways will be by means of power series. Thus we shall see that the series

$$1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots$$

converges for all z to define a function which is equal to e^z. Similarly, we shall extend the values of $\sin z$ and $\cos z$ by their usual series to complex valued functions of a complex variable, and we shall see that they have similar properties to the functions of a real variable which you already know.

First we shall learn to manipulate power series formally. In elementary calculus courses, we derived Taylor's formula with an error term. Here we are concerned with the full power series. In a way, we pick up where calculus left off. We study systematically sums, products, inverses, and composition of power series, and then relate the formal operations with questions of convergence.

II, §1. FORMAL POWER SERIES

We select at first a neutral letter, say T. In writing a formal power series

$$\sum_{n=0}^{\infty} a_n T^n = a_0 + a_1 T + a_2 T^2 + \cdots$$

what is essential are its "coefficients" $a_0, a_1, a_2, \ldots$ which we shall take as complex numbers. Thus the above series may be defined as the function

$$n \mapsto a_n$$

from the integers $\geqq 0$ to the complex numbers.

We could use other letters besides T, writing

$$f(T) = \sum a_n T^n,$$
$$f(r) = \sum a_n r^n,$$
$$f(z) = \sum a_n z^n.$$

In such notation, f does not denote a function, but a formal expression.

Also, as a matter of notation, we write a single term

$$a_n T^n$$

to denote the power series such that $a_k = 0$ if $k \neq n$. For instance, we would write

$$5T^3$$

for the power series

$$0 + 0T + 0T^2 + 5T^3 + 0T^4 + \cdots,$$

such that $a_3 = 5$ and $a_k = 0$ if $k \neq 3$.

By definition, we call a_0 the **constant term** of f.

If

$$f = \sum a_n T^n \qquad \text{and} \qquad g = \sum b_n T^n$$

are such formal power series, we define their **sum** to be

$$f + g = \sum c_n T^n, \qquad \text{where} \quad c_n = a_n + b_n.$$

We define their **product** to be

$$fg = \sum d_n T^n, \qquad \text{where} \quad d_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

The sum and product are therefore defined just as for polynomials. The first few terms of the product can be written as

$$fg = a_0 b_0 + (a_0 b_1 + a_1 b_0)T + (a_0 b_2 + a_1 b_1 + a_2 b_0)T^2 + \cdots.$$

If α is a complex number, we define

$$\alpha f = \sum (\alpha a_n) T^n$$

to be the power series whose n-th coefficient is αa_n. Thus we can multiply power series by numbers.

Just as for polynomials, one verifies that the sum and product are associative, commutative, and distributive. Thus in particular, if f, g, h are power series, then

$$f(g + h) = fg + fh \quad \text{(distributivity).}$$

We omit the proof, which is just elementary algebra.

The **zero power series** is the series such that $a_n = 0$ for all integers $n \geqq 0$.

Suppose a power series is of the form

$$f = a_r T^r + a_{r+1} T^{r+1} + \cdots,$$

and $a_r \neq 0$. Thus r is the smallest integer n such that $a_n \neq 0$. Then we call r the **order** of f, and write

$$r = \operatorname{ord} f.$$

If $\operatorname{ord} g = s$, so that

$$g = b_s T^s + b_{s+1} T^{s+1} + \cdots,$$

and $b_s \neq 0$, then by definition,

$$fg = a_r b_s T^{r+s} + \text{higher terms,}$$

and $a_r b_s \neq 0$. Hence

$$\boxed{\operatorname{ord} fg = \operatorname{ord} f + \operatorname{ord} g.}$$

A power series has order 0 if and only if it starts with a non-zero constant term. For instance, the geometric series

$$1 + T + T^2 + T^3 + \cdots$$

has order 0.

Let $f = \sum a_n T^n$ be a power series. We say that $g = \sum b_n T^n$ is an

inverse for f if

$$fg = 1.$$

In view of the relation for orders which we just mentioned, we note that if an inverse exists, then we must have

$$\operatorname{ord} f = \operatorname{ord} g = 0.$$

In other words, both f and g start with non-zero constant terms. The converse is true:

If f has a non-zero constant term, then f has an inverse as a power series.

Proof. Considering $a_0^{-1}f$ instead of f, we are reduced to the case when the constant term is equal to 1. We first note that the old geometric series gives us a formal inverse,

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

Written multiplicatively, this amounts to

$$\begin{aligned}(1-r)(1 + r + r^2 + \cdots) &= 1 + r + r^2 + \cdots - r(1 + r + r^2 + \cdots) \\ &= 1 + r + r^2 + \cdots - r - r^2 - \cdots \\ &= 1.\end{aligned}$$

Next, write

$$f = 1 - h, \qquad \text{where} \quad h = -(a_1 T + a_2 T^2 + \cdots).$$

To find the inverse $(1 - h)^{-1}$ is now easy, namely it is the power series

$$(*) \qquad \varphi = 1 + h + h^2 + h^3 + \cdots.$$

We have to verify that this makes sense. Any finite sum

$$1 + h + h^2 + \cdots + h^m$$

makes sense because we have defined sums and products of power series. Observe that the order of h^n is at least n, because h^n is of the form

$$(-1)^n a_1^n T^n + \text{higher terms.}$$

Thus in the above sum (*), if $m > n$, then the term h^m has all coefficients of order $\leqq n$ equal to 0. Thus we may define the n-th coefficient of φ to be the n-th coefficient of the finite sum

$$1 + h + h^2 + \cdots + h^n.$$

It is then easy to verify that

$$(1 - h)\varphi = (1 - h)(1 + h + h^2 + h^3 + \cdots)$$

is equal to

$$1 + \text{a power series of arbitrarily high order,}$$

and consequently is equal to 1. Hence we have found the desired inverse for f.

Example. Let

$$\cos T = 1 - \frac{T^2}{2!} + \frac{T^4}{4!} - \cdots$$

be the formal power series whose coefficients are the same as for the Taylor expansion of the ordinary cosine function in elementary calculus. We want to write down the first few terms of its (formal) inverse,

$$\frac{1}{\cos T}.$$

Up to terms of order 4, these will be the same as

$$\begin{aligned}\frac{1}{1 - \left(\frac{T^2}{2!} - \frac{T^4}{4!} + \cdots\right)} &= 1 + \left(\frac{T^2}{2!} - \frac{T^4}{4!} + \cdots\right) \\ &\quad + \left(\frac{T^2}{2!} - \frac{T^4}{4!} + \cdots\right)^2 + \cdots \\ &= 1 + \frac{T^2}{2!} - \frac{T^4}{4!} + \cdots + \frac{T^4}{(2!)^2} + \cdots \\ &= 1 + \frac{1}{2}T^2 + \left(\frac{-1}{24} + \frac{1}{4}\right)T^4 + \text{higher terms.}\end{aligned}$$

This gives us the coefficients of $1/\cos T$ up to order 4.

The substitution of h in the geometric series used to find an inverse can be generalized. Let

$$f = \sum a_n T^n$$

be a power series, and let

$$h(T) = c_1 T + \cdots$$

be a power series *whose constant term is* 0, so $\operatorname{ord} h \geqq 1$. Then we may "**substitute**" h in f to define a power series $f \circ h$ or $f(h)$, by

$$(f \circ h)(T) = f(h(T)) = f \circ h = a_0 + a_1 h + a_2 h^2 + a_3 h^3 + \cdots$$

in a natural way. Indeed, the finite sums

$$a_0 + a_1 h + \cdots + a_n h^n$$

are defined by the ordinary sum and product of power series. If $m > n$, then $a_m h^m$ has order $> n$; in other words, it is a power series starting with non-zero terms of order $> n$. Consequently we can define the power series $f \circ h$ as that series whose n-th coefficient is the n-th coefficient of

$$a_0 + a_1 h + \cdots + a_n h^n.$$

This composition of power series, like addition and multiplication, can therefore be computed by working only with polynomials. In fact, it is useful to discuss this approximation by polynomials a little more systematically.

We say that two power series $f = \sum a_n T^n$ and $g = \sum b_n T^n$ are **congruent mod** T^N and write $f \equiv g \pmod{T^N}$ if

$$a_n = b_n \qquad \text{for} \quad n = 0, \ldots, N - 1.$$

This means that the terms of order $\leqq N - 1$ coincide for the two power series. Given the power series f, there is a unique polynomial $P(T)$ of degree $\leqq N - 1$ such that

$$f(T) \equiv P(T) \pmod{T^N},$$

namely the polynomial

$$P(T) = a_0 + a_1 T + \cdots + a_{N-1} T^{N-1}.$$

If $f_1 \equiv f_2$ *and* $g_1 \equiv g_2 \pmod{T^N}$, *then*

$$f_1 + g_1 \equiv f_2 + g_2 \qquad \textit{and} \qquad f_1 g_1 \equiv f_2 g_2 \pmod{T^N}.$$

If h_1, h_2 *are power series with zero constant term, and*

$$h_1 \equiv h_2 \pmod{T^N},$$

then

$$f_1(h_1(T)) \equiv f_2(h_2(T)) \pmod{T^N}.$$

Proof. We leave the sum and product to the reader. Let us look at the proof for the composition $f_1 \circ h_1$ and $f_2 \circ h_2$. First suppose h has zero constant term. Let P_1, P_2 be the polynomials of degree $N-1$ such that

$$f_1 \equiv P_1 \qquad \text{and} \qquad f_2 \equiv P_2 \pmod{T^N}.$$

Then by hypothesis, $P_1 = P_2 = P$ is the same polynomial, and

$$f_1(h) \equiv P_1(h) = P_2(h) \equiv f_2(h) \pmod{T^N}.$$

Next let Q be the polynomial of degree $N-1$ such that

$$h_1(T) \equiv h_2(T) \equiv Q(T) \pmod{T^N}.$$

Write $P = a_0 + a_1 T + \cdots + a_{N-1} T^{N-1}$. Then

$$\begin{aligned} P(h_1) &= a_0 + a_1 h_1 + \cdots + a_{N-1} h_1^{N-1} \\ &\equiv a_0 + a_1 Q + \cdots + a_{N-1} Q^{N-1} \\ &\equiv a_0 + a_1 h_2 + \cdots + a_{N-1} h_2^{N-1} \\ &\equiv P(h_2) \pmod{T^N}. \end{aligned}$$

This proves the desired property, that $f_1 \circ h_1 \equiv f_2 \circ h_2 \pmod{T^N}$.

With these rules we can compute the coefficients of various operations between power series by reducing the computations to polynomial operations, which amount to high-school algebra. Indeed, two power series f, g are equal if and only if

$$f \equiv g \pmod{T^N}$$

for every positive integer N. Verifying that $f \equiv g \pmod{T^N}$ can be done by working entirely with polynomials of degree $< N$.

If f_1, f_2 are power series, then

$$(f_1 + f_2)(h) = f_1(h) + f_2(h),$$
$$(f_1 f_2)(h) = f_1(h) f_2(h), \quad \textit{and} \quad (f_1/f_2)(h) = f_1(h)/f_2(h)$$

if ord $f_2 = 0$. *If g, h have constant terms equal to* 0, *then*

$$f(g(h)) = (f \circ g)(h).$$

Proof. In each case, the proof is obtained by reducing the statement to the polynomial case, and seeing that the required properties hold for polynomials, which is standard. For instance, for the associativity of composition, given a positive integer N, let P, Q, R be polynomials of degree $\leqq N - 1$ such that

$$f \equiv P, \qquad g \equiv Q, \qquad h \equiv R \pmod{T^N}.$$

The ordinary theory of polynomials shows that

$$P(Q(R)) = (P \circ Q)(R).$$

The left-hand side is congruent to $f(g(h))$, and the right-hand side is congruent to $(f \circ g)(h) \pmod{T^N}$ by the properties which have already been proved. Hence

$$f(g(h)) \equiv (f \circ g)(h) \pmod{T^N}.$$

This is true for each N, whence $f(g(h)) = (f \circ g)(h)$, as desired.

In applications it is useful to consider power series which have a finite number of terms involving $1/z$, and this amounts also to considering arbitrary quotients of power series as follows.

Just as fractions m/n are formed with integers m, n and $n \neq 0$, we can form quotients

$$f/g = f(T)/g(T)$$

of power series such that $g \neq 0$. Two such quotients f/g and f_1/g_1 are regarded as equal if and only if

$$fg_1 = f_1 g,$$

which is exactly the condition under which we regard two rational numbers m/n and m_1/n_1 as equal. We have then defined for power series all the operations of arithmetic.

Let

$$f(T) = a_m T^m + a_{m+1} T^{m+1} + \cdots = \sum_{n \geqq m} a_n T^n$$

be a power series with $a_m \neq 0$. We may then write f in the form

$$f = a_m T^m (1 + h(T)),$$

where $h(T)$ has zero constant term. Consequently $1/f$ has the form

$$1/f = \frac{1}{a_m T^m} \frac{1}{1 + h(T)}.$$

We know how to invert $1 + h(T)$, say

$$(1 + h(T))^{-1} = 1 + b_1 T + b_2 T^2 + \cdots.$$

Then $1/f(T)$ has the shape

$$1/f = a_m^{-1} \frac{1}{T^m} + a_m^{-1} b_1 \frac{1}{T^{m-1}} + \cdots.$$

It is a power series with a finite number of terms having negative powers of T. In this manner, one sees that an arbitrary quotient can always be expressed as a power series of the form

$$\begin{aligned} f/g &= \frac{c_{-m}}{T^m} + \frac{c_{-m+1}}{T^{m-1}} + \cdots + c_0 + c_1 T + c_2 T^2 + \cdots \\ &= \sum_{n \geqq -m} c_n T^n. \end{aligned}$$

If $c_{-m} \neq 0$, then we call $-m$ the **order** of f/g. It is again verified as for power series without negative terms that if

$$\varphi = f/g \qquad \text{and} \qquad \varphi_1 = f_1/g_1,$$

then

$$\operatorname{ord} \varphi\varphi_1 = \operatorname{ord} \varphi + \operatorname{ord} \varphi_1.$$

Example. Find the terms of order $\leqq 3$ in the power series for $1/\sin T$. By definition,

$$\begin{aligned} \sin T &= T - T^3/3! + T^5/5! - \cdots \\ &= T(1 - T^2/3! + T^4/5! - \cdots). \end{aligned}$$

Hence

$$\frac{1}{\sin T} = \frac{1}{T}\frac{1}{1 - T^2/3! + T^4/5! + \cdots}$$

$$= \frac{1}{T}(1 + T^2/3! - T^4/5! + (T^2/3!)^2 + \text{higher terms}$$

$$= \frac{1}{T} + \frac{1}{3!}T + \left(\frac{1}{(3!)^2} - \frac{1}{5!}\right)T^3 + \text{higher terms.}$$

This does what we wanted.

II, §1. EXERCISES

We shall write the formal power series in terms of z because that's the way they arise in practice. The series for $\sin z$, $\cos z$, e^z, etc. are to be viewed as formal series.

1. Give the terms of order $\leqq 3$ in the power series:

(a) $e^z \sin z$ (b) $(\sin z)(\cos z)$ (c) $\dfrac{e^z - 1}{z}$

(d) $\dfrac{e^z - \cos z}{z}$ (e) $\dfrac{1}{\cos z}$ (f) $\dfrac{\cos z}{\sin z}$

(g) $\dfrac{\sin z}{\cos z}$ (h) $e^z/\sin z$

2. Let $f(z) = \sum a_n z^n$. Define $f(-z) = \sum a_n(-z)^n = \sum a_n(-1)^n z^n$. We define $f(z)$ to be **even** if $a_n = 0$ for n odd. We define $f(z)$ to be **odd** if $a_n = 0$ for n even. Verify that f is even if and only if $f(-z) = f(z)$ and f is odd if and only if $f(-z) = -f(z)$.

3. Define the **Bernoulli numbers** B_n by the power series

$$\frac{z}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} z^n.$$

Prove the recursion formula

$$\frac{B_0}{n!\,0!} + \frac{B_1}{(n-1)!\,1!} + \cdots + \frac{B_{n-1}}{1!\,(n-1)!} = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

Then $B_0 = 1$. Compute B_1, B_2, B_3, B_4. Show that $B_n = 0$ if n is odd $\neq 1$.

4. Show that

$$\frac{z}{2}\frac{e^{z/2} + e^{-z/2}}{e^{z/2} - e^{-z/2}} = \sum_{n=0}^{\infty} \frac{B_{2n}}{(2n)!} z^{2n}.$$

Replace z by $2\pi iz$ to show that

$$\pi z \cot \pi z = \sum_{n=0}^{\infty} (-1)^n \frac{(2\pi)^{2n}}{(2n)!} B_{2n} z^{2n}.$$

5. Express the power series for $\tan z$, $z/\sin z$, $z \cot z$, in terms of Bernoulli numbers.

6. **(Difference Equations).** Given complex numbers a_0, a_1, u_1, u_2 define a_n for $n \geqq 2$ by

$$a_n = u_1 a_{n-1} + u_2 a_{n-2}.$$

If we have a factorization

$$T^2 - u_1 T - u_2 = (T - \alpha)(T - \alpha'), \qquad \text{and} \quad \alpha \neq \alpha',$$

show that the numbers a_n are given by

$$a_n = A\alpha^n + B\alpha'^n$$

with suitable A, B. Find A, B in terms of a_0, a_1, α, α'. Consider the power series

$$F(T) = \sum_{n=0}^{\infty} a_n T^n.$$

Show that it represents a rational function, and give its partial fraction decomposition.

7. More generally, let $a_0, \ldots, a_{r-1}$ be given complex numbers. Let $u_1, \ldots, u_r$ be complex number such that the polynomial

$$P(T) = T^r - (u_1 T^{r-1} + \cdots + u_r)$$

has distinct roots $\alpha_1, \ldots, \alpha_r$. Define a_n for $n \geqq r$ by

$$a_n = u_1 a_{n-1} + \cdots + u_r a_{n-r}.$$

Show that there exist numbers $A_1, \ldots, A_r$ such that for all n,

$$a_n = A_1 \alpha_1^n + \cdots + A_r \alpha_r^n.$$

II, §2. CONVERGENT POWER SERIES

We first recall some terminology about series of complex numbers.

Let $\{z_n\}$ be a sequence of complex numbers. Consider the series

$$\sum_{n=1}^{\infty} z_n.$$

We define the **partial sum**

$$s_n = \sum_{k=1}^{n} z_k = z_1 + z_2 + \cdots + z_n.$$

We say that the series **converges** if there is some w such that

$$\lim_{n\to\infty} s_n = w$$

exists, in which case we say that w is equal to the **sum of the series**, that is,

$$w = \sum_{n=1}^{\infty} z_n.$$

If $A = \sum \alpha_n$ and $B = \sum \beta_n$ are two convergent series, with partial sums

$$s_n = \sum_{k=1}^{n} \alpha_k \qquad \text{and} \qquad t_n = \sum_{k=1}^{n} \beta_k,$$

then the sum and product converge. Namely,

$$A + B = \sum (\alpha_n + \beta_n)$$
$$AB = \lim_{n\to\infty} s_n t_n$$

Let $\{c_n\}$ be a series of real numbers $c_n \geqq 0$. If the partial sums

$$\sum_{k=1}^{n} c_k$$

are bounded, we recall from calculus that the series converges, and that the least upper bound of these partial sums is the limit.

Let $\sum \alpha_n$ be a series of complex numbers. We shall say that this series **converges absolutely** if the real positive series

$$\sum |\alpha_n|$$

converges. If a series converges absolutely, then it converges. Indeed, let

$$s_n = \sum_{k=1}^{n} \alpha_k$$

be the partial sums. Then for $m \leqq n$ we have

$$s_n - s_m = \alpha_{m+1} + \cdots + \alpha_n$$

whence

$$|s_n - s_m| \leqq \sum_{k=m+1}^{n} |\alpha_k|.$$

Assuming absolute convergence, given ϵ there exists N such that if n, $m \geqq N$, then the right-hand side of this last expression is $< \epsilon$, thereby proving that the partial sums form a Cauchy sequence, and hence that the series converges.

We have the usual test for convergence:

Let $\sum c_n$ be a series of real numbers $\geqq 0$ which converges. If $|\alpha_n| \leqq c_n$ for all n, then the series $\sum \alpha_n$ converges absolutely.

Proof. The partial sums

$$\sum_{k=1}^{n} c_k$$

are bounded by assumption, whence the partial sums

$$\sum_{k=1}^{n} |\alpha_k| \leqq \sum_{k=1}^{n} c_k$$

are also bounded, and the absolute convergence follows.

In the sequel we shall also assume some standard facts about absolutely convergent series, namely:

(i) *If a series $\sum \alpha_n$ is absolutely convergent, then the series obtained by any rearrangement of the terms is also absolutely convergent, and converges to the same limit.*

(ii) *If a double series*

$$\sum_{n=1}^{\infty} \left(\sum_{m=1}^{\infty} \alpha_{mn} \right)$$

is absolutely convergent, then the order of summation can be interchanged, and the series so obtained is absolutely convergent, and converges to the same value.

The proof is easily obtained by considering approximating partial sums (finite sums), and estimating the tail ends. We omit it.

We shall now consider series of functions, and deal with questions of uniformity.

Let S be a set, and f a bounded function on S. Then we define the

sup norm

$$\|f\|_S = \|f\| = \sup_{z \in S} |f(z)|,$$

where sup means least upper bound. It is a norm in the sense that for two functions f, g we have $\|f + g\| \leqq \|f\| + \|g\|$, and for any number c we have $\|cf\| = |c|\,\|f\|$. Also $f = 0$ if and only if $\|f\| = 0$.

Let $\{f_n\}$ $(n = 1, 2, \ldots)$ be a sequence of functions on S. We shall say that this sequence **converges uniformly** on S if there exists a function f on S satisfying the following properties. Given ϵ, there exists N such that if $n \geqq N$, then

$$\|f_n - f\| < \epsilon.$$

We say that $\{f_n\}$ is a **Cauchy sequence** (for the sup norm), if given ϵ, there exists N such that if $m, n \geqq N$, then

$$\|f_n - f_m\| < \epsilon.$$

In this case, for each $z \in S$, the sequence of complex numbers

$$\{f_n(z)\}$$

converges, because for each $z \in S$, we have the inequality

$$|f_n(z) - f_m(z)| \leqq \|f_n - f_m\|.$$

Theorem 2.1. *If a sequence $\{f_n\}$ of functions on S is Cauchy, then it converges uniformly.*

Proof. For each $z \in S$, let

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

Given ϵ, there exists N such that if $m, n \geqq N$, then

$$|f_n(z) - f_m(z)| < \epsilon, \qquad \text{for all} \quad z \in S.$$

Let $n \geqq N$. Given $z \in S$ select $m \geqq N$ sufficiently large (depending on z) such that

$$|f(z) - f_m(z)| < \epsilon.$$

Then

$$\begin{aligned} |f(z) - f_n(z)| &\leqq |f(z) - f_m(z)| + |f_m(z) - f_n(z)| \\ &< \epsilon + \|f_m - f_n\| \\ &< 2\epsilon. \end{aligned}$$

This is true for any z, and therefore $\|f - f_n\| < 2\epsilon$, which proves the theorem.

Remark. It is immediately seen that if the functions f_n in the theorem are bounded, then the limiting function f is also bounded.

Consider a series of functions, $\sum f_n$. Let

$$s_n = \sum_{k=1}^{n} f_k = f_1 + f_2 + \cdots + f_n$$

be the partial sum. We say that the **series converges uniformly** if the sequence of partial sums $\{s_n\}$ converges uniformly.

A series $\sum f_n$ is said to converge **absolutely** if for each $z \in S$ the series

$$\sum |f_n(z)|$$

converges.

The next theorem is sometimes called the **comparison test**.

Theorem 2.2. *Let $\{c_n\}$ be a sequence of real numbers $\geqq 0$, and assume that*

$$\sum c_n$$

converges. Let $\{f_n\}$ be a sequence of functions on S such that $\|f_n\| \leqq c_n$ for all n. Then $\sum f_n$ converges uniformly and absolutely.

Proof. Say $m \leqq n$. We have an estimate for the difference of partial sums,

$$\|s_n - s_m\| \leqq \sum_{k=m+1}^{n} \|f_k\| \leqq \sum_{k=m+1}^{n} c_k.$$

The assumption that $\sum c_k$ converges implies at once the uniform convergence of the partial sums. The argument also shows that the convergence is absolute.

Theorem 2.3. *Let S be a set of complex numbers, and let $\{f_n\}$ be a sequence of continuous functions on S. If this sequence converges uniformly, then the limit function f is also continuous.*

Proof. You should already have seen this theorem some time during a calculus course. We reproduce the proof for convenience. Let $\alpha \in S$. Select n so large that $\|f - f_n\| < \epsilon$. For this choice of n, using the continuity of f_n at α, select δ such that whenever $|z - \alpha| < \delta$ we have

$$|f_n(z) - f_n(\alpha)| < \epsilon.$$

Then

$$|f(z) - f(\alpha)| \leqq |f(z) - f_n(z)| + |f_n(z) - f_n(\alpha)| + |f_n(\alpha) - f(\alpha)|.$$

The first and third term on the right are bounded by $\|f - f_n\| < \epsilon$. The middle term is $< \epsilon$. Hence

$$|f(z) - f(\alpha)| < 3\epsilon,$$

and our theorem is proved.

We now consider the power series, where the functions f_n are

$$f_n(z) = a_n z^n,$$

with complex numbers a_n.

Theorem 2.4. *Let $\{a_n\}$ be a sequence of complex numbers, and let r be a number > 0 such that the series*

$$\sum |a_n| r^n$$

converges. Then the series $\sum a_n z^n$ converges absolutely and uniformly for $|z| \leqq r$.

Proof. Special case of the comparison test.

Example. For any $r > 0$, the series

$$\sum z^n/n!$$

converges absolutely and uniformly for $|z| \leqq r$. Indeed, let

$$c_n = r^n/n!.$$

Then

$$\frac{c_{n+1}}{c_n} = \frac{r^{n+1}}{(n+1)!} \frac{n!}{r^n} = \frac{r}{n+1}.$$

Take $n \geqq 2r$. Then the right-hand side is $\leqq 1/2$. Hence for all n sufficiently large, we have

$$c_{n+1} \leqq \tfrac{1}{2} c_n.$$

Therefore there exists some positive integer n_0 such that

$$c_n \leqq C/2^{n-n_0},$$

for some constant C and all $n \geqq n_0$. We may therefore compare our series with a geometric series to get the absolute and uniform convergence.

The series

$$\exp(z) = \sum_{n=0}^{\infty} z^n/n!$$

therefore defines a continuous function for all values of z. Similarly, the series

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}$$

and

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}$$

converge absolutely and uniformly for all $|z| \leqq r$. They give extensions of the sine and cosine functions to the complex numbers. We shall see later that $\exp(z) = e^z$ as defined in Chapter I, and that these series define the unique analytic functions which coincide with the usual exponential, sine, and cosine functions, respectively, when z is real.

Theorem 2.5. *Let $\sum a_n z^n$ be a power series. If it does not converge absolutely for all z, then there exists a number r such that the series converges absolutely for $|z| < r$ and does not converge absolutely for $|z| > r$.*

Proof. Suppose that the series does not converge absolutely for all z. Let r be the least upper bound of those numbers $s \geqq 0$ such that

$$\sum |a_n| s^n$$

converges. Then $\sum |a_n||z|^n$ diverges if $|z| > r$, and converges if $|z| < r$ by the comparison test, so our assertion is obvious.

The number r in Theorem 2.5 is called the **radius of convergence** of the power series. If the power series converges absolutely for all z, then we say that its **radius of convergence** is **infinity**. When the radius of convergence is 0, then the series converges absolutely only for $z = 0$.

If a power series has a non-zero radius of convergence, then it is called a **convergent** power series. If D is a disc centered at the origin and contained in the disc $D(0, r)$, where r is the radius of convergence, then we say that the power series **converges on** D.

The radius of convergence can be determined in terms of the coefficients. Let t_n be a sequence of real numbers. We recall that a **point of accumulation** of this sequence is a number t such that, given ϵ, there exist infinitely many indices n such that

$$|t_n - t| < \epsilon.$$

In other words, infinitely many points of the sequence lie in a given interval centered at t. An elementary property of real numbers asserts that every bounded sequence has a point of accumulation (Weierstrass–Balzano theorem).

Assume now that $\{t_n\}$ is a bounded sequence. Let S be the set of points of accumulation, so that S looks like Fig. 1.

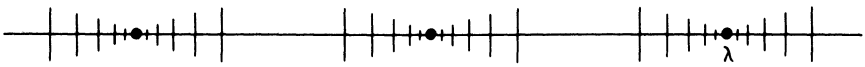


Figure 1

We define the **limit superior**, lim sup, of the sequence to be

$$\lambda = \limsup t_n = \text{least upper bound of } S.$$

Then the reader will verify at once that λ is itself a point of accumulation of the sequence, and is therefore the largest such point. Furthermore, λ has the following properties:

Given ϵ, there exist only finitely many n such that $t_n \geqq \lambda + \epsilon$. There exist infinitely many n such that

$$t_n \geqq \lambda - \epsilon.$$

Proof. If there were infinitely many n such that $t_n \geqq \lambda + \epsilon$, then the sequence $\{t_n\}$ would have a point of accumulation

$$\geqq \lambda + \epsilon > \lambda,$$

contrary to assumption. On the other hand, since λ itself is a point of accumulation, given the ϵ-interval about λ, there have to be infinitely many n such that t_n lies in this ϵ-interval, thus proving the second assertion.

We leave it to the reader to verify that *if a number λ has the above properties, then it is the* lim sup *of the sequence.*

For convenience, if $\{t_n\}$ is not bounded from above, we define its lim sup to be infinity, written ∞.

As an exercise, you should prove:

Let $\{t_n\}$, $\{s_n\}$ *be sequences of real numbers* $\geqq 0$. *Let*

$$t = \limsup t_n \quad \textit{and} \quad s = \limsup s_n.$$

Then

$$\limsup(t_n + s_n) \leqq t + s.$$

If $t \neq 0$, *then*

$$\limsup(t_n s_n) \leqq ts.$$

If $\lim_{n\to\infty} t_n$ *exists, then* $t = \lim_{n\to\infty} t_n$.

This last statement says that if the sequence has an ordinary limit, then that limit is the lim sup of the sequence.

The second statement is often applied in case one sequence has a lim sup, and the other sequence has a limit $\neq 0$. The hypothesis $t \neq 0$ is made only to allow the possibility that $s = \infty$, in which case ts is understood to be ∞. If $s \neq \infty$, and $t \neq \infty$, and $\lim t_n$ exists, then it is true unrestrictedly that

$$\limsup(t_n s_n) = ts.$$

Theorem 2.6. *Let* $\sum a_n z^n$ *be a power series, and let* r *be its radius of convergence. Then*

$$\frac{1}{r} = \limsup |a_n|^{1/n}.$$

If $r = 0$, *this relation is to be interpreted as meaning that the sequence* $\{|a_n|^{1/n}\}$ *is not bounded. If* $r = \infty$, *it is to be interpreted as meaning that* $\limsup |a_n|^{1/n} = 0$.

Proof. Let $t = \limsup |a_n|^{1/n}$. Suppose first that $t \neq 0, \infty$. Given $\epsilon > 0$, there exist only a finite number of n such that $|a_n|^{1/n} \geqq t + \epsilon$. Thus for all but a finite number of n, we have

$$|a_n| \leqq (t + \epsilon)^n,$$

whence the series $\sum a_n z^n$ converges absolutely if $|z| < 1/(t + \epsilon)$, by comparison with the geometric series. Therefore the radius of convergence r satisfies $r \geqq 1/(t + \epsilon)$ for every $\epsilon > 0$, whence $r \geqq 1/t$.

Conversely, given ϵ there exist infinitely many n such that $|a_n|^{1/n} \geqq t - \epsilon$, and therefore

$$|a_n| \geqq (t - \epsilon)^n.$$

Hence the series $\sum a_n z^n$ does not converge if $|z| = 1/(t - \epsilon)$, because its n-th term does not even tend to 0. Therefore the radius of convergence r satisfies $r \leqq 1/(t - \epsilon)$ for every $\epsilon > 0$, whence $r \leqq 1/t$. This concludes the proof in case $t \neq 0, \infty$.

The case when $t = 0$ or ∞ will be left to the reader. The above arguments apply, even with some simplifications.

Corollary 2.7. *If* $\lim |a_n|^{1/n} = t$ *exists, then* $r = 1/t$.

Proof. If the limit exists, then t is the only point of accumulation of the sequence $|a_n|^{1/n}$, and the theorem states that $t = 1/r$.

Corollary 2.8. *Suppose that* $\sum a_n z^n$ *has a radius of convergence greater than* 0. *Then there exists a positive number* C *such that if* $A > 1/r$ *then*

$$|a_n| \leqq CA^n \qquad \text{for all } n.$$

Proof. Let $s = 1/A$ so $0 < s < r$ at the beginning of the proof of the theorem.

In the next examples, we shall use a weak form of **Stirling's formula,** namely

$$n! = n^n e^{-n} u_n \qquad \text{where} \quad \lim u_n^{1/n} = 1.$$

You can prove this estimate by comparing the integral

$$\int_1^n \log x \, dx = n \log n - n + 1$$

with the upper and lower Riemann sums on the interval $[1, n]$, using the partition consisting of the integers from 1 to n. This is a very simple exercise in calculus. Exponentiating the inequalities given by the Riemann sums yields the weak form of Stirling's formula.

Let $\{a_n\}$, $\{b_n\}$ be two sequences of positive numbers. We shall write

$$a_n \equiv b_n \qquad \text{for} \quad n \to \infty$$

if for each n there exists a positive real number u_n such that $\lim u_n^{1/n} = 1$, and $a_n = b_n u_n$. If $\lim a_n^{1/n}$ exists, and $a_n \equiv b_n$, then $\lim b_n^{1/n}$ exists and is equal to $\lim a_n^{1/n}$. We can use this result in the following examples.

Example. The radius of convergence of the series $\sum n!\, z^n$ is 0. Indeed, we have $n! \equiv n^n e^{-n}$ and $(n!)^{1/n}$ is unbounded as $n \to \infty$.

Example. The radius of convergence of $\sum (1/n!)z^n$ is infinity, because $1/n! \equiv e^n/n^n$ so $(1/n!)^{1/n} \to 0$ as $n \to \infty$.

Example. The radius of convergence of $\sum (n!/n^n)z^n$ is e, because $n!/n^n \equiv e^{-n}$, so $\lim(n!/n^n)^{1/n} = e^{-1}$.

Ratio Test. *Let $\{a_n\}$ be a sequence of positive numbers, and assume that* $\lim a_{n+1}/a_n = A \geqq 0$. *Then* $\lim a_n^{1/n} = A$ *also.*

Proof. Suppose first $A > 0$ for simplicity. Given $\epsilon > 0$, let n_0 be such that $A - \epsilon \leqq a_{n+1}/a_n \leqq A + \epsilon$ if $n \geqq n_0$. Without loss of generality, we can assume $\epsilon < A$ so $A - \epsilon > 0$. Write

$$a_n = a_1 \prod_{k=1}^{n_0-1} \frac{a_{k+1}}{a_k} \prod_{k=n_0}^{n} \frac{a_{k+1}}{a_k}.$$

Then by induction, there exist constants $C_1(\epsilon)$ and $C_2(\epsilon)$ such that

$$C_1(\epsilon)(A-\epsilon)^{n-n_0} \leqq a_n \leqq C_2(\epsilon)(A+\epsilon)^{n-n_0},$$

Put $C_1'(\epsilon) = C_1(\epsilon)(A-\epsilon)^{-n_0}$ and $C_2'(\epsilon) = C_2(\epsilon)(A+\epsilon)^{n_0}$. Then

$$C_1'(\epsilon)^{1/n}(A-\epsilon) \leqq a_n^{1/n} \leqq C_2'(\epsilon)^{1/n}(A+\epsilon).$$

There exists $N \geqq n_0$ such that for $n \geqq N$ we have

$$C_1'(\epsilon)^{1/n} = 1 + \delta_1(n) \quad \text{where } |\delta_1(n)| \leqq \epsilon/(A-\epsilon),$$

and similarly $C_2'(\epsilon)^{1/n} = 1 + \delta_2(n)$ with $|\delta_2(n)| \leqq \epsilon/(A+\epsilon)$. Then

$$A - \epsilon + \delta_1(n)(A-\epsilon) \leqq a_n^{1/n} \leqq A + \epsilon + \delta_2(n)(A+\epsilon).$$

This shows that $|a_n^{1/n} - A| \leqq 2\epsilon$, and concludes the proof of the ratio test when $A > 0$. If $A = 0$, one can simply replace the terms on the left of the inequalities by 0 throughout.

Example (The Binomial Series). Let α be any complex number $\neq 0$. Define the **binomial coefficients** as usual,

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!},$$

and the **binomial series**

$$(1+T)^\alpha = B_\alpha(T) = \sum_{n=0}^{\infty} \binom{\alpha}{n} T^n.$$

By convention,

$$\binom{\alpha}{0} = 1.$$

The radius of convergence of the binomial series is 1 if α is not equal to an integer $\geqq 0$.

Proof. Under the stated assumption, none of the coefficients a_n are 0, and we have

$$|a_{n+1}/a_n| = \left|\frac{\alpha - n}{n+1}\right|.$$

The limit is 1 as $n \to \infty$, so we can apply the ratio test.

Warning. Let r be the radius of convergence of the series $f(z)$. *Nothing* has been said about possible convergence if $|z| = r$. Many cases can occur concerning convergence or non-convergence on this circle. See Exercises 6 and 8 for example.

II, §2. EXERCISES

1. Let $|\alpha| < 1$. Express the sum of the geometric series $\sum_{n=1}^{\infty} \alpha^n$ in its usual simple form.

2. Let r be a real number, $0 \leqq r < 1$. Show that the series

$$\sum_{n=0}^{\infty} r^n e^{in\theta} \quad \text{and} \quad \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta}$$

converge (θ is real). Express these series in simple terms using the usual formula for a geometric series.

3. Show that the usual power series for $\log(1+z)$ or $\log(1-z)$ from elementary calculus converges absolutely for $|z| < 1$.

4. Determine the radius of convergence for the following power series.

(a) $\sum n^n z^n$ (b) $\sum z^n/n^n$

(c) $\sum 2^n z^n$ (d) $\sum (\log n)^2 z^n$

(e) $\sum 2^{-n} z^n$ (f) $\sum n^2 z^n$

(g) $\sum \frac{n!}{n^n} z^n$ (h) $\sum \frac{(n!)^3}{(3n)!} z^n$

5. Let $f(z) = \sum a_n z^n$ have radius of convergence $r > 0$. Show that the following series have the same radius of convergence:
   (a) $\sum na_n z^n$ (b) $\sum n^2 a_n z^n$
   (c) $\sum n^d a_n z^n$ for any positive integer d (d) $\sum_{n \geqq 1} na_n z^{n-1}$

6. Give an example of a power series whose radius of convergence is 1, and such that the corresponding function is continuous on the closed unit disc. [*Hint*: Try $\sum z^n/n^2$.]

7. Let a, b be two complex numbers, and assume that b is not equal to any integer $\leqq 0$. Show that the radius of convergence of the series

$$\sum \frac{a(a+1)\cdots(a+n)}{b(b+1)\cdots(b+n)} z^n$$

is at least 1. Show that this radius can be ∞ in some cases.

8. Let $\{a_n\}$ be a decreasing sequence of positive numbers approaching 0. Prove that the power series $\sum a_n z^n$ is uniformly convergent on the domain of z such that

$$|z| \leqq 1 \quad \text{and} \quad |z - 1| \geqq \delta,$$

where $\delta > 0$. [*Hint*: For this problem and the next, use summation by parts, see Appendix, §1.]

9. **(Abel's Theorem).** Let $\sum_{n=0}^{\infty} a_n z^n$ be a power series with radius of convergence $\geqq 1$. Assume that the series $\sum_{n=0}^{\infty} a_n$ converges. Let $0 \leqq x < 1$. Prove that

$$\lim_{x \to 1} \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n.$$

Remark. This result amounts to proving an interchange of limits. If

$$s_n(x) = \sum_{k=1}^{n} a_k x^k,$$

then one wants to prove that

$$\lim_{n \to \infty} \lim_{x \to 1} s_n(x) = \lim_{x \to 1} \lim_{n \to \infty} s_n(x).$$

Cf. Theorem 3.5 of Chapter VII in my *Undergraduate Analysis*, Springer-Verlag, 1983.

10. Let $\sum a_n z^n$ and $\sum b_n z^n$ be two power series, with radius of convergence r and s, respectively. What can you say about the radius of convergence of the series:
    (a) $\sum (a_n + b_n) z^n$ (b) $\sum a_n b_n z^n$?

11. Let α, β be complex numbers with $|\alpha| < |\beta|$. Let

$$f(z) = \sum (3\alpha^n - 5\beta^n) z^n.$$

Determine the radius of convergence of $f(z)$.

12. Let $\{a_n\}$ be the sequence of real numbers defined by the conditions:

$$a_0 = 1,\ a_1 = 2, \qquad \text{and} \qquad a_n = a_{n-1} + a_{n-2} \qquad \text{for} \quad n \geqq 2.$$

Determine the radius of convergence of the power series

$$\sum_{n=0}^{\infty} a_n z^n.$$

[*Hint*: What is the general solution of a difference equation? Cf. Exercise 6 of §1.]

13. More generally, let u_1, u_2 be complex numbers such that the polynomial

$$P(T) = T^2 - u_1 T - u_2 = (T - \alpha_1)(T - \alpha_2)$$

has two distinct roots with $|\alpha_1| < |\alpha_2|$. Let a_0, a_1 be given, and let

$$a_n = u_1 a_{n-1} + u_2 a_{n-2} \qquad \text{for} \quad n \geqq 2.$$

What is the radius of convergence of the series $\sum a_n T^n$?

II, §3. RELATIONS BETWEEN FORMAL AND CONVERGENT SERIES

Sums and Products

Let $f = f(T)$ and $g = g(T)$ be formal power series. We may form their formal product and sum, $f + g$ and fg. If f converges absolutely for some complex number z, then we have the value $f(z)$, and similarly for $g(z)$.

Theorem 3.1. *If f, g are power series which converge absolutely on the disc $D(0, r)$, then $f + g$ and fg also converge absolutely on this disc. If α is a complex number, αf converges absolutely on this disc, and we have*

$$(f + g)(z) = f(z) + g(z), \qquad (fg)(z) = f(z)g(z),$$
$$(\alpha f)(z) = \alpha \cdot f(z)$$

for all z in the disc.

Proof. We give the proof for the product, which is the hardest. Let

$$f = \sum a_n T^n \qquad \text{and} \qquad g = b_n T^n,$$

so that

$$fg = \sum c_n T^n, \qquad \text{where} \quad c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Let $0 < s < r$. We know that there exists a positive number C such that for all n,

$$|a_n| \leqq C/s^n \qquad \text{and} \qquad |b_n| \leqq C/s^n.$$

Then

$$|c_n| \leqq \sum_{k=0}^{n} |a_k b_{n-k}| \leqq (n+1)\frac{C}{s^k}\frac{C}{s^{n-k}} = \frac{(n+1)C^2}{s^n}.$$

Therefore

$$|c_n|^{1/n} \leqq \frac{(n+1)^{1/n}C^{2/n}}{s}.$$

But $\lim_{n\to\infty} (n+1)^{1/n}C^{1/n} = 1$. Hence

$$\limsup |c_n|^{1/n} \leqq 1/s.$$

This is true for every $s < r$. It follows that $\limsup |c_n|^{1/n} \leqq 1/r$, thereby proving that the formal product converges absolutely on the same disc. We have also shown that the series of positive terms

$$\sum_{n=0}^{\infty} \sum_{k=0}^{n} |a_k||b_{n-k}||z|^n$$

converges.

Let

$$f_N(T) = a_0 + a_1 T + \cdots + a_N T^N,$$

and similarly, let $g_N(T)$ be the polynomial consisting of the terms of order $\leqq N$ in the power series for g. Then

$$f(z) = \lim_N f_N(z) \qquad \text{and} \qquad g(z) = \lim_N g_N(z).$$

Furthermore,

$$|(fg)_N(z) - f_N(z)g_N(z)| \leqq \sum_{n=N+1}^{\infty} \sum_{k=0}^{n} |a_k||b_{n-k}||z|^n.$$

In view of the convergence proved above, for N sufficiently large the right-hand side is arbitrarily small, and hence

$$f(z)g(z) = \lim_N f_N(z)g_N(z) = (fg)(z),$$

thereby proving the theorem for the product.

The previous theorem shows that a formal power series determines a function on the disc of absolute convergence. We can raise the question: If two formal power series f, g give rise to the same function on some neighborhood of 0, are they equal as formal power series? Subtracting g from f, this amounts to asking: If a power series determines the zero function on some disc centered at the origin, is it the zero series, i.e. are all its coefficients equal to 0? The answer is yes. In fact, more is true.

Theorem 3.2.

(a) *Let $f(T) = \sum a_n T^n$ be a non-constant power series, having a non-zero radius of convergence. If $f(0) = 0$, then there exists $s > 0$ such that $f(z) \neq 0$ for all z with $|z| \leqq s$, and $z \neq 0$.*
(b) *Let $f(T) = \sum a_n T^n$ and $g(T) = \sum b_n T^n$ be two convergent power series. Suppose that $f(x) = g(x)$ for all points x in an infinite set having 0 as a point of accumulation. Then $f(T) = g(T)$, that is $a_n = b_n$ for all n.*

Proof. We can write

$$\begin{aligned} f(z) &= a_m z^m + \text{higher terms}, \qquad \text{and } a_m \neq 0 \\ &= a_m z^m(1 + b_1 z + b_2 z^2 + \cdots) \\ &= a_m z^m(1 + h(z)), \end{aligned}$$

where $h(z) = b_1 z + b_2 z^2 + \cdots$ is a power series having a non-zero radius of convergence, and zero constant term. For all sufficiently small $|z|$, the value $|h(z)|$ is small, and hence

$$1 + h(z) \neq 0.$$

If $z \neq 0$, then $a_m z^m \neq 0$. This proves the first part of the theorem.

For part (b), let $h(t) = f(T) - g(T) = \sum (a_n - b_n)T^n$. We have $h(x) = 0$ for an infinite set of points x having 0 as point of accumulation. By part (a), this implies that $h(T)$ is the zero power series, so $a_n = b_n$ for all n, thus proving the theorem.

Example. There exists at most one convergent power series $f(T) = \sum a_n T^n$ such that for some interval $[-\epsilon, \epsilon]$ we have $f(x) = e^x$ for all x in $[-\epsilon, \epsilon]$. This proves the uniqueness of any power series extension of the exponential function to all complex numbers. Similarly, one has the uniqueness of the power series extending the sine and cosine functions.

Furthermore, let $\exp(z) = \sum z^n/n!$. Then

$$\exp(iz) = \sum (iz)^n/n!.$$

Summing over even n and odd n, we find that

$$\exp(iz) = C(z) + iS(z),$$

where $C(z)$ and $S(z)$ are the power series for the cosine and sine of z respectively. Hence $e^{i\theta}$ for real θ coincides with $\exp(i\theta)$ as given by the power series expansion.

Quite generally, if $g(T)$, $h(T)$ are power series with 0 constant term, then

$$\exp\bigl(g(T) + h(T)\bigr) = \bigl(\exp g(T)\bigr)\bigl(\exp h(T)\bigr).$$

Proof. On one hand, by definition,

$$\exp\bigl(g(T) + h(T)\bigr) = \sum_{n=0}^{\infty} \frac{\bigl(g(T) + h(T)\bigr)^n}{n!}$$

and on the other hand,

$$\begin{aligned}\bigl(\exp g(T)\bigr)\bigl(\exp h(T)\bigr) &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{g(T)^k h(T)^{n-k}}{k!(n-k)!} \\ &= \sum_{n=0}^{\infty} \frac{\bigl(g(T) + h(T)\bigr)^n}{n!} \qquad \text{qed.}\end{aligned}$$

In particular, for complex numbers z, w we have

$$\exp(z + w) = (\exp z)(\exp w),$$

because we can apply the above identity to $g(T) = zT$ and $h(T) = wT$, and then substitute $T = 1$. Thus we see that the exponential function e^z defined in Chapter I has the same values as the function defined by the usual power series $\exp(z)$. From now on, we make no distinction between e^z and $\exp(z)$.

Theorem 3.2 also allows us to conclude that any polynomial relation between the elementary functions which have a convergent Taylor expansion at the origin also holds for the extension of these functions as complex power series.

Example. We can now conclude that $\sin^2 z + \cos^2 z = 1$, where $\sin z = S(z)$, $\cos z = C(z)$ are defined by the usual power series. Indeed, the power series $S(z)^2 + C(z)^2$ has infinite radius of convergence, and has value 1 for all real z. Theorem 3.2 implies that there is at most one series having this property, and that is the series 1, as desired. It would be disagreeable to show directly that the formal power series for the sine and cosine satisfy this relation. It is easier to do it through elementary calculus as above.

Example. Let m be a positive integer. We have seen in §2 that the binomial series

$$B(z) = \sum \binom{\alpha}{n} z^n$$

with $\alpha = 1/m$ has a radius of convergence $\geqq 1$, and thus converges absolutely for $|z| < 1$. By elementary calculus, we have

$$B(x)^m = 1 + x$$

when x is real, and $|x| < 1$ (or even when $|x|$ is sufficiently small). Therefore $B(T)^m$ is the unique formal power series such that

$$B(x)^m = 1 + x$$

for all sufficiently small real x, and therefore we conclude that

$$B(T)^m = 1 + T.$$

In this manner, we see that we can take m-th roots

$$(1 + z)^{1/m}$$

by the binomial series when $|z| < 1$.

Quotients

In our discussion of formal power series, besides the polynomial relations, we dealt with quotients and also composition of series. We still have to relate these to the convergent case. It will be convenient to introduce a simple notation to estimate power series.

Let $f(T) = \sum a_n T^n$ be a power series. Let

$$\varphi(T) = \sum c_n T^n$$

be a power series with **real coefficients** $c_n \geqq 0$. We say that f is **dominated by** φ, and write

$$f \prec \varphi \qquad \text{or} \qquad f(T) \prec \varphi(T),$$

if $|a_n| \leqq c_n$ for all n. It is clear that if φ, ψ are power series with real coefficients $\geqq 0$ and if

$$f \prec \varphi, \qquad g \prec \psi,$$

then

$$f + g \prec \varphi + \psi \qquad \text{and} \qquad fg \prec \varphi\psi.$$

Theorem 3.3. *Suppose that f has a non-zero radius of convergence, and non-zero constant term. Let g be the formal power series which is inverse to f, that is, $fg = 1$. Then g also has a non-zero radius of convergence.*

Proof. Multiplying f by some constant, we may assume without loss of generality that the constant term is 1, so we write

$$f = 1 + a_1 T + a_2 T^2 + \cdots = 1 - h(T),$$

where $h(T)$ has constant term equal to 0. By Corollary 2.8, we know that there exists a number $A > 0$ such that

$$|a_n| \leqq A^n, \qquad n \geqq 1.$$

(We can take $C = 1$ by picking A large enough.) Then

$$\frac{1}{f(T)} = \frac{1}{1 - h(T)} = 1 + h(T) + h(T)^2 + \cdots.$$

But

$$h(T) \prec \sum_{n=1}^{\infty} A^n T^n = \frac{AT}{1 - AT}.$$

Therefore $1/f(T) = g(T)$ satisfies

$$g(T) \prec 1 + \frac{AT}{1 - AT} + \frac{(AT)^2}{(1 - AT)^2} + \cdots = \frac{1}{1 - \dfrac{AT}{1 - AT}}.$$

But

$$\frac{1}{1 - \dfrac{AT}{1 - AT}} = (1 - AT)(1 + 2AT + (2AT)^2 + \cdots)$$

$$\prec (1 + AT)(1 + 2AT + (2AT)^2 + \cdots).$$

Therefore $g(T)$ is dominated by a product of power series having non-zero radius of convergence, whence $g(T)$ itself a non-zero radius of convergence, as was to be shown.

Composition of Series

Theorem 3.4. *Let*

$$f(z) = \sum_{n \geqq 0} a_n z^n \qquad \textit{and} \qquad h(z) = \sum_{n \geqq 1} b_n z^n$$

be convergent power series, and assume that the constant term of h is 0. Assume that $f(z)$ is absolutely convergent for $|z| \leqq r$, with $r > 0$, and that $s > 0$ is a number such that

$$\sum |b_n| s^n \leqq r.$$

Let $g = f(h)$ be the formal power series obtained by composition,

$$g(T) = \sum_{n \geqq 0} a_n \left(\sum_{k=1}^{\infty} b_k T^k \right)^n.$$

Then g converges absolutely for $|z| \leqq s$, and for such z,

$$g(z) = f(h(z)).$$

Proof. Let $g(T) = \sum c_n T^n$. Then $g(T)$ is dominated by the series

$$g(T) \prec \sum_{n=0}^{\infty} |a_n| \left(\sum_{k=1}^{\infty} |b_k| T^k \right)^n$$

and by hypothesis, the series on the right converges absolutely for $|z| \leqq s$, so $g(z)$ converges absolutely for $|z| \leqq s$. Let

$$f_N(T) = a_0 + a_1 T + \cdots + a_{N-1} T^{N-1}$$

be the polynomial of degree $\leqq N - 1$ beginning the power series f. Then

$$f(h(T)) - f_N(h(T)) \prec \sum_{n=N}^{\infty} |a_n| \left(\sum_{k=1}^{\infty} |b_k| T^k \right)^n,$$

and $f(h(T)) = g(T)$ by definition. By the absolute convergence we conclude: Given ϵ, there exists N_0 such that if $N \geqq N_0$ and $|z| \leqq s$, then

$$|g(z) - f_N(h(z))| < \epsilon.$$

Since the polynomials f_N converge uniformly to the function f on the closed disc of radius r, we can pick N_0 sufficiently large so that for

$N \geqq N_0$ we have

$$|f_N(h(z)) - f(h(z))| < \epsilon.$$

This proves that

$$|g(z) - f(h(z))| < 2\epsilon,$$

for every ϵ, whence $g(z) - f(h(z)) = 0$, thereby proving the theorem.

Example. Let m be a positive integer, and let $h(z)$ be a convergent power series with zero constant term. Then we can form the m-th root

$$(1 + h(z))^{1/m}$$

by the binomial expansion, and this m-th root is a convergent power series whose m-th power is $1 + h(z)$.

Example. Define

$$f(w) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{w^n}{n}.$$

Readers should immediately recognize that the series on the right is the usual series of calculus for $\log(1 + w)$ when $w = x$ and x is real. This series converges absolutely for $|w| < 1$. We can therefore define $\log z$ for $|z - 1| < 1$ by

$$\log z = f(z - 1).$$

We leave it as Exercise 1 to verify that $\exp \log z = z$.

II, §3. EXERCISES

1. (a) Use the above definition of $\log z$ for $|z - 1| < 1$ to prove that $\exp \log z = z$. [*Hint*: What are the values on the left when $z = x$ is real?]

   (b) Let $z_0 \neq 0$. Let α be any complex number such that $\exp(\alpha) = z_0$. For $|z - z_0| < |z_0|$ define

$$\log z = f\left(\frac{z}{z_0} - 1\right) + \alpha = f\left(\frac{z - z_0}{z_0}\right) + \alpha.$$

   Prove that $\exp \log z = z$ for $|z - z_0| < |z_0|$.

   Warning. The above definitions in parts (a) and (b) may differ by a constant. Since you should have proved that $\exp \log z = z$ in both cases, and since $\exp(w_1) = \exp(w_2)$ if and only if there exists an integer k such that $w_1 = w_2 + 2\pi ik$, it follows that if we denote the two logs by $\log_1$ and $\log_2$, respectively, then $\log_1(z) = \log_2(z) + 2\pi ik$.

2. (a) Let $\exp(T) = \sum_{n=0}^{\infty} T^n/n!$ and $\log(1+T) = \sum_{k=1}^{\infty}(-1)^{k-1}T^k/k$. Show that

$$\exp\log(1+T) = 1+T \quad \text{and} \quad \log\exp(T) = T.$$

   (b) Let $h_1(T)$ and $h_2(T)$ be formal power series with 0 constant terms. Prove that $\log\big((1+h_1(T))(1+h_2(T))\big) = \log\big(1+h_1(T)\big) + \log\big(1+h_2(T)\big)$.

   (c) For complex numbers α, β show that $\log(1+T)^\alpha = \alpha\log(1+T)$ and

$$(1+T)^\alpha(1+T)^\beta = (1+T)^{\alpha+\beta}.$$

3. Prove that for all complex z we have

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

4. Show that the only complex numbers z such that $\sin z = 0$ are $z = k\pi$, where k is an integer. State and prove a similar statement for $\cos z$.

5. Find the power series expansion of $f(z) = 1/(z+1)(z+2)$, and find the radius of convergence.

6. The **Legendre polynomials** can be defined as the coefficients $P_n(\alpha)$ of the series expansion of

$$\begin{aligned} f(z) &= \frac{1}{(1-2\alpha z+z^2)^{1/2}} \\ &= 1 + P_1(\alpha)z + P_2(\alpha)z^2 + \cdots + P_n(\alpha)z^n + \cdots. \end{aligned}$$

   Calculate the first four Legendre polynomials.

II, §4. ANALYTIC FUNCTIONS

So far we have looked at power series expansions at the origin. Let f be a function defined in some neighborhood of a point z_0. We say that f is **analytic** at z_0 if there exists a power series

$$\sum_{n=0}^{\infty} a_n(z-z_0)^n$$

and some $r > 0$ such that the series converges absolutely for $|z-z_0| < r$, and such that for such z, we have

$$f(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^n.$$

Suppose f is a function on an open set U. We say that f is **analytic on** U if f is analytic at every point of U.

In the light of the uniqueness theorem for power series, Theorem 3.2, we see that the above power series expressing f in some neighborhood of z_0 is uniquely determined. We have

$$f(z_0) = 0 \qquad \text{if and only if} \qquad a_0 = 0.$$

A point z_0 such that $f(z_0) = 0$ is called a **zero** of f. Instead of saying that f is analytic at z_0, we also say that f **has a power series expansion at** z_0 (meaning that the values of $f(z)$ for z near z_0 are given by an absolutely convergent power series as above).

If S is an arbitrary set, not necessarily open, it is useful to make the convention that a function is **analytic** on S if it is the restriction of an analytic function on an open set containing S. This is useful, for instance, when S is a closed disc.

The theorem concerning sums, products, quotients and composites of convergent power series now immediately imply:

If f, g are analytic on U, so are $f + g$, fg. Also f/g is analytic on the open subset of $z \in U$ such that $g(z) \neq 0$.

If $g\colon U \to V$ is analytic and $f\colon V \to \mathbf{C}$ is analytic, then $f \circ g$ is analytic.

For this last assertion, we note that if $z_0 \in U$ and $g(z_0) = w_0$, so

$$g(z) = w_0 + \sum_{n \geqq 1} b_n (z - z_0)^n \quad \text{and} \quad f(w) = \sum_{n \geqq 0} a_n (w - w_0)^n$$

for w near w_0, then $g(z) - w_0$ is represented by a power series $h(z - z_0)$ without constant term, so that Theorem 3.4 applies: We can "substitute"

$$f(g(z)) = \sum a_n (g(z) - w_0)^n$$

to get the power series representation for $f(g(z))$ in a neighborhood of z_0.

The next theorem, although easy to prove, requires being stated. It gives us in practice a way of finding a power series expansion for a function at a point.

Theorem 4.1. *Let $f(z) = \sum a_n z^n$ be a power series whose radius of convergence is r. Then f is analytic on the open disc $D(0, r)$.*

Proof. We have to show that f has a power series expansion at an arbitrary point z_0 of the disc, so $|z_0| < r$. Let $s > 0$ be such that

$|z_0| + s < r$. We shall see that f can be represented by a convergent power series at z_0, converging absolutely on a disc of radius s centered at z_0.

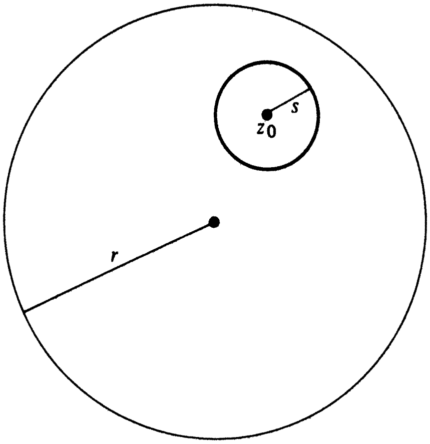


Figure 2

We write

$$z = z_0 + (z - z_0)$$

so that

$$z^n = \big(z_0 + (z - z_0)\big)^n.$$

Then

$$f(z) = \sum_{n=0}^{\infty} a_n \left(\sum_{k=0}^{n} \binom{n}{k} z_0^{n-k} (z - z_0)^k \right).$$

If $|z - z_0| < s$ then $|z_0| + |z - z_0| < r$, and hence the series

$$\sum_{n=0}^{\infty} |a_n| (|z_0| + |z - z_0|)^n = \sum_{n=0}^{\infty} |a_n| \left[\sum_{k=0}^{n} \binom{n}{k} |z_0|^{n-k} |z - z_0|^k \right]$$

converges. Then we can interchange the order of summation, to get

$$f(z) = \sum_{k=0}^{\infty} \left[\sum_{n=k}^{\infty} a_n \binom{n}{k} z_0^{n-k} \right] (z - z_0)^k,$$

which converges absolutely also, as was to be shown.

Example. Let us find the terms of order $\leqq 3$ in the power series expansion of the function

$$f(z) = z^2/(z + 2)$$

at the point $z_0 = 1$. We write

$$z = 1 + (z - 1), \qquad z + 2 = 3 + (z - 1).$$

Let $\equiv$ denote congruence $\bmod (z-1)^4$ (so disregard terms of order > 3). Then

$$\begin{aligned}
z^2 &= 1 + 2(z-1) + (z-1)^2 \\
z + 2 &= 3\left(1 + \frac{1}{3}(z-1)\right) \\
\frac{1}{z+2} &= \frac{1}{3}\frac{1}{1 + \frac{1}{3}(z-1)} \\
&= \frac{1}{3}\left(1 - \frac{1}{3}(z-1) + \frac{1}{3^2}(z-1)^2 - \frac{1}{3^3}(z-1)^3 + \cdots\right).
\end{aligned}$$

Hence

$$\begin{aligned}
\frac{z^2}{z+2} &\equiv (1 + 2(z-1) + (z-1)^2) \\
&\quad \times \frac{1}{3}\left(1 - \frac{1}{3}(z-1) + \frac{1}{3^2}(z-1)^2 - \frac{1}{3^3}(z-1)^3\right) \\
&= \frac{1}{3}\left[1 + \frac{5}{3}(z-1) + \left(\frac{1}{3} + \frac{1}{3^2}\right)(z-1)^2 + \left(-\frac{1}{3} + \frac{2}{3^2} - \frac{1}{3^3}\right)(z-1)^3\right].
\end{aligned}$$

These are the desired terms of the expansion.

Remark. Making a translation, the theorem shows that if f has a power series expansion on a disc $D(z_0, r)$, that is,

$$f(z) = \sum a_n (z - z_0)^n$$

for $|z - z_0| < r$, then f is analytic on this disc.

II, §4. EXERCISES

1. Find the terms of order $\leqq 3$ in the power series expansion of the function $f(z) = z^2/(z-2)$ at $z = 1$.

2. Find the terms of order $\leqq 3$ in the power series expansion of the function $f(z) = (z-2)/(z+3)(z+2)$ at $z = 1$.

II, §5. DIFFERENTIATION OF POWER SERIES

Let $D(0, r)$ be a disc of radius $r > 0$. A function f on the disc for which there exists a power series $\sum a_n z^n$ having a radius of convergence $\geqq r$ and such that

$$f(z) = \sum a_n z^n$$

for all z in the disc is said to **admit a power series expansion on this disc.** We shall now see that such a function is holomorphic, and that its derivative is given by the "obvious" power series.

Indeed, define the **formal derived series** to be

$$\sum n a_n z^{n-1} = a_1 + 2a_2 z + 3a_3 z^2 + \cdots .$$

Theorem 5.1. *If $f(z) = \sum a_n z^n$ has radius of convergence r, then:*

(i) *The series $\sum n a_n z^{n-1}$ has the same radius of convergence.*
(ii) *The function f is holomorphic on $D(0, r)$, and its derivative is equal to $\sum n a_n z^{n-1}$.*

Proof. By Theorem 2.6, we have

$$\limsup |a_n|^{1/n} = 1/r.$$

But

$$\limsup |n a_n|^{1/n} = \limsup n^{1/n} |a_n|^{1/n}.$$

Since $\lim n^{1/n} = 1$, the sequences

$$|n a_n|^{1/n} \quad \text{and} \quad |a_n|^{1/n}$$

have the same lim sup, and therefore the series $\sum a_n z^n$ and $\sum n a_n z^n$ have the same radius of convergence. Then

$$\sum n a_n z^{n-1} \quad \text{and} \quad \sum n a_n z^n$$

converge absolutely for the same values of z, so the first part of the theorem is proved.

As to the second, let $|z| < r$, and $\delta > 0$ be such that $|z| + \delta < r$. We consider complex numbers h such that

$$|h| < \delta.$$

We have

$$\begin{aligned} f(z + h) &= \sum a_n (z + h)^n \\ &= \sum a_n \big(z^n + n z^{n-1} h + h^2 P_n(z, h)\big), \end{aligned}$$

where $P_n(z, h)$ is a polynomial in z and h, with positive integer coefficients, in fact

$$P_n(z, h) = \sum_{k=2}^{n} \binom{n}{k} h^{k-2} z^{n-k}.$$

Note that we have the estimate:

$$|P_n(z, h)| \leqq \sum_{k=2}^{n} \binom{n}{k} \delta^{k-2} |z|^{n-k} = P_n(|z|, \delta).$$

Subtracting series, we find

$$f(z + h) - f(z) - \sum n a_n z^{n-1} h = h^2 \sum a_n P_n(z, h),$$

and since the series on the left is absolutely convergent, so is the series on the right. We divide by h to get

$$\frac{f(z + h) - f(z)}{h} - \sum n a_n z^{n-1} = h \sum a_n P_n(z, h).$$

For $|h| < \delta$, we have the estimate

$$\begin{aligned} |\sum a_n P_n(z, h)| &\leqq \sum |a_n| |P_n(z, h)| \\ &\leqq \sum |a_n| P_n(|z|, \delta). \end{aligned}$$

This last expression is fixed, independent of h. Hence

$$|h \sum a_n P_n(z, h)| \leqq |h| \sum |a_n| P_n(|z|, \delta).$$

As h approaches 0, the right-hand side approaches 0, and therefore

$$\lim_{h \to 0} |h \sum a_n P_n(z, h)| = 0.$$

This proves that f is differentiable, and that its derivative at z is given by the series $\sum n a_n z^{n-1}$, as was to be shown.

Remark. Conversely, we shall see after Cauchy's theorem that a function which is differentiable admits power series expansion—a very remarkable fact, characteristic of complex differentiability.

From the theorem, we see that the k-th derivative of f is given by the series

$$f^{(k)}(z) = k!\, a_k + h_k(z),$$

where h_k is a power series without constant term. Therefore we obtain the standard expression for the coefficients of the power series in terms of the derivatives, namely

$$\boxed{a_n = \frac{f^{(n)}(0)}{n!}.}$$

If we deal with the expansion at a point z_0, namely

$$f(z) = a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \cdots,$$

then we find

$$\boxed{a_n = \frac{f^{(n)}(z_0)}{n!}.}$$

It is utterly trivial that the formally integrated series

$$\sum_{n=0}^{\infty} \frac{a_n}{n+1} z^{n+1}$$

has radius of convergence at least r, because its coefficients are smaller in absolute value than the coefficients of f. Since the derivative of this integrated series is exactly the series for f, it follows from Theorem 6.1 that the integrated series has the same radius of convergence as f.

Let f be a function on an open set U. If g is a holomorphic function on U such that $g' = f$, then g is called a **primitive** for f. We see that a function which has a power series expansion on a disc always has a primitive on that disc. In other words, an analytic function has a local primitive at every point.

Example. The function $1/z$ is analytic on the open set U consisting of the plane from which the origin has been deleted. Indeed, for $z_0 \neq 0$, we have the power series expansion

$$\begin{aligned}\frac{1}{z} = \frac{1}{z_0 + z - z_0} &= \frac{1}{z_0} \frac{1}{\left(1 + (z - z_0)/z_0\right)} \\ &= \frac{1}{z_0}\left(1 - \frac{1}{z_0}(z - z_0) + \cdots\right)\end{aligned}$$

converging on some disc $|z - z_0| < r$. Hence $1/z$ has a primitive on such a disc, and this primitive may be called $\log z$.

II, §5. EXERCISES

In Exercises 1 through 5, also determine the radius of convergence of the given series.

1. Let

$$f(z) = \sum \frac{z^{2n}}{(2n)!}$$

Prove that $f''(z) = f(z)$.

2. Let

$$f(z) = \sum_{n=0}^{\infty} \frac{z^{2n}}{(n!)^2}.$$

Prove that

$$z^2 f''(z) + zf'(z) = 4z^2 f(z).$$

3. Let

$$f(z) = z - \frac{z^3}{3} + \frac{z^5}{5} - \frac{z^7}{7} + \cdots.$$

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

4. Let

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

Prove that

$$z^2 J''(z) + zJ'(z) + z^2 J(z) = 0.$$

5. For any positive integer k, let

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

Prove that

$$z^2 J_k''(z) + zJ_k'(z) + (z^2 - k^2)J_k(z) = 0.$$

6. (a) For $|z - 1| < 1$, show that the derivative of the function

$$\log z = \log\big(1 + (z - 1)\big) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(z-1)^n}{n}$$

is $1/z$.

(b) Let $z_0 \neq 0$. For $|z - z_0| < 1$, define $f(z) = \sum (-1)^{n-1}\big((z - z_0)/z_0\big)^n/n$. Show that $f'(z) = 1/z$.

II, §6. THE INVERSE AND OPEN MAPPING THEOREMS

Let f be an analytic function on an open set U, and let $f(U) = V$. We shall say that f is an **analytic isomorphism** if V is open and there exists an analytic function

$$g\colon V \to U$$

such that $f \circ g = \mathrm{id}_V$ and $g \circ f = \mathrm{id}_U$, in other words, f and g are inverse functions to each other.

We say that f is a **local analytic isomorphism** or is **locally invertible** at a point z_0 if there exists an open set U containing z_0 such that f is an analytic isomorphism on U.

Remark. The word "inverse" is used in two senses: the sense of §1, when we consider the reciprocal $1/f$ of a function f, and in the current sense, which may be called the **composition inverse**, i.e. an inverse for the composite of mappings. The context makes clear which is meant. In this section, we mean the composition inverse.

Theorem 6.1.

(a) *Let $f(T) = a_1 T +$ higher terms be a formal power series with $a_1 \neq 0$. Then there exists a unique power series $g(T)$ such that $f(g(T)) = T$. This power series also satisfies $g(f(T)) = T$.*

(b) *If f is a convergent power series, so is g.*

(c) *Let f be an analytic function on an open set U containing z_0. Suppose that $f'(z_0) \neq 0$. Then f is a local analytic isomorphism at z_0.*

Proof. We first deal with the formal power series problem (a), and we find first a formal inverse for $f(T)$. For convenience of notation below we write $f(T)$ in the form

$$f(T) = a_1 T - \sum_{n=2}^{\infty} a_n T^n.$$

We seek a power series

$$g(T) = \sum_{n=1}^{\infty} b_n T^n$$

such that

$$f(g(T)) = T.$$

The solution to this problem is given by solving the equations in terms of the coefficients of the power series

$$a_1 g(T) - a_2 g(T)^2 - \cdots = T.$$

These equations are of the form

$$a_1 b_n - P_n(a_2, \ldots, a_n, b_1, \ldots, b_{n-1}) = 0, \qquad \text{and} \qquad a_1 b_1 = 1 \quad \text{for } n = 1,$$

where P_n is a polynomial with positive integer coefficients (generalized binomial coefficients). In fact, one sees at once that

$$\begin{aligned} &P_n(a_2, \ldots, a_n, b_1, \ldots, b_{n-1}) \\ &\qquad = a_2 P_{2,n}(b_1, \ldots, b_{n-1}) + \cdots + a_n P_{n,n}(b_1, \ldots, b_{n-1}), \end{aligned}$$

where again $P_{k,n}$ is a polynomial with positive coefficients. In this manner we can solve recursively for the coefficients

$$b_1, b_2, \ldots$$

since b_n appears linearly with coefficient $a_1 \neq 0$ in these equations, and the other terms do not contain b_n. This shows that a formal inverse exists and is uniquely determined.

Next we prove that $g(f(T)) = T$. By what we have proved already, there exists a power series $h(T) = c_1 T +$ higher terms with $c_1 \neq 0$ such that $g(h(T)) = T$. Then using $f(g(T)) = T$ and $g(h(T)) = T$, we obtain:

$$g(f(T)) = g(f(g(h(T)))) = g(h(T)) = T,$$

which proves the desired formal relation.

Assume next that f is convergent.

We must now show that $g(z)$ is absolutely convergent on some disc. To simplify the number of symbols used, we assume that $a_1 = 1$. This loses no generality, because if we find a convergent inverse power series for $a_1^{-1} f(z)$, we immediately get the convergent inverse power series for $f(z)$ itself.

Let

$$f^*(T) = T - \sum_{n \geqq 2} a_n^* T^n$$

be a power series with a_n^* real $\geqq 0$ such that $|a_n| \leqq a_n^*$ for all n. Let $\varphi(T)$ be the formal inverse of $f^*(T)$, so

$$\varphi(T) = \sum_{n \geqq 1} c_n T^n, \qquad c_1 = 1.$$

Then we have

$$c_n - P_n(a_2^*, \ldots, a_n^*, c_1, \ldots, c_{n-1}) = 0$$

with those same polynomials P_n as before. By induction, it therefore follows that c_n is real $\geqq 0$, and also that

$$|b_n| \leqq c_n,$$

since $b_n = P_n(a_2, \ldots, b_{n-1})$. It suffices therefore to pick the series f^* so that it has an easily computed formal inverse φ which is easily verified to have a positive radius of convergence.

It is now a simple matter to carry out this idea, and we pick for f^* a geometric series. There exists $A > 0$ such that for all n we have

$$|a_n| \leqq A^n.$$

(We can omit a constant C in front of A^n by picking A sufficiently large.) Then

$$f^*(T) = T - \sum_{n \geqq 2} A^n T^n = T - \frac{A^2 T^2}{1 - AT}.$$

The power series $\varphi(T)$ is such that $f^*(\varphi(T)) = T$, namely

$$\varphi(T) - \frac{A^2 \varphi(T)^2}{1 - A\varphi(T)} = T,$$

which is equivalent with the quadratic equation

$$(A^2 + A)\varphi(T)^2 - (1 + AT)\varphi(T) + T = 0.$$

This equation has the solution

$$\varphi(T) = \frac{1 + AT - \sqrt{(1 + AT)^2 - 4T(A^2 + A)}}{2(A^2 + A)}.$$

The expression under the radical sign is of the form

$$(1 + AT)^2\left(1 - \frac{4T(A^2 + A)}{(1 + AT)^2}\right)$$

and its square root is given by

$$(1 + AT)\left(1 - \frac{4T(A^2 + A)}{(1 + AT)^2}\right)^{1/2}.$$

We use the binomial expansion to find the square root of a series of the form $1 + h(T)$ when $h(T)$ has zero constant term. It is now clear that $\varphi(T)$ is obtained by composition of convergent power series, and hence has a non-zero radius of convergence. This proves that the power series $g(T)$ also converges.

Finally, for (c), suppose first that $z_0 = 0$ and $f(z_0) = 0$, so f is analytic on an open set containing 0. This means that f has a convergent power series expansion at 0, so we view f as being defined on its open disc of convergence

$$f : D \to \mathbf{C}.$$

Let V_0 be an open disc centered at 0 such that V_0 is contained in the disc of convergence of g, and such that $g(V_0) \subset D$. Such a neighborhood of 0 exists simply because g is continuous. Let $U_0 = f^{-1}(V_0)$ be the set of all $z \in D$ such that $f(z) \subset V_0$. Let

$$f_0 \colon U_0 \to V_0$$

be the restriction of f to U_0. We claim that f_0 is an analytic isomorphism. Note that $g(V_0) \subset U_0$ because for $w \in V_0$ we have $f\big(g(w)\big) = w$ by Theorem 3.4, so we consider the restriction g_0 of g to V_0 as mapping

$$g_0 \colon V_0 \to U_0.$$

Again by Theorem 3.4, for $z \in U_0$ we have $g_0\big(f_0(z)\big) = z$, which proves that f_0 and g_0 are inverse to each other, and concludes the proof of Theorem 6.1(c) in case $z_0 = 0$ and $f(z_0) = 0$.

The general case is reduced to the above case by translation, as one says. Indeed, for an arbitrary f, with $f(z) = \sum a_n(z - z_0)^n$, change variables and let

$$w = z - z_0, \qquad F(w) = f(z) - f(z_0) = \sum_{n=1}^{\infty} a_n w^n.$$

Then we may apply the previous special case to F and find a local inverse G for F. Let $w_0 = f(z_0)$, and let

$$g(w) = G(w - w_0) + z_0.$$

Then g is a local inverse for f, thus finishing the proof of Theorem 6.1.

There are (at least) four ways of proving the inverse function theorem.

1. The way we have just gone through, by estimating the formal inverse to show that it converges.

2. Reproducing the real variable proof for real functions of class C^1. By the contraction principle, (shrinking lemma), one first shows that the map is locally surjective, and one constructs a local inverse, which is shown to be differentiable, and whose derivative satisfies, for $w = f(z)$, the relation

$$g'(w) = 1/f'(z).$$

The reader should be able to copy the proof from any standard book on analysis, (certainly from my *Undergraduate Analysis* [La 83]).

3. Assuming the theorem for C^∞ real functions. One can show (and we shall do so later when we discuss the real aspects of an analytic function) that an analytic function is C^∞, as a function of (x, y), writing

$$z = x + iy.$$

The hypothesis $f'(z_0) \neq 0$ (namely $a_1 \neq 0$) is then seen to amount to the property that the Jacobian of the real function of two variables has non-zero determinant, whence f has a C^∞ inverse locally by the real theorem. It is then an easy matter to show by the chain rule that this inverse satisfies the Cauchy–Riemann equations, and is therefore holomorphic, whence analytic by the theory which follows Cauchy's theorem.

4. Giving an argument based on more complex function theory, and carried out in Theorem 1.7 of Chapter VI.

All four methods are important, and are used in various contexts in analysis, both of functions of one variable, and functions of several variables.

Let U be an open set and let f be a function on U. We say that f is an **open mapping** if for every open subset U' of U the image $f(U')$ is open.

Theorem 6.1 shows that the particular type of function considered there, i.e. with non-zero first coefficient in the power series expansion, is locally open. We shall now consider arbitrary analytic functions, first at the origin.

Let

$$f(z) = \sum a_n z^n$$

be a convergent non-constant power series, and let $m = \operatorname{ord} f$, so that

$$\begin{aligned} f(z) &= a_m z^m + \text{higher terms}, \qquad a_m \neq 0. \\ &= a_m z^m (1 + h(z)), \end{aligned}$$

where $h(z)$ is convergent, and has zero constant term. Let a be a complex number such that $a^m = a_m$. Then we can write $f(z)$ in the form

$$f(z) = \big(az(1 + h_1(z))\big)^m,$$

where $h_1(z)$ is a convergent power series with zero constant term, obtained from the binomial expansion

$$\big(1 + h(z)\big)^{1/m} = 1 + h_1(z),$$

and

$$f_1(z) = az\big(1 + h_1(z)\big) = az + azh_1(z)$$

is a power series whose coefficient of z is $a \neq 0$. Theorem 6.1 therefore applies to $f_1(z)$, which is therefore locally open at the origin. We have

$$f(z) = f_1(z)^m.$$

Let U be an open disc centered at the origin on which f_1 converges. Then $f_1(U)$ contains an open disc V. The image of V under the map

$$w \mapsto w^m$$

is a disc. Hence $f(U)$ contains an open disc centered at the origin.

Theorem 6.2. *Let f be analytic on an open set U, and assume that for each point of U, f is not constant on a given neighborhood of that point. Then f is an open mapping.*

Proof. We apply the preceding discussion to the power series expansion of f at a point of U, so the proof is obvious in the light of what we have already done.

The construction in fact yielded the following statement which it is worthwhile extracting as a theorem.

Theorem 6.3. *Let f be analytic at a point z_0,*

$$f(z) = a_0 + \sum_{n=m}^{\infty} a_n(z - z_0)^n,$$

with $m \geqq 1$ and $a_m \neq 0$. Then there exists a local analytic isomorphism φ at 0 such that

$$f(z) = a_0 + \varphi(z - z_0)^m.$$

We interpret Theorem 6.3 as follows. Let

$$\psi\colon U \to V$$

be an analytic isomorphism. We write $w = \psi(z)$. We may view ψ as a change of coordinates, from the coordinate z to the coordinate w. In Theorem 6.3 we may therefore write

$$f(z) = a_0 + w^m,$$

where $w = \varphi(z - z_0)$. The expansion for f in terms of the coordinate w is therefore much simpler than in terms of the coordinate z.

We also get a criterion for a function to have an analytic inverse on a whole open set.

Theorem 6.4. *Let f be analytic on an open set U, and assume that f is injective. Let $V = f(U)$ be its image. Then $f\colon U \to V$ is an analytic isomorphism, and $f'(z) \neq 0$ for all $z \in U$.*

Proof. The function f between U and V is bijective, so we can define an inverse mapping $g\colon V \to U$. Let z_0 be a point of U, and let the power series expansion of f at z_0 be as in Theorem 6.3. If $m > 1$ then we see that f cannot be injective, because the m-th power function in a neighborhood of the origin is not injective (it wraps the disc m times around). Hence $m = 1$, and Theorem 6.1 now shows that the inverse function g is analytic at $f(z_0)$. This proves the theorem.

Example 1. Let $f(z) = 3 - 5z +$ higher terms. Then $f(0) = 3$, and

$$f'(0) = a_1 = -5 \neq 0.$$

Hence f is a local analytic isomorphism, or locally invertible, at 0.

Example 2. Let $f(z) = 2 - 2z + z^2$. We want to determine whether f is locally invertible at $z = 1$. We write the power series expansion of f at 1, namely

$$f(z) = 1 + (z - 1)^2 = 1 + a_2(z - 1)^2.$$

Here we have $a_1 = 0$. Hence f is not locally invertible at $z = 1$.

Example 3. Let $f(z) = \cos z$. Determine whether f is locally invertible at $z = 0$. In this case,

$$f(z) = 1 - \frac{z^2}{2} + \text{higher terms},$$

so $a_1 = 0$ and f is not locally invertible.

Example 4. Let $f(z) = z^3$. Then $f'(z) = 3z^2$ and $f'(0) = 0$. Thus f is not locally invertible at 0. On the other hand, $f'(z) \neq 0$ if $z \neq 0$. Hence if $z_0 \neq 0$ then f is locally invertible at z_0. However, let U be the open set obtained by deleting the origin from **C**. Then f is not invertible on U. (Why?)

II, §6. EXERCISES

Determine which of the following functions are local analytic isomorphism at the given point. Give the reason for your answer.

1. $f(z) = e^z$ at $z = 0$.

2. $f(z) = \sin(z^2)$ at $z = 0$.

3. $f(z) = (z - 1)/(z - 2)$ at $z = 1$.

4. $f(z) = (\sin z)^2$ at $z = 0$.

5. $f(z) = \cos z$ at $z = \pi$.

6. **Linear Differential Equations.** Prove:

   Theorem. *Let $a_0(z), \ldots, a_k(z)$ be analytic functions in a neighborhood of 0. Assume that $a_0(0) \neq 0$. Given numbers $c_0, \ldots, c_{k-1}$, there exists a unique analytic function f at 0 such that*

   $$D^n f(0) = c_n \qquad \textit{for} \quad n = 0, \ldots, k-1$$

   and such that

   $$a_0(z)D^k f(z) + a_1(z)D^{k-1}f(z) + \cdots + a_k(z)f(z) = 0.$$

   [*Hint*: First you may assume $a_0(z) = 1$ (why?). Then solve for f by a formal power series. Then prove this formal series converges.]

7. **Ordinary Differential Equations.** Prove:

   Theorem. *Let g be analytic at 0. There exists a unique analytic function f at 0 satisfying*

   $$f(0) = 0, \qquad \textit{and} \qquad f'(z) = g(f(z)).$$

   [*Hint*: Again find a formal solution, and then prove that it converges.]

[*Note*: You will find the above two problems worked out in the Appendix, §3, but please try to do them first before looking up the solutions.]

II, §7. THE LOCAL MAXIMUM MODULUS PRINCIPLE

This principle is an immediate application of the open mapping theorem, and so we give it here, to emphasize its direct dependence with the preceding section. On the other hand, we wait for a later chapter for less basic applications mostly for psychological reasons. We want to alternate the formal operations with power series and the techniques which will arise from Cauchy's theorem. The later chapter could logically be read almost in its entirety after the present section, however.

We say that a function f is **locally constant** at a point z_0 if there exists an open set D (or a disc) containing z_0 such that f is constant on D.

Theorem 7.1. *Let f be analytic on an open set U. Let $z_0 \in U$ be a maximum for $|f|$, that is,*

$$|f(z_0)| \geqq |f(z)|, \qquad \textit{for all} \quad z \in U.$$

Then f is locally constant at z_0.

Proof. The function f has a power series expansion at z_0,

$$f(z) = a_0 + a_1(z - z_0) + \cdots.$$

If f is not the constant $a_0 = f(z_0)$, then by Theorem 6.2 we know that f is an open mapping in a neighborhood of z_0, and therefore the image of f contains a disc $D(a_0, s)$ of radius $s > 0$, centered at a_0. Hence the set of numbers $|f(z)|$, for z in a neighborhood of z_0, contains an open interval around a_0, so $|f(z)| > |f(z_0)|$ for some z. Hence

$$|f(z_0)| = |a_0|$$

cannot be a maximum for f, a contradiction which proves the theorem.

Corollary 7.2. *Let f be analytic on an open set U, and let $z_0 \in U$ be a maximum for the real part* Re f, *that is,*

$$\operatorname{Re} f(z_0) \geqq \operatorname{Re} f(z), \qquad \textit{for all} \quad z \in U.$$

Then f is locally constant at z_0.

Proof. The function $e^{f(z)}$ is analytic on U, and if

$$f(z) = u(z) + iv(z)$$

is the expression of f in terms of its real and imaginary parts, then

$$|e^{f(z)}| = e^{u(z)}.$$

Hence a maximum for $\operatorname{Re} f$ is also a maximum for $|e^{f(z)}|$, and the corollary follows from the theorem.

The theorem is often applied when f is analytic on an open set U and is continuous at the boundary of U. Then a maximum for $|f(z)|$ necessarily occurs on the boundary of U. For this one needs that U is connected, and the relevant form of the theorem will be proved as Theorem 1.3 of the next chapter.

We shall give here one more example of the power of the maximum modulus principle, and postpone to a later chapter some of the other applications.

Theorem 7.3. *Let*

$$f(z) = a_0 + a_1 z + \cdots + a_d z^d$$

be a polynomial, not constant, and say $a_d \neq 0$. Then f has some complex zero, i.e. a number z_0 such that $f(z_0) = 0$.

Proof. Suppose otherwise, so that $1/f(z)$ is defined for all z, and defines an analytic function. Writing

$$f(z) = a_d z^d \left(\frac{a_0}{a_d z^d} + \frac{a_1 z}{a_d z^d} + \cdots + 1 \right),$$

one sees that

$$\lim_{|z| \to \infty} 1/f(z) = 0.$$

Let α be some complex number such that $f(\alpha) \neq 0$. Pick a positive number R large enough such that $|\alpha| < R$, and if $|z| \geqq R$, then

$$\frac{1}{|f(z)|} < \frac{1}{|f(\alpha)|}.$$

Let S be the closed disc of radius R centered at the origin. Then S is closed and bounded, and $1/|f(z)|$ is continuous on S, whence has a maximum on S, say at z_0. By construction, this point z_0 cannot be on the boundary of the disc, and must be an interior point. By the maximum modulus principle, we conclude that $1/f(z)$ is locally constant at z_0. This is obviously impossible since f itself is not locally constant, say from the expansion

$$f(z) = b_0 + b_1(z - z_0) + \cdots + b_d(z - z_0)^d,$$

with suitable coefficients $b_0, \ldots, b_d$ and $b_d \neq 0$. This proves the theorem.

CHAPTER III

Cauchy's Theorem, First Part

III, §1. HOLOMORPHIC FUNCTIONS ON CONNECTED SETS

Let $[a, b]$ be a closed interval of real numbers. By a **curve** γ (defined on this interval) we mean a function

$$\gamma: [a, b] \to \mathbf{C}$$

which we assume to be of class C^1.

Figure 1

We recall what this means. We write

$$\gamma(t) = \gamma_1(t) + i\gamma_2(t),$$

where γ_1 is the real part of γ, and γ_2 is its imaginary part. For instance, the curve

$$\gamma(\theta) = \cos\theta + i\sin\theta, \qquad 0 \leqq \theta \leqq 2\pi,$$

is the unit circle. **Of class** C^1 means that the functions $\gamma_1(t)$, $\gamma_2(t)$ have continuous derivatives in the ordinary sense of calculus. We have drawn a curve in Fig. 1. Thus a curve is a parametrized curve. We call $\gamma(a)$ the **beginning point**, and $\gamma(b)$ the **end point** of the curve. By a **point on the curve** we mean a point w such that $w = \gamma(t)$ for some t in the interval of definition of γ.

We define the derivative $\gamma'(t)$ in the obvious way, namely

$$\gamma'(t) = \gamma_1'(t) + i\gamma_2'(t).$$

It is easily verified as usual that the rules for the derivative of a sum, product, quotient, and chain rule are valid in this case, and we leave this as an exercise. In fact, prove systematically the following statements:

Let $F\colon [a, b] \to \mathbf{C}$ and $G\colon [a, b] \to \mathbf{C}$ be complex valued differentiable functions, defined on the same interval. Then:

$$(F + G)' = F' + G',$$
$$(FG)' = FG' + F'G,$$
$$(F/G)' = (GF' - FG')/G^2$$

(this quotient rule being valid only on the set where $G(t) \neq 0$).

Let $\psi\colon [c, d] \to [a, b]$ be a differentiable function. Then $\gamma \circ \psi$ is differentiable, and

$$(\gamma \circ \psi)'(t) = \gamma'(\psi(t))\psi'(t),$$

as illustrated on Fig. 2(i).

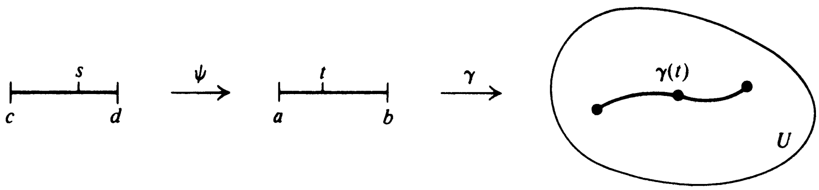


Figure 2(i)

Finally suppose γ is a curve in an open set U and

$$f\colon U \to \mathbf{C}$$

is a holomorphic function. Then the composite $f \circ \gamma$ is differentiable (as a function of the real variable t) and

$$(f \circ \gamma)'(t) = f'(\gamma(t))\gamma'(t),$$

as illustrated on Fig. 2(ii).

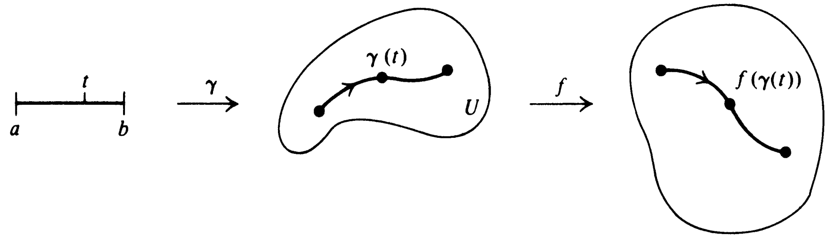


Figure 2(ii)

It is technically convenient to deal with a generalization of curves. By a **path** we shall mean a sequence of curves,

$$\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_n\}$$

(so each curve γ_j is C^1) such that the end point of γ_j is equal to the beginning point of γ_{j+1}. If γ_j is defined on the interval $[a_j, b_j]$, this means that

$$\gamma_j(b_j) = \gamma_{j+1}(a_{j+1}).$$

We have drawn a path on Fig. 3, where z_j is the end point of γ_j. We call $\gamma_1(a_1)$ the **beginning point** of γ, and $\gamma_n(b_n)$ the **end point** of γ. The path is said to **lie in an open set** U if each curve γ_j lies in U, i.e. for each t, the point $\gamma_j(t)$ lies in U.

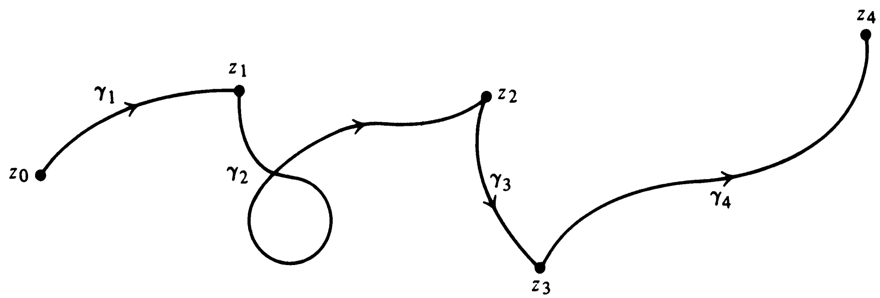


Figure 3

We define an open set U to be **connected** if given two points α and β in U, there exists a path $\{\gamma_1, \ldots, \gamma_n\}$ in U such that α is the beginning point of γ_1 and β is the end point of γ_n; in other words, if there exists a path in U which joins α to β. In Fig. 4 we have drawn an open set which is not connected. In Fig. 5 we have drawn a connected open set.

(The definition of connected applies of course equally well to a set which is not necessarily open. It is usually called **pathwise connected**, but for open sets, this coincides with another possible definition. See the appendix of this section.)

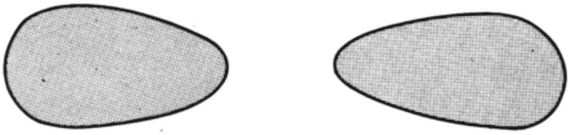


Figure 4

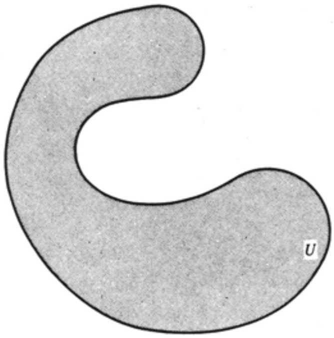


Figure 5

Theorem 1.1. *Let U be a connected open set, and let f be a holomorphic function on U. If $f' = 0$ then f is constant.*

Proof. Let α, β be two points in U, and suppose first that γ is a curve joining α to β, so that

$$\gamma(a) = \alpha \qquad \text{and} \qquad \gamma(b) = \beta.$$

The function

$$t \mapsto f(\gamma(t))$$

is differentiable, and by the chain rule, its derivative is

$$f'(\gamma(t))\gamma'(t) = 0.$$

Hence this function is constant, and therefore

$$f(\alpha) = f(\gamma(a)) = f(\gamma(b)) = f(\beta).$$

Next, suppose that $\gamma = \{\gamma_1, \ldots, \gamma_n\}$ is a path joining α to β, and let z_j be the end point of γ_j, putting

$$z_0 = \alpha, \qquad z_n = \beta.$$

By what we have just proved, we have

$$f(\alpha) = f(z_0) = f(z_1) = f(z_2) = \cdots = f(z_n) = f(\beta),$$

thereby proving the theorem.

If f is a function on an open set U and g is a holomorphic function on U such that $g' = f$, then we call g a **primitive** of f on U. Theorem 1.1 says that on a connected open set, a primitive of f is uniquely determined up to a constant, i.e. if g_1 and g_2 are two primitives, then $g_1 - g_2$ is constant, because the derivative of $g_1 - g_2$ is equal to 0.

In what follows we shall attempt to get primitives by integration. On the other hand, primitives can also be written down directly.

Example. For each integer $n \neq -1$, the function $f(z) = z^n$ has the usual primitive

$$\frac{z^{n+1}}{n+1}.$$

Let S be a set of points, and let $z_0 \in S$. We say that z_0 is **isolated in** S if there exists a disc $D(z_0, r)$ of some radius $r > 0$ such that $D(z_0, r)$ does not contain any point of S other than z_0. We say that S is **discrete** if every point of S is isolated.

Theorem 1.2. *Let U be a connected open set.*

(i) *If f is analytic on U and not constant, then the set of zeros of f on U is discrete.*

(ii) *Let f, g be analytic on U. Let S be a set of points in U which is not discrete (so some point of S is not isolated). Assume that $f(z) = g(z)$ for all z in S. Then $f = g$ on U.*

Proof. We observe that (ii) follows from (i). It suffices to consider the difference $f - g$. Therefore we set about to prove (i). We know from Theorem 3.2 of the preceding chapter that either f is locally constant and equal to 0 in the neighborhood of a zero z_0, or z_0 is an isolated zero.

Suppose that f is equal to 0 in the neighborhood of some point z_0. We have to prove that $f(z) = 0$ for all $z \in U$. Let S be the set of points z such that f is equal to 0 in a neighborhood of z. Then S is open. By

Theorem 1.6 below, it will suffice to prove that S is closed in U. Let z_1 be a point in the closure of S in U. Since f is continuous, it follows that $f(z_1) = 0$. If z_1 is not in S, then there exist points of S arbitrarily close to z_1, and by Theorem 3.2 of the preceding chapter, it follows that f is locally equal to 0 in a neighborhood of z_1. Hence in fact $z_1 \in S$, so S is closed in U. This concludes the proof.

Remarks. The argument using open and closed subsets of U applies in very general situations, and shows how to get a global statement on a connected set U knowing only a local property as in Theorem 3.2 of the preceding chapter.

It will be proved in Chapter V, §1, that a function is holomorphic if and only if it is analytic. Thus Theorem 1.2 will also apply to holomorphic functions.

The second part of Theorem 1.2 will be used later in the study of analytic continuation, but we make some comments here in anticipation. Let f be an analytic function defined on an open set U and let g be an analytic function defined on an open set V. Suppose that U and V have a non-empty intersection, as illustrated on Fig. 6. If U, V are connected, and if $f(z) = g(z)$ for all $z \in U \cap V$, i.e. if f and g are equal on the intersection $U \cap V$, then Theorem 1.2 tells us that g is the only possible analytic function on V having this property. In the applications, we shall be interested in extending the domain of definition of an analytic function f, and Theorem 1.2 guarantees the uniqueness of the extended function. We say that g is the **analytic continuation** of f to V.

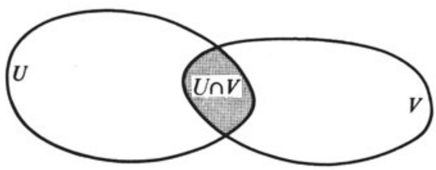


Figure 6

It is also appropriate here to formulate the global version of the **maximum modulus principle**.

Theorem 1.3. *Let U be a connected open set, and let f be an analytic function on U. If $z_0 \in U$ is a maximum point for $|f|$, that is*

$$|f(z_0)| \geqq |f(z)|$$

for all $z \in U$, then f is constant on U.

Proof. By Theorem 6.1 of the preceding chapter, we know that f is locally constant at z_0. Therefore f is constant on U by Theorem 1.2(ii) (compare the constant function and f). This concludes the proof.

Corollary 1.4. *Let U be a connected open set and U^c its closure. Let f be a continuous function on U^c, analytic and non-constant on U. If z_0 is a maximum for f on U^c, that is, $|f(z_0)| \geqq |f(z)|$ for all $z \in U^c$, then z_0 lies on the boundary of U^c.*

Proof. This comes from a direct application of Theorem 1.3.

Remark. If U^c is closed and bounded, then a continuous function has a maximum on U^c, so a maximum for f always exists in Corollary 1.4.

Appendix: Connectedness

The purpose of this appendix is to put together a couple of statements describing connectedness in various terms. Essentially we want to prove that two possible definitions of connectedness are equivalent. For purposes of this appendix, we use the words pathwise connected for the notion we have already defined. Let U be an open set in the complex numbers. We say that U is **topologically connected** if U cannot be expressed as a union $U = V \cup W$, where V, W are open, non-empty, and disjoint. We start with what amounts to a remark. Let S be a subset of U. We say that S is **closed in** U if given $z \in U$ and z in the closure of S, then $z \in S$.

Lemma 1.5. *Let S be a subset of an open set U. Then S is closed in U if and only if the complement of S in U is open, that is, $U - S$ is open. In particular, if S is both open and closed in U, then $U - S$ is also open and closed in U.*

Proof. Exercise 1.

Theorem 1.6. *Let U be an open set. Then U is pathwise connected if and only if U is topologically connected.*

Proof of Theorem 1.6. Assume that U is pathwise connected. We want to prove that U is topologically connected. Suppose not. Then $U = V \cup W$ where V, W are non-empty and open. Let $z_1 \in V$ and $z_2 \in W$. By assumption there exists a path $\gamma\colon [a, b] \to U$ such that $\gamma(a) = z_1$ and $\gamma(b) = z_2$. Let T be the set of $t \in [a, b]$ such that $\gamma(t) \in V$. Then T is not empty because $a \in T$, and T is bounded by b. Let c be the least upper bound of T. Then $c \neq b$. By definition of an upper bound, there exists a sequence of real numbers t_n, with $c < t_n \leqq b$ such that $\gamma(t_n) \in W$, and t_n converges to c. Since γ is continuous, it follows that $\gamma(c) = \lim \gamma(t_n)$, and since W is closed in U, it follows that $\gamma(c) \in W$. On the other hand, by definition of a least upper bound, there exists a sequence of real numbers s_n with $a \leqq s_n \leqq c$ such that s_n converges to c, and $\gamma(s_n) \in V$. Since γ is continuous, it follows that $\gamma(c) = \lim \gamma(s_n)$, and since V is closed in U, it follows that $\gamma(c) \in V$, which is a contradiction proving that U is topologically connected.

Conversely, assume U is topologically connected. We want to prove that U is pathwise connected. Let $z_0 \in U$. Let V be the set of points in U which can be joined to z_0 by a path in U. Then V is open. Indeed, suppose that there is a path in U joining z_0 to z_1. Since U is open, there exists a disc $D(z_1, r)$ of radius $r > 0$ contained in U. Then every element of this disc can be joined to z_1 by a line segment in the disc, and can therefore be joined to z_0 by a path in U, so V is open. We assert further that V is closed. To see this, let $\{z_n\}$ be a sequence in V converging to a point u in U. Since U is open, there exists a disc $D(u, r)$ of radius $r > 0$ contained in U. For some n the point z_n lies in $D(u, r)$. Then there is a line segment in $D(u, r)$ joining u and z_n, and so u can be joined by a path to z_0. This proves that V is closed. Hence V is both open and closed, and by assumption, $V = U$. This proves that U is pathwise connected, and concludes the proof of Theorem 1.6.

Warning. The equivalence of the two notions of connectedness for open sets may not be valid for other types of sets. For instance, consider the set consisting of the horizontal positive x-axis, together with vertical segments of length 1 above the points $1, 1/2, 1/3, \ldots, 1/n, \ldots$ and also above 0. Now delete the origin. The remaining set is topologically connected but not pathwise connected. Draw the picture! Also compare with inaccessible points as in Chapter X, §4.

III, §1. EXERCISES

1. Prove Lemma 1.5.

2. Let U be a bounded open connected set, $\{f_n\}$ a sequence of continuous functions on the closure of U, analytic on U. Assume that $\{f_n\}$ converges uniformly on the boundary of U. Prove that $\{f_n\}$ converges uniformly on U.

3. Let $a_1, \ldots, a_n$ be points on the unit circle. Prove that there exists a point z on the unit circle so that the product of the distances from z to the a_j is at least 1. (You may use the maximum principle.)

III, §2. INTEGRALS OVER PATHS

Let $F\colon [a, b] \to \mathbf{C}$ be a continuous function.

Write F in terms of its real and imaginary parts, say

$$F(t) = u(t) + iv(t).$$

Define the **indefinite integral** by

$$\int F(t)\,dt = \int u(t)\,dt + i\int v(t)\,dt.$$

Verify that integration by parts is valid (assuming that F' and G' exist and are continuous), namely

$$\int F(t)G'(t)\,dt = F(t)G(t) - \int G(t)F'(t)\,dt.$$

(The proof is the same as in ordinary calculus, from the derivative of a product.)

We define the **integral of** F over $[a, b]$ to be

$$\int_a^b F(t)\,dt = \int_a^b u(t)\,dt + i\int_a^b v(t)\,dt.$$

Thus the integral is defined in terms of the ordinary integrals of the real functions u and v. Consequently, by the fundamental theorem of calculus *the function*

$$t \mapsto \int_a^t F(s)\,ds$$

is differentiable, and its derivative is $F(t)$, because this assertion is true if we replace F by u and v, respectively.

Using simple properties of the integral of real-valued functions, one has the inequality

$$\left|\int_a^b F(t)\,dt\right| \leqq \int_a^b |F(t)|\,dt.$$

Work it out as Exercise 11.

Let f be a continuous function on an open set U, and suppose that γ is a curve in U, meaning that all values $\gamma(t)$ lie in U for $a \leqq t \leqq b$. We define the **integral of** f **along** γ to be

$$\boxed{\int_\gamma f = \int_a^b f(\gamma(t))\gamma'(t)\,dt.}$$

This is also frequently written

$$\int_\gamma f(z)\,dz.$$

Example 1. Let $f(z) = 1/z$. Let $\gamma(\theta) = e^{i\theta}$. Then

$$\gamma'(\theta) = ie^{i\theta}.$$

We want to find the value of the integral of f over the circle,

$$\int_\gamma \frac{1}{z}\,dz$$

so $0 \leqq \theta \leqq 2\pi$. By definition, this integral is equal to

$$\int_0^{2\pi} \frac{1}{e^{i\theta}} ie^{i\theta}\,d\theta = i\int_0^{2\pi} d\theta = 2\pi i.$$

As in calculus, we have defined the integral over parametrized curves. In practice, we sometimes describe a curve without giving an explicit parametrization. The context should always make it clear what is meant. Furthermore, one can also easily see that the integral is independent of the parametrization, in the following manner

Let

$$g\colon [a, b] \to [c, d]$$

be a C^1 function, such that $g(a) = c$, $g(b) = d$, and let

$$\psi\colon [c, d] \to \mathbf{C}$$

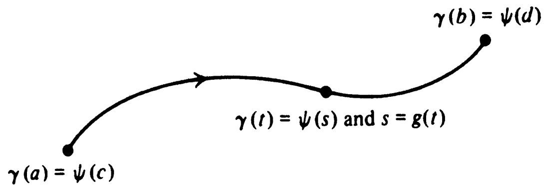


Figure 7

be a curve. Then we may form the composed curve

$$\gamma(t) = \psi(g(t)).$$

We find:

$$\begin{aligned}
\int_\gamma f &= \int_a^b f(\gamma(t))\gamma'(t)\,dt \\
&= \int_a^b f(\psi(g(t)))\psi'(g(t))g'(t)\,dt \\
&= \int_c^d f(\psi(s))\psi'(s)\,ds \\
&= \int_\psi f.
\end{aligned}$$

Thus the integral of f along the curve is independent of the parametrization.

If $\gamma = \{\gamma_1, \ldots, \gamma_n\}$ is a path, then we define

$$\int_\gamma f = \sum_{i=1}^{n} \int_{\gamma_i} f,$$

to be the sum of the integrals of f over each curve γ_i of the path.

Theorem 2.1. *Let f be continuous on an open set U, and suppose that f has a primitive g, that is, g is holomorphic and $g' = f$. Let α, β be two points of U, and let γ be a path in U joining α to β. Then*

$$\int_\gamma f = g(\beta) - g(\alpha),$$

and in particular, this integral depends only on the beginning and end point of the path. It is independent of the path itself.

Proof. Assume first that the path is a curve. Then

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

By the chain rule, the expression under the integral sign is the derivative

$$\frac{d}{dt} g(\gamma(t)).$$

Hence by ordinary calculus, the integral is equal to

$$g(\gamma(t))\Big|_a^b = g(\gamma(b)) - g(\gamma(a)),$$

which proves the theorem in this case. In general, if the path consists of curves $\gamma_1, \ldots, \gamma_n$, and z_j is the end point of γ_j, then by the case we have just settled, we find

$$\begin{aligned}\int_\gamma f &= g(z_1) - g(z_0) + g(z_2) - g(z_1) + \cdots + g(z_n) - g(z_{n-1})\\ &= g(z_n) - g(z_0),\end{aligned}$$

which proves the theorem.

Example 2. Let $f(z) = z^3$. Then f has a primitive, $g(z) = z^4/4$. Hence the integral of f from $2 + 3i$ to $1 - i$ over *any path* is equal to

$$\frac{(1-i)^4}{4} - \frac{(2+3i)^4}{4}.$$

Example 3. Let $f(z) = e^z$. Find the integral of f from 1 to $i\pi$ taken over a line segment. Here again $f'(z) = f(z)$, so f has a primitive. Thus the integral is independent of the path and equal to $e^{i\pi} - e^1 = -1 - e$.

By a **closed path**, we mean a path whose beginning point is equal to its end point. We may now give an important example of the theorem:

If f is a continuous function on U admitting a holomorphic primitive g, and γ is any closed path in U, then

$$\int_\gamma f = 0.$$

Example 4. Let $f(z) = z^n$, where n is an integer $\neq -1$. Then for any closed path γ (or any closed path not passing through the origin if n is negative), we have

$$\int_\gamma z^n\,dz = 0.$$

This is true because z^n has the primitive $z^{n+1}/(n+1)$. [When n is negative, we have to assume that the closed path does not pass through the origin, because the function is then not defined at the origin.]

Putting this together with Example 1, we have the following tabulation. Let C_R be the circle of radius R centered at the origin oriented counterclockwise. Let n be an integer. Then:

$$\int_{C_R} z^n\,dz = \begin{cases} 0 & \text{if } n \neq -1, \\ 2\pi i & \text{if } n = -1. \end{cases}$$

Of course, in Example 1 we did the computation when $R = 1$, but you can check that one gets the same value for arbitrary R. In the exercises, you can check similar values for the integral around a circle centered around any point z_0.

We shall see later that holomorphic functions are analytic. In that case, in the domain of convergence a power series

$$\sum a_n(z - z_0)^n$$

can be integrated term by term, and thus integrals of holomorphic functions are reduced to integrals of polynomials. This is the reason why there is no need here to give further examples.

Theorem 2.2. *Let U be a connected open set, and let f be a continuous function on U. If the integral of f along any closed path in U is equal to* 0, *then f has a primitive g on U, that is, a function g which is holomorphic such that $g' = f$.*

Proof. Pick a point z_0 in U and define

$$g(z) = \int_{z_0}^{z} f,$$

where the integral is taken along any path from z_0 to z in U. If γ, η are two such paths, and η^- is the reverse path of η (cf. Exercise 9), then $\{\gamma, \eta^-\}$ is a closed path, and by Exercise 9 we know that

$$\int_\gamma f = \int_\eta f.$$

Therefore the integral defining g is independent of the path from z_0 to z, and defines the function. We have

$$\frac{g(z + h) - g(z)}{h} = \frac{1}{h}\int_{z}^{z+h} f(\zeta)\,d\zeta,$$

and the integral from z to $z + h$ can be taken along a segment in U from z to $z + h$. Write

$$f(\zeta) = f(z) + \varphi(\zeta),$$

where $\lim_{\zeta \to z} \varphi(\zeta) = 0$ (this can be done by the continuity of f at z). Then

$$\begin{aligned} \frac{1}{h}\int_z^{z+h} f(\zeta)\,d\zeta &= \frac{1}{h}\int_z^{z+h} f(z)\,d\zeta + \frac{1}{h}\int_z^{z+h} \varphi(\zeta)\,d\zeta \\ &= f(z) + \frac{1}{h}\int_z^{z+h} \varphi(\zeta)\,d\zeta. \end{aligned}$$

The length of the interval from z to $z + h$ is $|h|$. Hence the integral on the right is estimated by (see below, Theorem 2.3)

$$\frac{1}{|h|}|h| \max |\varphi(\zeta)|,$$

where the max is taken for ζ on the interval. This max tends to 0 as $h \to 0$, and this proves the theorem.

Remarks. The reader should recognize Theorems 2.1 and 2.2 as being the exact analogues for (complex) differentiable functions of the standard theorems of advanced calculus concerning the relation between the existence of a primitive (potential function for a vector field), and the independence of the integral (of a vector field) from the path. We shall see later that a holomorphic function is infinitely complex differentiable, and therefore that f itself is analytic.

Let γ be a curve, $\gamma\colon [a, b] \to \mathbf{C}$, assumed of class C^1 as always. The **speed** is defined as usual to be $|\gamma'(t)|$, and the **length** $L(\gamma)$ is defined to be the integral of the speed,

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

If $\gamma = \{\gamma_1, \ldots, \gamma_n\}$ is a path, then by definition

$$L(\gamma) = \sum_{i=1}^{n} L(\gamma_i).$$

Let f be a bounded function on a set S. We let $\|f\|$ be the **sup norm**, written $\|f\|_S$ if the reference to S needs to be made for clarity, so that

$$\|f\| = \sup_{z \in S} |f(z)|$$

is the least upper bound of the values $|f(z)|$ for $z \in S$.

Let f be continuous on an open set U. By standard results of elementary real analysis, Theorem 4.3 of Chapter I, the image of a curve or a path γ is closed and bounded, i.e. compact. If the curve is in U, then the function

$$t \mapsto f(\gamma(t))$$

is continuous, and hence f is bounded on the image of γ. By the compactness of the image of γ, we can always find an open subset of U containing γ, on which f is bounded. If γ is defined on $[a, b]$, we let

$$\|f\|_\gamma = \max_{t \in [a,b]} |f(\gamma(t))|.$$

Theorem 2.3. *Let f be a continuous function on U. Let γ be a path in U. Then*

$$\left| \int_\gamma f \right| \leqq \|f\|_\gamma L(\gamma).$$

Proof. If γ is a curve, then

$$\begin{aligned} \left| \int_\gamma f \right| &= \left| \int_a^b f(\gamma(t))\gamma'(t)\, dt \right| \\ &\leqq \int_a^b |f(\gamma(t))|\,|\gamma'(t)|\, dt \\ &\leqq \|f\|_\gamma L(\gamma), \end{aligned}$$

as was to be shown. The statement for a path follows by taking an appropriate sum.

Theorem 2.4. *Let $\{f_n\}$ be a sequence of continuous functions on U, converging uniformly to a function f. Then*

$$\lim \int_\gamma f_n = \int_\gamma f.$$

If $\sum f_n$ is a series of continuous functions converging uniformly on U, *then*

$$\int_\gamma \sum f_n = \sum \int_\gamma f_n.$$

Proof. The first assertion is immediate from the inequality.

$$\left|\int_\gamma f_n - \int_\gamma f\right| \leqq \int_\gamma |f_n - f| \leqq \|f_n - f\| L(\gamma).$$

The second follows from the first because uniform convergence of a series is defined in terms of the uniform convergence of its partial sums,

$$s_n = f_1 + \cdots + f_n.$$

This proves the theorem.

Example 5. Let f be analytic on an open set containing the closed disc $\bar{D}(0, R)$ of radius R centered at the origin, except possibly at the origin. Suppose f has a power series expansion

$$f(z) = \frac{a_{-m}}{z^m} + \cdots + \frac{a_{-1}}{z} + a_1 z + \cdots + a_n z^n + \cdots$$

possibly with negative terms, such that the series with non-negative terms

$$\sum_{n=0}^{\infty} a_n z^n$$

has a radius of convergence $> R$. Let C_R be the circle of radius R centered at the origin. Then

$$\int_{C_R} f(z)\,dz = 2\pi i a_{-1}.$$

This is a special case of Theorem 2.4 and Example 4, by letting

$$f_n(z) = \sum_{k=-m}^{n} a_k z^k.$$

Each f_n is a finite sum, so the integral of f_n is the sum of the integrals of the individual terms, which were evaluated in Example 4.

III, §2. EXERCISES

1. (a) Given an arbitrary point z_0, let C be a circle of radius $r > 0$ centered at z_0, oriented counterclockwise. Find the integral

$$\int_C (z - z_0)^n \, dz$$

for all integers n, positive or negative.

(b) Suppose f has a power series expansion

$$f(z) = \sum_{k=-m}^{\infty} a_k (z - z_0)^k,$$

which is absolutely convergent on a disc of radius $> R$ centered at z_0. Let C_R be the circle of radius R centered at z_0. Find the integral

$$\int_{C_R} f(z) \, dz.$$

2. Find the integral of $f(z) = e^z$ from -3 to 3 taken along a semicircle. Is this integral different from the integral taken over the line segment between the two points?

3. Sketch the following curves with $0 \leqq t \leqq 1$.
   (a) $\gamma(t) = 1 + it$
   (b) $\gamma(t) = e^{-\pi it}$
   (c) $\gamma(t) = e^{\pi it}$
   (d) $\gamma(t) = 1 + it + t^2$

4. Find the integral of each one of the following functions over each one of the curves in Exercise 3.
   (a) $f(z) = z^3$
   (b) $f(z) = \bar{z}$
   (c) $f(z) = 1/z$

5. Find the integral

$$\int_\gamma z e^{z^2} \, dz$$

   (a) from the point i to the point $-i + 2$, taken along a straight line segment, and
   (b) from 0 to $1 + i$ along the parabola $y = x^2$.

6. Find the integral

$$\int_\gamma \sin z \, dz$$

from the origin to the point $1 + i$, taken along the parabola

$$y = x^2.$$

7. Let σ be a vertical segment, say parametrized by

$$\sigma(t) = z_0 + itc, \qquad -1 \leqq t \leqq 1,$$

where z_0 is a fixed complex number, and c is a fixed real number > 0. (Draw the picture.) Let $\alpha = z_0 + x$ and $\alpha' = z_0 - x$, where x is real positive. Find

$$\lim_{x \to 0} \int_\sigma \left(\frac{1}{z - \alpha} - \frac{1}{z - \alpha'} \right) dz.$$

(Draw the picture.) *Warning*: The answer is not 0!

8. Let $x > 0$. Find the limit:

$$\lim_{B \to \infty} \int_{-B}^{B} \left(\frac{1}{t + ix} - \frac{1}{t - ix} \right) dt.$$

9. Let $\gamma \colon [a, b] \to \mathbf{C}$ be a curve. Define the **reverse** or opposite curve to be

$$\gamma^- \colon [a, b] \to \mathbf{C}$$

such that $\gamma^-(t) = \gamma(a + b - t)$. Show that

$$\int_{\gamma^-} F = -\int_\gamma F.$$

10. Let $[a, b]$ and $[c, d]$ be two intervals (not reduced to a point). Show that there is a function $g(t) = rt + s$ such that g is strictly increasing, $g(a) = c$ and $g(b) = d$. Thus a curve can be parametrized by any given interval.

11. Let F be a continuous complex-valued function on the interval $[a, b]$. Prove that

$$\left| \int_a^b F(t)\, dt \right| \leqq \int_a^b |F(t)|\, dt.$$

[*Hint*: Let $P = [a = a_0, a_1, \ldots, a_n = b]$ be a partition of $[a, b]$. From the definition of integrals with Riemann sums, the integral

$$\int_a^b F(t)\, dt \quad \text{is approximated by the Riemann sum} \quad \sum_{k=0}^{n-1} F(a_k)(a_{k+1} - a_k)$$

whenever $\max(a_{k+1} - a_k)$ is small, and

$$\int_a^b |F(t)|\, dt \quad \text{is approximated by} \quad \sum_{k=0}^{n-1} |F(a_k)|(a_{k+1} - a_k).$$

The proof is concluded by using the triangle inequality.]

III, §3. LOCAL PRIMITIVE FOR A HOLOMORPHIC FUNCTION

Let U be a connected open set, and let f be holomorphic on U. Let $z_0 \in U$. We want to define a primitive for f on some open disc centered at z_0, i.e. locally at z_0. The natural way is to define such a primitive by an integral,

$$g(z) = \int_{z_0}^{z} f(\zeta)\, d\zeta,$$

taken along some path from z_0 to z. However, the integral may depend on the path.

It turns out that we may define g locally by using only a special type of path. Indeed, suppose U is a disc centered at z_0. Let $z \in U$. We select for a path from z_0 to z the edges of a rectangle as shown on Fig. 8.

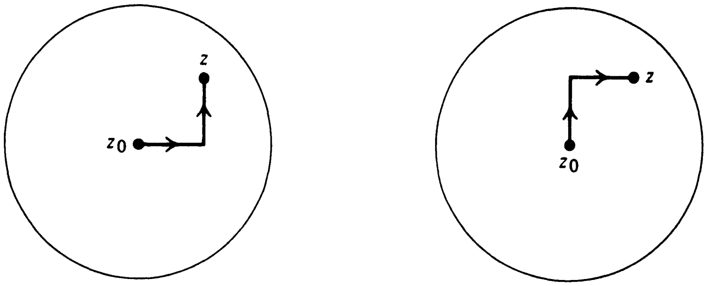


Figure 8

We then have restricted our choice of path to two possible choices as shown. We shall see that we get the same value for the integrals in the two cases. It will be shown afterwards that the integral then gives us a primitive.

By a **rectangle** R we shall mean a rectangle whose sides are vertical or horizontal, and R is meant as the set of points inside and on the boundary of the rectangle, so R is assumed to be closed. The path describing the boundary of the rectangle taken counterclockwise will be also called the **boundary of the rectangle**, and will be denoted by

$$\partial R.$$

If S is an arbitrary set of points, we say that a function f is **holomorphic on** S if it is holomorphic on some open set containing S.

Theorem 3.1 (Goursat). *Let R be a rectangle, and let f be a function holomorphic on R. Then*

$$\int_{\partial R} f = 0.$$

Proof. Decompose the rectangle into four rectangles by bisecting the sides, as shown on Fig. 9.

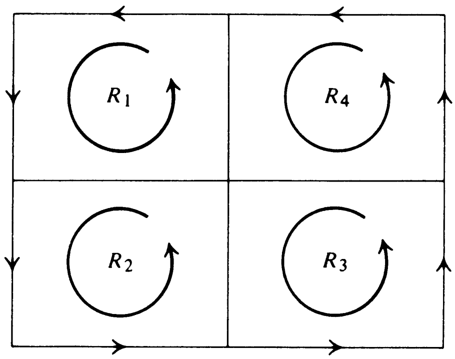


Figure 9

Then

$$\int_{\partial R} f = \sum_{i=1}^{4} \int_{\partial R_i} f.$$

Consequently,

$$\left| \int_{\partial R} f \right| \leqq \sum_{i=1}^{4} \left| \int_{\partial R_i} f \right|,$$

and there is one rectangle, say $R^{(1)}$, among R_1, R_2, R_3, R_4 such that

$$\left| \int_{\partial R^{(1)}} f \right| \geqq \frac{1}{4} \left| \int_{\partial R} f \right|.$$

Next we decompose $R^{(1)}$ into four rectangles, again bisecting the sides of $R^{(1)}$ as shown on Fig. 10.

For one of the four rectangles thus obtained, say $R^{(2)}$, we have the similar inequality

$$\left| \int_{\partial R^{(2)}} f \right| \geqq \frac{1}{4} \left| \int_{\partial R^{(1)}} f \right|.$$

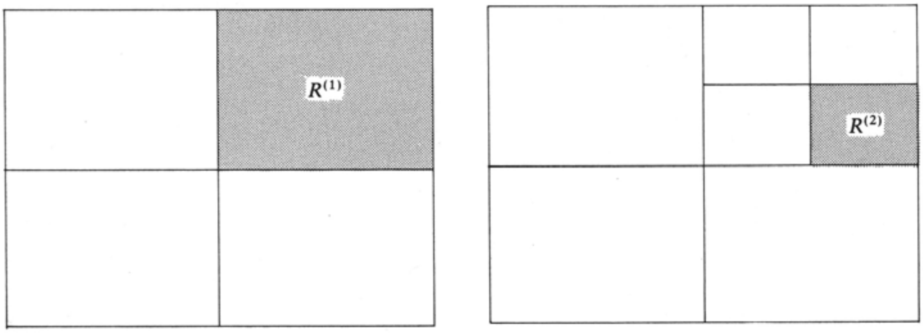


Figure 10

We continue in this way, to obtain a sequence of rectangles

$$R^{(1)} \supset R^{(2)} \supset R^{(3)} \supset \cdots$$

such that

$$\left| \int_{\partial R^{(n+1)}} f \right| \geqq \frac{1}{4} \left| \int_{\partial R^{(n)}} f \right|.$$

Then

$$\left| \int_{\partial R^{(n)}} f \right| \geqq \frac{1}{4^n} \left| \int_{\partial R} f \right|.$$

On the other hand, let L_n be the length of $\partial R^{(n)}$. Then

$$L_{n+1} = \frac{1}{2} L_n$$

so that by induction,

$$L_n = \frac{1}{2^n} L_0,$$

where $L_0 =$ length of ∂R.

We contend that the intersection

$$\bigcap_{n=1}^{\infty} R^{(n)}$$

consists of a single point z_0. Since the diameter of $R^{(n)}$ tends to 0 as n becomes large, it is immediate that there is at most one point in the intersection. Let α_n be the center of $R^{(n)}$. Then the sequence $\{\alpha_n\}$ is a

Cauchy sequence, because given ϵ, let N be such that the diameter of $R^{(N)}$ is less than ϵ. If $n, m \geqq N$, then α_n, α_m lie in $R^{(N)}$ and so

$$|\alpha_n - \alpha_m| \leqq \operatorname{diam} R^{(N)} < \epsilon.$$

Let $z_0 = \lim \alpha_n$. Then z_0 lies in each rectangle, because each rectangle is closed. Hence z_0 lies in the intersection of the rectangles $R^{(N)}$ for $N = 1, 2, \ldots$, as desired. (See also Theorem 4.2 of Chapter I.)

Since f is differentiable at z_0, there is a disc V centered at z_0 such that for all $z \in V$ we have

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

where

$$\lim_{z \to z_0} h(z) = 0.$$

If n is sufficiently large, then $R^{(n)}$ is contained in V, and then

$$\int_{\partial R^{(n)}} f(z)\,dz = \int_{\partial R^{(n)}} f(z_0)\,dz + f'(z_0) \int_{\partial R^{(n)}} (z - z_0)\,dz$$
$$+ \int_{\partial R^{(n)}} (z - z_0)h(z)\,dz.$$

By Example 4 of §2, we know that the first two integrals on the right of this equality sign are 0. Hence

$$\int_{\partial R^{(n)}} f = \int_{\partial R^{(n)}} (z - z_0)h(z)\,dz,$$

and we obtain the inequalities

$$\frac{1}{4^n}\left|\int_{\partial R} f\right| \leqq \left|\int_{\partial R^{(n)}} f\right| \leqq \left|\int_{\partial R^{(n)}} (z - z_0)h(z)\,dz\right|$$
$$\leqq \frac{1}{2^n} L_0 \operatorname{diam} R^{(n)} \sup|h(z)|,$$

where the sup is taken for all $z \in R^{(n)}$. But $\operatorname{diam} R^{(n)} = (1/2^n) \operatorname{diam} R$. This yields

$$\left|\int_{\partial R} f\right| \leqq L_0 \operatorname{diam} R \sup|h(z)|.$$

The right-hand side tends to 0 as n becomes large, and consequently

$$\int_{\partial R} f = 0,$$

as was to be shown.

We carry out the program outlined at the beginning of the section to find a primitive locally.

Theorem 3.2. *Let U be a disc centered at a point z_0. Let f be continuous on U, and assume that for each rectangle R contained in U we have*

$$\int_{\partial R} f = 0.$$

For each point z_1 in the disc, define

$$g(z_1) = \int_{z_0}^{z_1} f,$$

where the integral is taken along the sides of a rectangle R whose opposite vertices are z_0 and z_1. Then g is holomorphic on U and is a primitive for f, namely

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

Proof. We have

$$g(z_1 + h) - g(z_1) = \int_{z_1}^{z_1+h} f(z)\,dz$$

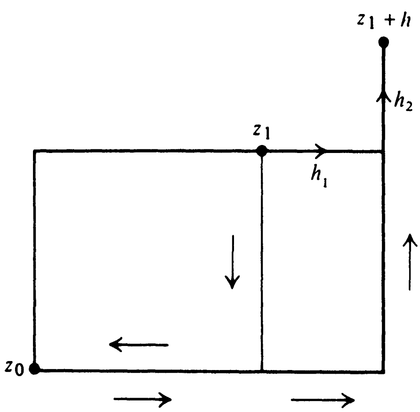


Figure 11

The integral between z_1 and $z_1 + h$ is taken over the bottom side h_1 and vertical side h_2 of the rectangle shown in Fig. 11. Since f is continuous at z_1, there exists a function $\psi(z)$ such that

$$\lim_{z \to z_1} \psi(z) = 0$$

and

$$f(z) = f(z_1) + \psi(z).$$

Then

$$\begin{aligned} g(z_1 + h) - g(z_1) &= \int_{z_1}^{z_1+h} f(z_1)\,dz + \int_{z_1}^{z_1+h} \psi(z)\,dz \\ &= hf(z_1) + \int_{z_1}^{z_1+h} \psi(z)\,dz. \end{aligned}$$

We divide by h and take the limit as $h \to 0$. The length of the path from z_1 to $z_1 + h$ is bounded by $|h_1| + |h_2|$. Hence we get a bound

$$\left| \frac{1}{h} \int_{z_1}^{z_1+h} \psi(z)\,dz \right| \leqq \frac{1}{|h|}(|h_1| + |h_2|) \sup|\psi(z)|,$$

where the sup is taken for z on the path of integration. The expression on the right therefore tends to 0 as $h \to 0$. Hence

$$\lim_{h \to 0} \frac{g(z_1 + h) - g(z_1)}{h} = f(z_1),$$

as was to be shown.

Knowing that a primitive for f exists on a disc U centered at z_0, we can now conclude that the integral of f along any path between z_0 and z in U is independent of the path, according to Theorem 2.1, and we find:

Theorem 3.3. *Let U be a disc and suppose that f is holomorphic on U. Then f has a primitive on U, and the integral of f along any closed path in U is 0.*

Remark. In Theorem 7.2 we shall prove that a holomorphic function is analytic. Applying this result to the function g in Theorem 3.2, we shall conclude that the function f in Theorem 3.2 is analytic. See Theorem 7.7.

III, §4. ANOTHER DESCRIPTION OF THE INTEGRAL ALONG A PATH

Knowing the existence of a local primitive for a holomorphic function allows us to describe its integral along a path in a way which makes no use of the differentiability of the path, and would apply to a continuous path as well. We start with curves.

Lemma 4.1. *Let $\gamma: [a, b] \to U$ be a continuous curve in an open set U. Then there is some positive number $r > 0$ such that every point on the curve lies at distance $\geqq r$ from the complement of U.*

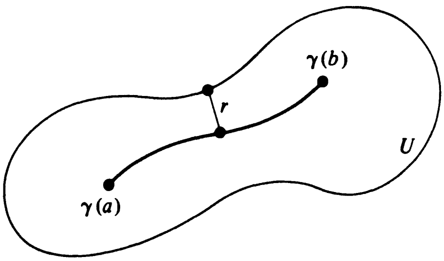


Figure 12

Proof. The image of γ is compact. Consider the function

$$\varphi(t) = \min_{w} |\gamma(t) - w|,$$

where the minimum is taken for all w in the complement of U. This minimum exists because it suffices to consider w lying inside some big circle. Then $\varphi(t)$ is easily verified to be a continuous function of t, whence φ has a minimum on $[a, b]$, and this minimum cannot be 0 because U is open. This proves our assertion.

Let $P = [a_0, \ldots, a_n]$ be a partition of the interval $[a, b]$. We also write P in the form

$$a = a_0 \leqq a_1 \leqq a_2 \leqq \cdots \leqq a_n = b.$$

Let $\{D_0, \ldots, D_n\}$ be a sequence of discs. We shall say that this sequence of discs is **connected by the curve along the partition** if D_i contains the image $\gamma([a_i, a_{i+1}])$. The following figure illustrates this.

One can always find a partition and such a connected sequence of discs. Indeed, let $\epsilon > 0$ be a positive number such that $\epsilon < r/2$ where r is as in Lemma 4.1. Since γ is uniformly continuous, there exists δ such

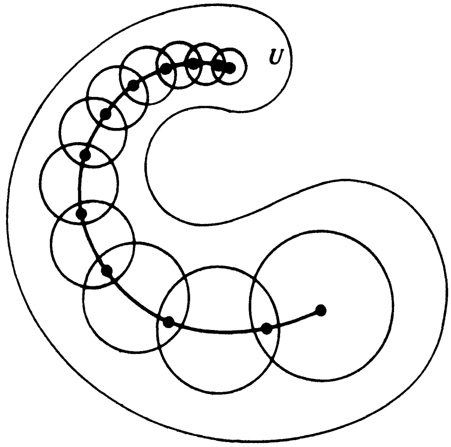


Figure 13

that if $t, s \in [a, b]$ and $|t - s| < \delta$, then $|\gamma(t) - \gamma(s)| < \epsilon$. We select an integer n and a partition P such that each interval $[a_i, a_{i+1}]$ has length $< \delta$. Then the image $\gamma([a_i, a_{i+1}])$ lies in a disc D_i centered at $\gamma(a_i)$ of radius ϵ, and this disc is contained in U.

Let f be holomorphic on U. Let $\gamma_i: [a_i, a_{i+1}] \to U$ be the restriction of γ to the smaller interval $[a_i, a_{i+1}]$. Then

$$\int_\gamma f = \sum_{i=0}^{n-1} \int_{\gamma_i} f.$$

Let $\gamma(a_i) = z_i$, and let g_i be a primitive of f on the disc D_i. If each γ_i is of class C^1 then we find:

$$\boxed{\int_\gamma f = \sum_{i=0}^{n-1} [g_i(z_{i+1}) - g_i(z_i)].}$$

Thus even though f may not have a primitive g on the whole open set U, its integral can nevertheless be expressed in terms of local primitives by decomposing the curve as a sum of sufficiently smaller curves. The same formula then applies to a path.

This procedure allows us to define the **integral of f along any continuous curve**; we do not need to assume any differentiability property of the curve. We need only apply the above procedure, but then we must show that the expression

$$\sum_{i=0}^{n-1} [g_i(z_{i+1}) - g_i(z_i)]$$

is independent of the choice of partition of the interval $[a, b]$ and of the choices of the discs D_i containing $\gamma([a_i, a_{i+1}])$. Then this sum can be taken as the definition of the integral

$$\int_\gamma f.$$

The reader interested only in applications may omit the following considerations. First we state formally this independence, repeating the construction.

Lemma 4.2. *Let $\gamma\colon [a, b] \to U$ be a continuous curve. Let*

$$a_0 = a \leqq a_1 \leqq a_2 \leqq \cdots \leqq a_n = b$$

be a partition of $[a, b]$ such that the image $\gamma([a_i, a_{i+1}])$ is contained in a disc D_i, and D_i is contained in U. Let f be holomorphic on U and let g_i be a primitive of f on D_i.

Let $z_i = \gamma(a_i)$. Then the sum

$$\sum_{i=0}^{n-1} [g_i(z_{i+1}) - g_i(z_i)]$$

is independent of the choices of partitions, discs D_i, and primitives g_i on D_i subject to the stated conditions.

Proof. First let us work with the given partition, but let B_i be another disc containing the image $\gamma([a_i, a_{i+1}])$, and B_i contained in U. Let h_i be a primitive of f on B_i. Then both g_i, h_i are primitives of f on the intersection $B_i \cap D_i$, which is open and connected. Hence there exists a constant C_i such that $g_i = h_i + C_i$ on $B_i \cap D_i$. Therefore the differences are equal:

$$g_i(z_{i+1}) - g_i(z_i) = h_i(z_{i+1}) - h_i(z_i).$$

Thus we have proved that given the partition, the value of the sum is independent of the choices of primitives and choices of discs.

Given two partitions, we can always find a common refinement, as in elementary calculus. Recall that a partition

$$Q = [b_0, \ldots, b_m]$$

is called a **refinement** of the partition P if every point of P is among the points of Q, that is if each a_j is equal to some b_i. Two partitions always have a common refinement, which we obtain by inserting all the points of one partition into the other. Furthermore, we can obtain a refinement of a partition by inserting one point at a time. Thus it suffices to prove that if the partition Q is a refinement of the partition P obtained by inserting one point, then Lemma 4.2 is valid in this case. So we can suppose that Q is obtained by inserting some point c in some interval $[a_k, a_{k+1}]$ for some k, that is Q is the partition

$$[a_0, \ldots, a_k, c, a_{k+1}, \ldots, a_n].$$

We have already shown that given a partition, the value of the sum as in the statement of the lemma is independent of the choice of discs and primitives as described in the lemma. Hence for this new partition Q, we can take the same discs D_i for all the old intervals $[a_i, a_{i+1}]$ when $i \neq k$, and we take the disc D_k for the intervals $[a_k, c]$ and $[c, a_{k+1}]$. Similarly, we take the primitive g_i on D_i as before, and g_k on D_k. Then the sum with respect to the new partition is the same as for the old one, except that the single term

$$g_k(z_{k+1}) - g_k(z_k)$$

is now replaced by two terms

$$g_k(z_{k+1}) - g_k(\gamma(c)) + g_k(\gamma(c)) - g_k(z_k).$$

This does not change the value, and concludes the proof of Lemma 4.2.

For any continuous path $\gamma\colon [a, b] \to U$ we may thus **define**

$$\int_\gamma f = \sum_{i=0}^{n-1} [g_i(\gamma(a_{i+1})) - g_i(\gamma(a_i))]$$

for any partition $[a_0, a_1, \ldots, a_n]$ of $[a, b]$ such that $\gamma([a_i, a_{i+1}])$ is contained in a disc D_i, $D_i \subset U$, and g_i is a primitive of f on D_i. We have just proved that the expression on the right-hand side is independent of the choices made, and we had seen previously that if γ is piecewise C^1 then the expression on the right-hand side gives the same value as the definition used in §2. It is often convenient to have the additional flexibility provided by arbitrary continuous paths.

Remark. The technique of propagating discs along a curve will again be used in the chapter on holomorphic continuation along a curve.

As an application, we shall now see that if two paths lie "close together", and have the same beginning point and the same end point, then the integrals of f along the two paths have the same value. We must define precisely what we mean by "close together". After a reparametrization, we may assume that the two paths are defined over the same interval $[a, b]$. We say that they are **close together** if there exists a partition

$$a = a_0 \leqq a_1 \leqq a_2 \leqq \cdots \leqq a_n = b,$$

and for each $i = 0, \ldots, n-1$ there exists a disc D_i contained in U such that the images of each segment $[a_i, a_{i+1}]$ under the two paths γ, η are contained in D_i, that is,

$$\gamma([a_i, a_{i+1}]) \subset D_i \qquad \text{and} \qquad \eta([a_i, a_{i+1}]) \subset D_i.$$

Lemma 4.3. *Let γ, η be two continuous paths in an open set U, and assume that they have the same beginning point and the same end point. Assume also that they are close together. Let f be holomorphic on U. Then*

$$\int_\gamma f = \int_\eta f.$$

Proof. We suppose that the paths are defined on the same interval $[a, b]$, and we choose a partition and discs D_i as above. Let g_i be a primitive of f on D_i. Let

$$z_i = \gamma(a_i) \qquad \text{and} \qquad w_i = \eta(a_i).$$

We illustrate the paths and their partition in Fig. 14.

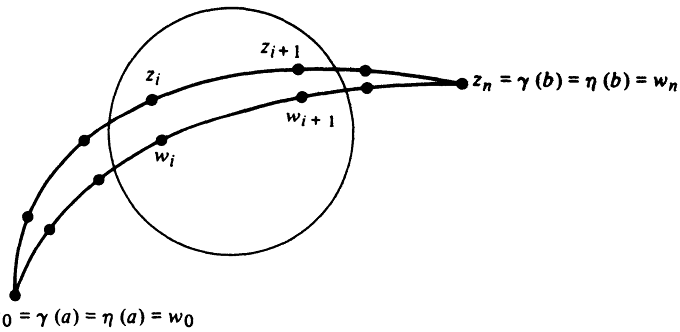


Figure 14

But g_{i+1} and g_i are primitives of f on the connected open set $D_{i+1} \cap D_i$, so $g_{i+1} - g_i$ is constant on $D_{i+1} \cap D_i$. But $D_{i+1} \cap D_i$ contains z_{i+1} and w_{i+1}. Consequently

$$g_{i+1}(z_{i+1}) - g_{i+1}(w_{i+1}) = g_i(z_{i+1}) - g_i(w_{i+1}).$$

Then we find

$$\begin{aligned}\int_\gamma f - \int_\eta f &= \sum_{i=0}^{n-1} [g_i(z_{i+1}) - g_i(z_i) - (g_i(w_{i+1}) - g_i(w_i))]\\ &= \sum_{i=0}^{n-1} [(g_i(z_{i+1}) - g_i(w_{i+1})) - (g_i(z_i) - g_i(w_i))]\end{aligned}$$

$$= g_{n-1}(z_n) - g_{n-1}(w_n) - (g_0(z_0) - g_0(w_0))$$
$$= 0,$$

because the two paths have the same beginning point $z_0 = w_0$, and the same end point $z_n = w_n$. This proves the lemmas.

One can also formulate an analogous lemma for closed paths.

Lemma 4.4. *Let γ, η be closed continuous paths in the open set U, say defined on the same interval $[a, b]$. Assume that they are close together. Let f be holomorphic on U. Then*

$$\int_\gamma f = \int_\eta f.$$

Proof. The proof is the same as above, except that the reason why we find 0 in the last step is now slightly different. Since the paths are closed, we have

$$z_0 = z_n \quad \text{and} \quad w_0 = w_n,$$

as illustrated in Fig. 15. The two primitives g_{n-1} and g_0 differ by a constant on some disc contained in U and containing z_0, w_0. Hence the last expression obtained in the proof of Lemma 4.3 is again equal to 0, as was to be shown.

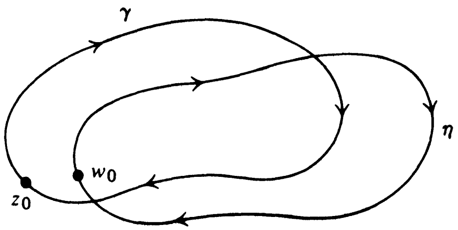


Figure 15

III, §5. THE HOMOTOPY FORM OF CAUCHY'S THEOREM

Let γ, η be two paths in an open set U. After a reparametrization if necessary, we assume that they are defined over the same interval $[a, b]$. We shall say that γ is **homotopic** to η if there exists a continuous function

$$\psi : [a, b] \times [c, d] \to U$$

defined on a rectangle $[a, b] \times [c, d]$, such that

$$\psi(t, c) = \gamma(t) \quad \text{and} \quad \psi(t, d) = \eta(t)$$

for all $t \in [a, b]$.

For each number s in the interval $[c, d]$, we may view the function ψ_s such that

$$\psi_s(t) = \psi(t, s)$$

as a continuous curve, defined on $[a, b]$, and we may view the family of continuous curves ψ_s as a deformation of the path γ to the path η. The picture is drawn on Fig. 16. The paths have been drawn with the same end points because that's what we are going to use in practice. Formally, we say that the homotopy ψ **leaves the end points fixed** if we have

$$\psi(a, s) = \gamma(a) \quad \text{and} \quad \psi(b, s) = \gamma(b)$$

for all values of s in $[c, d]$. *In the sequel it will be always understood that when we speak of a homotopy of paths having the same end points, then the homotopy leaves the end points fixed.*

Similarly, when we speak of a homotopy of closed paths, **we assume always that each path ψ_s is a closed path.** These additional requirements are now regarded as part of the definition of homotopy and will not be repeated each time.

Theorem 5.1. *Let γ, η be paths in an open set U having the same beginning point and the same end point. Assume that they are homotopic in U. Let f be holomorphic on U. Then*

$$\int_\gamma f = \int_\eta f.$$

Theorem 5.2. *Let γ, η be closed paths in U, and assume that they are*

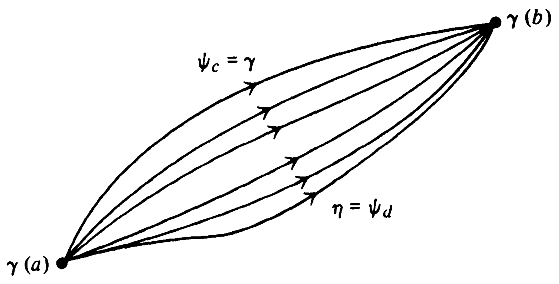


Figure 16

homotopic in U. Let f be holomorphic on U. Then

$$\int_\gamma f = \int_\eta f.$$

In particular, if γ is homotopic to a point in U, then

$$\int_\gamma f = 0.$$

Either of these statements may be viewed as a form of Cauchy's theorem. We prove Theorem 5.2 in detail, and leave Theorem 5.1 to the reader; the proof is entirely similar using Lemma 4.3 instead of Lemma 4.4 from the preceding section. The idea is that the homotopy gives us a finite sequence of paths close to each other in the sense of these lemmas, so that the integral of f over each successive path is unchanged.

The formal proof runs as follows. Let

$$\psi \colon [a, b] \times [c, d] \to U$$

be the homotopy. The image of ψ is compact, and hence has distance > 0 from the complement of U. By uniform continuity we can therefore find partitions

$$a = a_0 \leqq a_1 \leqq \cdots \leqq a_n = b,$$

$$c = c_0 \leqq c_1 \leqq \cdots \leqq c_m = d$$

of these intervals, such that if

$$S_{ij} = \text{small rectangle } [a_i, a_{i+1}] \times [c_j, c_{j+1}]$$

then the image $\psi(S_{ij})$ is contained in a disc D_{ij} which is itself contained in U. Let ψ_j be the continuous curve defined by

$$\psi_j(t) = \psi(t, c_j), \qquad j = 0, \ldots, m.$$

Then the continuous curves ψ_j, ψ_{j+1} are close together, and we can apply the lemma of the preceding section to conclude that

$$\int_{\psi_j} f = \int_{\psi_{j+1}} f.$$

Since $\psi_0 = \gamma$ and $\psi_m = \eta$, we see that the theorem is proved.

Remark. It is usually not difficult, although sometimes it is tedious, to exhibit a homotopy between continuous curves. Most of the time, one can achieve this homotopy by simple formulas when the curves are given explicitly.

Example. Let z, w be two points in the complex numbers. The segment between z, w, denoted by $[z, w]$, is the set of points

$$z + t(w - z), \qquad 0 \leqq t \leqq 1,$$

or equivalently,

$$(1 - t)z + tw, \qquad 0 \leqq t \leqq 1.$$

A set S of complex numbers is called **convex**, if, whenever z, $w \in S$, then the segment $[z, w]$ is also contained in S. We observe that a disc and a rectangle are convex.

Lemma 5.3. *Let S be a convex set, and let γ, η be continuous closed curves in S. Then γ, η are homotopic in S.*

Proof. We define

$$\psi(t, s) = s\gamma(t) + (1 - s)\eta(t).$$

It is immediately verified that each curve ψ_s defined by $\psi_s(t) = \psi(t, s)$ is a closed curve, and ψ is continuous. Also

$$\psi(t, 0) = \eta(t) \qquad \text{and} \qquad \psi(t, 1) = \gamma(t),$$

so the curves are homotopic. Note that the homotopy is given by a linear function, so if γ, η are smooth curves, that is C^1 curves, then each curve ψ_s is also of class C^1.

We say that an open set U is **simply connected** if it is connected and if every closed path in U is homotopic to a point. By Lemma 5.3, a convex open set is simply connected. Other examples of simply connected open sets will be given in the exercises. Simply connected open sets will be used in an essential way in the next section.

Remark. The technique used in this section, propagating along curves, will again be used in the theory of analytic continuation in Chapter XI, §1, which actually could be read immediately as a continuation of this section.

III, §5. EXERCISES

1. A set S is called **star-shaped** if there exists a point z_0 in S such that the line segment between z_0 and any point z in S is contained in S. Prove that a star-shaped set is simply connected, that is, every closed path is homotopic to a point.

2. Let U be the open set obtained from $\mathbf{C}$ by deleting the set of real numbers $\geqq 0$. Prove that U is simply connected.

3. Let V be the open set obtained from $\mathbf{C}$ by deleting the set of real numbers $\leqq 0$. Prove that V is simply connected.

4. (a) Let U be a simply connected open set and let f be an analytic function on U. Is $f(U)$ simply connected?
   (b) Let H be the upper half-plane, that is, the set of complex numbers $z = x + iy$ such that $y > 0$. Let $f(z) = e^{2\pi i z}$. What is the image $f(H)$? Is $f(H)$ simply connected?

III, §6. EXISTENCE OF GLOBAL PRIMITIVES. DEFINITION OF THE LOGARITHM

In §3 we constructed locally a primitive for a holomorphic function by integrating. We now have the means of constructing primitives for a much wider class of open sets.

Theorem 6.1. *Let f be holomorphic on a simply connected open set U. Let $z_0 \in U$. For any point $z \in U$ the integral*

$$g(z) = \int_{z_0}^{z} f(\zeta)\, d\zeta$$

is independent of the path in U from z_0 to z, and g is a primitive for f, namely $g'(z) = f(z)$.

Proof. Let γ_1, γ_2 be two paths in U from z_0 to z. Let γ_2^- be the reverse path of γ_2, from z to z_0. Then

$$\gamma = \{\gamma_1, \gamma_2^-\}$$

is a closed path, and by the first form of Cauchy's theorem,

$$\int_{\gamma_1} f + \int_{\gamma_2^-} f = \int_{\gamma} f = 0.$$

Since the integral of f over γ_2^- is the negative of the integral of f over γ_2, we have proved the first assertion.

As to the second, to prove the differentiability of g at a point z_1, if z is near z_1, then we may select a path from z_0 to z by passing through z_1, that is

$$g(z) = g(z_1) + \int_{z_1}^{z} f,$$

and we have already seen that this latter integral defines a local primitive for f in a neighborhood of z_1. Hence

$$g'(z) = f(z),$$

as desired.

Example. Let U be the plane from which a ray starting from the origin has been deleted. Then U is simply connected.

Proof. Let γ be any closed path in U. For simplicity, suppose the ray is the negative x-axis, as on Fig. 17. Then the path may be described in terms of polar coordinates,

$$\gamma(t) = r(t)e^{i\theta(t)}, \qquad a \leqq t \leqq b,$$

with $-\pi < \theta(t) < \pi$. We define the homotopy by

$$\psi(t, u) = r\big(ua + (1-u)t\big)e^{i\theta(t)(1-u)}, \qquad 0 \leqq u \leqq 1.$$

Geometrically, we are folding back the angle towards 0, and we are contracting the distance $r(t)$ towards $r(a)$. It is clear that ψ has the desired property.

Remark. You could also note that the open set U is star-shaped (proof?), and so if you did Exercise 1 of §5, you don't need the above argument to show that U is simply connected.

Example (Definition of the Logarithm). Let U be a simply connected open set not containing 0. Pick a point $z_0 \in U$. Let w_0 be a complex

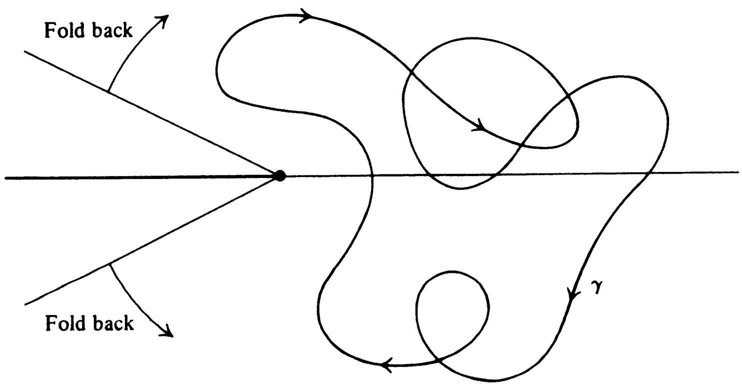


Figure 17

number such that

$$e^{w_0} = z_0.$$

(Any two such numbers differ by an integral multiple of $2\pi i$.) Define

$$\log z = w_0 + \int_{z_0}^{z} \frac{1}{\zeta}\, d\zeta.$$

Then $\log z$ (which depends on the choice of z_0 and w_0 only) is a primitive for $1/z$ on U, and any other primitive differs from this one by a constant.

Let $L_0(1 + w) = \sum (-1)^{n-1} w^n/n$ be the usual power series for the log in a neighborhood of 1. If z is near z_0, then the function

$$F(z) = w_0 + L_0\big(1 + (z - z_0)/z_0\big)$$

defines an analytic function. By Exercise 6 of Chapter II, §5, we have $F'(z) = 1/z$. Hence there exists a constant K such that for all z near z_0 we have $\log z = F(z) + K$. Since both $\log z_0 = w_0$ and $F(z_0) = w_0$, it follows that $K = 0$, so

$$\log z = F(z) \qquad \text{for } z \text{ near } z_0.$$

Consequently, by Exercise 1 of Chapter II, §3, we find that

$$e^{\log z} = z \qquad \text{for } z \text{ near } z_0.$$

Furthermore, given $z_1 \in U$, we have

$$\int_{z_0}^{z} = \int_{z_0}^{z_1} + \int_{z_1}^{z},$$

so by a similar argument, we see that $\log z$ is analytic on U. The two analytic functions $e^{\log z}$ and z are equal near z_0. Since U is connected, they are equal on U by Theorem 1.2(ii), and the equation $e^{\log z} = z$ remains valid for all $z \in U$.

If $L(z)$ is a primitive for $1/z$ on U such that $e^{L(z)} = z$, then there exists an integer k such that

$$L(z) = \log z + 2\pi i k.$$

Indeed, if we let $g(z) = L(z) - \log z$, then $e^{g(z)} = 1$, so $g(z) = 2\pi i k$ for some integer k.

Example. Let V be the open set obtained by deleting the negative real axis from $\mathbf{C}$, and write a complex number $z \in V$ in the form

$$z = re^{i\theta} \qquad \text{with} \qquad -\pi < \theta < \pi.$$

We can select some $z_0 \in V$ with

$$z_0 = r_0 e^{i\theta_0}.$$

For a positive real number r we let $\log r$ be the usual real logarithm, and we let

$$\log z_0 = \log r_0 + i\theta_0.$$

Then V is simply connected, and for all $z \in V$ we have

$$\log z = \log r + i\theta \qquad \text{with} \quad -\pi < \theta < \pi.$$

For a numerical example, we have

$$1 - i = re^{i\theta} = \sqrt{2}e^{i(-\pi/4)},$$

so

$$\log(1 - i) = \tfrac{1}{2}\log 2 - \frac{i\pi}{4}.$$

Example. On the other hand, let U be the open set obtained by deleting the positive real axis from $\mathbf{C}$, i.e. $U = \mathbf{C} - \mathbf{R}_{\geqq 0}$. Take $0 < \theta < 2\pi$. For this determination of the logarithm, let us find $\log(1 - i)$. We write

$$1 - i = re^{i\theta} = \sqrt{2}e^{i7\pi/4}.$$

Then

$$\log(1 - i) = \tfrac{1}{2}\log 2 + \frac{i7\pi}{4}.$$

We see concretely how the values of the logarithm depend on the choice of open set and the choice of a range for the angle.

Definition of z^α for any complex α. By using the logarithm, we can define z under the following conditions.

Let U be simply connected not containing 0. Let α be a complex number $\neq 0$. Fix a determination of the log on U. With respect to this determination, we define

$$z^\alpha = e^{\alpha \log z}.$$

Then z^α is analytic on U.

Example. Let U be the open set obtained by deleting the positive real axis from the complex plane. We define the log to have the values

$$\log re^{i\theta} = \log r + i\theta,$$

where $0 < \theta < 2\pi$. This is also called a **principal value** for the log in that open set. Then

$$\log i = i\pi/2 \qquad \text{and} \qquad \log(-i) = 3\pi i/2.$$

In this case,

$$i^i = e^{i \log i} = e^{\pi i^2/2} = e^{-\pi/2}.$$

Definition of $\log f(z)$. Let U be a simply connected open set and let f be an analytic function on U such that $f(z) \neq 0$ for all $z \in U$. We want to define $\log f(z)$. If we had this logarithm, obeying the same formalism as in ordinary calculus, then we should have

$$\frac{d}{dz} \log f(z) = \frac{1}{f(z)} f'(z) = \frac{f'(z)}{f(z)}.$$

Conversely, this suggests the correct definition. Select a point $z_0 \in U$. Let w_0 be a complex number such that $\exp(w_0) = f(z_0)$. Since f is assumed to be without zeros on U, the function f'/f is analytic on U. Therefore we can define an analytic function L_f on U by the integral

$$L_f(z) = w_0 + \int_{z_0}^{z} \frac{f'(\zeta)}{f(\zeta)} d\zeta.$$

The function L_f depends on the choice of z_0 and w_0, and we shall determine the extent of this dependence in a moment. The integral can be taken along any path in U from z_0 to z because U is assumed to be simply connected. From the definition, we get the derivative

$$L_f'(z) = f'(z)/f(z).$$

This derivative is independent of the choice of z_0 and w_0, so choosing a different z_0 and w_0 changes L_f at most by an additive constant which we shall prove is an integral multiple of $2\pi i$. We claim that

$$\exp L_f(z) = f(z).$$

To prove this formula, abbreviate $L_f(z)$ by $L(z)$, and differentiate

$e^{-L(z)}f(z)$. We find:

$$\begin{aligned}\frac{d}{dz}e^{-L(z)}f(z) &= e^{-L(z)}(-L'(z))f(z) + e^{-L(z)}f'(z)\\ &= e^{-L(z)}\left(-\frac{f'(z)}{f(z)}f(z) + f'(z)\right)\\ &= 0.\end{aligned}$$

Therefore $e^{-L(z)}f(z)$ is constant on U since U is connected. By definition

$$e^{-L(z_0)}f(z_0) = e^{-w_0}f(z_0) = 1,$$

so the constant is 1, and we have proved that $\exp L_f(z) = f(z)$ for $z \in U$.

If we change the choice of z_0 and w_0 such that $e^{w_0} = f(z_0)$, then the new value for $L_f(z)$ which we obtain is simply

$$\log f(z) + 2\pi ik \quad \text{for some integer } k,$$

because the exponential of both values gives $f(z)$.

Remark. The integral for $\log f(z)$ which we wrote down cannot be written in the form

$$\int_{f(z_0)}^{f(z)} \frac{1}{\zeta}\,d\zeta,$$

because even though U is simply connected, the image $f(U)$ may not be simply connected, as you can see in Exercise 7. Of course, if $\gamma\colon [a, b] \to U$ is a path from z_0 to z, then we may form the composite path $f \circ \gamma\colon [a, b] \to \mathbf{C}$. Then we could take the integral

$$\int_{f(z_0),\, f\circ\gamma}^{f(z)} \frac{1}{\zeta}\,d\zeta$$

along the path $f \circ \gamma$. In this case, by the chain rule,

$$\begin{aligned}\int_{f(z_0),\, f\circ\gamma}^{f(z)} \frac{1}{\zeta}\,d\zeta &= \int_{f(\gamma(a)),\, f\circ\gamma}^{f(\gamma(b))} \frac{1}{\zeta}\,d\zeta\\ &= \int_a^b \frac{1}{f(\gamma(t))} f'(\gamma(t))\gamma'(t)\,dt\\ &= \int_{z_0,\gamma}^{z} \frac{f'(\zeta)}{f(\zeta)}\,d\zeta,\end{aligned}$$

which is the integral that was used to define $L_f(z)$.

III, §6. EXERCISES

1. Compute the following values when the log is defined by its principal value on the open set U equal to the plane with the positive real axis deleted.
   (a) $\log i$ (b) $\log(-i)$ (c) $\log(-1+i)$
   (d) i^i (e) $(-i)^i$ (f) $(-1)^i$
   (g) $(-1)^{-i}$ (h) $\log(-1-i)$

2. Compute the values of the same expressions as in Exercise 1 (except (f) and (g)) when the open set consists of the plane from which the negative real axis has been deleted. Then take $-\pi < \theta < \pi$.

3. Let U be the plane with the negative real axis deleted. Let $y > 0$. Find the limit
$$\lim_{y \to 0} [\log(a+iy) - \log(a-iy)]$$
where $a > 0$, and also where $a < 0$.

4. Let U be the plane with the positive real axis deleted. Find the limit
$$\lim_{y \to 0} [\log(a+iy) - \log(a-iy)]$$
where $a < 0$, and also where $a > 0$.

5. Over what kind of open sets could you define an analytic function $z^{1/3}$, or more generally $z^{1/n}$ for any positive integer n? Give examples, taking the open set to be as "large" as possible.

6. Let U be a simply connected open set. Let f be analytic on U and assume that $f(z) \neq 0$ for all $z \in U$. Show that there exists an analytic function g on U such that $g^2 = f$. Does this last assertion remain true if 2 is replaced by an arbitrary positive integer n?

7. Let U be the upper half plane, consisting of all complex numbers $z = x + iy$ with $y > 0$. Let $\varphi(z) = e^{2\pi i z}$. Prove that $\varphi(U)$ is the open unit disc from the origin has been deleted.

8. Let U be the open set obtained by deleting 0 and the negative real axis from the complex numbers. For an integer $m \geqq 1$ define
$$L_{-m}(z) = \left(\log z - \left(1 + \frac{1}{2} + \cdots + \frac{1}{m}\right)\right)\frac{z^m}{m!}.$$
Show that $L'_{-m}(z) = L_{-m+1}(z)$, and that $L'_{-1}(z) = \log z$. Thus L_{-m} is an m-fold integral of the logarithm.

III, §7. THE LOCAL CAUCHY FORMULA

We shall next give an application of the homotopy Theorem 5.2 to prove that a holomorphic function is analytic. The property of being analytic is local: it means that a function has a power series expansion at every point (absolutely convergent on a disc of positive radius centered at the point).

Theorem 7.1 (Local Cauchy Formula). *Let $\bar{D}$ be a closed disc of positive radius, and let f be holomorphic on $\bar{D}$ (that is, on an open disc U containing $\bar{D}$). Let γ be the circle which is the boundary of $\bar{D}$. Then for every $z_0 \in D$ we have*

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

Proof. Let C_r be the circle of radius r centered at z_0, as illustrated on Fig. 18.

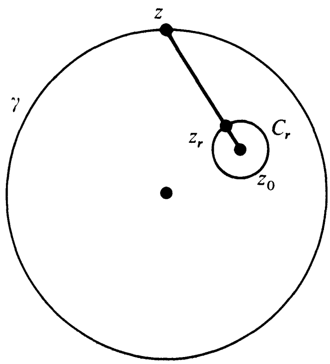


Figure 18

Then for small r, γ and C_r are homotopic. The idea for constructing the homotopy is to shrink γ toward C_r along the rays emanating from z_0. The formula can easily be given. Let z_r be the point of intersection of a line through z and z_0 with the circle of radius r, as shown on Fig. 18. Then

$$z_r = z_0 + r\frac{z - z_0}{|z - z_0|}.$$

Let $\gamma(t)$ $(0 \leqq t \leqq 2\pi)$ parametrize the circle γ. Substituting $\gamma(t)$ for z we obtain

$$\gamma(t)_r = z_0 + r\frac{\gamma(t) - z_0}{|\gamma(t) - z_0|}.$$

Now define the homotopy by letting

$$h(t, u) = u\gamma(t)_r + (1 - u)\gamma(t) \qquad \text{for} \quad 0 \leqq u \leqq 1.$$

Let U be an open disc containing $\bar{D}$, and let U_0 be the open set obtained by removing z_0 from U. Then $h(t, u) \in U_0$, that is, z_0 does not lie in the image of h, because the segment between z and the point z_r lies entirely outside the open disc of radius r centered at z_0. Thus γ is homotopic to C_r in U_0.

Let

$$g(z) = \frac{f(z) - f(z_0)}{z - z_0}$$

for $z \in D$ and $z \neq z_0$. Then g is holomorphic on the open set U_0. By Theorem 5.2, we get

$$\int_\gamma g(z)\,dz = \int_{C_r} g(z)\,dz.$$

Since f is differentiable at z_0, it follows that g is bounded in a neighborhood of z_0. Let B be a bound, so let $|g(z)| \leqq B$ for all z sufficiently close to z_0. Then for r sufficiently small we get

$$\left| \int_{C_r} g(z)\,dz \right| \leqq B(\text{length of } C_r) = B2\pi r,$$

and the right side approaches 0 as r approaches 0. Hence we conclude that

$$\int_\gamma g(z)\,dz = 0.$$

But then

$$\begin{aligned}\int_\gamma \frac{f(z)}{z - z_0}\,dz = \int_\gamma \frac{f(z_0)}{z - z_0}\,dz &= f(z_0)\int_\gamma \frac{1}{z - z_0}\,dz.\\ &= f(z_0)\int_{C_r} \frac{1}{z - z_0}\,dz\\ &= f(z_0)2\pi i.\end{aligned}$$

This proves the theorem.

Theorem 7.2. *Let f be holomorphic on an open set U. Then f is analytic on U.*

Proof. We must show that f has a power series expansion at every point z_0 of U. Because U is open, for each $z_0 \in U$ there is some $R > 0$ such that the closed disc $\bar{D}(z_0, R)$ centered at z_0 and of radius R is contained in U. We are therefore reduced to proving the following theorem, which will give us even more information concerning the power series expansion of f at z_0.

Theorem 7.3. *Let f be holomorphic on a closed disc $\bar{D}(z_0, R)$, $R > 0$. Let C_R be the circle bounding the disc. Then f has a power series expansion*

$$f(z) = \sum a_n(z - z_0)^n$$

whose coefficients a_n are given by the formula:

$$a_n = \frac{1}{n!} f^{(n)}(z_0) = \frac{1}{2\pi i} \int_{C_R} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}}\, d\zeta.$$

Furthermore, if $\|f\|_R$ denotes the sup norm *of f on the circle C_R, then we have the estimate*

$$|a_n| \leqq \|f\|_R / R^n.$$

In particular, the radius of convergence of the series is $\geqq R$.

Proof. By Theorem 7.1, for all z inside the circle C_R, we have

$$f(z) = \frac{1}{2\pi i} \int_{C_R} \frac{f(\zeta)}{\zeta - z}\, d\zeta.$$

Let $0 < s < R$. Let $D(z_0, s)$ be the disc of radius s centered at z_0. We shall see that f has a power series expansion on this disc. We write

$$\frac{1}{\zeta - z} = \frac{1}{\zeta - z_0 - (z - z_0)} = \frac{1}{\zeta - z_0}\left(\frac{1}{1 - \dfrac{z - z_0}{\zeta - z_0}}\right)$$

$$= \frac{1}{\zeta - z_0}\left(1 + \frac{z - z_0}{\zeta - z_0} + \left(\frac{z - z_0}{\zeta - z_0}\right)^2 + \cdots\right).$$

This geometric series converges absolutely and uniformly for $|z - z_0| \leqq s$ because

$$\left|\frac{z - z_0}{\zeta - z_0}\right| \leqq s/R < 1.$$

The function f is bounded on γ. By Theorem 2.4 of Chapter III, we can therefore integrate term by term, and we find

$$f(z) = \sum_{n=0}^{\infty} \frac{1}{2\pi i} \int_{\gamma} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}}\, d\zeta \cdot (z - z_0)^n$$

$$= \sum_{n=0}^{\infty} a_n (z - z_0)^n,$$

where

$$a_n = \frac{1}{2\pi i} \int_\gamma \frac{f(\zeta)}{(\zeta - z_0)^{n+1}}\, d\zeta.$$

This proves that f is analytic, and gives us the coefficients of its power series expansion on the disc of radius R.

In particular, we now see that **a function is analytic if and only if it is holomorphic**. The two words will be used interchangeably from now on.

There remains only to estimate the integral to get an estimate for the coefficients. The estimate is taken as usual, equal to the product of the sup norm of the expression under the integral sign, and the length of the curve which is $2\pi R$. For all ζ on the circle, we have

$$|\zeta - z_0| = R,$$

so the desired estimate falls out. Taking the n-th root of $|a_n|$, we conclude at once that the radius of convergence is at least R.

Remark. From the statement about the radius of convergence in Theorem 7.3 we now see that if R is the radius of convergence of a power series, then its analytic function does not extend to a disc of radius $> R$; otherwise the given power series would have a larger radius of convergence, and would represent this analytic function on the bigger disc. For example, let

$$f(z) = e^z/(z - 1).$$

Then f is analytic except at $z = 1$. From the theorem, we conclude:

The radius of convergence of the power series for f at the origin is 1.

The radius of convergence of the power series for f at 2 is 1.

The radius of convergence of the power series for f at 5 is 4.

The radius of convergence of the power series for f at -3 is 4.

A function f is called **entire** if it is holomorphic on all of $\mathbf{C}$. We also conclude from the above remark and the theorem that if a function is entire, then its power series converges for all $z \in \mathbf{C}$, in other words the radius of convergence is ∞.

Corollary 7.4. *Let f be an entire function, and let $\|f\|_R$ be its* sup norm *on the circle of radius R. Suppose that there exists a constant C and a positive integer k such that*

$$\|f\|_R \leqq CR^k$$

for arbitrarily large R. Then f is a polynomial of degree $\leqq k$.

Proof. Exercise 3, but we carry out one important special case explicitly:

Theorem 7.5 (Liouville's Theorem). *A bounded entire function is constant.*

Proof. If f is bounded, then $\|f\|_R$ is bounded for all R. In the preceding theorem, we let R tend to infinity, and conclude that the coefficients are all equal to 0 if $n \geqq 1$. This proves Liouville's theorem.

We have already proved that a polynomial always has a root in the complex numbers. We give here the more usual proof as a corollary of Liouville's theorem.

Corollary 7.6. *A polynomial over the complex numbers which does not have a root in* $\mathbf{C}$ *is constant.*

Proof. Let $f(z)$ be a non-constant polynomial,

$$f(z) = a_n z^n + \cdots + a_0,$$

with $a_n \neq 0$. Suppose that $f(z) \neq 0$ for all z. Then the function

$$g(z) = 1/f(z)$$

is defined for all z and analytic on $\mathbf{C}$. On the other hand, writing

$$f(z) = a_n z^n (1 + b_1/z + \cdots + b_n/z^n)$$

with appropriate constants $b_1, \ldots, b_n$ we see that $|f(z)|$ is large when $|z|$ is large, and hence that $|g(z)| \to 0$ as $|z| \to \infty$. For sufficiently large radius R, $|g(z)|$ is small for z outside the closed disc of radius R, and $|g(z)|$ has a maximum on this disc since the disc is compact. Hence g is a bounded entire function, and therefore constant by Liouville's theorem. This is obviously a contradiction, proving that f must have a zero somewhere in $\mathbf{C}$.

We end this section by pointing out that the main argument of Theorem 7.3 can be used essentially unchanged to define an analytic function and its derivatives by means of an integral, as follows.

Theorem 7.7. *Let γ be a path in an open set U and let g be a continuous function on γ (i.e. on the image $\gamma([a, b])$) if γ is defined on $[a, b]$). If z is not on γ, define*

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

Then f is analytic on the complement of γ in U, and its derivatives are given by

$$f^{(n)}(z) = n! \int_\gamma \frac{g(\zeta)}{(\zeta - z)^{n+1}}\, d\zeta.$$

Proof. Let $z_0 \in U$ and z_0 not on γ. Since the image of γ is compact, there is a minimum distance between z_0 and points on γ. Select $0 < R < \operatorname{dist}(z_0, \gamma)$, and take R also small enough that the closed disc $\bar{D}(z_0, R)$ is contained in U. Now we are essentially in the situation of Theorem 7.3. We may repeat the arguments of the proof. We select $0 < s < R$, and we simply replace f by g inside the integral sign. We expand $1/(\zeta - z)$ by means of the geometric series, and proceed without any further change to see that f has a power series expansion $f = \sum a_n(z - z_0)^n$, where now the coefficients a_n are given by

$$a_n = \int_\gamma \frac{g(\zeta)}{(\zeta - z_0)^{n+1}}\, d\zeta.$$

We know from Chapter II, §5 that $a_n = f^{(n)}(z_0)/n!$, which gives us the proof of Theorem 7.7.

There is also another way of looking at Theorem 7.7. Indeed, from the formula for f, it is natural to think that one can differentiate with respect to z under the integral sign. This differentiation will be justified in Theorem A3, §6, Chapter VIII, which the reader may wish to look at now. Then one gets the integral formula also for the derivatives.

From Theorem 7.7 we obtain a bound for the derivative of an analytic function in terms of the function itself. This is of course completely different from what happens for real differentiable functions.

Corollary 7.8. *Let f be analytic on a closed disc $\bar{D}(z_0, R)$, $R > 0$. Let $0 < R_1 < R$. Denote by $\|f\|_R$ the sup norm of f on the circle of radius R. Then for $z \in \bar{D}(z_0, R_1)$ we have*

$$|f^{(n)}(z)| \leqq \frac{n!\, R}{(R - R_1)^{n+1}} \|f\|_R.$$

Proof. This is immediate by using Theorem 7.1, and putting $g = f$ inside the integral, with a factor of $1/2\pi i$ in front. The factor R in the numerator comes from the length of the circle in the integral. The 2π in the denominator cancels the 2π in the numerator, coming from the formula for the length of the circle.

Note that if R_1 is close to R, then the denominator may be correspondingly large. On the other hand, suppose $R_1 = R/2$. Then the estimate reads

$$|f^{(n)}(z)| \leqq \frac{n!\,2^{n+1}}{R^n}\|f\|_R,$$

which is thus entirely in terms of f, n, and R.

Finally we return to reconsider Theorem 3.2 in light of the fact that a holomorphic function is analytic.

Theorem 7.9 (Morera's Theorem). *Let U be an open set in $\mathbf{C}$ and let f be continuous on U. Assume that the integral of f along the boundary of every closed rectangle contained in U is 0. Then f is analytic.*

Proof. By Theorem 3.2, we know that f has a local primitive g at every point on U, and hence that g is holomorphic. By Theorem 7.2, we conclude that g is analytic, and hence that $g' = f$ is analytic, as was to be shown.

We have now come to the end of a chain of ideas linking complex differentiability and power series expansions. The next two chapters treat different applications, and can be read in any order, but we have to project the book in a totally ordered way on the page axis, so we have to choose an order for them. The next chapter will study more systematically a global version of Cauchy's formula and winding numbers, which amounts to studying the relation between an integral and the winding number which we already encountered in some way via the logarithm. After that in Chapter V, we return to analytic considerations and estimates.

III, §7. EXERCISES

1. Find the integrals over the unit circle γ:

(a) $\displaystyle\int_\gamma \frac{\cos z}{z}\,dz$ (b) $\displaystyle\int_\gamma \frac{\sin z}{z}\,dz$ (c) $\displaystyle\int_\gamma \frac{\cos(z^2)}{z}\,dz$

2. Write out completely the proof of Theorem 7.6 to see that all the steps in the proof of Theorem 7.3 apply.

3. Prove Corollary 7.4.

CHAPTER IV

Winding Numbers and Cauchy's Theorem

We wish to give a general global criterion when the integral of a holomorphic function along a closed path is 0. In practice, we meet two types of properties of paths: (1) properties of homotopy, and (2) properties having to do with integration, relating to the number of times a curve "winds" around a point, as we already saw when we evaluated the integral

$$\int \frac{1}{\zeta - z}\, d\zeta$$

along a circle centered at z. These properties are of course related, but they also exist independently of each other, so we now consider those conditions on a closed path γ when

$$\int_\gamma f = 0$$

for all holomorphic functions f, and also describe what the value of this integral may be if not 0.

We shall give two proofs for the global version of Cauchy's theorem. Artin's proof depends only on Goursat's theorem for the integral of a holomorphic function around a rectangle, and a self-contained topological lemma, having only to do with paths and not holomorphic functions. Dixon's proof uses some of the applications to holomorphic functions which bypass the topological considerations.

In this chapter, paths are again assumed to be piecewise C^1, and curves are again C^1.

IV, §1. THE WINDING NUMBER

In an example of Chapter III, §2, we found that

$$\frac{1}{2\pi i}\int_\gamma \frac{1}{z}\,dz = 1,$$

if γ is a circle around the origin, oriented counterclockwise. It is therefore reasonable to define for any closed path γ its **winding number with respect to a point** α to be

$$W(\gamma, \alpha) = \frac{1}{2\pi i}\int_\gamma \frac{1}{z-\alpha}\,dz,$$

provided the path does not pass through α. If γ is a curve defined on an interval $[a, b]$, then this integral can be written in the form

$$\int_\gamma \frac{1}{z-\alpha}\,dz = \int_a^b \frac{\gamma'(t)}{\gamma(t)-\alpha}\,dt.$$

Intuitively, the integral of $1/(z-\alpha)$ should be called $\log(z-\alpha)$, but it depends on the path. Later, we shall analyze this situation more closely, but for the moment, we need only the definition above without dealing with the log formally, although the interpretation in terms of the log is suggestive.

The definition of the winding number would be improper if the following lemma were not true.

Lemma 1.1. *If γ is a closed path, then $W(\gamma, \alpha)$ is an integer.*

Proof. Let $\gamma = \{\gamma_1, \ldots, \gamma_n\}$ where each γ_i is a curve defined on an interval $[a_i, b_i]$. After a reparametrization of each curve if necessary, we may assume without loss of generality that $b_i = a_{i+1}$ for $i = 1, \ldots, n-1$. Then γ is defined and continuous on an interval $[a, b]$, where $a = a_1$, $b = b_n$, and γ is differentiable on each open interval $]a_i, b_i[$, (at the end points, γ is merely right and left differentiable). Let

$$F(t) = \int_a^t \frac{\gamma'(t)}{\gamma(t)-\alpha}\,dt.$$

Then F is continuous on $[a, b]$ and differentiable for $t \neq a_i, b_i$. Its derivative is

$$F'(t) = \frac{\gamma'(t)}{\gamma(t)-\alpha}.$$

(Intuitively, $F(t) = \log(\gamma(t) - \alpha)$ except for the dependence of path and a constant of integration, but this suggests our next step.) We compute the derivative of another function:

$$\frac{d}{dt}e^{-F(t)}(\gamma(t) - \alpha) = e^{-F(t)}\gamma'(t) - F'(t)e^{-F(t)}(\gamma(t) - \alpha) = 0.$$

Hence there is a constant C such that $e^{-F(t)}(\gamma(t) - \alpha) = C$, so

$$\gamma(t) - \alpha = Ce^{F(t)}.$$

Since γ is a closed path, we have $\gamma(a) = \gamma(b)$, and

$$Ce^{F(b)} = \gamma(b) - \alpha = \gamma(a) - \alpha = Ce^{F(a)}.$$

Since $\gamma(a) - \alpha \neq 0$ we conclude that $C \neq 0$, so that

$$e^{F(a)} = e^{F(b)}.$$

Hence there is an integer k such that

$$F(b) = F(a) + 2\pi ik.$$

But $F(a) = 0$, so $F(b) = 2\pi ik$, thereby proving the lemma.

The winding number of the curve in Fig. 1 with respect to α is equal to 2.

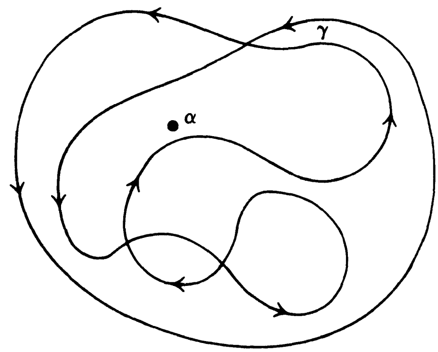


Figure 1

Lemma 1.2. *Let γ be a path. Then the function of α defined by*

$$\alpha \mapsto \int_\gamma \frac{1}{z - \alpha}\,dz$$

for α not on the path, is a continuous function of α.

Proof. Given α_0 not on the path, we have to see that

$$\int_\gamma \left(\frac{1}{z-\alpha} - \frac{1}{z-\alpha_0}\right) dz$$

tends to 0 as α tends to α_0. This integral is estimated as follows. The function $t \mapsto |\alpha_0 - \gamma(t)|$ is continuous and not 0, hence it has a minimum, the minimum distance between α_0 and the path, say

$$\min_t |\alpha_0 - \gamma(t)| = s.$$

If α is sufficiently close to α_0, then $|\alpha - \gamma(t)| \geqq s/2$, as illustrated in Fig. 2.

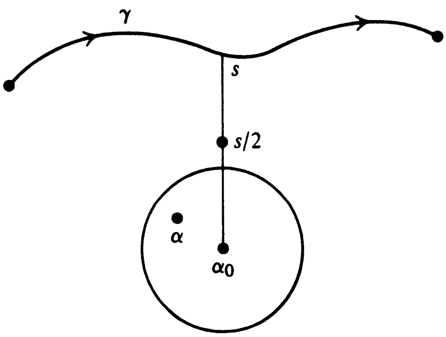


Figure 2

We have

$$\frac{1}{z-\alpha} - \frac{1}{z-\alpha_0} = \frac{\alpha - \alpha_0}{(z-\alpha)(z-\alpha_0)}$$

whence the estimate

$$\left|\frac{1}{z-\alpha} - \frac{1}{z-\alpha_0}\right| \leqq \frac{1}{s^2/4}|\alpha - \alpha_0|.$$

Consequently, we get

$$\left|\int_\gamma \left(\frac{1}{z-\alpha} - \frac{1}{z-\alpha_0}\right) dz\right| \leqq \frac{1}{s^2/4}|\alpha - \alpha_0| L(\gamma).$$

The right-hand side tends to 0 as α tends to α_0, and the continuity is proved.

Lemma 1.3. *Let γ be a closed path. Let S be a connected set not intersecting γ. Then the function*

$$\alpha \mapsto \frac{1}{2\pi i} \int_\gamma \frac{1}{z - \alpha}\, dz$$

is constant for α in S. If S is not bounded, then this constant is 0.

Proof. We know from Lemma 1.1 that the integral is the winding number, and is therefore an integer. If a function takes its values in the integers, and is continuous, then it is constant on any curve, and consequently constant on a connected set. If S is not bounded, then for α arbitrarily large, the integrand has arbitrarily small absolute value, that is,

$$\frac{1}{|z - \alpha|}$$

is arbitrarily small, and estimating the integral shows that it must be equal to 0, as desired.

Example. Let U be the open set in Fig. 3. Then the set of points not in U consists of two connected components, one inside U and the other unbounded. Let γ be the closed curve shown in the figure, and let α_1 be the point inside γ, whereas α_2 is the point outside U, in the unbounded connected region. Then

$$W(\gamma, \alpha_1) = 1, \qquad \text{but} \qquad W(\gamma, \alpha_2) = 0.$$

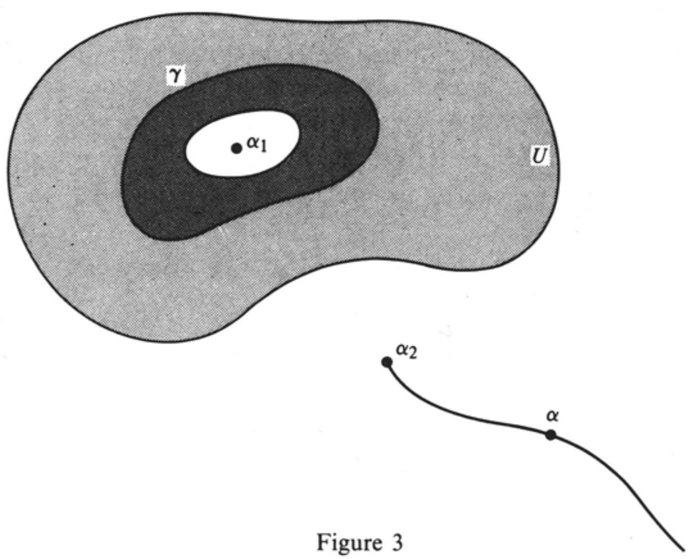


Figure 3

We have drawn a curve extending from α_2 towards infinity, such that $W(\gamma, \alpha) = 0$ for α on this curve, according to the argument of Lemma 1.3.

IV, §2. THE GLOBAL CAUCHY THEOREM

Let U be an open set. Let γ be a closed path in U. We want to give conditions that

$$\int_\gamma f = 0$$

for every holomorphic function f on U. We already know from the example of a winding circle that if the path winds around some point outside of U (in this example, the center of the circle), then definitely we can find functions whose integral is not equal to 0, and even with the special functions

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

where α is a point not in U. The remarkable fact about Cauchy's theorem is that it will tell us this is the only obstruction possible to having

$$\int_\gamma f = 0$$

for all possible functions f. In other words, the functions

$$\frac{1}{z-\alpha}, \qquad \alpha \notin U,$$

suffice to determine the behavior of $\int_\gamma f$ for all possible functions. With this in mind, we want to give a name to those closed paths in U having the property that they do not wind around points in the complement of U. The name we choose is homologous to 0, for historical reasons. Thus formally, we say that a closed path γ in U is **homologous to 0 in U** if

$$\int_\gamma \frac{1}{z - \alpha}\, dz = 0$$

for every point α not in U, or in other words, more briefly,

$$W(\gamma, \alpha) = 0$$

for every such point.

Similarly, let γ, η be closed paths in U. We say that they are **homologous in** U if

$$W(\gamma, \alpha) = W(\eta, \alpha)$$

for every point α in the complement of U. It will also follow from Cauchy's theorem that if γ and η are homologous, then

$$\int_{\gamma} f = \int_{\eta} f$$

for all holomorphic functions f on U.

Theorem 2.1.

(i) *If γ, η are closed paths in U and are homotopic, then they are homologous.*
(ii) *If γ, η are closed paths in U and are close together then they are homologous.*

Proof. The first statement follows from Theorem 5.2 of the preceding chapter because the function $1/(z - \alpha)$ is analytic on U for $\alpha \notin U$. The second statement is a special case of Lemma 4.4 of the preceding chapter.

Next we draw some examples of homologous paths.

In Fig. 4, the curves γ and η are **homologous**. Indeed, if α is a point inside the curves, then the winding number is 1, and if α is a point in the connected part going to infinity, then the winding number is 0.

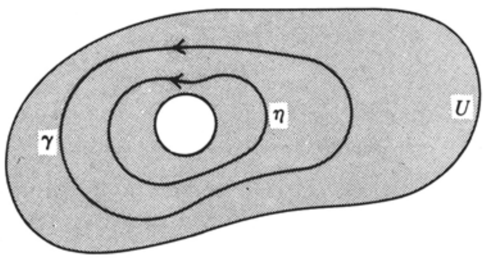


Figure 4

In Fig. 5 the path indicated is supposed to go around the top hole counterclockwise once, then around the bottom hole counterclockwise once, then around the top in the opposite direction, and then around the bottom in the opposite direction. This path is homologous to 0, but not homotopic to a point.

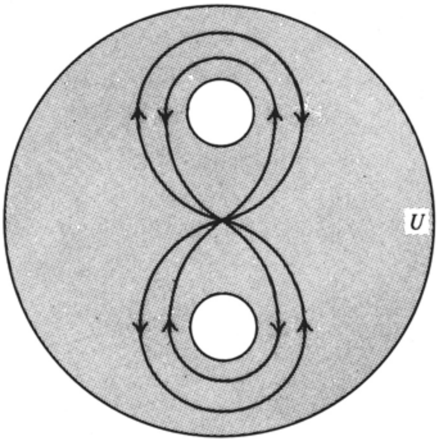


Figure 5

In Fig. 6, we are dealing with a simple closed curve, whose inside is contained in U, and the figure is intended to show that γ can be deformed to a point, so that γ is homologous to 0.

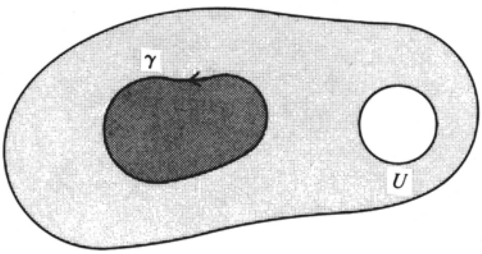


Figure 6

Given an open set U, we wish to determine in a simple way those closed paths which are not homologous to 0. For instance, the open set U might be as in Fig. 7, with three holes in it, at points z_1, z_2, z_3, so these points are assumed not to be in U.

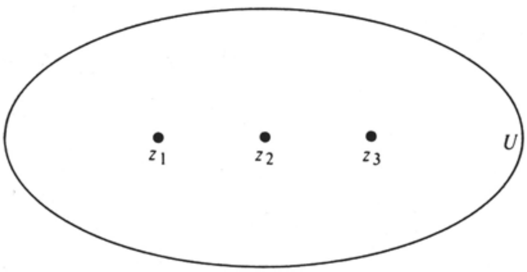


Figure 7

Let γ be a closed path in U, and let f be holomorphic on U. We illustrate γ in Fig. 8.

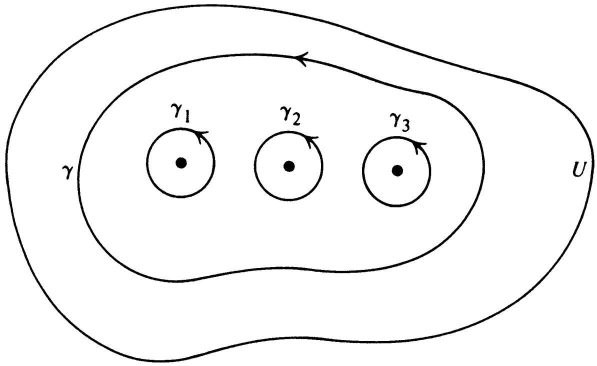


Figure 8

In that figure, we see that γ winds around the three points, and winds once. Let γ_1, γ_2, γ_3 be small circles centered at z_1, z_2, z_3 respectively, and oriented counterclockwise, as shown on Fig. 8. Then it is reasonable to expect that

$$\int_\gamma f = \int_{\gamma_1} f + \int_{\gamma_2} f + \int_{\gamma_3} f.$$

This will in fact be proved after Cauchy's theorem. We observe that taking γ_1, γ_2, γ_3 together does not constitute a "path" in the sense we have used that word, because, for instance, they form a disconnected set. However, it is convenient to have a terminology for a formal sum like $\gamma_1 + \gamma_2 + \gamma_3$, and to give it a name η, so that we can write

$$\int_\gamma f = \int_\eta f.$$

The name that is standard is the name **chain**. Thus let, in general, $\gamma_1, \ldots, \gamma_n$ be curves, and let $m_1, \ldots, m_n$ be integers which need not be positive. A formal sum

$$\gamma = m_1\gamma_1 + \cdots + m_n\gamma_n = \sum_{i=1}^{n} m_i\gamma_i$$

will be called a **chain**. If each curve γ_i is a curve in an open set U, we call γ a **chain in** U. We say that the chain is **closed** if it is a finite sum of closed paths. If γ is a chain as above, we define

$$\int_\gamma f = \sum m_i \int_{\gamma_i} f.$$

If $\gamma = \sum m_i\gamma_i$ is a closed chain, where each γ_i is a closed path, then its winding number with respect to a point α not on the chain is defined as before,

$$W(\gamma, \alpha) = \frac{1}{2\pi i}\int_\gamma \frac{1}{z-\alpha}\,dz.$$

If γ, η are closed chains in U, then we have

$$W(\gamma + \eta, \alpha) = W(\gamma, \alpha) + W(\eta, \alpha).$$

We say that γ is **homologous to** η in U, and write $\gamma \sim \eta$, if

$$W(\gamma, \alpha) = W(\eta, \alpha)$$

for every point $\alpha \notin U$. We say that γ is **homologous to** 0 **in** U and write $\gamma \sim 0$ if

$$W(\gamma, \alpha) = 0$$

for every point $\alpha \notin U$.

Example. Let γ be the curve illustrated in Fig. 9, and let U be the plane from which three points z_1, z_2, z_3 have been deleted. Let γ_1, γ_2, γ_3 be small circles centered at z_1, z_2, z_3 respectively, oriented counterclockwise. Then it will be shown after Cauchy's theorem that

$$\gamma \sim \gamma_1 + 2\gamma_2 + 2\gamma_3,$$

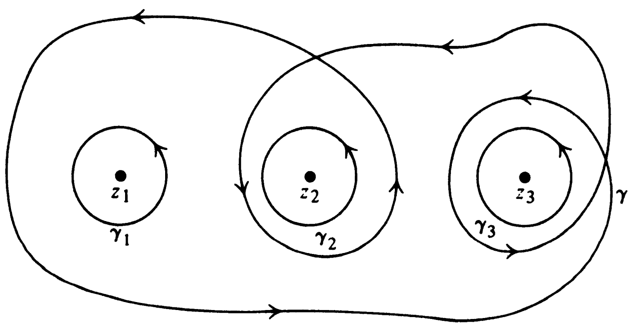


Figure 9

so that for any function f holomorphic on U, we have

$$\int_\gamma f = \int_{\gamma_1} f + 2\int_{\gamma_2} f + 2\int_{\gamma_3} f.$$

The above discussion and definition of chain provided motivation for what follows. We now go back to the formal development, and state the global version of Cauchy's theorem.

Theorem 2.2 (Cauchy's Theorem). *Let γ be a closed chain in an open set U, and assume that γ is homologous to 0 in U. Let f be holomorphic in U. Then*

$$\int_\gamma f = 0.$$

A proof will be given in the next section. Observe that all we shall need of the holomorphic property is the existence of a primitive locally at every point of U, which was proved in the preceding chapter.

Corollary 2.3. *If γ, η are closed chains in U and γ, η are homologous in U, then*

$$\int_\gamma f = \int_\eta f.$$

Proof. Apply Cauchy's theorem to the closed chain $\gamma - \eta$.

Before giving the proof of Cauchy's theorem, we state two important applications, showing how one reduces integrals along complicated paths to integrals over small circles.

Theorem 2.4.

(a) *Let U be an open set and γ a closed chain in U such that γ is homologous to 0 in U. Let $z_1, \ldots, z_n$ be a finite number of distinct points of U. Let γ_i $(i = 1, \ldots, n)$ be the boundary of a closed disc $\bar{D}_i$ contained in U, containing z_i, and oriented counterclockwise. We assume that $\bar{D}_i$ does not intersect $\bar{D}_j$ if $i \neq j$. Let*

$$m_i = W(\gamma, z_i).$$

Let U^* be the set obtained by deleting $z_1, \ldots, z_n$ from U. Then γ is homologous to $\sum m_i\gamma_i$ in U^*.

(b) *Let f be holomorphic on U^*. Then*

$$\int_\gamma f = \sum_{i=1}^n m_i \int_{\gamma_i} f.$$

Proof. Let $C = \gamma - \sum m_i\gamma_i$. Let α be a point outside U. Then

$$W(C, \alpha) = W(\gamma, \alpha) - \sum m_i W(\gamma_i, \alpha) = 0$$

because α is outside every small circle γ_i. If $\alpha = z_k$ for some k, then $W(\gamma_i, z_k) = 1$ if $i = k$ and 0 if $i \neq k$ by Lemma 1.3. Hence

$$W(C, z_k) = W(\gamma, z_k) - m_k = 0.$$

This proves that C is homologous to 0 in U^*. We apply Theorem 2.2 to conclude the proof.

The theorem is illustrated in Fig. 10. We have

$$\gamma \sim -\gamma_1 - 2\gamma_2 - \gamma_3 - 2\gamma_4,$$

and

$$\int_\gamma f = -\int_{\gamma_1} f - 2\int_{\gamma_2} f - \int_{\gamma_3} f - 2\int_{\gamma_4} f.$$

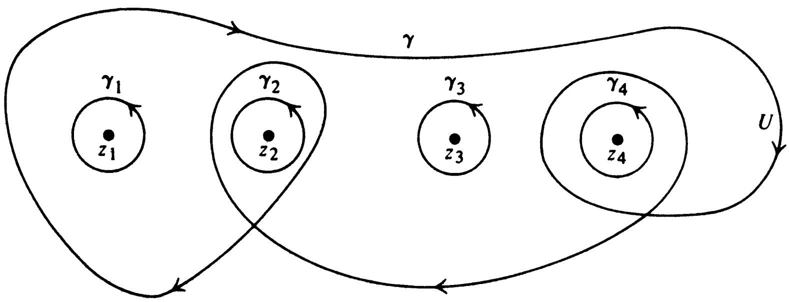


Figure 10

The theorem will be applied in many cases when U is a disc, say centered at the origin, and γ is a circle in U. Then certainly γ is homotopic to a point in U, and therefore homologous to 0 in U. Let $z_1, \ldots, z_n$ be points inside the circle, as on Fig. 11. Then Theorem 2.4 tells us that

$$\boxed{\int_\gamma f = \sum_{i=1}^n \int_{C_i} f,}$$

where C_i is a small circle around z_i. (Circles throughout are assumed oriented counterclockwise unless otherwise specified.)

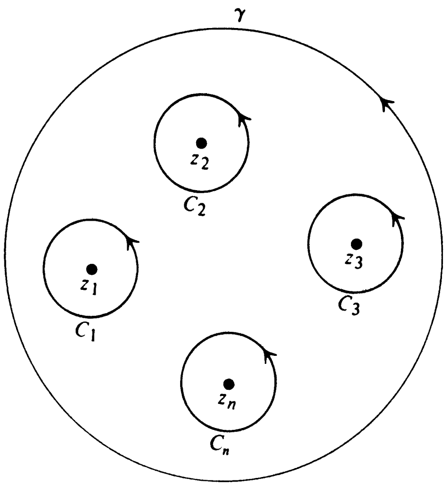


Figure 11

In Example 5 of Chapter III, §2, we gave explicitly the values of the integrals around small circles in terms of the power series expansion of f around the points $z_1, \ldots, z_n$. We may also state the global version of **Cauchy's formula**.

Theorem 2.5 (Cauchy's Formula). *Let γ be a closed chain in U, homologous to 0 in U. Let f be analytic on U, let z_0 be in U and not on γ. Then*

$$\frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z - z_0}\,dz = W(\gamma, z_0)f(z_0).$$

Proof. We base this proof on Theorems 2.2 and 2.4. An independent proof will be given below. By assumption, in a neighborhood of z_0, we have a power series expansion

$$f(z) = a_0 + a_1(z - z_0) + \text{higher terms}, \qquad \text{with } a_0 = f(z_0).$$

Let C_r be the circle of radius r centered at z_0 for a small value of r. By Theorem 2.4, the integral over γ can be replaced by the integral over C_r times the appropriate winding number, that is

$$\frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z - z_0}\,dz = W(\gamma, z_0)\frac{1}{2\pi i}\int_{C_r} \sum_{n=0}^{\infty} a_n(z - z_0)^{n-1}\,dz = W(\gamma, z_0)a_0,$$

because we can integrate term by term by Theorem 2.4 of Chapter III, and we can apply Example 5 or Exercise 1 of Chapter III, §2, to conclude the proof.

Example. Using Theorem 2.5, we find the integral

$$\int_\gamma \frac{e^z}{z}\,dz$$

taken over a path γ not passing through the origin, and having winding number 1 with respect to 0, that is, $W(\gamma, 0) = 1$. We let $U = \mathbf{C}$. Then γ is homologous to 0 in U, and in fact γ is homotopic to a point. Hence Theorem 2.5 applies by letting $z_0 = 0$, and we find

$$\int_\gamma \frac{e^z}{z}\,dz = 2\pi i e^0 = 2\pi i.$$

Remark 1. We have shown that Theorem 2.2 (Cauchy's theorem) implies Theorem 2.5 (Cauchy's formula). Conversely, it is easily seen that Cauchy's formula implies Cauchy's theorem. Namely, we let z_0 be a point in U not on γ, and we let

$$F(z) = (z - z_0)f(z).$$

Applying Cauchy's formula to F yields

$$\frac{1}{2\pi i}\int_\gamma f(z)\,dz = \frac{1}{2\pi i}\int_\gamma \frac{F(z)}{z - z_0}\,dz = F(z_0)W(\gamma, z_0) = 0,$$

as desired.

Remark 2. In older texts, Cauchy's theorem is usually stated for the integral over a simple closed curve, in the following form:

Let U be an open set, f holomorphic on U and let γ be a simple closed curve whose interior is contained in U. Then

$$\int_\gamma f = 0.$$

It was realized for a long time that it is rather hard to prove that a simple closed curve decomposes the plane into two regions, its interior and exterior. It is not even easy to define what is meant by "interior" or "exterior" a priori. In fact, the theorem would be that the plane from which one deletes the curve consists of two connected sets. For all points in one of the sets the winding number with respect to the curve is 1, and for all points in the other, the winding number is 0. In any case, these general results are irrelevant in the applications. Indeed, both in theoretical work and in practical applications, the statement of Cauchy's theorem as we gave it is quite efficient. In special cases, it is usually immediate to define the "interior" and "exterior" having the above property, for instance for circles or rectangles. One can apply Theorem 2.2 without appealing to any complicated result about general closed curves.

Dixon's Proof of Theorem 2.5 (Cauchy's Formula)

The proof we gave of Theorem 2.5 was based on Theorem 2.2 via Theorem 2.4. We shall now reproduce Dixon's proof of Theorem 2.5, which is direct, and is based only on Cauchy's formula for a circle and Liouville's theorem. Those results were proved in Chapter III, §7. Dixon's proof goes as follows.

We define a function g on $U \times U$ by:

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

For each w, the function $z \mapsto g(z, w)$ is analytic. Furthermore, g is continuous on $U \times U$. This is obvious for points off the diagonal, and if (z_0, z_0) is on the diagonal, then for (z, w) close to (z_0, z_0)

$$g(z, w) - g(z_0, z_0) = \frac{1}{w - z} \int_z^w [f'(\zeta) - f'(z_0)]\, d\zeta.$$

The integral can be taken along the line segment from z to w. Estimating the right-hand side, we see that $1/|w - z|$ cancels the length of the interval, and the expression under the integral sign tends to 0 by the continuity of f', as (z, w) approaches (z_0, z_0). Thus g is continuous.

Let V be the open set of complex numbers z not on γ such that $W(\gamma, z) = 0$. By the hypothesis of Cauchy's theorem, we know that V contains the complement of U. Hence $\mathbf{C} = U \cup V$. We now define a function h on $\mathbf{C}$ by two integrals:

$$h(z) = \frac{1}{2\pi i} \int_\gamma g(z, w)\, dw \quad \text{if} \quad z \in U,$$

$$h(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{w - z}\, dw \quad \text{if} \quad z \in V.$$

We note that for $z \in U \cap V$, the two definitions of h coincide. We shall prove that h is a bounded entire function, whence constant by Liouville's theorem, whence equal to 0 by letting z tend to infinity for $z \in V$, and using the definition of h. It is then clear that for $z \in U$ the first integral being zero immediately implies Cauchy's formula

$$\frac{1}{2\pi i}\int_{\gamma}\frac{f(w)}{w - z}\,dw = f(z)W(\gamma, z).$$

We have already seen in Remark 1 that Cauchy's formula implies Cauchy's theorem.

There remains therefore to prove that h is an analytic function and is bounded. We first prove that h is analytic. It is immediate that h is analytic on V. Hence it suffices to prove that h is analytic on U. So let $z_0 \in U$. From the uniform continuity of g on compact subsets of $U \times U$ it follows at once that h is continuous. To prove that h is analytic, by Theorem 3.2 of Chapter III, and the fact that a holomorphic function is analytic, it suffices to prove that in some disc centered at z_0, the integral of h around the boundary of any rectangle contained in the disc is 0. But we have

$$\begin{aligned}\int_{\partial R} h(z)\,dz &= \frac{1}{2\pi i}\int_{\partial R}\int_{\gamma} g(z, w)\,dw\,dz \\ &= \frac{1}{2\pi i}\int_{\gamma}\int_{\partial R} g(z, w)\,dz\,dw.\end{aligned}$$

Since for each w, the function $z \mapsto g(z, w)$ is analytic, we obtain the value 0, thereby concluding the proof that h is analytic.

As for the boundedness, suppose that z lies outside a large circle. Then

$$\int_{\gamma}\frac{f(z)}{w - z}\,dw = f(z)\int_{\gamma}\frac{1}{w - z}\,dw = 0$$

because the winding number of γ with respect to z is 0 by Lemma 1.3 of Chapter IV. Furthermore, if $|z| \to \infty$ then

$$\int_{\gamma}\frac{f(w)}{w - z}\,dw \to 0.$$

It follows that h is bounded outside a large circle, whence bounded since h is analytic. This concludes the proof.

IV, §2. EXERCISES

1. (a) Show that the association $f \mapsto f'/f$ (where f is holomorphic) sends products to sums.
   (b) If $P(z) = (z - a_1)\cdots(z - a_n)$, where $a_1, \ldots, a_n$ are the roots, what is P'/P?
   (c) Let γ be a closed path such that none of the roots of P lie on γ. Show that

$$\frac{1}{2\pi i}\int_\gamma (P'/P)(z)\,dz = W(\gamma, a_1) + \cdots + W(\gamma, a_n).$$

2. Let $f(z) = (z - z_0)^m h(z)$, where h is analytic on an open set U, and $h(z) \neq 0$ for all $z \in U$. Let γ be a closed path homologous to 0 in U, and such that z_0 does not lie on γ. Prove that

$$\frac{1}{2\pi i}\int_\gamma \frac{f'(z)}{f(z)}\,dz = W(\gamma, z_0)m.$$

3. Let U be a simply connected open set and let $z_1, \ldots, z_n$ be points of U. Let $U^* = U - \{z_1, \ldots, z_n\}$ be the set obtained from U by deleting the points $z_1, \ldots, z_n$. Let f be analytic on U^*. Let γ_k be a small circle centered at z_k and let

$$a_k = \frac{1}{2\pi i}\int_{\gamma_k} f(\zeta)\,d\zeta.$$

   Let $h(z) = f(z) - \sum a_k/(z - z_k)$. Prove that there exists an analytic function H on U^* such that $H' = h$.

Note. The train of thought of the above exercises will be pursued systematically in Chapter VI, Theorem 1.5.

IV, §3. ARTIN'S PROOF

In this section we prove Theorem 2.2 by making greater use of topological considerations. We reduce Theorem 2.2 to a theorem which involves only the winding number, and not the holomorphic function f, and we state this result as Theorem 3.2. The application to the holomorphic function will then be immediate by applying some results of Chapter III. We have already found that integrating along sides of a rectangle works better than over arbitrary curves. We pursue this idea. A path will be said to be **rectangular** if every curve of the path is either a horizontal segment or a vertical segment. We shall see that every path is homologous with a rectangular path, and in fact we prove:

Lemma 3.1. *Let γ be a path in an open set U. Then there exists a rectangular path η with the same end points, and such that γ, η are close together in U in the sense of Chapter* III, §4. *In particular, γ and η are* *homologous in U, and for any holomorphic function f on U we have*

$$\int_\gamma f = \int_\eta f.$$

Proof. Suppose γ is defined on an interval $[a, b]$. We take a partition of the interval,

$$a = a_0 \leqq a_1 \leqq a_2 \leqq \cdots \leqq a_n = b$$

such that the image of each small interval

$$\gamma([a_i, a_{i+1}])$$

is contained in a disc D_i on which f has a primitive. Then we replace the curve γ on the interval $[a_i, a_{i+1}]$ by the rectangular curve drawn on Fig. 12. This proves the lemma.

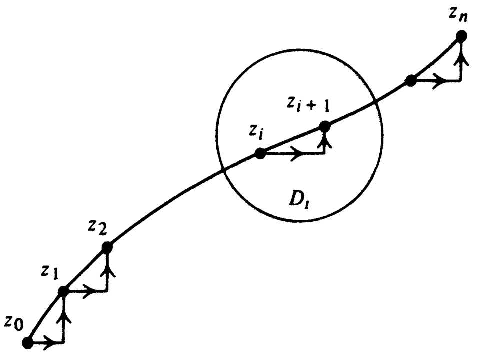


Figure 12

In the figure, we let $z_i = \gamma(a_i)$.

If γ is a closed path, then it is clear that the rectangular path constructed in the lemma is also a closed path, looking like this:

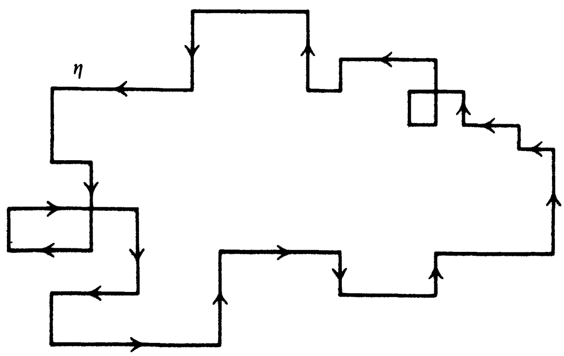


Figure 13

The lemma reduces the proof of Cauchy's theorem to the case when γ is a rectangular closed chain. We shall now reduce Cauchy's theorem to the case of rectangles by stating and proving a theorem having nothing to do with holomorphic functions. We need a little more terminology.

Let γ be a curve in an open set U, defined on an interval $[a, b]$. Let

$$a = a_0 \leqq a_1 \leqq a_2 \leqq \cdots \leqq a_n = b$$

be a partition of the interval. Let

$$\gamma_i\colon [a_i, a_{i+1}] \to U$$

be the restriction of γ to the smaller interval $[a_i, a_{i+1}]$. Then we agree to call the chain

$$\gamma_1 + \gamma_2 + \cdots + \gamma_n$$

a **subdivision** of γ. Furthermore, if η_i is obtained from γ_i by another parametrization, we again agree to call the chain

$$\eta_1 + \eta_2 + \cdots + \eta_n$$

a **subdivision of** γ. For any practical purposes, the chains γ and

$$\eta_1 + \eta_2 + \cdots + \eta_n$$

do not differ from each other. In Fig. 14 we illustrate such a chain γ and a subdivision $\eta_1 + \eta_2 + \eta_3 + \eta_4$.

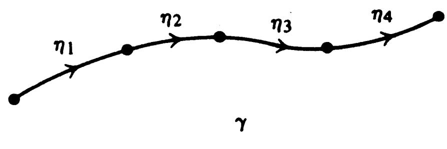


Figure 14

Similarly, if $\gamma = \sum m_i\gamma_i$ is a chain, and $\{\eta_{ij}\}$ is a subdivision of γ_i, we call

$$\sum_i \sum_j m_i\eta_{ij}$$

a **subdivision of** γ.

Theorem 3.2. *Let γ be a rectangular closed chain in U, and assume that γ is homologous to 0 in U, i.e.*

$$W(\gamma, \alpha) = 0$$

for every point α not in U. Then there exist rectangles $R_1, \ldots, R_N$

contained in U, such that if ∂R_i is the boundary of R_i oriented counterclockwise, then a subdivision of γ is equal to

$$\sum_{i=1}^{N} m_i \cdot \partial R_i$$

for some integers m_i.

Lemma 3.1 and Theorem 3.2 make Cauchy's Theorem 2.2 obvious because we know that for any holomorphic function f on U, we have

$$\int_{\partial R_i} f = 0$$

by Goursat's theorem. Hence the integral of f over the subdivision of γ is also equal to 0, whence the integral of f over γ is also equal to 0.

We now prove the theorem. Given the rectangular chain γ, we draw all vertical and horizontal lines passing through the sides of the chain, as illustrated on Fig. 15.

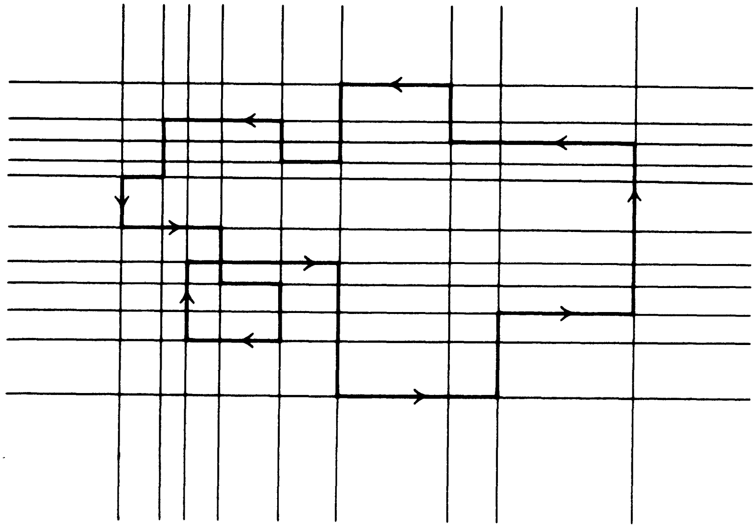


Figure 15

Then these vertical and horizontal lines decompose the plane into rectangles, and rectangular regions extending to infinity in the vertical and horizontal direction. Let R_i be one of the rectangles, and let α_i be a point inside R_i. Let

$$m_i = W(\gamma, \alpha_i).$$

For some rectangles we have $m_i = 0$, and for some rectangles, we have $m_i \neq 0$. We let $R_1, \ldots, R_N$ be those rectangles such that $m_1, \ldots, m_N$ are not 0, and we let ∂R_i be the boundary of R_i for $i = 1, \ldots, N$, oriented counterclockwise. We shall prove the following two assertions:

1. *Every rectangle R_i such that $m_i \neq 0$ is contained in U.*
2. *Some subdivision of γ is equal to*

$$\sum_{i=1}^{N} m_i \partial R_i.$$

This will prove the desired theorem.

Assertion 1. By assumption, α_i must be in U, because $W(\gamma, \alpha) = 0$ for every point α outside of U. Since the winding number is constant on connected sets, it is constant on the interior of R_i, hence $\neq 0$, and the interior of R_i is contained in U. If a boundary point of R_i is on γ, then it is in U. If a boundary point of R_i is not on γ, then the winding number with respect to γ is defined, and is equal to $m_i \neq 0$ by continuity (Lemma 3.2). This proves that the whole rectangle R_i, including its boundary, is contained in U, and proves the first assertion.

Assertion 2. We now replace γ by an appropriate subdivision. The vertical and horizontal lines cut γ in various points. We can then find a subdivision η of γ such that every curve occurring in η is some side of a rectangle, or the finite side of one of the infinite rectangular regions. The subdivision η is the sum of such sides, taken with appropriate multiplicities. If a finite side of an infinite rectangle occurs in the subdivision, after inserting one more horizontal or vertical line, we may assume that this side is also the side of a finite rectangle in the grid. Thus without loss of generality, we may assume that every side of the subdivision is also the side of one of the finite rectangles in the grid formed by the horizontal and vertical lines.

It will now suffice to prove that

$$\eta = \sum m_i \partial R_i.$$

Suppose $\eta - \sum m_i \partial R_i$ is not the 0 chain. Then it contains some horizontal or vertical segment σ, so that we can write

$$\eta - \sum m_i \partial R_i = m\sigma + C^*,$$

where m is an integer, and C^* is a chain of vertical and horizontal segments other than σ. Then σ is the side of a finite rectangle R_k. We take σ with the orientation arising from the counterclockwise orientation of the boundary of the rectangle R_k. Then the closed chain

$$C = \eta - \sum m_i \partial R_i - m \partial R_k$$

does not contain σ. Let α_k be a point interior to R_k, and let α' be a point near σ but on the opposite side from α_k, as shown on the figure.

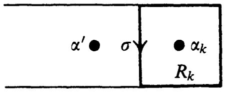


Figure 16

Since $\eta - \sum m_i \partial R_i - m \partial R_k$ does not contain σ, the points α_k and α' are connected by a line segment which does not intersect C. Therefore

$$W(C, \alpha_k) = W(C, \alpha').$$

But $W(\eta, \alpha_k) = m_k$ and $W(\partial R_i, \alpha_k) = 0$ unless $i = k$, in which case $W(\partial R_k, \alpha_k) = 1$. Similarly, if α' is inside some finite rectangle R_j, so $\alpha' = \alpha_j$, we have

$$W(\partial R_k, \alpha_j) = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

If α' is in an infinite rectangle, then $W(\partial R_k, \alpha') = 0$. Hence:

$$W(C, \alpha_k) = W\left(\eta - \sum m_i \partial R_i - m \partial R_k, \alpha_k\right) = m_k - m_k - m = -m;$$

$$W(C, \alpha') = W\left(\eta - \sum m_i \partial R_i - m \partial R_k, \alpha'\right) = 0.$$

This proves that $m = 0$, and concludes the proof that $\eta - \sum m_i \partial R_i = 0$.

CHAPTER V

Applications of Cauchy's Integral Formula

In this chapter, we return to the ideas of Theorem 7.3 of Chapter III, which we interrupted to discuss some topological considerations about winding numbers. We come back to analysis. We shall give various applications of the fact that the derivative of an analytic function can be expressed as an integral. This is completely different from real analysis, where the derivative of a real function often is less differentiable than the function itself. In complex analysis, one can exploit the phenomenon in various ways. For instance, in real analysis, a uniform limit of a sequence of differentiable functions may be only continuous. However, in complex analysis, we shall see that a uniform limit of analytic functions is analytic.

We shall also study a point where a function is analytic near the point, but not necessarily at the point itself. Such points are the isolated singular points of the function, and the behavior of the function can be described rather accurately near these points.

V, §1. UNIFORM LIMITS OF ANALYTIC FUNCTIONS

We first prove a general theorem that the uniform limit of analytic functions is analytic. This will allow us to define analytic functions by uniformly convergent series, and we shall give several examples, in text and in the exercises.

Theorem 1.1. *Let $\{f_n\}$ be a sequence of holomorphic functions on an open set U. Assume that for each compact subset K of U the sequence converges uniformly on K, and let the limit function be f. Then f is holomorphic.*

Proof. Let $z_0 \in U$, and let $\bar{D}_R$ be a closed disc of radius R centered at z_0 and contained in U. Then the sequence $\{f_n\}$ converges uniformly on $\bar{D}_R$. Let C_R be the circle which is the boundary of $\bar{D}_R$. Let $\bar{D}_{R/2}$ be the closed disc of radius $R/2$ centered at z_0. Then for $z \in \bar{D}_{R/2}$ we have

$$f_n(z) = \frac{1}{2\pi i}\int_{C_R} \frac{f_n(\zeta)}{\zeta - z}\,d\zeta,$$

and $|\zeta - z| \geqq R/2$. Since $\{f_n\}$ converges uniformly, for $|z - z_0| \leqq R/2$, we get

$$f(z) = \frac{1}{2\pi i}\int_{C_R} \frac{f(\zeta)}{\zeta - z}\,d\zeta.$$

By Theorem 7.7 of Chapter III it follows that f is holomorphic on a neighborhood of z_0. Since this is true for every z_0 in U, we have proved what we wanted.

Theorem 1.2. *Let $\{f_n\}$ be a sequence of analytic functions on an open set U, converging uniformly on every compact subset K of U to a function f. Then the sequence of derivatives $\{f_n'\}$ converges uniformly on every compact subset K, and* $\lim f_n' = f'$.

Proof. The proof will be left as an exercise to the reader. [*Hint*: Cover the compact set with a finite number of closed discs contained in U, and of sufficiently small radius. Cauchy's formula expresses the derivative f_n' as an integral, and one can argue as in the previous theorem.]

Example. Let

$$f(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}.$$

We shall prove that this function is holomorphic for $\operatorname{Re} z > 1$. Each term

$$f_n(z) = n^{-z} = e^{-z\log n}$$

is an entire function. Let $z = x + iy$. We have

$$|e^{-z\log n}| = |e^{-x\log n}e^{-iy\log n}| = n^{-x}.$$

Let $c > 1$. For $x \geqq c$ we have $|n^{-z}| \leqq n^{-c}$ and the series

$$\sum_{n=1}^{\infty} \frac{1}{n^c}$$

converges for $c > 1$. Hence the series $\sum f_n(z)$ converges uniformly and absolutely for $\operatorname{Re} z \geqq c$, and therefore defines a holomorphic function for $\operatorname{Re} z > c$. This is true for every $c > 1$, and hence f is holomorphic for $\operatorname{Re} z > 1$.

In the same example, we have

$$f_n'(z) = \frac{-\log n}{n^z}.$$

By Theorem 1.2, it follows that

$$f'(z) = \sum_{n=1}^{\infty} \frac{-\log n}{n^z}$$

in this same region.

V, §1. EXERCISES

1. Let f be analytic on an open set U, let $z_0 \in U$ and $f'(z_0) \neq 0$. Show that

$$\frac{2\pi i}{f'(z_0)} = \int_C \frac{1}{f(z) - f(z_0)}\, dz,$$

where C is a small circle centered at z_0.

2. Weierstrass' theorem for a real interval $[a, b]$ states that a continuous function can be uniformly approximated by polynomials. Is this conclusion still true for the closed unit disc, i.e. can every continuous function on the disc be uniformly approximated by polynomials?

3. Let $a > 0$. Show that each of the following series represents a holomorphic function:

(a) $\displaystyle\sum_{n=1}^{\infty} e^{-an^2 z}$ for $\operatorname{Re} z > 0$;

(b) $\displaystyle\sum_{n=1}^{\infty} \frac{e^{-anz}}{(a+n)^2}$ for $\operatorname{Re} z > 0$;

(c) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{(a+n)^z}$ for $\operatorname{Re} z > 1$.

4. Show that each of the two series converges uniformly on each closed disc $|z| \leqq c$ with $0 < c < 1$:

$$\sum_{n=1}^{\infty} \frac{nz^n}{1 - z^n} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{z^n}{(1 - z^n)^2}.$$

5. Prove that the two series in Exercise 4 are actually equal. [*Hint*: Write each one in a double series and reverse the order of summation.]

6. **Dirichlet Series.** Let $\{a_n\}$ be a sequence of complex numbers. Show that the series $\sum a_n/n^s$, if it converges absolutely for some complex s, converges absolutely in a right half-plane $\mathrm{Re}(s) > \sigma_0$, and uniformly in $\mathrm{Re}(s) > \sigma_0 + \varepsilon$ for every $\epsilon > 0$. Show that the series defines an analytic function in this half plane. The number σ_0 is called the **abscissa of convergence**.

The next exercises give expressions and estimates for an analytic function in terms of integrals.

7. Let f be analytic on a closed disc $\bar{D}$ of radius $b > 0$, centered at z_0. Show that

$$\frac{1}{\pi b^2} \iint_D f(x + iy)\, dy\, dx = f(z_0).$$

[*Hint*: Use polar coordinates and Cauchy's formula. Without loss of generality, you may assume that $z_0 = 0$. Why?]

8. Let D be the unit disc and let S be the unit square, that is, the set of complex numbers z such that $0 < \mathrm{Re}(z) < 1$ and $0 < \mathrm{Im}(z) < 1$. Let $f: D \to S$ be an analytic isomorphism such that $f(0) = (1 + i)/2$. Let u, v be the real and imaginary parts of f respectively. Compute the integral

$$\iint_D \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right] dx\, dy.$$

9. (a) Let f be an analytic isomorphism on the unit disc D, and let

$$f(z) = \sum_{n=1}^{\infty} a_n z^n$$

be its power series expansion. Prove that

$$\text{area } f(D) = \pi \sum_{n=1}^{\infty} n|a_n|^2.$$

(b) Suppose that f is an analytic isomorphism on the closed unit disc $\bar{D}$, and that $|f(z)| \geqq 1$ if $|z| = 1$, and $f(0) = 0$. Prove that area $f(D) \geqq \pi$.

10. Let f be analytic on the unit disc D and assume that $\iint_D |f|^2\, dx\, dy$ exists. Let

$$f(z) = \sum_{n=0}^{\infty} a_n z^n.$$

Prove that

$$\frac{1}{2\pi} \iint_D |f(z)|^2\, dx\, dy = \sum_{n=0}^{\infty} |a_n|^2/(2n + 2).$$

For the next exercise, recall that a norm $\|\ \|$ on a space of functions associates to each function f a real number $\geqq 0$, satisfying the following conditions:

N 1. We have $\|f\| = 0$ if and only if $f = 0$.

N 2. If c is a complex number, then $\|cf\| = |c|\,\|f\|$.

N 3. $\|f + g\| \leqq \|f\| + \|g\|$.

11. Let A be the closure of a bounded open set in the plane. Let f, g be continuous functions on A. Define their scalar product

$$\langle f, g \rangle = \iint_A f(z)\overline{g(z)}\, dy\, dx$$

and define the associated L^2-**norm** by its square,

$$\|f\|_2^2 = \iint_A |f(z)|^2\, dy\, dx.$$

Show that $\|f\|_2$ does define a norm. Prove the **Schwarz inequality**

$$|\langle f, g \rangle| \leqq \|f\|_2 \|g\|_2 .$$

On the other hand, define

$$\|f\|_1 = \iint_A |f(z)|\, dy\, dx.$$

Show that $f \mapsto \|f\|_1$ is a norm on the space of continuous functions on A, called the L^1-**norm**. This is just preliminary. Prove:

(a) Let $0 < s < R$. Prove that there exist constants C_1, C_2 having the following property. If f is analytic on a closed disc $\bar{D}$ of radius R, then

$$\|f\|_s \leqq C_1 \|f\|_{1,R} \leqq C_2 \|f\|_{2,R},$$

where $\|\ \|_s$ is the sup norm on the closed disc of radius s, and the L^1, L^2 norms refer to the integral over the disc of radius R.

(b) Let $\{f_n\}$ be a sequence of holomorphic functions on an open set U, and assume that this sequence is L^2-Cauchy. Show that it converges uniformly on compact subsets of U.

12. Let U, V be open discs centered at the origin. Let $f = f(z, w)$ be a *continuous* function on the product $U \times V$, such that for each w the function $z \mapsto f(z, w)$ and for each z the function $w \mapsto f(z, w)$ are analytic on U and V, respectively. Show that f has a power series expansion

$$f(z, w) = \sum a_{mn} z^m w^n$$

which converges absolutely and uniformly for $|z| \leqq r$ and $|w| \leqq r$, for some positive number r. [*Hint*: Apply Cauchy's formula for derivatives twice, with respect to the two variables to get an estimate for the coefficients a_{mn}.] Generalize to several variables instead of two variables.

Note. This exercise is really quite trivial, although it is not generally realized that it is so. *The point is that the function f is assumed to be continuous.* If that assumption is not made, the situation becomes much more difficult to handle, and the result is known as Hartogs' theorem. In practice, continuity is indeed satisfied.

V, §2. LAURENT SERIES

By a **Laurent series**, we mean a series

$$f(z) = \sum_{n=-\infty}^{\infty} a_n z^n.$$

Let A be a set of complex numbers. We say that the Laurent series **converges absolutely** (resp. uniformly) on A if the two series

$$f^+(z) = \sum_{n \geqq 0} a_n z^n \qquad \text{and} \qquad f^-(z) = \sum_{n<0} a_n z^n$$

converge absolutely (resp. uniformly) on A. If that is the case, then $f(z)$ is regarded as the sum,

$$f(z) = f^+(z) + f^-(z).$$

Let r, R be positive numbers with $0 \leqq r < R$. We shall consider the annulus A consisting of all complex numbers z such that

$$r \leqq |z| \leqq R.$$

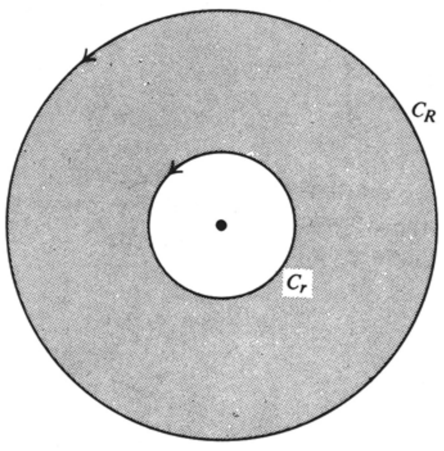


Figure 1

Theorem 2.1. *Let A be the above annulus, and let f be a holomorphic function on A. Let $r < s < S < R$. Then f has a Laurent expansion*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n z^n$$

which converges absolutely and uniformly on $s \leqq |z| \leqq S$. Let C_R and C_r be the circles of radius R and r, respectively. Then the coefficients a_n are obtained by the usual formula:

$$a_n = \frac{1}{2\pi i} \int_{C_R} \frac{f(\zeta)}{\zeta^{n+1}} \, d\zeta \qquad \text{if} \quad n \geqq 0,$$

$$a_n = \frac{1}{2\pi i} \int_{C_r} \frac{f(\zeta)}{\zeta^{n+1}} \, d\zeta \qquad \text{if} \quad n < 0.$$

Proof. For some $\epsilon > 0$ we may assume (by the definition of what it means for f to be holomorphic on the closed annulus) that f is holomorphic on the open annulus U of complex numbers z such that

$$r - \epsilon < |z| < R + \epsilon.$$

The chain $C_R - C_r$ is homologous to 0 on U, because if a point lies in the outer part then its winding number is zero by the usual Lemma 1.3 of Chapter IV, and if the point lies in the disc inside the annulus, then its winding number is 0. Cauchy's formula then implies that for z in the annulus,

$$f(z) = \frac{1}{2\pi i} \int_{C_R} \frac{f(\zeta)}{\zeta - z} \, d\zeta - \frac{1}{2\pi i} \int_{C_r} \frac{f(\zeta)}{\zeta - z} \, d\zeta.$$

We may now prove the theorem. The first integral is handled just as in the ordinary case of the derivation of Cauchy's formula, and the second is handled in a similar manner as follows. We write

$$\zeta - z = -z\left(1 - \frac{\zeta}{z}\right).$$

Then

$$\left|\frac{\zeta}{z}\right| \leqq r/s < 1,$$

so the geometric series converges,

$$\frac{1}{z} \frac{1}{1 - \zeta/z} = \frac{1}{z}\left(1 + \frac{\zeta}{z} + \left(\frac{\zeta}{z}\right)^2 + \cdots\right).$$

We can then integrate term by term, and the desired expansion falls out. To show the uniqueness of the coefficients, integrate the series $\sum a_n s^n e^{in\theta}$ against $e^{-ik\theta}$ for a given integer k, term by term from 0 to 2π. All terms drop out except for $n = k$, showing that the k-th coefficient is determined by f.

An example of a function with a Laurent series with infinitely many negative terms is given by $e^{1/z}$, that is, by substituting $1/z$ in the ordinary exponential series.

If an annulus is centered at a point z_0, then one obtains a Laurent series at z_0 of the form

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n.$$

Example. We want to find the Laurent series for

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

for $0 < |z| < 1$. We write f in partial fractions:

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

Then for one term we get the geometric series,

$$\frac{1}{z-1} = -\frac{1}{1-z} = -(1 + z + z^2 + \cdots)$$

whence

$$f(z) = -\frac{1}{z} - 1 - z - z^2 - \cdots.$$

On the other hand, suppose we want the Laurent series for $|z| > 1$. Then we write

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

whence

$$f(z) = \frac{1}{z^2} + \frac{1}{z^3} + \frac{1}{z^4} + \cdots.$$

V, §2. EXERCISES

1. Prove that the Laurent series can be differentiated term by term in the usual manner to give the derivative of f on the annulus.

2. Let f be holomorphic on the annulus A, defined by $0 < r \leqq |z| \leqq R$.

Prove that there exist functions f_1, f_2 such that f_1 is holomorphic for $|z| \leqq R$, f_2 is holomorphic for $|z| \geqq r$ and

$$f = f_1 + f_2$$

on the annulus.

3. Is there a polynomial $P(z)$ such that $P(z)e^{1/z}$ is an entire function? Justify your answer. What is the Laurent expansion of $e^{1/z}$ for $|z| \neq 0$?

4. Expand the function

$$f(z) = \frac{z}{1 + z^3}$$

(a) in a series of positive powers of z, and
(b) in a series of negative powers of z.
In each case, specify the region in which the expansion is valid.

5. Give the Laurent expansions for the following functions:
(a) $z/(z + 2)$ for $|z| > 2$ (b) $\sin 1/z$ for $z \neq 0$
(c) $\cos 1/z$ for $z \neq 0$ (d) $\dfrac{1}{(z - 3)}$ for $|z| > 3$

6. Prove the following expansions:

(a) $e^z = e + e \sum_{n=1}^{\infty} \dfrac{1}{n!}(z - 1)^n$

(b) $1/z = \sum_{n=0}^{\infty} (-1)^n(z - 1)^n$ for $|z - 1| < 1$

(c) $1/z^2 = 1 + \sum_{n=1}^{\infty} (n + 1)(z + 1)^n$ for $|z + 1| < 1$

7. Expand (a) $\cos z$, (b) $\sin z$ in a power series about $\pi/2$.

8. Let $f(z) = \dfrac{1}{(z - 1)(z - 2)}$. Find the Laurent series for f:
(a) In the disc $|z| < 1$.
(b) In the annulus $1 < |z| < 2$.
(c) In the region $2 < |z|$.

9. Find the Laurent series for $(z + 1)/(z - 1)$ in the region (a) $|z| < 1$; (b) $|z| > 1$.

10. Find the Laurent series for $1/z^2(1 - z)$ in the regions:
(a) $0 < |z| < 1$; (b) $|z| > 1$.

11. Find the power series expansion of

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

around the point $z = 1$, and find the radius of convergence of this series.

12. Find the Laurent expansion of

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

for $1 < |z| < 2$.

13. Obtain the first four terms of the Laurent series expansion of

$$f(z) = \frac{e^z}{z(z^2+1)}$$

valid for $0 < |z| < 1$.

*14. Assume that f is analytic in the upper half plane, and that f is periodic of period 1. Show that f has an expansion of the form

$$f = \sum_{-\infty}^{\infty} c_n e^{2\pi i n z},$$

where

$$c_n = \int_0^1 f(x+iy)e^{-2\pi i n(x+iy)}\,dx,$$

for any value of $y > 0$. [*Hint*: Show that there is an analytic function f^* on a disc from which the origin is deleted such that

$$f^*(e^{2\pi i z}) = f(z).$$

What is the Laurent series for f^*? Abbreviate $q = e^{2\pi i z}$.

*15. Assumptions being as in Exercise 14, suppose in addition that there exists $y_0 > 0$ such that $f(z) = f(x+iy)$ is bounded in the domain $y \geqq y_0$. Prove that the coefficients c_n are equal to 0 for $n < 0$. Is the converse true? Proof?

V, §3. ISOLATED SINGULARITIES

Let z_0 be a complex number and let D be an open disc centered at z_0. Let U be the open set obtained by removing z_0 from D. A function f which is analytic on U is said to have an **isolated singularity** at z_0. We suppose this is the case.

Removable Singularities

Theorem 3.1. *If f is bounded in some neighborhood of z_0, then one can define $f(z_0)$ in a unique way such that the function is also analytic at z_0.*

Proof. Say $z_0 = 0$. By §2, we know that f has a Laurent expansion

$$f(z) = \sum_{n \geqq 0} a_n z^n + \sum_{n < 0} a_n z^n$$

for $0 < |z| < R$. We have to show $a_n = 0$ if $n < 0$. Let $n = -m$ with $m > 0$. We have

$$a_{-m} = \frac{1}{2\pi i} \int_{C_r} f(\zeta)\zeta^{m-1}\, d\zeta,$$

for any circle C_r of small radius r. Since f is assumed bounded near 0 it follows that the right-hand side tends to 0 as r tends to 0, whence $a_{-m} = 0$, as was to be shown. (The uniqueness is clear by continuity.)

In the case of Theorem 3.1 it is customary to say that z_0 is a **removable singularity**.

Poles

Suppose the Laurent expansion of f in the neighborhood of a singularity z_0 has only a finite number of negative terms,

$$f(z) = \frac{a_{-m}}{(z - z_0)^m} + \cdots + a_0 + a_1(z - z_0) + \cdots,$$

and $a_{-m} \neq 0$. Then f is said to have a **pole of order** (or multiplicity m) at z_0. However, we still say that the **order of f at z_0 is** $-m$, that is,

$$\operatorname{ord}_{z_0} f = -m,$$

because we want the formula

$$\operatorname{ord}_{z_0}(fg) = \operatorname{ord}_{z_0} f + \operatorname{ord}_{z_0} g$$

to be true. This situation is characterized as follows:

f has a pole of order m at z_0 if and only if $f(z)(z - z_0)^m$ is holomorphic at z_0 and has no zero at z_0.

The proof is immediate and is left to the reader.

If g is holomorphic at z_0 and $g(z_0) \neq 0$, then the function f defined by

$$f(z) = (z - z_0)^{-m} g(z)$$

in a neighborhood of z_0 from which z_0 is deleted, has a pole of order m. We abide by the convention that a pole is a zero of negative order.

A pole of order 1 is said to be a **simple pole**.

Examples. The function $1/z$ has a simple pole at the origin.

The function $1/\sin z$ has a simple pole at the origin. This comes from the power series expansion, since

$$\sin z = z(1 + \text{higher terms}),$$

and

$$\frac{1}{\sin z} = \frac{1}{z}(1 + \text{higher terms})$$

by inverting the series $1/(1 - h) = 1 + h + h^2 + \cdots$ for $|h| < 1$.

Let f be defined on an open set U except at a discrete set of points S which are poles. Then we say that f is **meromorphic** on U. If z_0 is such a point, then there exists an integer m such that $(z - z_0)^m f(z)$ is holomorphic in a neighborhood of z_0. Thus f is the quotient of two holomorphic functions in the neighborhood of a point. We say that f is **meromorphic at a point** z_0 if f is meromorphic on some open set U containing z_0.

Example. Let $P(z)$ be a polynomial. Then $f(z) = 1/P(z)$ is a meromorphic function. This is immediately seen by factoring $P(z)$ into linear factors.

Example. A meromorphic function can be defined often by a uniformly convergent series. For instance, let

$$f(z) = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{z}{z^2 - n^2}.$$

We claim that f is meromorphic on $\mathbf{C}$ and has simple poles at the integers, but is holomorphic elsewhere.

We prove that f has these properties inside every disc of radius R centered at the origin. Let $R > 0$ and let $N > 2R$. Write

$$f(z) = g(z) + h(z),$$

where

$$g(z) = \frac{1}{z} + \sum_{n=1}^{N} \frac{z}{z^2 - n^2} \qquad \text{and} \qquad h(z) = \sum_{n=N+1}^{\infty} \frac{z}{z^2 - n^2}.$$

Then g is a rational function, and is therefore meromorphic on **C**. Furthermore, from its expression as a finite sum, we see that g has simple poles at the integers n such that $|n| \leqq N$.

For the infinite series defining h, we apply Theorem 1.1 and prove that the series is uniformly convergent. For $|z| < R$ we have the estimate

$$\left|\frac{z}{z^2 - n^2}\right| \leqq \frac{R}{n^2 - R^2} = \frac{1}{n^2} \frac{R}{1 - (R/n)^2}.$$

The denominator satisfies

$$1 - (R/n)^2 \geqq \tfrac{3}{4}$$

for $n > N > 2R$. Hence

$$\left|\frac{z}{z^2 - n^2}\right| \leqq \frac{4R}{3n^2} \qquad \text{for} \quad n \geqq 2R.$$

Therefore the series for h converges uniformly in the disc $|z| < R$, and h is holomorphic in this disc. This proves the desired assertion.

Essential Singularities

If the Laurent series has a infinite number of negative terms, then we say that z_0 is an **essential singularity** of f.

Example. The function $f(z) = e^{1/z}$ has an essential singularity at $z = 0$ because its Laurent series is

$$\sum_{n=0}^{\infty} \frac{1}{z^n n!}.$$

Theorem 3.2 (Casorati–Weierstrass). *Let* 0 *be an essential singularity of the function* f, *and let* D *be a disc centered at* 0 *on which* f *is holomorphic except at* 0. *Let* U *be the complement of* 0 *in* D. *Then* $f(U)$ *is dense in the complex numbers. In other words, the values of* f *on* U *come arbitrarily close to any complex number.*

Proof. Suppose the theorem is false. There exists a complex number α and a positive number $s > 0$ such that

$$|f(z) - \alpha| > s \qquad \text{for all } z \in U.$$

The function

$$g(z) = \frac{1}{f(z) - \alpha}$$

is then holomorphic on U, and bounded on the disc D. Hence 0 is a removable singularity of g, and g may be extended to a holomorphic function on all of D. It then follows that $1/g(z)$ has at most a pole at 0, which means that $f(z) - \alpha$ has at most a pole, contradicting the hypothesis that $f(z)$ has an essential singularity (infinitely many terms of negative order in its Laurent series). This proves the theorem.

Actually, it was proved by Picard that f not only comes arbitrarily close to every complex number, but takes on every complex value except possibly one. The function $e^{1/z}$ omits the value 0, so it is necessary to allow for this one omission. See Chapter XI, §3 and Chapter XII, §2.

We recall that an **analytic isomorphism**

$$f: U \to V$$

from one open set to another is an analytic function such that there exists another analytic function

$$g: V \to U$$

satisfying

$$f \circ g = \mathrm{id}_V \qquad \text{and} \qquad g \circ f = \mathrm{id}_U,$$

where id is the identity function. An **analytic automorphism** of U is an analytic isomorphism of U with itself.

Using the Casorati–Weierstrass theorem, we shall prove:

Theorem 3.3. *The only analytic automorphisms of* $\mathbf{C}$ *are the functions of the form* $f(z) = az + b$, *where* a, b *are constants,* $a \neq 0$.

Proof. Let f be an analytic automorphism of $\mathbf{C}$. After making a translation by $-f(0)$, we may assume without loss of generality that $f(0) = 0$. We then have to prove that $f(z) = az$ for some constant a. Let

$$h(z) = f(1/z) \qquad \text{for} \quad z \neq 0.$$

Then h is defined for all complex numbers except for the origin. We first prove that h cannot have an essential singularity at 0. Since f is a local analytic isomorphism at 0, f gives a bijection between an open neighborhood of 0 with some open neighborhood of 0. Since f is also an analytic isomorphism of **C**, it follows that there exists $\delta > 0$ and $c > 0$ such that if $|w| > 1/\delta$ then $|f(w)| > c$. Let $z = 1/w$ or $w = 1/z$. Then $|h(z)| > c$ for $|z| < \delta$. If 0 is an essential singularity, this contradicts the Casorati–Weierstrass theorem.

Let $f(z) = \sum a_n z^n$ so $h(z) = \sum a_n(1/z)^n$. Since 0 is not an essential singularity of h, it follows that the series for h, hence for f, has only a finite number of terms, and

$$f(z) = a_0 + a_1 z + \cdots + a_N z^N$$

is a polynomial of degree N for some N. If f has two distinct roots, then f cannot be injective, contradicting the fact that f has an inverse function. Hence f has only one root, and

$$f(z) = a(z - z_0)^N$$

for some z_0. If $N > 1$, it is then clear that f is not injective so we must have $N = 1$, and the theorem is proved.

V, §3. EXERCISES

1. Show that the following series define a meromorphic function on **C** and determine the set of poles, and their orders.

(a) $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n+z)}$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{z^2+n^2}$ (c) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{(z+n)^2}$

(d) $\displaystyle\sum_{n=1}^{\infty} \frac{\sin nz}{n!(z^2+n^2)}$ (e) $\displaystyle\frac{1}{z} + \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \left[\frac{1}{z-n} + \frac{1}{n}\right]$

2. Show that the function

$$f(z) = \sum_{n=1}^{\infty} \frac{z^2}{n^2 z^2 + 8}$$

is defined and continuous for the real values of z. Determine the region of the complex plane in which this function is analytic. Determine its poles.

3. Show that the series

$$\sum_{n=1}^{\infty} \left(\frac{z+i}{z-i}\right)^n$$

defines an analytic function on a disc of radius 1 centered at $-i$.

4. Let $\{z_n\}$ be a sequence of distinct complex numbers such that

$$\sum \frac{1}{|z_n|^3} \text{ converges.}$$

Prove that the series

$$\sum_{n=1}^{\infty} \left(\frac{1}{(z - z_n)^2} - \frac{1}{z_n^2} \right)$$

defines a meromorphic function on $\mathbf{C}$. Where are the poles of this function?

5. Let f be meromorphic on $\mathbf{C}$ but not entire. Let $g(z) = e^{f(z)}$. Show that g is not meromorphic on $\mathbf{C}$.

6. Let f be a non-constant entire function, i.e. a function analytic on all of $\mathbf{C}$. Show that the image of f is dense in $\mathbf{C}$.

7. Let f be meromorphic on an open set U. Let

$$\varphi: V \to U$$

be an analytic isomorphism. Suppose $\varphi(z_0) = w_0$, and f has order n at w_0. Show that $f \circ \varphi$ has order n at z_0. In other words, the order is invariant under analytic isomorphisms. [Here n is a positive or negative integer.]

8. A meromorphic function f is said to be **periodic** with period w if $f(z + w) = f(z)$ for all $z \in \mathbf{C}$. Let f be a meromorphic function, and suppose f is periodic with three periods w_1, w_2, w_3 which are linearly independent over the rational numbers. Prove that f is constant. [*Hint*: Prove that there exist elements w which are integral linear combinations of w_1, w_2, w_3 and arbitrarily small in absolute value.] The exponential function is an example of a singly periodic function. Examples of doubly periodic functions will be given in Chapter XIV.

9. Let f be meromorphic on $\mathbf{C}$, and suppose

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

Prove that f is a rational function. (You cannot assume as given that f has only a finite number of poles.)

10. **(The Riemann Sphere).** Let S be the union of $\mathbf{C}$ and a single other point denoted by ∞, and called **infinity**. Let f be a function on S. Let $t = 1/z$, and define

$$g(t) = f(1/t)$$

for $t \neq 0$, ∞. We say that f has an **isolated singularity** (resp. is **meromorphic** resp. is **holomorphic**) **at infinity** if g has an isolated singularity (resp. is meromorphic, resp. is holomorphic) at 0. The order of g at 0 will also be called the **order of** f **at infinity**. If g has a removable singularity at 0, and so can be defined as a holomorphic function in a neighborhood of 0, then we say that f is **holomorphic at infinity**.

We say that f is meromorphic on S if f is meromorphic on $\mathbf{C}$ and is also meromorphic at infinity. We say that f is holomorphic on S if f is holomorphic on $\mathbf{C}$ and is also holomorphic at infinity.

Prove:

The only meromorphic functions on S are the rational functions, that is, quotients of polynomials. The only holomorphic functions on S are the constants. If f is holomorphic on $\mathbf{C}$ and has a pole at infinity, then f is a polynomial.

In this last case, how would you describe the order of f at infinity in terms of the polynomial?

11. Let f be a meromorphic function on the Riemann sphere, so a rational function by Exercise 10. Prove that

$$\sum_P \operatorname{ord}_P f = 0,$$

where the sum is taken over all points P which are either points of $\mathbf{C}$, or $P = \infty$.

12. Let P_i $(i = 1, \ldots, r)$ be points of $\mathbf{C}$ or ∞, and let m_i be integers such that

$$\sum_{i=1}^{r} m_i = 0.$$

Prove that there exists a meromorphic function f on the Riemann sphere such that

$$\operatorname{ord}_{P_i} f = m_i \qquad \text{for} \quad i = 1, \ldots, r$$

and $\operatorname{ord}_P f = 0$ if $P \neq P_i$.

CHAPTER VI

Calculus of Residues

We have established all the theorems needed to compute integrals of analytic functions in terms of their power series expansions. We first give the general statements covering this situation, and then apply them to examples.

VI, §1. THE RESIDUE FORMULA

Let

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n$$

have a Laurent expansion at a point z_0. We call a_{-1} the **residue** of f at z_0, and write

$$a_{-1} = \operatorname{Res}_{z_0} f.$$

Theorem 1.1. *Let z_0 be an isolated singularity of f, and let C be a small circle oriented counterclockwise, centered at z_0 such that f is holomorphic on C and its interior, except possibly at z_0. Then*

$$\int_C f(\zeta)\, d\zeta = 2\pi i a_{-1} = 2\pi i \operatorname{Res}_{z_0} f.$$

Proof. Since the series for $f(\zeta)$ converges uniformly and absolutely for ζ on the circle, we may integrate it term by term. The integral of

$(\zeta - z_0)^n$ over the circle is equal to 0 for all values of n except possibly when $n = -1$, in which case we know that the value is $2\pi i$, cf. Examples 1 and 4 of Chapter III, §2. This proves the theorem.

From this local result, we may then deduce a global result for more general paths, by using the reduction of Theorem 2.4, Chapter IV.

Theorem 1.2 (Residue Formula). *Let U be an open set, and γ a closed chain in U such that γ is homologous to 0 in U. Let f be analytic on U except at a finite number of points $z_1, \ldots, z_n$. Let $m_i = W(\gamma, z_i)$. Then*

$$\int_\gamma f = 2\pi\sqrt{-1} \sum_{i=1}^{n} m_i \cdot \mathrm{Res}_{z_i} f.$$

Proof. Immediate by plugging Theorem 1.1 in the above mentioned theorem of Chapter IV.

Theorem 1.2 is used most often when U is simply connected, in which case every closed path is homologous to 0 in U, and the hypothesis on γ need not be mentioned explicitly. In the applications, U will be a disc, or the inside of a rectangle, where the simple connectedness is obvious.

Remark. The notation $\sqrt{-1}$ is the standard device used when we don't want to confuse the complex number i with an index i.

We shall give examples how to find residues.

A pole of a function f is said to be **simple** if it is of order 1, in which case the power series expansion of f is of type

$$f(z) = \frac{a_{-1}}{z - z_0} + a_0 + \text{higher terms},$$

and $a_{-1} \neq 0$.

Lemma 1.3.

(a) *Let f have a simple pole at z_0, and let g be holomorphic at z_0. Then*

$$\mathrm{Res}_{z_0}(fg) = g(z_0)\,\mathrm{Res}_{z_0}(f).$$

(b) *Suppose $f(z_0) = 0$ but $f'(z_0) \neq 0$. Then $1/f$ has a pole of order 1 at z_0 and the residue of $1/f$ at z_0 is $1/f'(z_0)$.*

Proof. (a) Let $f(z) = a_{-1}/(z - z_0) +$ higher terms. Say $z_0 = 0$ for simplicity of notation. We have

$$f(z)g(z) = \left(\frac{a_{-1}}{z} + \cdots\right)(b_0 + b_1 z + \cdots)$$

$$= \frac{a_{-1}b_0}{z} + \text{higher terms,}$$

so our assertion is clear.

(b) Let $f(z_0) = 0$ but $f'(z_0) \neq 0$. Then $f(z) = a_1(z - z_0) +$ higher terms, and $a_1 \neq 0$. Say $z_0 = 0$ for simplicity. Then

$$f(z) = a_1 z(1 + h) \qquad \text{with} \quad \operatorname{ord} h \geqq 1,$$

so

$$\frac{1}{f(z)} = \frac{1}{a_1 z}(1 - h + h^2 - \cdots) = \frac{1}{a_1 z} + \text{higher terms,}$$

so $\operatorname{res}(1/f) = a_1^{-1} = 1/f'(0)$, as was to be shown.

Remark. Part (a) of the lemma merely repeats what you should have seen before, to make this chapter more systematic.

Example. We give an example for part (b) of the lemma. Let $f(z) = \sin z$. Then f has a simple zero at $z = \pi$, because $f'(z) = \cos z$ and $f'(\pi) = -1 \neq 0$. Hence $1/f(z)$ has a simple pole at $z = \pi$, and

$$\operatorname{res}_\pi \frac{1}{\sin z} = \frac{1}{\cos \pi} = -1.$$

Example. Find the residue of $f(z) = \dfrac{z^2}{z^2 - 1}$ at $z = 1$.

To do this, we write

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

Note that $g(z) = z^2/(z + 1)$ is holomorphic at 1, and that the residue of $1/(z - 1)$ is 1. Hence

$$\operatorname{Res}_1 f = g(1) = 1/2.$$

Example. Find the residue of $(\sin z)/z^2$ at $z = 0$. We have

$$\frac{\sin z}{z^2} = \frac{1}{z^2}\left(z - \frac{z^3}{3!} + \cdots\right)$$
$$= \frac{1}{z} + \text{higher terms.}$$

Hence the desired residue is 1.

Example. Find the residue of $f(z) = \dfrac{z^2}{(z+1)(z-1)^2}$ at $z = 1$.

We note that the function

$$g(z) = \frac{z^2}{z+1}$$

is holomorphic at $z = 1$, and has an expansion of type

$$g(z) = b_0 + b_1(z-1) + \text{higher terms.}$$

Then

$$f(z) = \frac{g(z)}{(z-1)^2} = \frac{b_0}{(z-1)^2} + \frac{b_1}{(z-1)} + \cdots$$

and therefore the residue of f at 1 is b_1, which we must now find. We write $z = 1 + (z-1)$, so that

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

Inverting by the geometric series gives

$$\frac{z^2}{z+1} = \frac{1}{2}\left(1 + \frac{3}{2}(z-1) + \cdots\right).$$

Therefore

$$f(z) = \frac{1}{2(z-1)^2} + \frac{3/4}{z-1} + \cdots$$

whence $\text{Res}_1 f = 3/4$.

Example. Let C be a circle centered at 1, of radius 1. Let

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

Find $\int_C f$.

The function f has only two singularities, at 1 and -1, and the circle is contained in a disc of radius > 1, centered at 1, on which f is holomorphic except at $z = 1$. Hence the residue formula and the preceding example give us

$$\int_C f = 2\pi i \cdot \tfrac{3}{4}.$$

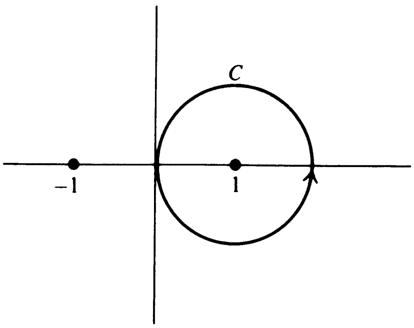


Figure 1

If C is the boundary of the rectangle as shown on Fig. 2, then we also find

$$\int_C f = 2\pi i \cdot \tfrac{3}{4}.$$

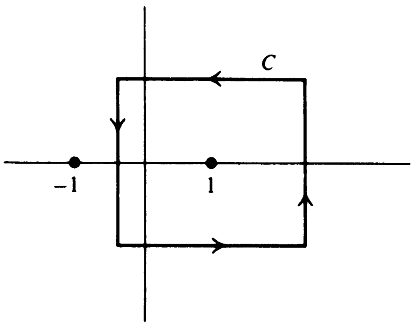


Figure 2

Example. Let $f(z) = z^2 - 2z + 3$. Let C be a rectangle as shown on Fig. 3, oriented clockwise. Find

$$\int_C \frac{1}{f(z)}\, dz.$$

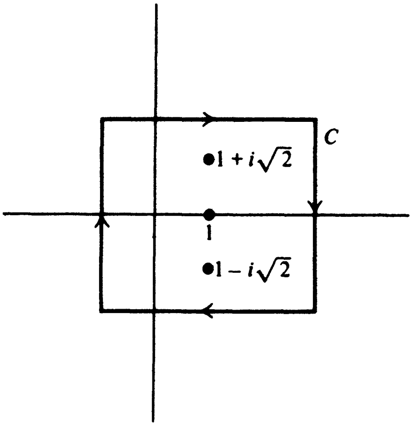


Figure 3

The roots of $f(z)$ are found by the quadratic formula to be

$$\frac{2 \pm \sqrt{-8}}{2}$$

and so are $z_1 = 1 + i\sqrt{2}$ and $z_2 = 1 - i\sqrt{2}$. The rectangle goes around these two points, in the clockwise direction. The residue of $1/f(z)$ at z_1 is $1/(z_1 - z_2)$ because f has a simple pole at z_1. The residue of $1/f(z)$ at z_2 is $1/(z_2 - z_1)$ for the same reason. The desired integral is equal to

$$-2\pi i(\text{sum of the residues}) = 0.$$

Example. Let f be the same function as in the preceding example, but now find the integral of $1/f$ over the rectangle as shown on Fig. 4. The rectangle is oriented clockwise. In this case, we have seen that the residue at $1 - i\sqrt{2}$ is

$$\frac{1}{z_2 - z_1} = \frac{1}{-2i\sqrt{2}}.$$

Therefore the integral over the rectangle is equal to

$$-2\pi i(\text{residue}) = -2\pi i/(-2i\sqrt{2}) = \pi/\sqrt{2}.$$

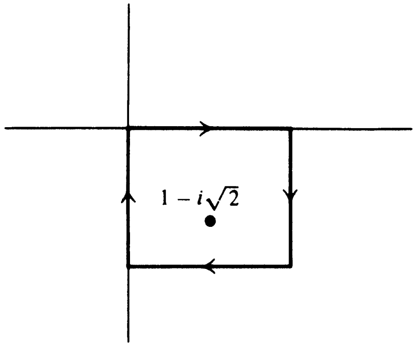


Figure 4

Next we give an example which has theoretical significance, besides computational significance.

Example. Let f have a power series expansion with only a finite number of negative terms (so at most a pole), say at the origin,

$$f(z) = a_m z^m + \text{higher terms}, \qquad a_m \neq 0,$$

and m may be positive or negative. Then we can write

$$f(z) = a_m z^m (1 + h(z)),$$

where $h(z)$ is a power series with zero constant term. For any two functions f, g we know the derivative of the product,

$$(fg)' = f'g + fg',$$

so that dividing by fg yields

$$\boxed{\frac{(fg)'}{fg} = \frac{f'}{f} + \frac{g'}{g}.}$$

Therefore we find for $f(z) = (a_m z^m)(1 + h(z))$,

$$\frac{f'(z)}{f(z)} = \frac{m}{z} + \frac{h'(z)}{1 + h(z)}$$

and $h'(z)/(1 + h(z))$ is holomorphic at 0. Consequently, we get:

Lemma 1.4. *Let f be meromorphic at 0. Then*

$$\operatorname{Res}_0 f'/f = \operatorname{ord}_0 f$$

and for any point z_0 where f has at most a pole,

$$\operatorname{Res}_{z_0} f'/f = \operatorname{ord}_{z_0} f.$$

Theorem 1.5. *Let γ be a closed chain in U, homologous to 0 in U. Let f be meromorphic on U, with only a finite number of zeros and poles, say at the points $z_1, \ldots, z_n$, none of which lie on γ. Let $m_i = W(\gamma, z_i)$. Then*

$$\int_\gamma f'/f = 2\pi\sqrt{-1} \sum m_i \operatorname{ord}_{z_i} f.$$

Proof. This is immediate by plugging the statement of the lemma into the residue formula.

In applications, γ is frequently equal to a circle C, or a rectangle, and the points $z_1, \ldots, z_n$ are inside C. Suppose that the zeros of f inside C are

$$a_1, \ldots, a_r$$

and the poles are

$$b_1, \ldots, b_s.$$

Then in the case,

$$\boxed{\int_C f'/f = 2\pi\sqrt{-1}\left(\sum_{i=1}^{r} \operatorname{ord}_{a_i} f - \sum_{j=1}^{s} \operatorname{mul}_{b_j} f\right).}$$

We follow our convention whereby the multiplicity of a pole is the negative of the order of f at the pole, so that

$$\operatorname{mul}_{b_j} f = -\operatorname{ord}_{b_j} f$$

by definition.

If one counts zeros and poles with their multiplicities, one may rephrase the above formula in the more suggestive fashion:

Let C be a simple closed curve, and let f be meromorphic on C and its interior. Assume that f has no zero or pole on C. Then

$$\boxed{\int_C f'/f = 2\pi i \text{ (number of zeros } - \text{ number of poles),}}$$

where number of zeros = number of zeros of f in the interior of C, and number of poles = number of poles of f in the interior of C.

Of course, we have not proved that a simple closed curve has an "interior". The theorem is applied in practice only when the curve is so explicitly given (as with a circle or rectangle) that it is clear what "interior" is meant.

Besides, one can (not so artificially) formalize what is needed of the notion of "interior" so that one can use the standard language. Let γ be a closed path. We say that γ has an **interior** if $W(\gamma, \alpha) = 0$ or 1 for every complex number α which does not lie on γ. Then the set of points α such that $W(\gamma, \alpha) = 1$ will be called the **interior** of γ. It's that simple.

Theorem 1.6 (Rouché's Theorem). *Let γ be a closed path homologous to 0 in U and assume that γ has an interior. Let f, g be analytic on U, and*

$$|f(z) - g(z)| < |f(z)|$$

for z on γ. Then f and g have the same number of zeros in the interior of γ.

Proof. Note that the assumption implies automatically that f, g have no zero on γ. We have

$$\left|\frac{g(z)}{f(z)} - 1\right| < 1$$

for z on γ. Then the values of the function g/f are contained in the open disc with center 1 and radius 1. Let $F = g/f$. Then $F \circ \gamma$ is a closed path contained in that disc, and therefore

$$W(F \circ \gamma, 0) = 0$$

because 0 lies outside the disc. If γ is defined on $[a, b]$ then

$$\begin{aligned} 0 = W(F \circ \gamma, 0) &= \int_{F \circ \gamma} \frac{1}{z}\, dz = \int_a^b \frac{F'(\gamma(t))}{F(\gamma(t))} \gamma'(t)\, dt \\ &= \int_\gamma F'/F \\ &= \int_\gamma g'/g - f'/f. \end{aligned}$$

What we want now follows from Theorem 1.5, as desired.

Example. Let $P(z) = z^8 - 5z^3 + z - 2$. We want to find the number of roots of this polynomial inside the unit circle. Let

$$f(z) = -5z^3.$$

For $|z| = 1$ it is immediate that

$$|f(z) - P(z)| = |-z^8 - z + 2| < |f(z)| = 5.$$

Hence f and P have the same number of zeros inside the unit circle, and this number is clearly equal to 3. (Remember, you have to count multiplicities, and the equation

$$5z^3 = 0$$

has one zero with multiplicity 3.)

We shall use Rouché's theorem to give an alternative treatment of the inverse function theorem, not depending on solving for an inverse power series as was done in Chapter II, §5.

Theorem 1.7. *Let f be analytic in a neighborhood of a point z_0, and assume $f'(z_0) \neq 0$. Then f is a local analytic isomorphism at z_0.*

Proof. Making translations, we may assume without loss of generality that $z_0 = 0$ and $f(z_0) = 0$, so that

$$f(z) = \sum_{n=m}^{\infty} a_n z^n,$$

and $m \geqq 1$. Since $f'(0) = a_1$ we have $m = 1$ and $a_1 \neq 0$. Dividing by a_1 we may assume without loss of generality that $a_1 = 1$. Thus f has the power series expansion

$$f(z) = z + h(z),$$

where $h(z)$ is divisible by z^2. In particular, if we restrict the values of z to some sufficiently small disc around 0, then there is a constant $K > 0$ such that

$$|f(z) - z| \leqq K|z|^2.$$

Let C_r be the circle of radius r, and let $|\alpha| < r/2$. Let r be sufficiently small, and let

$$f_\alpha(z) = f(z) - \alpha \qquad \text{and} \qquad g_\alpha(z) = z - \alpha.$$

We have the inequality

$$|f_\alpha(z) - g_\alpha(z)| = |f(z) - z| \leqq K|z|^2.$$

If z is on C_r, that is $|z| = r$, then

$$K|z|^2 = Kr^2 < |z - \alpha| = |g_\alpha(z)|$$

because $|z - \alpha| > r/2$ and $Kr^2 < r/2$ (for instance, taking $r < 1/2K$). By Rouché's theorem, f_α and g_α have the same number of zeros inside C_r, and since g_α has exactly one zero, it follows that f_α has exactly one zero. This means that the equation

$$f(z) = \alpha$$

has exactly one solution inside C_r if $|\alpha| < r/2$.

Let U be the set of points z inside C_r such that

$$|f(z)| < r/2.$$

Then U is open because f is continuous, and we have just shown that

$$f\colon U \to D(0, r/2)$$

is a bijection of U with the disc of radius $r/2$. The argument we have given also shows that f is an open mapping, and hence the inverse function

$$\varphi\colon D(0, r/2) \to U$$

is continuous. There remains only to prove that φ is analytic. As in freshman calculus, we write the Newton quotient

$$\frac{\varphi(w) - \varphi(w_1)}{w - w_1} = \frac{z - z_1}{f(z) - f(z_1)}.$$

Fix w_1 with $|w_1| < r/2$, and let w approach w_1. Since φ is continuous, it must be that $z = \varphi(w)$ approaches $z_1 = \varphi(w_1)$. Thus the right-hand side approaches

$$1/f'(z_1),$$

provided we took r so small that $f'(z_1) \neq 0$ for all z_1 inside the circle of radius r, which is possible by the continuity of f' and the fact that $f'(0) \neq 0$. This proves that φ is holomorphic, whence analytic and concludes the proof of the theorem.

Residues of Differentials

Let $f(T) = \sum a_n T^n$ be a power series with a finite number of negative terms. We defined the **residue** to be a_{-1}. This was convenient for a number of applications, but in some sense so far it constituted an incomplete treatment of the situation with residues because this definition did not take into account the chain rule when computing integrals by means of residues. We shall now fill the remaining gap.

Let U be an open set in **C**. We define a **meromorphic differential** on U to be an expression of the form

$$\omega = f(z)\, dz$$

where f is meromorphic on U. Let $z_0 \in U$. Then f has a power series expansion at z_0, say

$$f(z) = \sum_{n=m}^{\infty} a_n (z - z_0)^n = f_0(z - z_0). \tag{1}$$

Often one wants to make a change of coordinate. Thus we define a function w to be a **local coordinate** at z_0 if w has a zero of order 1 at z_0. Then w is a local isomorphism at z_0, and there is a power series h such that

$$z - z_0 = h(w) = c_1 w + \text{higher terms} \qquad \text{with} \quad c_1 \neq 0. \tag{2}$$

Then substituting (2) in (1) we obtain

$$f(z)\, dz = f_0(h(w)) h'(w)\, dw = g(w)\, dw, \tag{3}$$

where $g(w) = f_0(h(w))h'(w)$ also has a power series expansion in terms of w. We denote the coefficients of this power series by b_n, so that

$$g(w) = \sum_{n=m}^{\infty} b_n w^n. \tag{4}$$

Since $h(w)$ has order 1, $h'(0) = c_1 \neq 0$, it follows that the power series for g also has order m. Of course, the coefficients for the power series of g seem to be complicated expressions in the coefficients for the power series for f. However, it turns out that the really important coefficient, namely the residue b_{-1}, has a remarkable invariance property, stated in the next theorem.

Theorem 1.8. *Let w be a local coordinate at z_0. Let ω be a meromorphic differential in a neighborhood of z_0, and write $\omega = f(z)\, dz =$*

$g(w)\,dw$, where f, g are meromorphic functions, with the power series expansions as in (1) *and* (4) *above. Then the residues of the power series for f and g are equal, that is*

$$b_{-1} = a_{-1}.$$

Proof. Let γ be a small circle around z_0 in the z-plane. Let $w = \varphi(z)$. Then

$$b_{-1} = \frac{1}{2\pi i}\int_{\varphi\circ\gamma} g(w)\,dw = \frac{1}{2\pi i}\int_{\gamma} f(z)\,dz = a_{-1},$$

which proves the theorem.

In light of Theorem 1.8, we define the residue of a meromorphic differential $f(z)\,dz$ at a point z_0 as follows. We let w be a local coordinate at z_0. (Thus w may be $z - z_0$, but there are plenty of other local coordinates.) We write the differential as a power series in w,

$$\omega = g(w)\,dw \qquad \text{with} \qquad g(w) = \sum b_n w^n,$$

and we define the **residue of the differential** to be

$$\operatorname{res}_{z_0}(\omega) = b_{-1}.$$

This value b_{-1} is independent of the choice of coordinate at z_0. Using residues of differentials rather than residues of power series will be especially appropriate when the change of variables formula enters into consideration, for example in Exercises 35 and 36 below, when we deal with residues "at infinity" using the change of coordinate $w = 1/z$.

Remark. We could have defined a meromorphic differential on U also as an expression of the form $f\,dg$ where f and g are meromorphic. If w is a local coordinate at z_0, then both f and g have power series expansions in terms of w, so

$$f\,dg = f(w)\frac{dg}{dw}\,dw.$$

However, if U is an open set and f is a meromorphic function on U, not constant, then note that $d\log f$ is a meromorphic differential on U, because

$$d\log f(z) = \frac{f'(z)}{f(z)}\,dz.$$

Even though $\log f$ itself is not well defined on U, because of the ambiguity arising from the constant of integration, taking the derivative eliminates this constant, so that the differential itself is well defined.

VI, §1. EXERCISES

Find the residues of the following functions at 0.

1. $(z^2 + 1)/z$
2. $(z^2 + 3z - 5)/z^3$
3. $z^3/(z - 1)(z^4 + 2)$
4. $(2z + 1)/z(z^3 - 5)$
5. $(\sin z)/z^4$
6. $(\sin z)/z^5$
7. $(\sin z)/z^6$
8. $(\sin z)/z^7$
9. e^z/z
10. e^z/z^2
11. e^z/z^3
12. e^z/z^4
13. $z^{-2}\log(1 + z)$
14. $e^z/\sin z$

Find the residues of the following functions at 1.

15. $1/(z^2 - 1)(z + 2)$
16. $(z^3 - 1)(z + 2)/(z^4 - 1)^2$
17. Factor the polynomial $z^n - 1$ into factors of degree 1. Find the residue at 1 of $1/(z^n - 1)$.
18. Let $z_1, \ldots, z_n$ be distinct complex numbers. Let C be a circle around z_1 such that C and its interior do not contain z_j for $j > 1$. Let

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

Find

$$\int_C \frac{1}{f(z)}\,dz.$$

19. Find the residue at i of $1/(z^4 - 1)$. Find the integral

$$\int_C \frac{1}{(z^4 - 1)}\,dz$$

where C is a circle of radius 1/2 centered at i.

20. (a) Find the integral

$$\int_C \frac{1}{z^2 - 3z + 5}\,dz,$$

where C is a rectangle oriented clockwise, as shown on the figure.

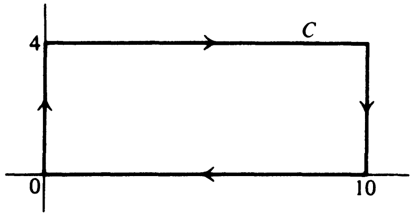


Figure 5

(b) Find the integral $\int_C 1/(z^2 + z + 1)\, dz$ over the same C.

(c) Find the integral $\int_C 1/(z^2 - z + 1)\, dz$ over this same C.

21. (a) Let $z_1, \ldots, z_n$ be distinct complex numbers. Determine explicitly the partial fraction decomposition (i.e. the numbers a_i):

$$\frac{1}{(z - z_1)\cdots(z - z_n)} = \frac{a_1}{z - z_1} + \cdots + \frac{a_n}{z - z_n}.$$

(b) Let $P(z)$ be a polynomial of degree $\leqq n - 1$, and let $a_1, \ldots, a_n$ be distinct complex numbers. Assume that there is a partial fraction decomposition of the form

$$\frac{P(z)}{(z - a_1)\cdots(z - a_n)} = \frac{c_1}{z - a_1} + \cdots + \frac{c_n}{z - a_n}.$$

Prove that

$$c_1 = \frac{P(a_1)}{(a_1 - a_2)\cdots(a_1 - a_n)},$$

and similarly for the other coefficients c_j.

22. Let f be analytic on an open disc centered at a point z_0, except at the point itself where f has a simple pole with residue equal to an integer n. Show that there is an analytic function g on the disc such that $f = g'/g$, and

$$g(z) = (z - z_0)^n h(z), \qquad \text{where } h \text{ is analytic and } \quad h(z_0) \neq 0.$$

(To make life simpler, you may assume $z_0 = 0$.)

23. (a) Let f be a function which is analytic on the upper half plane, and on the real line. Assume that there exist numbers $B > 0$ and $c > 0$ such that

$$|f(\zeta)| \leq \frac{B}{|\zeta|^c}$$

for all ζ. Prove that for any z in the upper half plane, we have the integral formula

$$\boxed{f(z) = \frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{f(t)}{t - z}\, dt.}$$

[*Hint*: Consider the integral over the path shown on the figure, and take the limit as $R \to \infty$.]

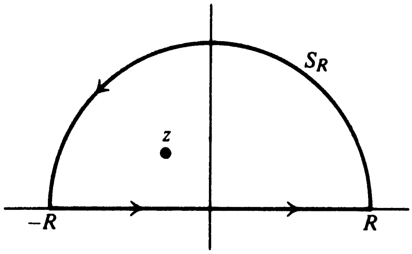


Figure 6

The path consists of the segment from $-R$ to R on the real axis, and the semicircle S_R as shown.

(b) By using a path similar to the previous one, but slightly raised over the real axis, and taking a limit, prove that the formula is still true if instead of assuming that f is analytic on the real line, one merely assumes that f is continuous on the line, but otherwise satisfies the same hypotheses as before.

24. Determine the poles and find the residues of the following functions.

(a) $1/\sin z$ (b) $1/(1 - e^z)$ (c) $z/(1 - \cos z)$.

25. Show that

$$\int_{|z|=1} \frac{\cos e^{-z}}{z^2}\, dz = 2\pi i \cdot \sin 1.$$

26. Find the integrals, where C is the circle of radius 8 centered at the origin.

(a) $\displaystyle\int_C \frac{1}{\sin z}\, dz$ (b) $\displaystyle\int_C \frac{1}{1 - \cos z}\, dz$

(c) $\displaystyle\int_C \frac{1 + z}{1 - e^z}\, dz$ (d) $\displaystyle\int_C \tan z\, dz$

(e) $\displaystyle\int_C \frac{1 + z}{1 - \sin z}\, dz$

27. Let f be holomorphic on and inside the unit circle, $|z| \leqq 1$, except for a pole of order 1 at a point z_0 on the circle. Let $f = \sum a_n z^n$ be the power series for f on the open disc. Prove that

$$\lim_{n\to\infty} \frac{a_n}{a_{n+1}} = z_0.$$

28. Let a be real > 1. Prove that the equation $ze^{a-z} = 1$ has a single solution with $|z| \leqq 1$, which is real and positive.

29. Let U be a connected open set, and let D be an open disc whose closure is contained in U. Let f be analytic on U and not constant. Assume that the absolute value $|f|$ is constant on the boundary of D. Prove that f has at least one zero in D. [*Hint*: Consider $g(z) = f(z) - f(z_0)$ with $z_0 \in D$.]

30. Let f be a function analytic inside and on the unit circle. Suppose that $|f(z) - z| < |z|$ on the unit circle.
    (a) Show that $|f'(1/2)| \leqq 8$.
    (b) Show that f has precisely one zero inside of the unit circle.

31. Determine the number of zeros of the polynomial

$$z^{87} + 36z^{57} + 71z^4 + z^3 - z + 1$$

inside the circle
    (a) of radius 1,
    (b) of radius 2, centered at the origin.
    (c) Determine the number of zeros of the polynomial

$$2z^5 - 6z^2 + z + 1 = 0$$

in the annulus $1 \leqq |z| \leqq 2$.

32. Let f, h be analytic on the closed disc of radius R, and assume that $f(z) \neq 0$ for z on the circle of radius R. Prove that there exists $\epsilon > 0$ such that $f(z)$ and $f(z) + \epsilon h(z)$ have the same number of zeros inside the circle of radius R. Loosely speaking, we may say that f and a small perturbation of f have the same number of zeros inside the circle.

33. Let $f(z) = a_n z^n + \cdots + a_0$ be a polynomial with $a_n \neq 0$. Use Rouché's theorem to show that $f(z)$ and $a_n z^n$ have the same number of zeros in a disc of radius R for R sufficiently large.

34. (a) Let f be analytic on the closed unit disc. Assume that $|f(z)| = 1$ if $|z| = 1$, and f is not constant. Prove that the image of f contains the unit disc.
    (b) Let f be analytic on the closed unit disc $\bar{D}$. Assume that there exists some point $z_0 \in D$ such that $|f(z_0)| < 1$, and that $|f(z)| \geqq 1$ if $|z| = 1$. Prove that $f(D)$ contains the unit disc.

35. Let $P_n(z) = \sum_{k=0}^{n} z^k/k!$. Given R, prove that P_n has no zeros in the disc of radius R for all n sufficiently large.

36. Let $z_1, \ldots, z_n$ be distinct complex numbers contained in the disc $|z| < R$. Let f be analytic on the closed disc $\bar{D}(0, R)$. Let

$$Q(z) = (z - z_1) \cdots (z - z_n).$$

Prove that

$$P(z) = \frac{1}{2\pi i} \int_{C_R} f(\zeta) \frac{1 - Q(z)/Q(\zeta)}{\zeta - z} \, d\zeta$$

is a polynomial of degree $n - 1$ having the same value as f at the points $z_1, \ldots, z_n$.

37. Let f be analytic on $\mathbf{C}$ with the exception of a finite number of isolated singularities which may be poles. Define the **residue at infinity**

$$\mathbf{res}_\infty f(z)\, dz = -\frac{1}{2\pi i} \int_{|z|=R} f(z)\, dz$$

for R so large that f has no singularities in $|z| \geqq R$.
(a) Show that $\mathrm{res}_\infty f(z)\, dz$ is independent of R.
(b) Show that the sum of the residues of f at all singularities and the residue at infinity is equal to 0.

38. **Cauchy's Residue Formula on the Riemann Sphere.** Recall Exercise 2 of Chapter V, §3 on the Riemann sphere. By a (meromorphic) **differential** ω on the Riemann sphere S, we mean an expression of the form

$$\omega = f(z)\, dz,$$

where f is a rational function. For any point $z_0 \in \mathbf{C}$ the **residue** of ω at z_0 is defined to be the usual residue of $f(z)\, dz$ at z_0. For the point ∞, we write $t = 1/z$,

$$dt = -\frac{1}{z^2}\, dz \qquad \text{and} \qquad dz = -\frac{1}{t^2}\, dt,$$

so we write

$$\omega = f(1/t)\left(-\frac{1}{t^2}\right) dt = -\frac{1}{t^2} f(1/t)\, dt.$$

The **residue of** ω **at infinity** is then defined to be the residue of $-\frac{1}{t^2} f(1/t)\, dt$ at $t = 0$. Prove:
(a) $\sum$ residues $\omega = 0$, if the sum is taken over all points of $\mathbf{C}$ and also infinity.
(b) Let γ be a circle of radius R centered at the origin in $\mathbf{C}$. If R is sufficiently large, show that

$$\frac{1}{2\pi i} \int_\gamma f(z)\, dz = -\text{residue of } f(z)\, dz \text{ at infinity.}$$

(Instead of a circle, you can also take a simple closed curve such that all the poles of f in $\mathbf{C}$ lie in its interior.)

(c) If R is arbitrary, and f has no pole on the circle, show that

$$\frac{1}{2\pi i}\int_\gamma f(z)\,dz = -\sum \text{residues of } f(z)\,dz \text{ outside the circle, including the residue at } \infty.$$

[**Note**: In dealing with (a) and (b), you can either find a direct algebraic proof of (a), as in Exercise 38 and deduce (b) from it, or you can prove (b) directly, using a change of variables $t = 1/z$, and then deduce (a) from (b). You probably should carry both ideas out completely to understand fully what's going on.]

39. (a) Let $P(z)$ be a polynomial. Show directly from the power series expansions of $P(z)\,dz$ that $P(z)\,dz$ has 0 residue in $\mathbf{C}$ and at infinity.
  (b) Let α be a complex number. Show that $dz/(z-\alpha)$ has residue -1 at infinity.
  (c) Let m be an integer $\geqq 2$. Show that $dz/(z-\alpha)^m$ has residue 0 at infinity and at all complex numbers.
  (d) Let $f(z)$ be a rational function. The theorem concerning the **partial fraction decomposition** of f states that f has an expression

$$f(z) = \sum_{i=1}^{r}\sum_{m=1}^{n_i} \frac{a_{ij}}{(z-\alpha_i)^m} + P(z)$$

where $\alpha_1, \ldots, \alpha_r$ are the roots of the denominator of f, a_{ij} are constants, and P is some polynomial. Using this theorem, give a direct (algebraic) proof of Exercise 37(a).

40. Let $a, b \in \mathbf{C}$ with $|a|$ and $|b| < R$. Let C_R be the circle of radius R. Evaluate

$$\int_{C_R} \frac{z\,dz}{\sqrt{(z-a)(z-b)}}.$$

The square root is chosen so that the integrand is continuous for $|z| > R$ and has limit 1 as $|z| \to \infty$.

VI, §2. EVALUATION OF DEFINITE INTEGRALS

Let $f(x)$ be a continuous function of a real variable x. We want to compute

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{A\to\infty}\int_{-A}^{0} f(x)\,dx + \lim_{B\to\infty}\int_{0}^{B} f(x)\,dx.$$

We shall use the following method. We let γ be the closed path as indicated on Fig. 7, consisting of a segment on the real line, and a semicircle.

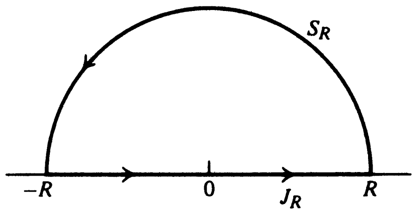


Figure 7

We suppose that $f(x)$ is the restriction to the line of a function f on the upper half plane, meromorphic and having only a finite number of poles. We let J_R be the segment from $-R$ to R, and let S_R be the semicircle. If we can prove that

$$\lim_{R\to\infty} \int_{S_R} f = 0$$

then by the residue formula, we obtain

$$\boxed{\int_{-\infty}^{\infty} f(x)\,dx = 2\pi i \sum \text{residues of } f \text{ in the upper half plane.}}$$

For this method to work, it suffices to know that $f(z)$ goes sufficiently fast to 0 when $|z|$ becomes large, so that the integral over the semicircle tends to 0 as the radius R becomes large. It is easy to state conditions under which this is true.

Theorem 2.1. *Suppose that there exists a number $B > 0$ such that for all $|z|$ sufficiently large, we have*

$$|f(z)| \leqq B/|z|^2.$$

Then

$$\lim_{R\to\infty} \int_{S_R} f = 0$$

and the above formula is valid.

Proof. The integral is estimated by the sup norm of f, which is B/R^2 by assumption, multiplied by the length of the semicircle, which is πR. Since $\pi B/R$ tends to 0 as $R \to \infty$, our theorem is proved.

Remark. We really did not need an R^2, only R^{1+a} for some $a > 0$, so the theorem could be correspondingly strengthened.

Example. Let us compute

$$\int_{-\infty}^{\infty} \frac{1}{x^4 + 1}\,dx.$$

The function $1/(z^4 + 1)$ is meromorphic on $\mathbf{C}$, and its poles are at the zeros of $z^4 + 1$, that is the solutions of $z^4 = -1$, which are

$$e^{\pi ik/4}, \qquad k = 1, -1, 3, -3.$$

Let $f(z) = z^4 + 1$. Since $f'(z) = 4z^3 \neq 0$ unless $z = 0$, we conclude that all the zeros of f are simple. The two zeros in the upper half plane are

$$z_1 = e^{\pi i/4} \qquad \text{and} \qquad z_2 = e^{3\pi i/4}.$$

The residues of $1/f(z)$ at these points are $1/f'(z_1)$, $1/f'(z_2)$, respectively, by Lemma 1.3(b), and

$$f'(z_1) = 4z_1^3 = 4e^{3\pi i/4}, \qquad f'(z_2) = 4z_2^3 = 4e^{\pi i/4}.$$

The estimate

$$\left|\frac{1}{z^4 + 1}\right| \leqq B/R^4$$

is satisfied for some constant B when $|z| = R$. Hence the theorem applies, and

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{1}{x^4 + 1}\,dx &= 2\pi i(\tfrac{1}{4}e^{-3\pi i/4} + \tfrac{1}{4}e^{-\pi i/4})\\ &= \frac{\pi i}{2}e^{-\pi i/4}(e^{-2\pi i/4} + 1)\\ &= \frac{\pi i}{2}\left(\frac{1-i}{\sqrt{2}}\right)(1 - i)\\ &= \frac{\pi}{\sqrt{2}}.\end{aligned}$$

The estimate for $1/(z^4 + 1)$ on a circle of radius R presented no subtlety. We give an example where the estimate takes into account a different phenomenon, and a different path. The fact that the integral over the part going to infinity like the semicircle tends to 0 will be due to a more conditional convergence, and the evaluation of an integral explicitly.

Fourier Transforms

Integrals of the form discussed in the next examples are called **Fourier transforms**, and the technique shows how to evaluate them.

Theorem 2.2. *Let f be meromorphic on $\mathbf{C}$, having only a finite number of poles, not lying on the real axis. Suppose that there is a constant K such that*

$$|f(z)| \leqq K/|z|$$

for all sufficiently large $|z|$. Let $a > 0$. Then

$$\boxed{\int_{-\infty}^{\infty} f(x)e^{iax}\,dx = 2\pi i \sum \textit{residues of } e^{iaz}f(z) \textit{ in the upper half plane.}}$$

Proof. For simplicity, take $a = 1$. We integrate over any rectangle as shown on Fig. 8, taking $T = A + B$. Taking A, $B > 0$ sufficiently large, it suffices to prove that the integral over the three sides other than the bottom side tend to 0 as A, B tend to infinity.

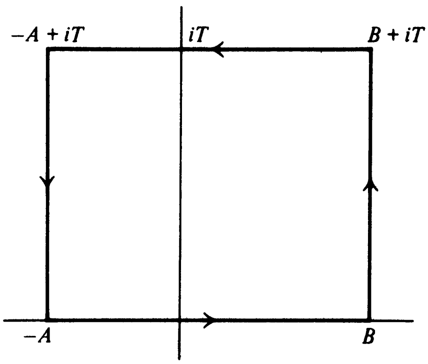


Figure 8

Note that

$$e^{iz} = e^{i(x+iy)} = e^{ix}e^{-y}.$$

In absolute value this is e^{-y}, and tends to 0 rapidly as y tends to infinity. We show that the integral over the top tends to 0. Parametrizing the top by $x + iT$, with $-A \leqq x \leqq B$, we find

$$-\int_{\text{top}} e^{iz}f(z)\,dz = \int_{-A}^{B} e^{ix}e^{-T}f(x + iT)\,dx$$

and in absolute value, this is less than

$$e^{-T}\int_{-A}^{B}|f(x+iT)|\,dx \leqq e^{-T}\frac{K}{T}(A+B).$$

Having picked $T = A + B$ shows that this integral becomes small as A, B become large, as desired.

For the right-hand side, we pick the parametrization

$$B + iy, \qquad \text{with} \quad 0 \leqq y \leqq T,$$

and we find that the right-hand side integral is bounded by

$$\left|\int_0^T e^{iB}e^{-y}f(B+iy)\,dy\right| \leqq \frac{K}{B}\int_0^T e^{-y}\,dy = \frac{K}{B}(1-e^{-T}),$$

which tends to 0 as B becomes large. A similar estimate shows that the integral over the left side tends to 0, and proves what we wanted.

Next we show an adjustment of the above techniques when the function may have some singularity on the real axis. We do this by an example.

Example. Let us compute

$$\begin{aligned} I &= \int_0^\infty \frac{\sin x}{x}\,dx = \frac{1}{2}\int_{-\infty}^{\infty}\frac{\sin x}{x}\,dx. \\ &= \frac{1}{2i}\lim_{\epsilon\to 0}\left[\int_{-\infty}^{-\epsilon}\frac{e^{ix}}{x}dx + \int_{\epsilon}^{\infty}\frac{e^{ix}}{x}dx\right]. \end{aligned}$$

Note that the integral I converges, although not absolutely. It is an oscillatory integral. The estimate for convergence comes from integration by parts, and is left to the reader. We can then use the technique of complex analysis to evaluate the integral. We use the closed path C as shown on Fig. 9. To compute such an integral, one has to show that both limits exist, and then one can deal with the more symmetric expression

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{R\to\infty}\int_{-R}^{R} f(x)\,dx.$$

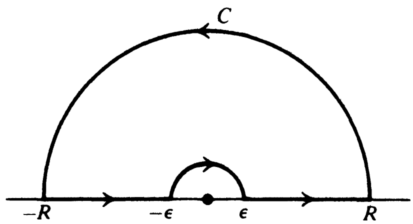


Figure 9

Let $S(\epsilon)$ be the small semicircle from ϵ to $-\epsilon$, oriented counterclockwise, and let $S(R)$ be the big semicircle from R to $-R$ similarly oriented. The function e^{iz}/z has no pole inside C, and consequently

$$0 = \int_C e^{iz}/z\,dz = \int_{S(R)} + \int_{-R}^{-\epsilon} - \int_{S(\epsilon)} + \int_{\epsilon}^{R} e^{iz}/z\,dz.$$

Hence

$$\begin{aligned}\int_{-R}^{-\epsilon} + \int_{\epsilon}^{R} e^{iz}/z\,dz &= \int_{S(\epsilon)} e^{iz}/z\,dz - \int_{S(R)} e^{iz}/z\,dz \\ &= I_{S(\epsilon)} - I_{S(R)}.\end{aligned}$$

We now assert that

$$\lim_{R\to\infty} I_{S(R)} = 0.$$

Proof. We have for $z = R(\cos\theta + i\sin\theta)$,

$$I_{S(R)} = \int_0^{\pi} \frac{e^{iR\cos\theta}e^{-R\sin\theta}}{Re^{i\theta}} Rie^{i\theta}\,d\theta$$

so that

$$|I_{S(R)}| \leqq \int_0^{\pi} e^{-R\sin\theta}\,d\theta = 2\int_0^{\pi/2} e^{-R\sin\theta}\,d\theta.$$

But if $0 \leqq \theta \leqq \pi/2$, then $\sin\theta \geqq 2\theta/\pi$ (any similar estimate would do), and hence

$$\begin{aligned}|I_{S(R)}| &\leqq 2\int_0^{\pi/2} e^{-2R\theta/\pi}\,d\theta \\ &= \frac{\pi}{R}(1 - e^{-R})\end{aligned}$$

by freshman calculus. This proves our assertion.

There remains to evaluate the limit of $I_{S(\epsilon)}$ as $\epsilon \to 0$. We state this as a general lemma.

Lemma. *Let g have a simple pole at* 0. *Then*

$$\lim_{\epsilon\to 0} \int_{S(\epsilon)} g(z)\,dz = \pi i\,\mathrm{Res}_0(g).$$

Proof. Write

$$g(z) = \frac{a}{z} + h(z),$$

where h is holomorphic at 0. Then the integral of h over $S(\epsilon)$ approaches 0 as $\epsilon \to 0$ because the length of $S(\epsilon)$ approaches 0 and h is bounded near the origin. A direct integration of a/z shows that the integral of a/z over the semicircle is equal to $\pi i a$. This proves the lemma.

We may therefore put everything together to find the value

$$\boxed{\int_0^\infty \frac{\sin x}{x}\, dx = \pi/2.}$$

Trigonometric Integrals

We wish to evaluate an integral of the form

$$\int_0^{2\pi} Q(\cos\theta, \sin\theta)\, d\theta,$$

where Q is a rational function of two variables, $Q = Q(x, y)$, which we assume is continuous on the unit circle. Since

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad \text{and} \quad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i},$$

we see that these expressions are equal to

$$\frac{z + 1/z}{2} \quad \text{and} \quad \frac{z - 1/z}{2i},$$

respectively, when z lies on the unit circle, $z = e^{i\theta}$. It is therefore natural to consider the function

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

(the denominator iz is put there for a purpose which will become apparent in a moment). This function f is a rational function of z, and in view of our assumption on Q, it has no pole on the unit circle.

Theorem 2.3. *Let $Q(x, y)$ be a rational function which is continuous when $x^2 + y^2 = 1$. Let $f(z)$ be as above. Then*

$$\boxed{\int_0^{2\pi} Q(\cos\theta, \sin\theta)\, d\theta = 2\pi i \sum \textit{residues of } f \textit{ inside the unit circle.}}$$

Proof. Let C be the unit circle. Then

$$\int_C f(z)\, dz = 2\pi i \sum \text{residues of } f \text{ inside the circle.}$$

On the other hand, by definition the integral on the left is equal to

$$\int_0^{2\pi} f(e^{i\theta}) ie^{i\theta}\, d\theta = \int_0^{2\pi} Q(\cos\theta, \sin\theta)\, d\theta,$$

as desired. [The term iz in the denominator of f was introduced to cancel $ie^{i\theta}$ at this point.]

Example. Let us compute the integral

$$I = \int_0^{2\pi} \frac{1}{a + \sin\theta}\, d\theta$$

where a is real > 1. By the theorem,

$$I = 2\pi \sum \text{residues of } \frac{2i}{z^2 + 2iaz - 1} \text{ inside circle.}$$

The only pole inside the circle is at

$$z_0 = -ia + i\sqrt{a^2 - 1}$$

and the residue is

$$\frac{i}{z_0 + ia} = \frac{1}{\sqrt{a^2 - 1}}.$$

Consequently,

$$I = \frac{2\pi}{\sqrt{a^2 - 1}}.$$

Mellin Transforms

We give a final example introducing new complications. Integrals of type

$$\int_0^\infty f(x)x^a \frac{dx}{x}$$

are called **Mellin transforms** (they can be viewed as functions of a). We wish to show how to evaluate them. *We assume that $f(z)$ is analytic on* **C** *except for a finite number of poles, none of which lies on the positive real axis $0 < x$, and we also assume that a is not an integer.* Then under appropriate conditions on the behavior of f near 0, and when x becomes large, we can show that the following formula holds:

$$\boxed{\int_0^\infty f(x)x^a \frac{dx}{x} = -\frac{\pi e^{-\pi i a}}{\sin \pi a} \sum \text{residues of } f(z)z^{a-1} \text{ at the poles of } f, \text{ excluding the residue at } 0.}$$

We comment right away on what we mean by z^{a-1}, namely z^{a-1} is defined as

$$z^{a-1} = e^{(a-1)\log z},$$

where the log is defined on the simply connected set equal to the plane from which the axis $x \geqq 0$ has been deleted. We take the value for the log such that if $z = re^{i\theta}$ and $0 < \theta < 2\pi$, then

$$\log z = \log r + i\theta.$$

Then, for instance,

$$\log i = \pi i/2 \qquad \text{and} \qquad \log(-i) = 3\pi i/2.$$

Precise sufficient conditions under which the formula is true are given in the next theorem. They involve suitable estimates for the function f near 0 and infinity.

Theorem 2.4. *The formula given for the integral*

$$\int_0^\infty f(x)x^a \frac{dx}{x}, \qquad \textit{with} \quad a > 0,$$

is valid under the following conditions:

1. *There exists a number $b > a$ such that*

$$|f(z)| \ll 1/|z|^b \quad \textit{for} \quad |z| \to \infty.$$

2. *There exists a number b' with $0 < b' < a$ such that*

$$|f(z)| \ll 1/|z|^{b'} \quad \textit{for} \quad |z| \to 0.$$

The symbol $\ll$ means that the left-hand side is less than or equal to some constant times the right-hand side.

For definiteness, we carry out the arguments on a concrete example, and let the reader verify that the arguments work under the conditions stated in Theorem 2.4.

Example. We shall evaluate for $0 < a < 2$:

$$\int_0^\infty \frac{1}{x^2+1} x^a \frac{dx}{x} = \frac{\pi \cos a\pi/2}{\sin a\pi}.$$

We choose the closed path C as on Fig. 10. Then C consists of two line segments L^+ and L^-, and two pieces of semicircles $S(R)$ and $-S(\epsilon)$, if we take $S(\epsilon)$ oriented in counterclockwise direction. The angle φ which the two segments L^+ and L^- make with the positive real axis will tend to 0.

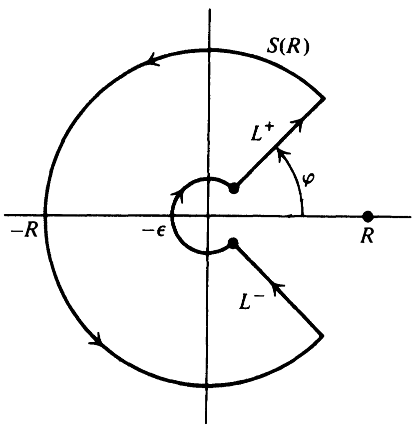


Figure 10

We let

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

Then $g(z)$ has only simple poles at $z = i$ and $z = -i$, where the residues are found to be:

$$\text{at } i\colon \frac{1}{2i}e^{(a-1)\log i} = \frac{1}{2i}e^{(a-1)\pi i/2},$$

$$\text{at } -i\colon -\frac{1}{2i}e^{(a-1)\log(-i)} = -\frac{1}{2i}e^{(a-1)3\pi i/2}.$$

The sum of the residues inside C is therefore equal to

$$\frac{1}{2i}\left(e^{(a-1)\pi i/2} - e^{(a-1)3\pi i/2}\right) = -e^{a\pi i}\cos(a\pi/2),$$

after observing that $e^{\pi i/2} = i$, $e^{-3\pi i/2} = i$, and factoring out $e^{a\pi i}$ from the sum.

The residue formula yields

$$2\pi i \sum \text{residues} = I_{S(R)} + I_{L^-} - I_{S(\epsilon)} + I_{L^+},$$

where I_X denotes the integral of $f(z)$ over the path X. We shall prove:

The integrals $I_{S(R)}$ and $I_{S(\epsilon)}$ tend to 0 as R becomes large and ϵ becomes small, independently of the angle φ.

Proof. First estimate the integral over $S(R)$. When comparing functions of R, it is useful to use the following notation. Let $F(R)$ and $G(R)$ be functions of R, and assume that $G(R)$ is > 0 for all R sufficiently large. We write

$$F(R) \ll G(R) \qquad (\text{for } R \to \infty)$$

if there exists a constant K such that

$$|F(R)| \leqq KG(R)$$

for all R sufficiently large.

With this notation, using $z^{a-1} = e^{(a-1)\log z}$, and

$$|\log z| \leqq \log R + \theta \leqq \log R + 2\pi,$$

we find

$$|z^{a-1}| = |e^{(a-1)\log z}| \ll R^{a-1}.$$

Consequently from $|1/(z^2 + 1)| \ll 1/R^2$ for $|z| = R$, we find

$$\left| \int_{S(R)} f(z)z^{a-1}\, dz \right| \ll 2\pi R \frac{1}{R^2} \max|z^{a-1}| \ll R^a/R^2.$$

Since we assumed that $a < 2$, the quotient R^a/R^2 approaches 0 as R becomes large, as desired. The estimate is independent of φ.

We use a similar estimating notation for functions of ϵ,

$$F(\epsilon) \ll G(\epsilon) \qquad (\text{for } \epsilon \to 0)$$

if there exists a constant K such that

$$|F(\epsilon)| \leqq KG(\epsilon)$$

for all $\epsilon > 0$ sufficiently small. With this notation, for $|z| = \epsilon$, we have

$$|z^{a-1}| = |e^{(a-1)\log z}| \ll \epsilon^{a-1}.$$

Hence

$$\left| \int_{S(\epsilon)} f(z)z^{a-1}\, dz \right| \ll 2\pi\epsilon\, \epsilon^{a-1} \ll \epsilon^a.$$

Again since we assumed that $a > 0$, the right-hand side approaches 0 and ϵ tends to 0, as desired. The estimate is independent of φ.

There remains to analyze the sums of the integrals over L^+ and L^-. We parametrize L^+ by

$$z(r) = re^{i\varphi}, \qquad \epsilon \leqq r \leqq R,$$

so that $\log z(r) = \log r + i\varphi$. Then

$$\begin{aligned}\int_{L^+} f(z)e^{(a-1)\log z}\, dz &= \int_\epsilon^R f(re^{i\varphi})e^{(a-1)(\log r + i\varphi)}e^{i\varphi}\, dr \\ &= \int_\epsilon^R f(re^{i\varphi})e^{(a-1)i\varphi}e^{i\varphi}r^{a-1}\, dr \\ &\to \int_\epsilon^R f(x)x^{a-1}\, dx \qquad \text{as} \quad \varphi \to 0.\end{aligned}$$

On the other hand, $-L^-$ is parametrized by

$$z(r) = re^{i(2\pi - \varphi)}, \qquad \epsilon \leqq r \leqq R,$$

and $\log z(r) = \log r + i(2\pi - \varphi)$. Consequently,

$$\int_{L^-} f(z)e^{(a-1)\log z}\,dz = -\int_\epsilon^R f(re^{-i\varphi})r^{a-1}e^{(a-1)(2\pi - i\varphi)}e^{i(2\pi - i\varphi)}\,dr$$
$$= -\int_\epsilon^R f(re^{i\varphi})r^{a-1}e^{ai(2\pi-\varphi)}\,dr$$
$$\to -\int_\epsilon^R f(x)x^{a-1}e^{2\pi ia}\,dx \qquad \text{as} \quad \varphi \to 0.$$

Hence as $\varphi \to 0$, we find

$$\int_{L^+} + \int_{L^-} f(z)z^{a-1}\,dz \to \int_\epsilon^R f(x)x^{a-1}(1 - e^{2\pi ia})\,dx$$
$$= e^{\pi ia}\int_\epsilon^R f(x)x^{a-1}(e^{-\pi ia} - e^{\pi ia})\,dx$$
$$= 2ie^{\pi ia}\sin \pi a \int_\epsilon^R f(x)x^{a-1}\,dx.$$

Let $C = C(R, \epsilon, \varphi)$ denote the path of integration. We obtain

$$\int_{C(R,\epsilon,\varphi)} f(z)z^{a-1}\,dz = 2\pi i \sum \text{residues of } f(z)z^{a-1} \text{ except at } 0$$
$$= I_{S(R,\varphi)} + I_{S(\epsilon,\varphi)} + E(R, \epsilon, \varphi)$$
$$+ 2ie^{\pi ia}\sin \pi a \int_\epsilon^R f(x)x^{a-1}\,dx.$$

The expression $E(R, \epsilon, \varphi)$ denotes a term which goes to 0 as φ goes to 0, and we have put subscripts on the integrals along the arcs of the circle to show that they depend on R, ϵ, φ. We divide by $2ie^{\pi ia}\sin \pi a$, and let φ tend to 0. Then $E(R, \epsilon, \varphi)$ approaches 0. Consequently,

$$\int_\epsilon^R f(x)x^{a-1}\,dx - \frac{\pi e^{-\pi ia}}{\sin \pi a}\sum = \lim_{\varphi\to 0}\frac{I_{S(R,\varphi)} + I_{S(\epsilon,\varphi)}}{2ie^{\pi ia}\sin \pi a}.$$

The right-hand side has been seen to be uniformly small, independently of φ, and tends to 0 when $R \to \infty$ and $\epsilon \to 0$. Taking the limits as $R \to \infty$ and $\epsilon \to 0$ proves what we wanted.

Finally, we observe that in situations of contour integrals as we just considered, it is often the practice to draw the limit contour as in Fig. 11.

It is then understood that the value for the log when integrating over the segment from ϵ to R from left to right, and the value for the log when integrating over the segment from R to ϵ, are different, arising from the analytic expressions for the log with values $\theta = 0$ for the first and $\theta = 2\pi$ for the second.

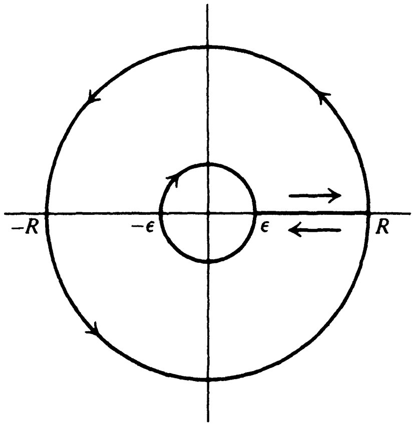


Figure 11

For the Mellin transform of the gamma function, which provides an interesting special concrete case of the considerations of this section, see Exercise 7 of Chapter XV, §2.

VI, §2. EXERCISES

Find the following integrals.

1. (a) $\displaystyle\int_{-\infty}^{\infty} \frac{1}{x^6+1}\,dx = 2\pi/3$

   (b) Show that for a positive integer $n \geqq 2$,

$$\int_0^{\infty} \frac{1}{1+x^n}\,dx = \frac{\pi/n}{\sin \pi/n}.$$

   [*Hint*: Try the path from 0 to R, then R to $Re^{2\pi i/n}$, then back to 0, or apply a general theorem.]

2. (a) $\displaystyle\int_{-\infty}^{\infty} \frac{x^2}{x^4+1}\,dx = \pi\sqrt{2}/2$ (b) $\displaystyle\int_0^{\infty} \frac{x^2}{x^6+1}\,dx = \pi/6$

3. Show that

$$\int_{-\infty}^{\infty} \frac{x-1}{x^5-1}\,dx = \frac{4\pi}{5}\sin\frac{2\pi}{5}.$$

4. Evaluate

$$\int_\gamma \frac{e^{-z^2}}{z^2}\,dz,$$

where γ is:
(a) the square with vertices $1+i$, $-1+i$, $-1-i$, $1-i$.
(b) the ellipse defined by the equation

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1.$$

(The answer is 0 in both cases.)

5. (a) $\displaystyle\int_{-\infty}^{\infty}\frac{e^{iax}}{x^2+1}\,dx=\pi e^{-a}$ if $a>0$

(b) For any real number $a>0$,

$$\int_{-\infty}^{\infty}\frac{\cos x}{x^2+a^2}\,dx=\pi e^{-a}/a.$$

[*Hint*: This is the real part of the integral obtained by replacing $\cos x$ by e^{ix}.]

6. Let $a, b>0$. Let $T\geqq 2b$. Show that

$$\left|\frac{1}{2\pi i}\int_{-T}^{T}\frac{e^{iaz}}{z-ib}\,dz-e^{-ba}\right|\leqq\frac{1}{Ta}(1-e^{-Ta})+e^{-Ta}.$$

Formulate a similar estimate when $a<0$.

7. Let $c>0$ and $a>0$. Taking the integral over the vertical line, prove that

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{a^z}{z}\,dz=\begin{cases}0 & \text{if } a<1,\\ \frac{1}{2} & \text{if } a=1,\\ 1 & \text{if } a>1.\end{cases}$$

If $a=1$, the integral is to be interpreted as the limit

$$\int_{c-i\infty}^{c+i\infty}=\lim_{T\to\infty}\int_{c-iT}^{c+iT}.$$

[*Hint*: If $a>1$, integrate around a rectangle with corners $c-Ai$, $c+Bi$, $-X+Bi$, $-X-Ai$, and let $X\to\infty$. If $a<1$, replace $-X$ by X.]

8. (a) Show that for $a>0$ we have

$$\int_{-\infty}^{\infty}\frac{\cos x}{(x^2+a^2)^2}\,dx=\frac{\pi(1+a)}{2a^3e^a}.$$

(b) Show that for $a > b > 0$ we have

$$\int_0^\infty \frac{\cos x}{(x^2 + a^2)(x^2 + b^2)}\,dx = \frac{\pi}{a^2 - b^2}\left(\frac{1}{be^b} - \frac{1}{ae^a}\right).$$

9. $\displaystyle\int_0^\infty \frac{\sin^2 x}{x^2}\,dx = \pi/2$. [*Hint*: Consider the integral of $(1 - e^{2ix})/x^2$.]

10. $\displaystyle\int_{-\infty}^\infty \frac{\cos x}{a^2 - x^2}\,dx = \frac{\pi \sin a}{a}$ for $a > 0$. The integral is meant to be interpreted as the limit:

$$\lim_{B\to\infty}\lim_{\delta\to 0}\int_{-B}^{-a-\delta} + \int_{-a+\delta}^{a-\delta} + \int_{a+\delta}^{B}.$$

11. $\displaystyle\int_{-\infty}^\infty \frac{\cos x}{e^x + e^{-x}}\,dx = \frac{\pi}{e^{\pi/2} + e^{-\pi/2}}$. Use the indicated contour:

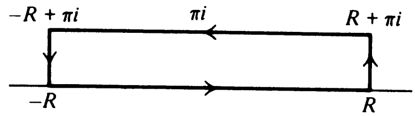


Figure 12

12. $\displaystyle\int_0^\infty \frac{x \sin x}{x^2 + a^2}\,dx = \frac{1}{2}\pi e^{-a}$ if $a > 0$.

13. $\displaystyle\int_{-\infty}^\infty \frac{e^{ax}}{e^x + 1}\,dx = \frac{\pi}{\sin \pi a}$ for $0 < a < 1$.

14. (a) $\displaystyle\int_0^\infty \frac{(\log x)^2}{1 + x^2}\,dx = \pi^3/8$. Use the contour

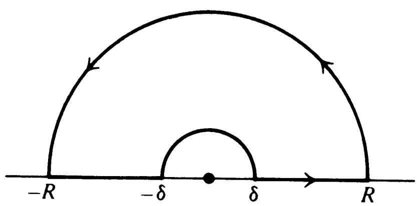


Figure 13

(b) $\displaystyle\int_0^\infty \frac{\log x}{(x^2 + 1)^2}\,dx = -\pi/4$.

15. (a) $\displaystyle\int_0^\infty \frac{x^a}{1 + x}\frac{dx}{x} = \frac{\pi}{\sin \pi a}$ for $0 < a < 1$.

(b) $\displaystyle\int_0^\infty \frac{x^a}{1+x^3}\frac{dx}{x} = \frac{\pi}{3\sin(\pi a/3)}$ for $0 < a < 3$.

16. Let f be a continuous function, and suppose that the integral

$$\int_0^\infty f(x)x^a \frac{dx}{x}$$

is absolutely convergent. Show that it is equal to the integral

$$\int_{-\infty}^\infty f(e^t)e^{at}\,dt.$$

If we put $g(t) = f(e^t)$, this shows that a Mellin transform is essentially a Fourier transform, up to a change of variable.

17. $\displaystyle\int_0^{2\pi} \frac{1}{1+a^2-2a\cos\theta}\,d\theta = \frac{2\pi}{1-a^2}$ if $0 < a < 1$. The answer comes out the negative of that if $a > 1$.

18. $\displaystyle\int_0^\pi \frac{1}{1+\sin^2\theta}\,d\theta = \pi/\sqrt{2}$.

19. $\displaystyle\int_0^\pi \frac{1}{3+2\cos\theta}\,d\theta = \pi/\sqrt{5}$.

20. $\displaystyle\int_0^\pi \frac{a\,d\theta}{a^2+\sin^2\theta} = \int_0^{2\pi} \frac{a\,d\theta}{1+2a^2-\cos\theta} = \frac{\pi}{\sqrt{1+a^2}}$.

21. $\displaystyle\int_0^{\pi/2} \frac{1}{(a+\sin^2\theta)^2}\,d\theta = \frac{\pi(2a+1)}{4(a^2+a)^{3/2}}$ for $a > 0$.

22. $\displaystyle\int_0^{2\pi} \frac{1}{2-\sin\theta}\,d\theta = 2\pi/\sqrt{3}$.

23. $\displaystyle\int_0^{2\pi} \frac{1}{(a+b\cos\theta)^2}\,d\theta = \frac{2\pi a}{(a^2-b^2)^{3/2}}$ for $0 < b < a$.

24. Let n be an even integer. Find

$$\int_0^{2\pi} (\cos\theta)^n\,d\theta$$

by the method of residues.

CHAPTER VII

Conformal Mappings

In this chapter we consider a more global aspect of analytic functions, describing geometrically what their effect is on various regions. Especially important are the analytic isomorphisms and automorphisms of various regions, of which we consider many examples.

Throughout the chapter, we use the words isomorphisms and automorphisms, omitting the word analytic, as there will be no others under consideration. We recall that an **isomorphism**

$$f: U \to V$$

is an analytic map which has an inverse analytic map

$$g: V \to U,$$

that is, $f \circ g = \mathrm{id}_V$ and $g \circ f = \mathrm{id}_U$. We say that f is an **automorphism** if $U = V$. We let $\mathrm{Aut}(U)$ be the set of automorphisms of U.

The main general theorem concerning isomorphisms is the **Riemann mapping theorem**:

If U is a simply connected open set which is not the whole plane, then there exists an isomorphism of U with the unit disc.

The general proof will be postponed to a later chapter. In the present chapter, we are concerned with specific examples, where the mapping can be exhibited concretely, in a simple manner.

It will also be useful to the reader to recall some simple algebraic formalism about isomorphisms and automorphisms, listed in the following properties.

Let $f: U \to V$ and $g: V \to W$ be two isomorphisms. Then

$$g \circ f: U \to W$$

is an isomorphism.

Let f, $g: U \to V$ be isomorphisms. Then there exists an automorphism h of V such that $g = h \circ f$.

Let $f: U \to V$ be an isomorphism. Then there is a bijection

$$\mathrm{Aut}(U) \to \mathrm{Aut}(V)$$

given by

$$\varphi \mapsto f \circ \varphi \circ f^{-1}.$$

The proofs are immediate in all cases. For instance, an inverse for $g \circ f$ is given by $f^{-1} \circ g^{-1}$ as one sees at once by composing these two maps in either direction. For the second statement, we have $h = g \circ f^{-1}$. As to the third, if φ is an automorphism of U then $f \circ \varphi \circ f^{-1}$ is an automorphism of V, because it is an isomorphism of V with itself. Similarly, if ψ is an automorphism of V then $f^{-1} \circ \psi \circ f$ is an automorphism of U. The reader will then verify that the associations

$$\varphi \mapsto f \circ \varphi \circ f^{-1} \quad \text{and} \quad \psi \mapsto f^{-1} \circ \psi \circ f$$

give maps between $\mathrm{Aut}(U)$ and $\mathrm{Aut}(V)$ which are inverse to each other, and hence establish the stated bijection between $\mathrm{Aut}(U)$ and $\mathrm{Aut}(V)$. The association $\varphi \mapsto f \circ \varphi \circ f^{-1}$ is called **conjugation by** f. It shows that if we know the set of automorphisms of U, then we also know the set of automorphisms of V if V is isomorphic to U: It is obtained by conjugation.

We also recall a result from Chapter II, Theorem 6.4.

Let f be analytic on an open set U. If f is injective, and $V = f(U)$ is its image, then

$$f: U \to V$$

is an analytic isomorphism, and in particular, $f'(z) \neq 0$ for all z in U.

This result came from the decomposition $f(z) = \varphi(z)^m$, where φ is a local analytic isomorphism in the neighborhood of a point z_0 in U, cf. Theorem 5.4.

Note that if f is analytic and $f'(z) \neq 0$ for all z in U, then we cannot conclude that f is injective. For instance let U be the open set obtained by deleting the origin from the plane, and let $f(z) = z^3$. Then $f'(z) \neq 0$ for all z in U, but f is not injective.

On the other hand, by restricting the open set U suitably, the map $z \mapsto z^3$ does become an isomorphism. For instance, let U be the sector consisting of all complex numbers $z = re^{i\theta}$ with $r > 0$ and $0 < \theta < \pi/3$. Then $z \mapsto z^3$ is an analytic isomorphism on U. What is its image?

VII, §1. SCHWARZ LEMMA

Let D be the unit disc of complex numbers z with $|z| < 1$.

Theorem 1.1. *Let $f: D \to D$ be an analytic function of the unit disc into itself such that $f(0) = 0$. Then:*

(i) *We have $|f(z)| \leqq |z|$ for all $z \in D$.*
(ii) *If for some $z_0 \neq 0$ we have $|f(z_0)| = |z_0|$, then there is some complex number α of absolute value 1 such that*

$$f(z) = \alpha z.$$

Proof. Let

$$f(z) = a_1 z + \cdots$$

be the power series for f. The constant term is 0 because we assumed $f(0) = 0$. Then $f(z)/z$ is holomorphic, and

$$\left|\frac{f(z)}{z}\right| < 1/r \qquad \text{for} \quad |z| = r < 1,$$

consequently also for $|z| \leqq r$ by the maximum modulus principle. Letting r tend to 1 proves the first assertion. If furthermore we have

$$\left|\frac{f(z_0)}{z_0}\right| = 1$$

for some z_0 in the unit disc, then again by the maximum modulus principle $f(z)/z$ cannot have a maximum unless it is constant, and therefore there is a constant α such that $f(z)/z = \alpha$, whence the second assertion also follows.

In the above statement of the Schwarz lemma, the function was normalized to map the unit disc into itself. The lemma obviously implies analogous statements when the functions satisfies a bound

$$|f(z)| \leqq B \quad \text{on a disc} \quad |z| < R, \qquad \text{and} \qquad f(0) = 0.$$

The conclusion is then that

$$|f(z)| \leqq B|z|/R,$$

and equality occurs at some point only if $f(z) = \frac{B}{R}\alpha z$, where α is a complex number of absolute value 1.

The following statement dealing with $f'(0)$ rather than the function itself will be considered as part of the Schwarz lemma.

Theorem 1.2. *Let $f: D \to D$ be an analytic function of the unit disc into itself such that $f(0) = 0$. Let*

$$f(z) = a_1 z + \textit{higher terms}.$$

Then $|f'(0)| = |a_1| \leqq 1$, and if $|a_1| = 1$, then $f(z) = a_1 z$.

Proof. Since $f(0) = 0$, the function $f(z)/z$ is analytic at $z = 0$, and

$$\frac{f(z)}{z} = a_1 + \text{higher terms}.$$

Letting z approach 0 and using the first part of Theorem 1.1 shows that $|a_1| \leqq 1$. Next suppose $|a_1| = 1$ and

$$f(z) = a_1 z + a_m z^m + \text{higher terms}$$

with $a_m \neq 0$ and $m \geqq 2$. Then

$$\frac{f(z)}{z} = a_1 + a_m z^{m-1} + \text{higher terms}.$$

Pick a value of z such that $a_m z^{m-1} = a_1$. There is a real number $C > 0$ such that for all small real $t > 0$ we have

$$\frac{f(tz)}{tz} = a_1 + a_m t^{m-1} z^{m-1} + h,$$

where $|h| \leqq Ct^m$. Since $|a_1| = 1$, it follows that

$$\left|\frac{f(tz)}{tz}\right| = |a_1(1 + t^{m-1}) + h| > 1$$

for t sufficiently small. This contradicts the first part of Theorem 1.1, and concludes the proof.

VII, §2. ANALYTIC AUTOMORPHISMS OF THE DISC

As an application of the Schwarz lemma, we shall determine all analytic automorphisms of the disc. First we give examples of such functions.

To begin with, we note that if φ is real, the map

$$z \mapsto e^{i\varphi}z$$

is interpreted geometrically as rotation counterclockwise by an angle φ. Indeed, if $z = re^{i\theta}$, then

$$e^{i\varphi}z = re^{i(\theta+\varphi)}.$$

Thus for example, the map $z \mapsto iz$ is a counterclockwise rotation by 90° (that is, $\pi/2$).

Let α be a complex number with $|\alpha| < 1$, and let

$$g_\alpha(z) = g(z) = \frac{\alpha - z}{1 - \bar{\alpha}z}.$$

Then g is analytic on the closed disc $|z| \leqq 1$. Furthermore, if $|z| = 1$, then $z = e^{i\theta}$ for some real θ, and

$$g(z) = \frac{\alpha - e^{i\theta}}{e^{i\theta}(e^{-i\theta} - \bar{\alpha})}.$$

Up to the factor $e^{i\theta}$ which has absolute value 1, the denominator is equal to the complex conjugate of the numerator, and hence

$$\textit{if } |z| = 1 \textit{ then } |g(z)| = 1.$$

We can argue by the maximum modulus principle, that if $|z| \leqq 1$, then $|g(z)| \leqq 1$. By the open mapping theorem, it follows that if $|z| < 1$ then $|g(z)| < 1$. Furthermore, g_α has an inverse function. As a trivial exercise, prove that

$$g_\alpha \circ g_\alpha = \mathrm{id}.$$

Therefore g_α is its own inverse function on the unit disc, and thus g_α gives an analytic automorphism of the unit disc with itself.

Observe that $g(\alpha) = 0$. We now prove that up to rotations there are no other automorphisms of the unit disc.

Theorem 2.1. *Let $f: D \to D$ be an analytic automorphism of the unit disc and suppose $f(\alpha) = 0$. Then there exists a real number φ such that*

$$f(z) = e^{i\varphi}\frac{\alpha - z}{1 - \bar{\alpha}z}.$$

Proof. Let $g = g_\alpha$ be the above automorphism. Then $f \circ g^{-1}$ is an automorphism of the unit disc, and maps 0 on 0, i.e. it has a zero at 0. It now suffices to prove that the function $h(w) = f(g^{-1}(w))$ is of the form

$$h(w) = e^{i\varphi}w$$

to conclude the proof of the theorem.

The first part of Schwarz lemma tells us that

$$|h(z)| \leqq |z| \qquad \text{if} \quad |z| < 1.$$

Since the inverse function h^{-1} also has a zero at the origin, we also get the inequality in the opposite direction, that is,

$$|z| \leqq |h(z)|,$$

and the second part of Schwarz lemma now implies that $h(z) = e^{i\varphi}z$, thereby proving our theorem.

Corollary 2.2. *If f is an automorphism of the disc which leaves the origin fixed, i.e. $f(0) = 0$, then $f(z) = e^{i\varphi}z$ for some real number φ, so f is a rotation.*

Proof. Let $\alpha = 0$ in the theorem.

VII, §2. EXERCISES

1. Let f be analytic on the unit disc D, and assume that $|f(z)| < 1$ on the disc. Prove that if there exist two distinct points a, b in the disc which are fixed points, that is, $f(a) = a$ and $f(b) = b$, then $f(z) = z$.

2. **(Schwarz–Pick Lemma).** Let $f: D \to D$ be a holomorphic map of the disc into itself. Prove that for all $a \in D$ we have

$$\frac{|f'(a)|}{1 - |f(a)|^2} \leqq \frac{1}{1 - |a|^2}.$$

[*Hint*: Let g be an automorphism of D such that $g(0) = a$, and let h be an automorphism which maps $f(a)$ on 0. Let $F = h \circ f \circ g$. Compute $F'(0)$ and apply the Schwarz lemma.]

3. Let α be a complex number, and let h be an isomorphism of the disc $D(\alpha, R)$ with the unit disc such that $h(z_0) = 0$. Show that

$$h(z) = \frac{R(z - z_0)}{R^2 - (z - \alpha)(\bar{z}_0 - \bar{\alpha})} e^{i\theta}$$

for some real number θ.

4. What is the image of the half strips as shown on the figure, under the mapping $z \mapsto iz$? Under the mapping $z \mapsto -iz$?

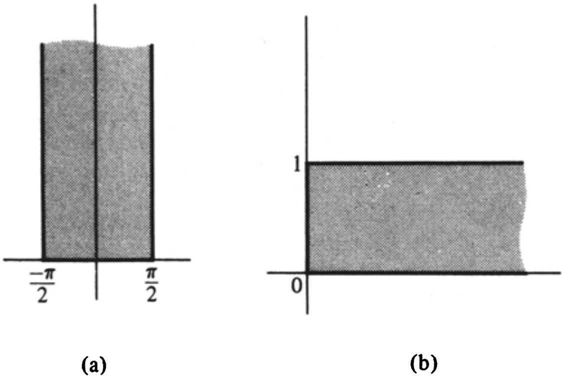


Figure 1

5. Let α be real, $0 \leqq \alpha < 1$. Let U_α be the open set obtained from the unit disc by deleting the segment $[\alpha, 1]$, as shown on the figure.
   (a) Find an isomorphism of U_α with the unit disc from which the segment $[0, 1]$ has been deleted.

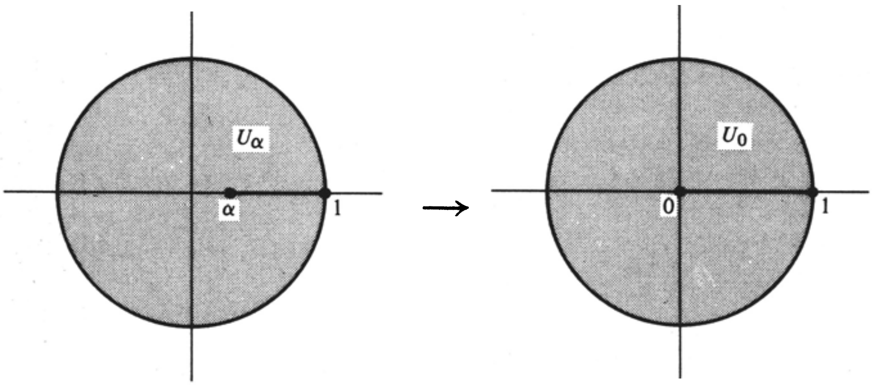


Figure 2

(b) For an isomorphism of U_0 with the upper half of the disc. Also find an isomorphism of U_α with this upper half disc.
[*Hint*: What does $z \mapsto z^2$ do to the upper half disc?]

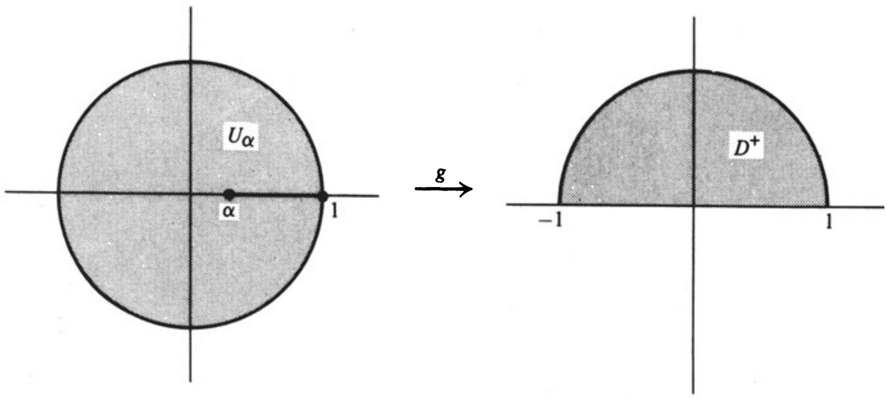


Figure 3

VII, §3. THE UPPER HALF PLANE

Theorem 3.1. *Let H be the upper half plane. The map*

$$\boxed{f\colon z \mapsto \frac{z-i}{z+i}}$$

is an isomorphism of H with the unit disc.

Proof. Let $w = f(z)$ and $z = x + iy$. Then

$$f(z) = \frac{x + (y-1)i}{x + (y+1)i}.$$

Since z is in H, $y > 0$, it follows that $(y-1)^2 < (y+1)^2$ whence

$$x^2 + (y-1)^2 = |z - i|^2 < x^2 + (y+1)^2 = |z + i|^2$$

and therefore

$$|z - i| < |z + i|,$$

so f maps the upper half plane into the unit disc. Since

$$\boxed{w = \frac{z-i}{z+i},}$$

we can solve for z in terms of w, because $wz + wi = z - i$, so that

$$\boxed{z = -i\frac{w+1}{w-1}.}$$

Write $w = u + iv$, with real u, v. By computing directly the real part of $(w+1)/(w-1)$, and so the imaginary part of

$$-i\frac{w+1}{w-1}$$

you will find that this imaginary part is > 0 if $|w| < 1$. Hence the map

$$h\colon w \mapsto -i\frac{w+1}{w-1}$$

sends the unit disc into the upper half plane. Since by construction f and h are inverse to each other, it follows that they are inverse isomorphisms of the upper half plane and the disc, as was to be shown.

Example. We wish to give an isomorphism of the first quadrant with the unit disc. Since we know that the upper half plane is isomorphic to the unit disc, it suffices to exhibit an isomorphism of the first quadrant with the upper half plane. The map

$$z \mapsto z^2$$

achieves this.

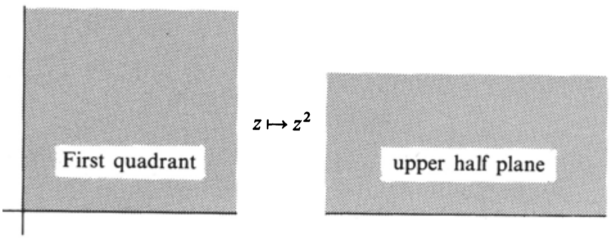


Figure 4

If $f: H \to D$ is the isomorphism of the upper half plane with the unit disc then

$$z \mapsto f(z^2)$$

is the desired isomorphism of the first quadrant with the unit disc. Thus the function

$$z \mapsto \frac{z^2 - i}{z^2 + i}$$

gives an isomorphism of the first quadrant with the unit disc.

The existence of an isomorphism $f\colon H \to D$ of H with the unit disc also in some sense determines the automorphisms of H. By the general formalism of isomorphisms and automorphisms, we know that

$$\operatorname{Aut}(H) = f^{-1}\operatorname{Aut}(D)f,$$

meaning that every automorphism of H is of the form $f^{-1} \circ \varphi \circ f$ with some automorphism φ of D. The question is to give a more explicit description of $\operatorname{Aut}(H)$. In the exercises, you will develop a proof of the following theorem.

Theorem 3.2. *Let*

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

be a 2×2 *real matrix with determinant* 1. *Let* f_M *be the mapping such that*

$$f_M(z) = \frac{az + b}{cz + d} \quad \textit{for} \quad z \in H.$$

Then f_M *is an automorphism of* H, *and every automorphism of* H *is of the form* f_M *for some such matrix* M. *Furthermore, two such automorphisms* $f_{M'}$ *and* f_M *are equal if and only if* $M' = \pm M$.

VII, §3. EXERCISES

Let

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

be a 2×2 matrix of real numbers, such that $ad - bc > 0$. For $z \in H$, define

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

1. Show that

$$\operatorname{Im} f_M(z) = \frac{(ad - bc)y}{|cz + d|^2}.$$

2. Show that f_M gives a map of H into H.

3. Let $\mathrm{GL}_2^+(\mathbf{R})$ denote the set of all real 2×2 matrices with positive determinant. Then $\mathrm{GL}_2^+(\mathbf{R})$ is closed under multiplication and taking multiplicative inverses, so $\mathrm{GL}_2^+(\mathbf{R})$ is called a group. Show that if M, $M' \in \mathrm{GL}_2^+(\mathbf{R})$, then

$$f_{MM'} = f_M \circ f_{M'}.$$

This is verified by brute force. Then verify that if I is the unit matrix,

$$f_I = \mathrm{id} \quad \text{and} \quad f_{M^{-1}} = (f_M)^{-1}.$$

Thus every analytic map f_M of H has an analytic inverse, actually in $\mathrm{GL}_2^+(\mathbf{R})$, and in particular f_M is an automorphism of H.

4. (a) If $c \in \mathbf{R}$ and cM is the usual scalar multiplication of a matrix by a number, show that $f_{cM} = f_M$. In particular, let $\mathrm{SL}_2(\mathbf{R})$ denote the subset of $\mathrm{GL}_2^+(\mathbf{R})$ consisting of the matrices with determinant 1. Then given $M \in \mathrm{GL}_2^+(\mathbf{R})$, one can find $c > 0$ such that $cM \in \mathrm{SL}_2(\mathbf{R})$. Hence as far as studying analytic automorphisms of H are concerned, we may concern ourselves only with $\mathrm{SL}_2(\mathbf{R})$.
   (b) Conversely, show that if $f_M = f_{M'}$ for M, $M' \in \mathrm{SL}_2(\mathbf{R})$, then

$$M' = \pm M.$$

5. (a) Given an element $z = x + iy \in H$, show that there exists an element $M \in \mathrm{SL}_2(\mathbf{R})$ such that $f_M(i) = z$.
   (b) Given $z_1, z_2 \in H$, show that there exists $M \in \mathrm{SL}_2(\mathbf{R})$ such that $f_M(z_1) = z_2$. In light of (b), one then says that $\mathrm{SL}_2(\mathbf{R})$ **acts transitively** on H.

6. Let K denote the subset of elements $M \in \mathrm{SL}_2(\mathbf{R})$ such that $f_M(i) = i$. Show that if $M \in K$, then there exists a real θ such that

$$M = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

All Automorphisms of the Upper Half Plane

Do the following exercises after you have read the beginning of §5. In particular, note that Exercise 3 generalizes to fractional linear maps. Indeed, if M, M' denote any complex non-singular 2×2 matrices, and F_M, $F_{M'}$ are the corresponding fractional linear maps, then

$$F_{MM'} = F_M \circ F_{M'}.$$

Hence if I is the unit 2×2 matrix, then

$$F_I = \text{id} \qquad \text{and} \qquad F_{M^{-1}} = F_M^{-1}.$$

7. Let $f\colon H \to D$ be the isomorphism of the text, that is

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

Note that f is represented as a fractional linear map, $f = F_M$ where M is the matrix

$$M = \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

Of course, this matrix does not have determinant 1.

Let K be the set of Exercises 5. Let $\mathrm{Rot}(D)$ denote the set of rotations of the unit disc, i.e. $\mathrm{Rot}(D)$ consists of all automorphisms

$$R_\theta\colon w \mapsto e^{i\theta}w \qquad \text{for } w \in D.$$

Show that $fKf^{-1} = \mathrm{Rot}(D)$, meaning that $\mathrm{Rot}(D)$ consists of all elements $f \circ f_M \circ f^{-1}$ with $M \in K$.

8. Finally prove the theorem:

Theorem. *Every automorphism of H is of the form f_M for some $M \in \mathrm{SL}_2(\mathbf{R})$.*

[*Hint*: Proceed a follows. Let $g \in \mathrm{Aut}(H)$. There exists $M \in \mathrm{SL}_2(\mathbf{R})$ such that

$$f_M(g(i)) = i.$$

By Exercise 6, we have $f_M \circ g \in K$, say $f_M \circ g = h \in K$, and therefore

$$g = f_M^{-1} \circ h \in \mathrm{SL}_2(\mathbf{R}),$$

thus concluding the proof.]

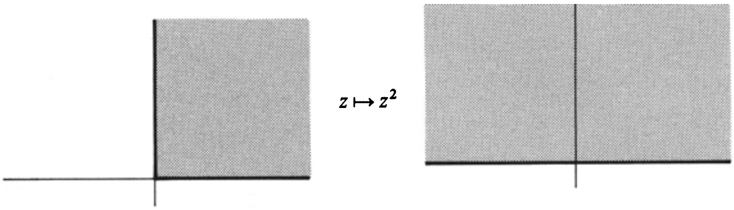
From the Upper Half Plane to the Punctured Disc

9. Let $f(z) = e^{2\pi iz}$. Show that f maps the upper half plane on the inside of a disc from which the center has been deleted. Given $B > 0$, let $H(B)$ be that part of the upper half plane consisting of those complex numbers $z = x + iy$ with $y \geqq B$. What is the image of $H(B)$ under f? Is f an isomorphism? Why? How would you restrict the domain of definition of f to make it an isomorphism?

VII, §4. OTHER EXAMPLES

We give the examples by pictures which illustrate various isomorphisms.

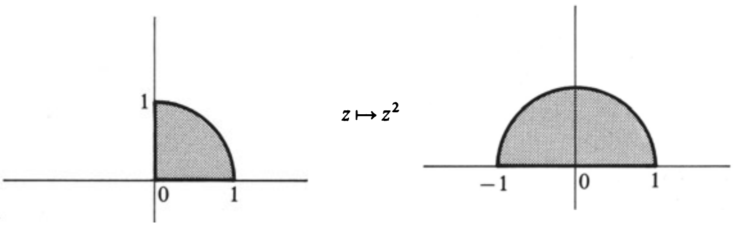
Example 1.



Isomorphism between first quadrant and upper half plane

Figure 5

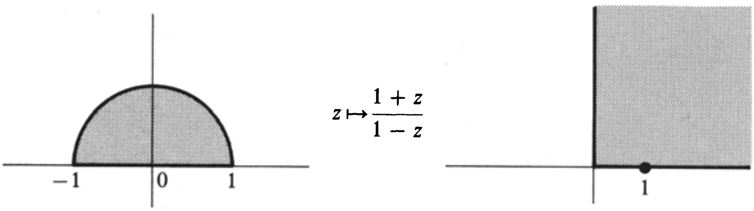
Example 2.



Isomorphism between quarter disc and half disc

Figure 6

Example 3.

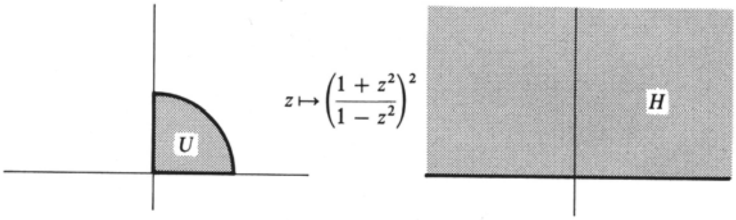


Upper half disc with first quadrant

Figure 7

Example 4. By composing the above isomorphisms, we get new ones. For instance, let U be the portion of the unit disc lying inside the first quadrant as in Example 2. We want to get an isomorphism of U with the upper half plane.

All we have to do is to compose the isomorphisms of Examples 2, 3, and 1 in that order. Thus an isomorphism of U with H is given by the formula in the picture.

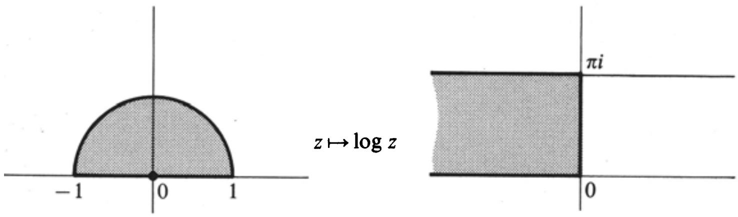


Quarter disc with upper half plane

Figure 8

The next three examples concern the logarithm.

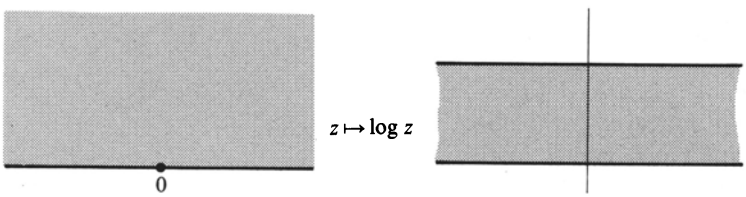
Example 5.



Upper half disc with a half strip

Figure 9

Example 6.



Upper half plane with a full strip

Figure 10

Example 7.

Plane with deleted positive real axis with a full strip

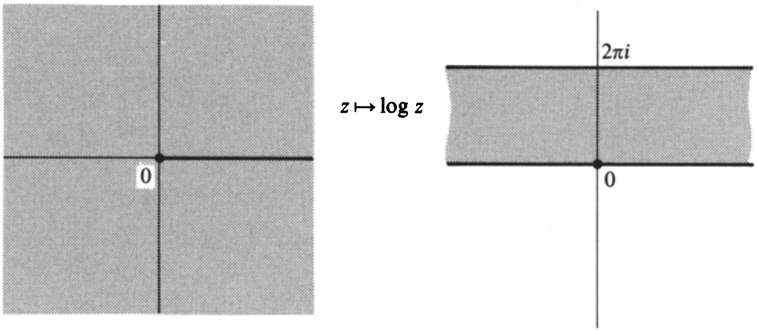


Figure 11

In the applications to fluid dynamics, we shall see in the next chapter that it is important to get isomorphisms of various regions with the upper half plane in order to be able to describe the flow lines. In particular, certain regions are obtained by placing obstacles inside simpler regions. We give several examples of this phenomenon. These will allow us to get an isomorphism from a strip containing a vertical obstacle with the upper half plane.

Example 8.

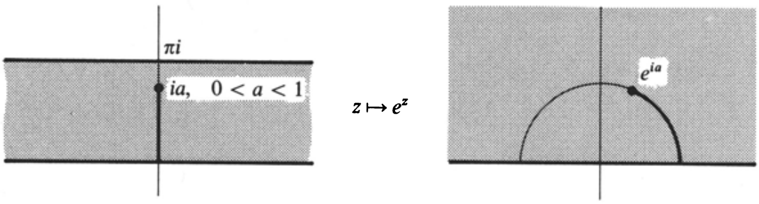


Figure 12

Example 9.

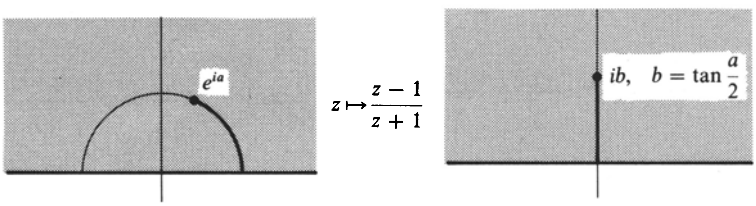


Figure 13

Example 10.

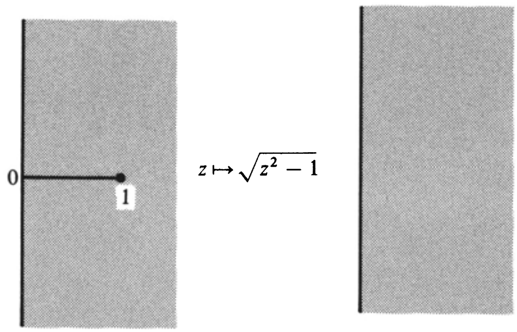


Figure 14

Remark. By composing the isomorphisms of Examples 8 and 9, using a dilation, and a rotation, and finally the isomorphism of Example 10, we get an isomorphism of the strip containing a vertical obstacle with the right half plane:

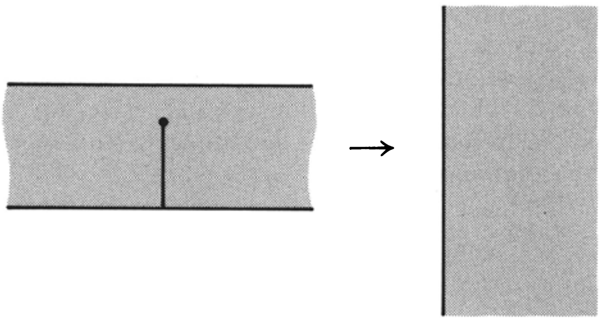


Figure 15

Another rotation would then yield the upper half plane.

Example 11. In this example, the obstacle is a bump rather than a vertical line segment. *We claim that the map*

$$z \mapsto z + \frac{1}{z}$$

is an isomorphism of the open set U lying inside the upper half plane, above the unit circle, with the upper half plane.

The isomorphism is shown on the following figure.

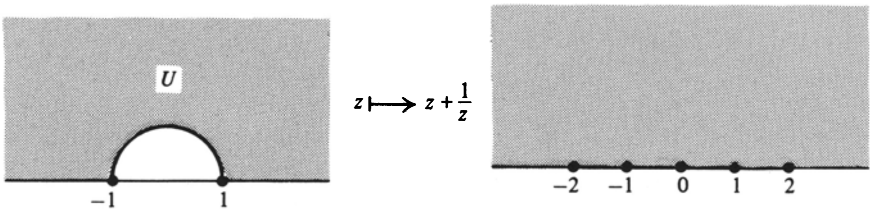


Figure 16

Proof. Let $w = z + 1/z$ so that

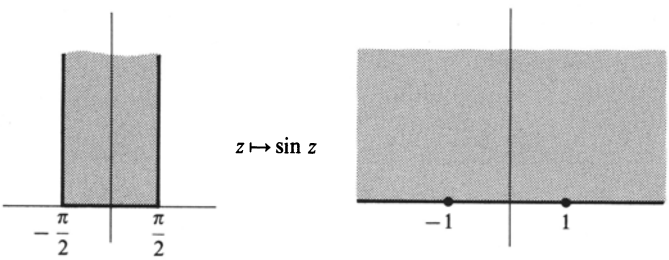
$$w = x\left(1 + \frac{1}{x^2 + y^2}\right) + iy\left(1 - \frac{1}{x^2 + y^2}\right).$$

If $z \in U$, then $|z| > 1$ so $\operatorname{Im} w > 0$ and $w \in H$. The quadratic equation

$$z^2 - zw + 1 = 0$$

has two distinct roots except for $w = \pm 2$. Given $w \in H$, any root $z = x + iy$ has the property that either $y > 0$ and $x^2 + y^2 > 1$, or $y < 0$ and $x^2 + y^2 < 1$. Since the product of the two roots is 1 (from the quadratic equation), and hence the product of their absolute values is also 1, it follows that not both roots can have absolute value > 1 or both have absolute value < 1. Hence exactly one root lies in U, so the map is both surjective and injective, as desired.

Example 12.



Upper half strip with upper half plane

Figure 17

The sine function maps the interval $[-\pi/2, \pi/2]$ on the interval $[-1, 1]$.

Let us look also at what the sine does to the right vertical boundary, which consists of all points $\pi/2 + it$ with $t \geqq 0$. We know that

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}.$$

Hence

$$\sin\left(\frac{\pi}{2} + it\right) = \frac{e^{i\pi/2}e^{iit} - e^{-i\pi/2}e^{-iit}}{2i}$$
$$= \frac{e^t + e^{-t}}{2}.$$

As t ranges from 0 to infinity, $\sin(\pi/2 + it)$ therefore ranges from 1 to infinity, so the image of the vertical half line is the part of the real axis lying to the right of 1. You can show similarly that the image of the left vertical boundary is the part of the real axis to the left of -1. Thus we see precisely what the mapping $z \mapsto \sin z$ does to the boundary of the region.

For the convenience of the reader, we also discuss the mapping on the interior of the region. Let $z = x + iy$ with

$$-\frac{\pi}{2} \leqq x \leqq \frac{\pi}{2} \quad \text{and} \quad 0 < y.$$

From the definition of $\sin z$, letting $w = \sin z = u + iv$, we find

$$(1) \qquad u = \sin x \cosh y \quad \text{and} \quad v = \cos x \sinh y,$$

where

$$\cosh y = \frac{e^y + e^{-y}}{2} \quad \text{and} \quad \sinh y = \frac{e^y - e^{-y}}{2}.$$

From (1) we get

$$(2) \qquad \frac{u^2}{\cosh^2 y} + \frac{v^2}{\sinh^2 y} = 1,$$

$$(3) \qquad \frac{u^2}{\sin^2 x} - \frac{v^2}{\cos^2 x} = 1.$$

If we fix a value of $y > 0$, then the line segment

$$x + iy \quad \text{with} \quad -\frac{\pi}{2} \leqq x \leqq \frac{\pi}{2}$$

gets mapped onto the upper half of an ellipse in the w-plane, as shown on the figure. Note that for the given intervals, we have $u \geqq 0$.

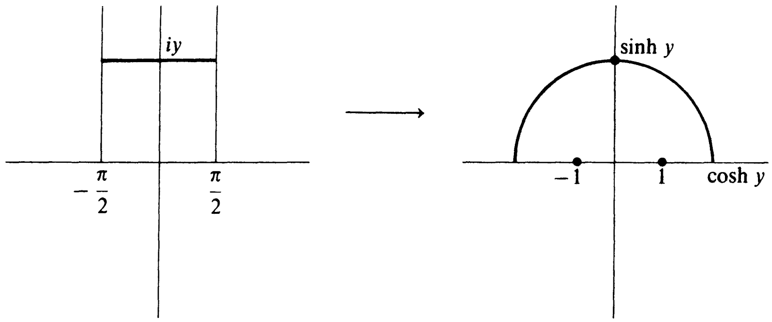


Figure 18

Geometrically speaking, as y increases from 0 to infinity, the ellipses expand and fill out the upper half plane.

One can also determine the image of vertical lines, fixing x and letting y vary, so half lines of the form $x + iy$ with $y > 0$ and x fixed. Equation (3) shows that the image of such half lines are upper parts of hyperbolas. It is the right upper part if $x > 0$ and the left upper part if $x < 0$. Since an analytic map with non-zero derivative is conformal, these hyperbolas are perpendicular to the above ellipses because vertical lines are perpendicular to horizontal lines.

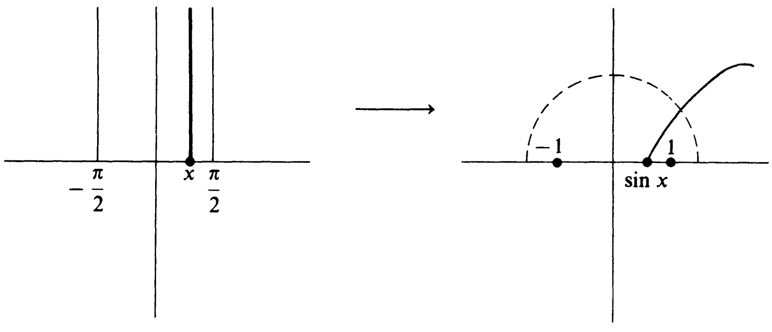


Figure 19

In the above examples, readers may have noticed that to a large extent knowledge of the behavior of the mapping function on the boundary of an open set determines the effect of the mapping inside this boundary, whatever inside means. We shall now give very general theorems which make this notion precise. Thus to verify that a certain mapping gives an isomorphism between open sets, it frequently suffices to know what the mapping does to the boundary. We shall give two criteria for this to happen.

Theorem 4.1. *Let U be a bounded connected open set, $\bar{U}$ its closure. Let f be a continuous function on $\bar{U}$, analytic on U. Suppose that f is not constant, and maps the boundary of U into the unit circle, so*

$$|f(z)| = 1 \qquad \text{for all } z \in \text{boundary of } U.$$

Then f maps U into the unit disc D and $f\colon U \to D$ is surjective.

Proof. That f maps U into D follows from the maximum modulus principle. Suppose there exists some $\alpha \in D$ but α is not in the image of f. Let g_α be the automorphism of D interchanging 0 and α, given in Chapter VII, §2, so g_α extends continuously to the closed unit disc. Then $g_\alpha \circ f$ satisfies the same hypotheses as f, but 0 (instead of α) is not in the image of $g_\alpha \circ f$. Then $1/g_\alpha \circ f$ is analytic on U, and continuous on $\bar{U}$, mapping the boundary of $\bar{U}$ into the unit circle. If $z_0 \in U$ so $|g_\alpha \circ f(z_0)| < 1$, then

$$|1/g_\alpha \circ f(z_0)| > 1,$$

contradicting the maximum modulus principle, and concluding the proof.

Remark. Suppose that f is injective in the first place. Then the above (which is very simple to prove) gives a criterion for f to be an isomorphism, entirely in terms of the effect of f on the boundary, which can usually be verified more simply than checking that f has an inverse map.

For the next theorems, we recall that the interior of a path (if it has one) was defined in Chapter VI, §1.

Lemma 4.2. *Let γ be a piecewise C^1 closed path in an open set U of* $\mathbf{C}$. *Suppose that γ has an interior, denoted by* $\text{Int}(\gamma)$. *Then the union*

$$S = \text{Int}(\gamma) \cup \gamma$$

(*identifying γ with its image in U*) *is compact.*

Proof. To see this, we have to prove that S is closed and bounded. Let $\langle z_n \rangle$ be a sequence in S, converging to some point in $\mathbf{C}$. If the sequence contains infinitely many points of $\text{Int}(\gamma)$, then for such points $W(\gamma, z_n) = 1$, so by continuity, the limit lies in $\text{Int}(\gamma)$ or on γ. If the sequence contains infinitely many points on γ, then the limit lies on γ (because γ is compact). This proves that S is closed. Furthermore $\text{Int}(\gamma)$ is bounded, because for all complex numbers α with $|\alpha|$ sufficiently large, $W(\gamma, \alpha) = 0$ since γ is homotopic to a point in a disc containing γ but not α. This proves the lemma.

Theorem 4.3. *Let γ be a piecewise C^1 closed path in a connected open set U of* $\mathbf{C}$. *Assume γ homologous to 0 in U. Let f be analytic non-*

constant on U. Assume that γ and $f \circ \gamma$ have interiors and $f \circ \gamma$ does not intersect $f(\operatorname{Int}\gamma)$. *Then f is injective on* $\operatorname{Int}(\gamma)$, *and so induces an isomorphism of* $\operatorname{Int}(\gamma)$ *with its image. If in addition the interior of $f \circ \gamma$ is connected, then*

$$f\colon \operatorname{Int}(\gamma) \to \operatorname{Int}(f \circ \gamma)$$

is an isomorphism.

Proof. For $\alpha \in \operatorname{Int}(\gamma)$, let $f_\alpha(z) = f(z) - f(\alpha)$. Then by the chain rule:

$$\begin{aligned} W(f \circ \gamma), f(\alpha)) &= \frac{1}{2\pi i}\int_{f\circ\gamma} \frac{1}{\zeta - f(\alpha)}\, d\zeta \\ &= \frac{1}{2\pi i}\int_a^b \frac{f'(\gamma(t))\gamma'(t)}{f(\gamma(t)) - f(\alpha)}\, dt \qquad \text{if } \gamma\colon [a, b] \to U \\ &= \frac{1}{2\pi i}\int_\gamma f'_\alpha/f_\alpha(z)\, dz \\ &\geqq 1 \end{aligned}$$

because f_α has a zero in $\operatorname{Int}(\gamma)$. By the definition of an interior and the assumption that $f \circ \gamma$ has an interior, it follows that this final value of $W(f \circ \gamma, f(\alpha))$ is precisely 1, so f_α has only one zero, which is simple, and in addition we also see that $f(\alpha) \in \operatorname{Int}(f \circ \gamma)$. Hence

$$f\colon \operatorname{Int}(\gamma) \to \operatorname{Int}(f \circ \gamma)$$

is injective. This proves the first assertion.

Suppose next that the interior of $f \circ \gamma$ is connected. The image of $\operatorname{Int}(\gamma)$ under f is an open subset, since we have shown that f maps $\operatorname{Int}(\gamma)$ into the interior of $f \circ \gamma$. So it suffices to prove that $f(\operatorname{Int}(\gamma))$ is closed in the interior of $f \circ \gamma$. Let $\{z_n\}$ be a sequence of points in $\operatorname{Int}(\gamma)$, such that $\{f(z_n)\}$ converges to some point w in $\operatorname{Int}(f \circ \gamma)$. We have to show that $w \in f(\operatorname{Int}(\gamma))$. Passing to a subsequence of $\{z_n\}$ if necessary, and using the fact that the union of $\operatorname{Int}(\gamma)$ and γ is compact, we may assume without loss of generality that $\{z_n\}$ itself converges, either to a point in the interior $\operatorname{Int}(\gamma)$, or to a point of γ itself. If $\{z_n\}$ converges to α, and $\alpha \in \operatorname{Int}(\gamma)$, then $f(\alpha) = w$ by continuity of f, and we are done. If α lies on γ, then again by continuity $f(\alpha) = w$, which is impossible since w is not on $f \circ \gamma$. This concludes the proof.

Readers should go through the examples given in the text to see that the mapping property in Theorem 4.3 also applies to unbounded domains in those cases. Cf. Exercises 7 and 8 below for specific instances. Readers may also look at Exercise 2 of Chapter X, §1, which illustrates several aspects of the way the mapping on the boundary affects the mapping inside, including aspects of reflection, to be discussed in greater detail in Chapter IX.

VII, §4. EXERCISES

1. (a) In each one of the examples, prove that the stated mapping is an isomorphism on the figures as shown. Also determine what the mapping does to the boundary lines. Thick lines should correspond to each other.
   (b) In Example 10, give the explicit formula giving an isomorphism of the strip containing a vertical obstacle with the right half plane, and also with the upper half plane. Note that counterclockwise rotation by $\pi/2$ is given by multiplication with i.

2. (a) Show that the function $z \mapsto z + 1/z$ is an analytic isomorphism of the region outside the unit circle onto the plane from which the segment $[-2, 2]$ has been deleted.
   (b) What is the image of the unit circle under this mapping? Use polar coordinates.

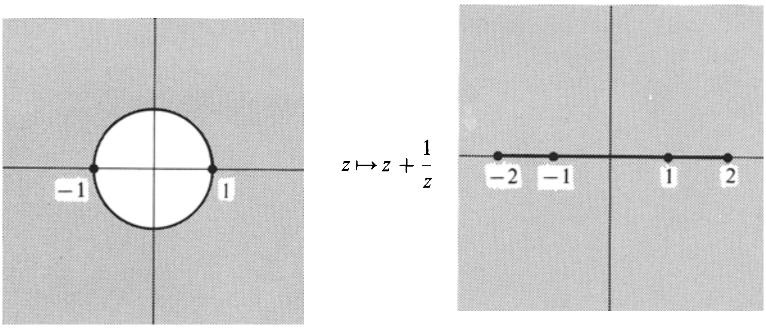


Figure 20

   (c) In polar coordinates, if $w = z + 1/z = u + iv$, then

$$u = \left(r + \frac{1}{r}\right)\cos\theta \qquad \text{and} \qquad v = \left(r - \frac{1}{r}\right)\sin\theta.$$

   Show that the circle $r = c$ with $c > 1$ maps to an ellipse with major axis $c + 1/c$ and minor axis $c - 1/c$. Show that the radial lines $\theta = c$ map onto quarters of hyperbolas.

3. Let U be the upper half plane from which the points of the closed unit disc are removed, i.e. U is the set of z such that $\mathrm{Im}(z) > 0$ and $|z| > 1$. Give an explicit isomorphism of U with the upper half disc D^+ (the set of z such that $|z| < 1$ and $\mathrm{Im}(z) > 0$).

4. Let a be a real number. Let U be the open set obtained from the complex plane by deleting the infinite segment $[a, \infty[$. Find explicitly an analytic isomorphism of U with the unit disc. Give this isomorphism as a composite of simpler ones. [*Hint*: Try first to see what $\sqrt{z}$ does to the set obtained by deleting $[0, \infty[$ from the plane.]

5. (a) Show that the function $w = \sin z$ can be decomposed as the composite of two functions:

$$w = \frac{\zeta - \zeta^{-1}}{2i} = f(\zeta) \qquad \text{and} \qquad \zeta = e^{iz} = g(z).$$

(b) Let U be the open upper half strip in Example 12. Let $g(U) = V$. Describe V explicitly and show that $g\colon U \to V$ is an isomorphism. Show that g extends to a continuous function on the boundary of U and describe explicitly the image of this boundary under g.

(c) Let $W = f(V)$. Describe W explicitly and show that $f\colon V \to W$ is an isomorphism. Again describe how f extends continuously to the boundary of V and describe explicitly the image of this boundary under f.

In this way you can recover the fact that $w = \sin z$ gives an isomorphism of the upper half strip with the upper half plane by using this decomposition into simpler functions which you have already studied.

6. In Example 12, show that the vertical imaginary axis above the real line is mapped onto itself by $z \mapsto \sin z$, and that this function gives an isomorphism of the half strip with the first quadrant as shown on the figure.

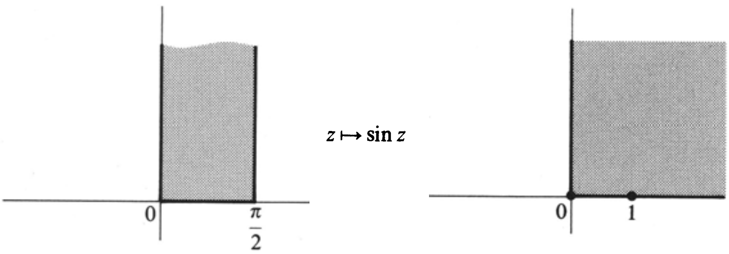


Figure 21

7. Let $w = u + iv = f(z) = z + \log z$ for z in the upper half plane H. Prove that f gives an isomorphism of H with the open set U obtained from the upper half plane by deleting the infinite half line of numbers

$$u + i\pi \qquad \text{with} \quad u \leqq -1.$$

Remark. In the next chapter, we shall see that the isomorphism f allows us to determine the flow lines of a fluid as shown on Fig. 23. These flow lines in the (u, v)-plane correspond to the rays $\theta = \text{constant}$ in the (x, y)-plane. In other words, they are the images under f of the rays $\theta = \text{constant}$.

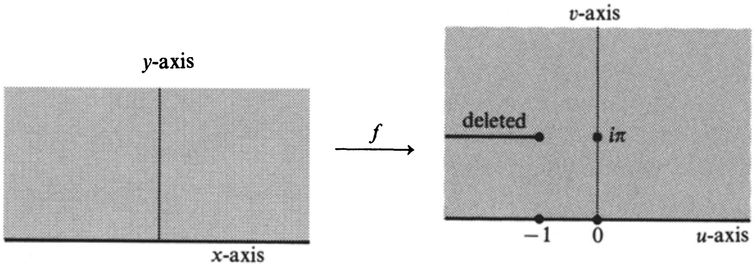
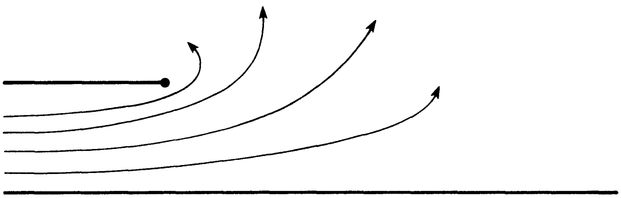


Figure 22



Flow lines in the (u, v)*-plane*

Figure 23

[*Hint*: Use Theorem 4.3 applied to the path consisting of the following pieces:

The segment from R to ϵ (R large, ϵ small >0).

The small semicircle in the upper half plane, from ϵ to $-\epsilon$.

The segment from $-\epsilon$ to $-R$.

The large semicircle in the upper half plane from $-R$ to R.

Note that if we write $z = re^{i\theta}$, then $f(z) = re^{i\theta} + \log r + i\theta$.]

8. Give another proof of Example 11 using Theorem 4.3.

VII, §5. FRACTIONAL LINEAR TRANSFORMATIONS

Let a, b, c, d be complex numbers such that $ad - bc \neq 0$. We may arrange these numbers as a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Let

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

We call F a **fractional linear map**, or **transformation**. We have already encountered functions of this type, and now we study them more systematically. First observe that if we multiply a, b, c, d by the same non-zero complex number λ, then the matrix

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

gives rise to the same map, because we can cancel λ in the fraction:

$$\frac{\lambda az + \lambda b}{\lambda cz + \lambda d} = \frac{az + b}{cz + d}.$$

It is an exercise to prove the converse, that if two matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$$

of complex numbers with $ad - bc \neq 0$ and $a'd' - b'c' \neq 0$ give the same fractional linear map, then there is a complex number λ such that

$$a' = \lambda a, \qquad b' = \lambda b, \qquad c' = \lambda c, \qquad d' = \lambda d.$$

We shall now see that F gives an isomorphism. Note that

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

The function F is not defined at $z = -d/c$, but is defined at all other complex numbers, and the formula for its derivative then shows that $F'(z) \neq 0$ for all complex numbers $z \neq -d/c$.

The function F has an inverse. Indeed, let

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

We can solve for z in terms of w by simple algebra. Cross multiplying yields

$$czw + dw = az + b,$$

whence

$$z = \frac{dw - b}{-cw + a}.$$

Thus the inverse function is associated with the matrix

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Observe that the inverse function, which we denote by F^{-1}, is not defined at $z = a/c$. Thus F gives an isomorphism of $\mathbf{C}$ from which $-d/c$ has been deleted with $\mathbf{C}$ from which a/c has been deleted.

To have a uniform language to deal with the "exceptional" points $z = -d/c$ and $z = a/c$, we agree to the following conventions.

Let S be the Riemann sphere, i.e. the set consisting of $\mathbf{C}$ and a single point ∞ which we call **infinity**. We extend the definition of F to S by defining

$$F(\infty) = a/c \qquad \text{if} \quad c \neq 0,$$

$$F(\infty) = \infty \qquad \text{if} \quad c = 0.$$

Also we define

$$F(-d/c) = \infty \qquad \text{if} \quad c \neq 0.$$

These definitions are natural, for if we write

$$F(z) = \frac{a + b/z}{c + d/z},$$

and let $|z| \to \infty$ then this fraction approaches a/c as a limit.

We may then say that F gives a bijection of S with itself.

We now define other maps as follows:

$T_b(z) = z + b$, called **translation** by b;

$J(z) = 1/z$, called **inversion** through the unit circle;

$M_a(z) = az$ for $a \neq 0$, called **multiplication** by a.

Observe that translations, reflections, or multiplications are fractional linear maps. Translations should have been encountered many times previously. As for inversion, note:

If $|z| = 1$ then $1/z = \bar{z}$ and $|1/z| = 1$ also. Thus an inversion maps the unit circle onto itself.

If $|z| > 1$ then $|1/z| < 1$ and vice versa, so an inversion interchanges the region outside the unit disc with the region inside the unit disc. Note that 0 and ∞ correspond to each other under the inversion.

Multiplication by a complex number a can be viewed as a dilation together with a rotation, by writing $a = re^{i\theta}$.

Thus each one of these particular linear maps has a simple geometric interpretation as above.

Theorem 5.1. *Given a fractional linear map F, there exist complex numbers α, β, γ such that either $F = \alpha z + \beta$, or*

$$F(z) = T_\gamma \circ M_\alpha \circ J \circ T_\beta.$$

Proof. Suppose $c = 0$. Then $F(z) = (az + b)/d$ and $F = T_\beta \circ M_\alpha$, with $\beta = b/d$, $\alpha = a/d$. Suppose this is not the case, so $c \neq 0$. We divide a, b, c, d by c and using these new numbers gives the same map F, so without loss of generality we may assume $c = 1$. We let $\beta = d$. We must solve

$$\frac{az + b}{z + d} = \frac{\alpha}{z + d} + \gamma,$$

or in other words, $az + b = \alpha + \gamma z + \gamma d$. We let $\gamma = a$, and then solve for $\alpha = b - ad \neq 0$ to conclude the proof.

The theorem shows that any fractional linear map is a composition of the simple maps listed above: translations, multiplication, or inversion.

Now let us define a **straight line on the Riemann sphere** S to consist of an ordinary line together with ∞.

Theorem 5.2. *A fractional linear transformation maps straight lines and circles onto straight lines and circles. (Of course, a circle may be mapped onto a line and vice versa.)*

Proof. By Theorem 5.1 it suffices to prove the assertion in each of the three cases of the simple maps. The assertion is obvious for translations and multiplications (which are rotations followed by dilations). There remains to deal with the inversion.

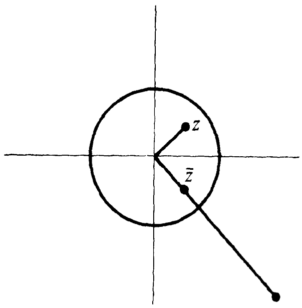


Figure 24

Let $w = 1/z = u + iv$, so that

$$u = \frac{x}{x^2 + y^2} \quad \text{and} \quad v = \frac{-y}{x^2 + y^2}.$$

The equation of a circle or straight line in the (u, v) real plane has the form

$$A(u^2 + v^2) + Bu + Cv = D$$

with some real numbers A, B, C, D such that not all A, B, C are equal to 0. Substituting the values for u, v in terms of x, y we find that under the inverse mapping the equation is equivalent to

$$A + Bx - Cy = D(x^2 + y^2),$$

which is the equation of a circle or a straight line. This proves the theorem.

As Exercise 2, you will prove that if F, G are fractional linear maps, then so is $F \circ G$. We shall use such compositions in the next considerations.

By a **fixed point** of F we mean a point z_0 such that $F(z_0) = z_0$.

Example. The point ∞ is a fixed point of the map

$$F(z) = az + b.$$

Proposition 5.3. *Let F be fractional linear map. If ∞ is a fixed point of F, then there exist complex numbers a, b such that $F(z) = az + b$.*

Proof. Let $F(z) = (az + b)/(cz + d)$. If $c \neq 0$ then $F(\infty) = a/c$ which is not ∞. By hypothesis, it follows that $c = 0$, in which case the assertion is obvious.

Theorem 5.4. *Given any three distinct points z_1, z_2, z_3 on the Riemann sphere, and any three distinct points w_1, w_2, w_3, there exists a unique fractional linear map F such that*

$$F(z_i) = w_i \quad \textit{for} \quad i = 1, 2, 3.$$

Proof. We proceed stepwise, and first prove uniqueness. Let F, G be fractional linear maps which have the same effect on three points. Then $F \circ G^{-1}$ has three fixed points, and it suffices to prove the following lemma.

Lemma 5.5. *Let F be a fractional linear map. If F has three fixed points, then F is the identity.*

Proof. Suppose first that one fixed point is ∞. By Proposition 5.3, we know that $F(z) = az + b$. Suppose $z_1 \in \mathbf{C}$ and z_1 is a fixed point. Then $az_1 + b = z_1$ so $(1 - a)z_1 = b$. If $a \neq 1$ then we see that $z_1 = b/(1 - a)$ is the only fixed point in **C**. If $a = 1$, then $f(z) = z + b$, and $z + b = z$ if and only if $b = 0$, so f is the identity mapping.

Suppose next that ∞ is not a fixed point, so $c \neq 0$. Let z be a fixed point, so that

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

Cross multiplying shows that z is a root of the quadratic equation

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

This equation has at most two roots, so we see that f has at most two fixed points, which are the roots of this equation. Of course the roots are given explicitly by the quadratic formula.

One can give an easy formula for the map F of Theorem 5.4. Note that the function $z \mapsto z - z_1$ sends z_1 to 0. Then

$$z \mapsto \frac{z - z_1}{z - z_2}$$

sends z_1 to 0 and z_2 to ∞. To send z_3 to 1, all we have to do is multiply by the right factor, and thus we obtain:

Theorem 5.6. *The function*

$$z \mapsto \frac{z - z_1}{z - z_2} \frac{z_3 - z_2}{z_3 - z_1}$$

is the unique function such that $F(z_1) = 0$, $F(z_2) = \infty$, $F(z_3) = 1$. *If* $w = F(z)$ *is the function such that* $F(z_i) = w_i$ *for* $i = 1, 2, 3$ *then* w *and* z *are related by the formula*

$$\boxed{\frac{w - w_1}{w - w_2} \frac{w_3 - w_2}{w_3 - w_1} = \frac{z - z_1}{z - z_2} \frac{z_3 - z_2}{z_3 - z_1}.}$$

This final equation can be used to find F explicitly in special cases.

Example. Find the map F in Theorem 5.4 such that

$$F(1) = i, \qquad F(i) = -1, \qquad F(-1) = 1.$$

By the formula,

$$\frac{w-i}{w+1}\frac{1+1}{1-i}=\frac{z-1}{z-i}\frac{-1-i}{-1-1},$$

or in other words,

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

We can solve for w in terms of z to give

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

To check the computation, substitute $z=1$, $z=i$, $z=-1$ in this expression to see that you get the desired values i, -1, and 1, respectively.

Warning. I find it pointless to memorize the formula in Theorem 5.6 relating z and w. However, the comments before Theorem 5.6 tell you how to reconstruct this formula in an easy way if you don't have it for reference in front of you.

VII, §5. EXERCISES

1. Give explicitly a fractional linear map which sends a given complex number z_1 to ∞. What is the simplest such map which sends 0 to ∞?

2. **Composition of Fractional Linear Maps.** Show that if F, G are fractional linear maps, then so is $F \circ G$.

3. Find fractional linear maps which map:
   (a) $1, i, -1$ on $i, -1, 1$
   (b) $i, -1, 1$ on $-1, -i, 1$
   (c) $-1, -i, 1$ on $-1, 0, 1$
   (d) $-1, 0, 1$ on $-1, i, 1$
   (e) $1, -1, i,$ on $1, i, -1$

4. Find fractional linear maps which map:
   (a) $0, 1, \infty$ on $1, \infty, 0$
   (b) $0, 1, \infty$ on $1, -1, i$
   (c) $0, 1, \infty$ on $-1, 0, 1$
   (d) $0, 1, \infty$ on $-1, -i, 1$

5. Let F and G be two fractional linear maps, and assume that $F(z)=G(z)$ for all complex numbers z (or even for three distinct complex numbers z). Show

that if

$$F(z) = \frac{az + b}{cz + d} \quad \text{and} \quad G(z) = \frac{a'z + b'}{c'z + d'}$$

then there exists a complex number λ such that

$$a' = \lambda a, \qquad b' = \lambda b, \qquad c' = \lambda c, \qquad d' = \lambda d.$$

Thus the matrices representing F and G differ by a scalar.

6. Consider the fractional linear map

$$F(z) = \frac{z - i}{z + i}.$$

What is the image of the real line **R** under this map? (You have encountered this map as an isomorphism between the upper half plane and the unit disc.)

7. Let F be the fractional linear map $F(z) = (z - 1)/(z + 1)$. What is the image of the real line under this map? (Cf. Example 9 of §4.)

8. Let $F(z) = z/(z - 1)$ and $G(z) = 1/(1 - z)$. Show that the set of all possible fractional linear maps which can be obtained by composing F and G above repeatedly with each other in all possible orders in fact has six elements, and give a formula for each one of these elements. [*Hint*: Compute F^2, F^3, G^2, G^3, $F \circ G$, $G \circ F$, etc.]

9. Let $F(z) = (z - i)/(z + i)$. What is the image under F of the following sets of points:
   (a) The upper half line it, with $t \geqq 0$.
   (b) The circle of center 1 and radius 1.
   (c) The horizontal line $i + t$, with $t \in \mathbf{R}$.
   (d) The half circle $|z| = 2$ with $\operatorname{Im} z \geqq 0$.
   (e) The vertical line $\operatorname{Re} z = 1$ and $\operatorname{Im} z \geqq 0$.

10. Find fractional linear maps which map:
    (a) $0, 1, 2$ to $1, 0, \infty$
    (b) $i, -1, 1$ to $1, 0, \infty$
    (c) $0, 1, 2$ to $i, -1, 1$

11. Let $F(z) = (z + 1)/(z - 1)$. Describe the image of the line $\operatorname{Re}(z) = c$ for a real number c. (Distinguish $c = 1$ and $c \neq 1$. In the second case, the image is a circle. Give its center and radius.)

12. Let z_1, z_2, z_3, z_4 be distinct complex numbers. Define their **cross ratio** to be

$$[z_1, z_2, z_3, z_4] = \frac{(z_1 - z_3)(z_2 - z_4)}{(z_2 - z_3)(z_1 - z_4)}.$$

   (a) Let F be a fractional linear map. Let $z_i' = F(z_i)$ for $i = 1, \ldots, 4$. Show that the cross ratio of z_1', z_2', z_3', z_4' is the same as the cross ratio of

z_1, z_2, z_3, z_4. It will be easy if you do this separately for translations, inversions, and multiplications.

(b) Prove that the four numbers lie on the same straight line or on the same circle if and only if their cross ratio is a real number.

(c) Let z_1, z_2, z_3, z_4 be distinct complex numbers. Assume that they lie on the same circle, in that order. Prove that

$$|z_1 - z_3||z_2 - z_4| = |z_1 - z_2||z_3 - z_4| + |z_2 - z_3||z_4 - z_1|.$$

Fixed Points and Linear Algebra

13. Find the fixed points of the following functions:

(a) $f(z) = \dfrac{z-3}{z+1}$ (b) $f(z) = \dfrac{z-4}{z+2}$

(c) $f(z) = \dfrac{z-i}{z+1}$ (d) $f(z) = \dfrac{2z-3}{z+1}$

For the next two exercises, we assume that you know the terminology of eigenvalues from an elementary course in linear algebra.

14. Let M be a 2×2 complex matrix with non-zero determinant,

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{and} \quad ad - bc \neq 0.$$

Define $M(z) = (az + b)/(cz + d)$ as in the text for $z \neq -d/c$ $(c \neq 0)$. If $z = -d/c$ $(c \neq 0)$ we put $M(z) = \infty$. We define $M(\infty) = a/c$ if $c \neq 0$, and ∞ if $c = 0$.

(a) If L, M are two complex matrices as above, show directly that

$$L(M(z)) = (LM)(z)$$

for $z \in \mathbf{C}$ or $z = \infty$. Here LM is the product of matrices from linear algebra.

(b) Let λ, λ' be the eigenvalues of M viewed as a linear map on $\mathbf{C}^2$. Let

$$W = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \quad \text{and} \quad W' = \begin{pmatrix} w_1' \\ w_2' \end{pmatrix}.$$

be the corresponding eigenvectors, so

$$MW = \lambda W \quad \text{and} \quad MW' = \lambda' W'$$

By a **fixed point** of M on $\mathbf{C}$ we mean a complex number z such that $M(z) = z$. Assume that M has two distinct fixed points in $\mathbf{C}$. Show that these fixed points are $w = w_1/w_2$ and $w' = w_1'/w_2'$.

(c) Assume that $|\lambda| < |\lambda'|$. Given $z \neq w$, show that

$$\lim_{k\to\infty} M^k(z) = w'.$$

Note. The iteration of the fractional linear map is sometimes called a **dynamical system**. Under the assumption in (c), one says that w' is an **attracting point** for the map, and that w is a **repelling** point.

CHAPTER VIII

Harmonic Functions

In this chapter we return to the connection between analytic functions and functions of a real variable, analyzing an analytic function in terms of its real part.

The first two sections, §1 and §2, are completely elementary and could have been covered in Chapter I. They combine well with the material in the preceding chapter, as they deal with the same matter, pursued to analyze the real part of analytic isomorphism more closely.

In §3 and §4 we deal with those aspects of harmonic functions having to do with integration and some form of Cauchy's formula. We shall characterize harmonic functions as real parts of analytic functions, giving an explicit integral formula for the associated analytic function (uniquely determined except for a pure imaginary constant).

VIII, §1. DEFINITION

A function $u = u(x, y)$ is called **harmonic** if it is real valued having continuous partial derivatives of order one and two, and satisfying

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

One usually defines the **Laplace** (differential) **operator**

$$\Delta = \left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2,$$

and so u is harmonic if and only if $\Delta u = 0$ (and u is of class C^2). Suppose f is analytic on an open set U. We know that f is infinitely complex differentiable. By the considerations of Chapter I, §6, it follows that its real and imaginary parts $u(x, y)$ and $v(x, y)$ are C^∞, and satisfy the Cauchy–Riemann equations

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

Consequently, taking the partial derivatives of these equations and using the known fact that $\dfrac{\partial}{\partial x}\dfrac{\partial}{\partial y} = \dfrac{\partial}{\partial y}\dfrac{\partial}{\partial x}$ yields:

Theorem 1.1. *The real part of an analytic function is harmonic.*

Example. Let $r = \sqrt{x^2 + y^2}$. Then $\log r$ is harmonic, being the real part of the complex log.

We introduce the **differential operators**

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

The reason for this notation is apparent if we write

$$x = \frac{1}{2}(z + \bar{z}) \quad \text{and} \quad y = \frac{1}{2i}(z - \bar{z}).$$

We want the chain rule to hold. Working formally, we see that the following equations must be satisfied.

$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial z} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial z} = \frac{1}{2}\frac{\partial f}{\partial x} + \frac{1}{2i}\frac{\partial f}{\partial y},$$

$$\frac{\partial f}{\partial \bar{z}} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \bar{z}} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \bar{z}} = \frac{1}{2}\frac{\partial f}{\partial x} - \frac{1}{2i}\frac{\partial f}{\partial y}.$$

This shows that it is reasonable to define $\partial/\partial z$ and $\partial/\partial \bar{z}$ as we have done. It is then immediately clear that u, v satisfy the Cauchy–Riemann equations if and only if

$$\frac{\partial f}{\partial \bar{z}} = 0.$$

(Carry out in detail.) Thus:

f is analytic if and only if $\partial f/\partial \bar{z} = 0$.

Since taking partial derivatives commutes, one also finds immediately

$$\boxed{4\frac{\partial}{\partial z}\frac{\partial}{\partial \bar{z}} = 4\frac{\partial}{\partial \bar{z}}\frac{\partial}{\partial z} = \Delta.}$$

(Do Exercise 1.) Also with the notation $\partial/\partial z$, if $u = \mathrm{Re}(f)$, we can write the complex derivative in the form

$$\boxed{f'(z) = 2\frac{\partial u}{\partial z} = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}.} \tag{1}$$

In Chapter I, §6, we had introduced the **associated vector field**

$$\bar{F}(x, y) = \big(u(x, y), -v(x, y)\big).$$

Recall that in calculus courses, one defines a **potential function** for $\bar{F}$ to be a function φ such that

$$\frac{\partial \varphi}{\partial x} = u \quad \text{and} \quad \frac{\partial \varphi}{\partial y} = -v.$$

Theorem 1.2. *Let g be a primitive for f on U, that is, $g' = f$. Write g in terms of its real and imaginary parts,*

$$g = \varphi + i\psi.$$

Then φ is a potential function for $\bar{F}$.

Proof. Go back to Chapter I, §6. By definition, $g' = u + iv$. The first computation of that section shows that

$$\frac{\partial \varphi}{\partial x} = u \quad \text{and} \quad \frac{\partial \varphi}{\partial y} = -v,$$

as desired.

We shall prove shortly that any harmonic function is locally the real part of an analytic function. *In that light, the problem of finding a primitive for an analytic function is equivalent to the problem of finding a potential function for its associated vector field.*

The expression (1) for the complex derivative suggests that we tabulate independently the integral in terms of real and imaginary parts, in a more general context as follows. *Let u be a real C^1 function on some open set.* Let γ be a piecewise C^1 path in the open set. Then setting $dz = dx + i\,dy$, we get

$$(2) \qquad \int_\gamma 2\frac{\partial u}{\partial z}\,dz = \int_\gamma \left(\frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy\right) + i\int_\gamma \left(\frac{\partial u}{\partial x}\,dy - \frac{\partial u}{\partial y}\,dx\right).$$

The first integral is the usual integral from calculus, of $\operatorname{grad} u = (\partial u/\partial x, \partial u/\partial y)$ along the curve. If we fix end points for the path, this integral is independent of the path since $\operatorname{grad} u$ admits a potential function, namely u itself. The second integral can be written as

$$(3) \qquad \int_\gamma \left(\frac{\partial u}{\partial x}\,dy - \frac{\partial u}{\partial y}\,dx\right) = \int_\gamma (\operatorname{grad} u) \cdot N_\gamma$$

where N_γ is the normal to the curve. If $\gamma(t) = \big(x(t), y(t)\big)$, then by definition,

$$N_\gamma(t) = \left(\frac{dy}{dt}, -\frac{dx}{dt}\right).$$

Thus we find (without any need for Cauchy–Riemann):

For a real C^1 function u, we have

$$(4) \qquad \int_\gamma 2\frac{\partial u}{\partial z}\,dz = \int_\gamma du + i\int_\gamma (\operatorname{grad} u) \cdot N_\gamma.$$

If γ is a closed path, then

$$\int_\gamma 2\frac{\partial u}{\partial z}\,dz = i\int_\gamma (\operatorname{grad} u) \cdot N_\gamma.$$

This kind of decomposition has not been relevant up to now, but when dealing explicitly with the real or imaginary part of an analytic function as we shall now do, such formulas come to the fore.

The next theorem gives us the **uniqueness of a harmonic function with prescribed boundary value.**

Theorem 1.3. *Let U be a bounded open set. Let u, v be two continuous functions on the closure $\overline{U}$ of U, and assume that u, v are harmonic on U. Assume that $u = v$ on the boundary of U. Then $u = v$ on U.*

Proof. Subtracting the two harmonic functions having the same boundary value yields a harmonic function with boundary value 0. Let u be such a function. We have to prove that $u = 0$. Suppose there is a point

$(x_0, y_0) \in U$ such that $u(x_0, y_0) > 0$. Let

$$\psi(x, y) = u(x, y) + \epsilon x^2 \qquad \text{for} \quad (x, y) \in \overline{U}.$$

We use repeatedly the fact that $|x|$ is bounded for $(x, y) \in \overline{U}$. Then

$$\psi(x_0, y_0) > 0$$

for ϵ small enough, and ψ is continuous on $\overline{U}$, so ψ has a maximum on $\overline{U}$. For ϵ small, the maximum of ψ is close to the maximum of u itself, and in particular is positive. But $u(x, y) = 0$ for (x, y) on the boundary, and ϵx^2 is small on the boundary for ϵ small. Hence the maximum of ψ must be an interior point (x_1, y_1). It follows that

$$D_1^2\psi(x_1, y_1) \leqq 0 \qquad \text{and} \qquad D_2^2\psi(x_1, y_1) \leqq 0.$$

But

$$(D_1^2 + D_2^2)u = 0 \qquad \text{so} \qquad (D_1^2 + D_2^2)\psi(x_1, y_1) = 2\epsilon > 0.$$

This contradiction proves the uniqueness.

Remarks. In practice, the above uniqueness is weak for two reasons. First, many natural domains are not bounded, and second the function may be continuous on the boundary except at a finite number of points. In examples below, we shall see some physical situations with discontinuities in the temperature function. Hence it is useful to have a more general theorem, which can be obtained as follows.

As to the unboundedness of the domain, it is usually possible to find an isomorphism of a given open set with a bounded open set such that the boundary curves correspond to each other. We shall see an example of this in the Riemann mapping theorem, which gives such isomorphisms with the unit disc. Thus the lack of boundedness of the domain may not be serious.

As to discontinuities on the boundary, let us pick for concreteness the unit disc D. Let u, v be two functions on $\overline{D}$ which are harmonic on the interior D, and which are continuous on the unit circle except at a finite number of points, where they are not defined. Suppose that u and v are equal on the boundary except at those exceptional points. Then the function $u - v$ is harmonic on the open disc D, and is continuous with value 0 on the boundary except at a finite number of points where it is not defined. Thus for the uniqueness, we need the following generalization of Theorem 1.3.

Theorem 1.4. *Let u be a bounded function on the closed unit disc $\overline{D}$. Assume that u is harmonic on D and continuous on the unit circle except at a finite number of points. Assume that u is equal to 0 on the unit* *circle except at a finite number of points. Then $u = 0$ on the open disc D.*

The situation is similar to that of removable singularities in Chapter V, §3, but technically slightly more difficult to deal with. We omit the proof. If one does not assume that u is bounded, then the conclusion of the theorem is not true in general. See Exercise 6.

Application: Perpendicularity

Recall from the calculus of several variables (actually two variables) that

$$\boxed{\operatorname{grad} u = (D_1 u, D_2 u) = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right).}$$

Let c be a number. The equation

$$u(x, y) = c$$

is interpreted as the equation of the level curve, consisting of those points at which u takes the constant value c. If u is interpreted as a potential function, these curves are called curves of **equipotential**. If u is interpreted as temperature, these curves are called **isothermal** curves. Except for such fancy names, they are just level curves of the function u. From calculus, you should know that grad $u(x, y)$ is perpendicular (orthogonal) to the curve at that point, as illustrated on the figure.

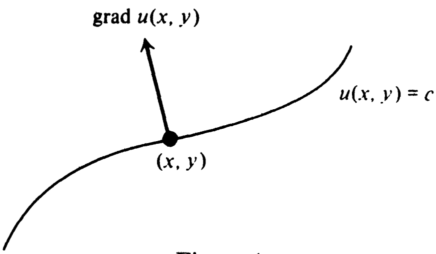


Figure 1

Let $A = (a_1, a_2)$ and $B = (b_1, b_2)$ be vectors. You should know their dot product,

$$A \cdot B = a_1 b_1 + a_2 b_2,$$

and you should know that A is perpendicular to B if and only if their dot product is equal to 0.

Using this and the chain rule, we recall the proof that the gradient is perpendicular to the level curve. We suppose that the level curve is parametrized, i.e. given in the form $\gamma(t)$ for t in some interval. Then we have

$$u(\gamma(t)) = c \qquad \text{for all } t.$$

Differentiating with respect to t yields by the chain rule

$$\operatorname{grad} u(\gamma(t)) \cdot \gamma'(t) = 0,$$

which proves what we wanted.

The following statement is an immediate consequence of the Cauchy–Riemann equations.

Let $f = u + iv$ be analytic. Then grad u *and* grad v *are perpendicular.*

Indeed, we take the dot product of

$$\left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) \quad \text{and} \quad \left(\frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}\right)$$

and apply the Cauchy–Riemann equations to find the value 0.

Two curves $u = c$ and $v = c'$ are said to be **perpendicular** at a point (x, y) if grad u is perpendicular to grad v at (x, y). Hence the above statement is interpreted as saying:

The level curves of the real part and imaginary part of an analytic function are perpendicular (or in other words, intersect orthogonally).

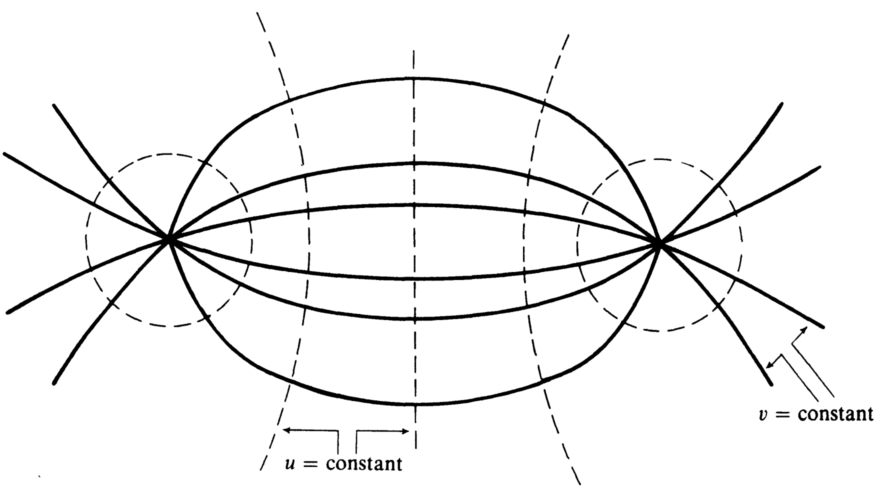


Figure 2

In the case when u is given as the potential function arising from two point sources of electricity, then the level curves for u and v look like Fig. 2.

Application: Flow Lines

For each point (x, y) in the plane, we have an associated vector

$$(x, y) \mapsto \operatorname{grad} u(x, y).$$

This association defines what is called a **vector field**, which we may visualize as arrows shown on Fig. 3.

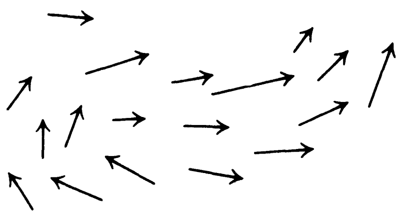


Figure 3

Let us abbreviate

$$G(x, y) = \operatorname{grad} u(x, y).$$

An **integral curve** for the vector field G is a curve γ such that

$$\gamma'(t) = G(\gamma(t)).$$

This means that the tangent vector at every point of the curve is the prescribed vector by G. Such an integral curve is shown on Fig. 4.

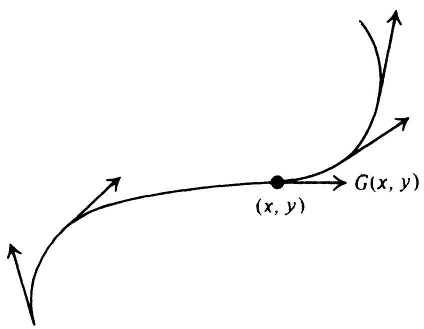


Figure 4

If we interpret the vector field G as a field of forces, then an integral curve is the path over which a bug will travel, when submitted to such a force.

Suppose that $f = u + iv$ is analytic, as usual. We have seen that the level curves of v are orthogonal to the level curves of u. Thus the level curves of v have the same direction as the gradient of u. It can be shown from the uniqueness of the solutions of differential equations that the level curves of v are precisely the integral curves of the vector field $G = \text{grad}\, u$. Thus interpreting u as temperature, for instance, we may say:

If $u = \text{Re}\, f$ and $v = \text{Im}\, f$, where f is analytic, then the heat flow of the temperature function u occurs along the level curves of v.

Finally, let U be simply connected, and let

$$f: U \to H$$

be an isomorphism of U with the upper half plane. We write $f = u + iv$ as usual. The curves

$$v = \text{constant}$$

in H are just horizontal straight lines. The level curves of v in U therefore correspond to these straight lines under the function f.

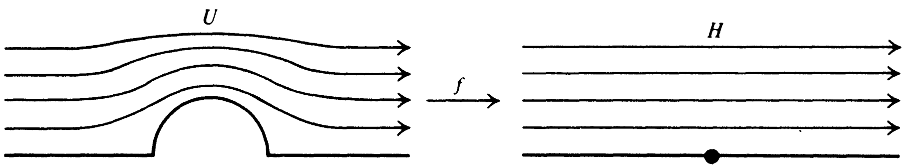


Figure 5

Consider the example of Chapter VII, §4 given by

$$f(z) = z + 1/z.$$

You should have worked out that this gives an isomorphism as shown on the figure. We interpret the right-hand figure as that of a fluid flowing horizontally in the upper half plane, without obstacle. The bump provided by the semicircle in the left-hand figure provides an obstacle to the flow in the open set U.

The nature of the physical world is such that the flow lines on the left are exactly the lines corresponding to the horizontal lines on the right *under the mapping function f! Thus the flow lines on the left are exactly the level curves $v = $ constant in U.*

This shows how an isomorphism

$$f\colon U \to H$$

can be applied to finding flow lines. The same principle could be applied to a similar obstacle as in Fig. 6. The open set U is defined here as that portion of the upper half plane obtained by deleting the vertical segment $(0, y)$ with $0 < y < \pi$ from the upper half plane. As an exercise, determine the isomorphism f to find the flow lines in U. Cf. Exercises 9, 10 and the examples of Chapter VII, §4.

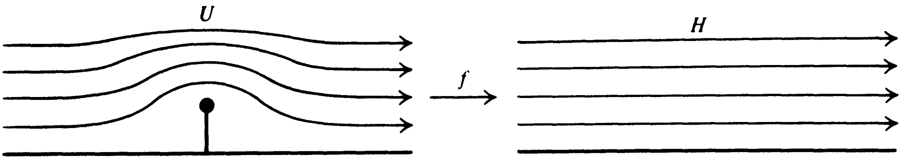


Figure 6

VIII, §1. EXERCISES

1. (a) Let $\Delta = \left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2$. Verify that $\Delta = 4\frac{\partial}{\partial z}\frac{\partial}{\partial \bar{z}}$.
   (b) Let f be a complex function on $\mathbf{C}$ such that both f and f^2 are harmonic. Show that f is holomorphic or $\bar{f}$ is holomorphic.
2. Let f be analytic, and $\bar{f} = u - iv$ the complex conjugate function. Verify that $\partial \bar{f}/\partial z = 0$.
3. Let $f\colon U \to V$ be an analytic isomorphism, and let φ be a harmonic function on V, which is the real part of an analytic function. Prove that the composite function $\varphi \circ f$ is harmonic.
4. Prove that the imaginary part of an analytic function is harmonic.
5. Prove the uniqueness statement in the following context. Let U be an open set contained in a strip $a \leqq x \leqq b$, where a, b are fixed numbers, and as usual $z = x + iy$. Let u be a continuous function on $\bar{U}$, harmonic on U. Assume that u is 0 on the boundary of U, and

$$\lim u(x, y) = 0$$

as $y \to \infty$ or $y \to -\infty$, uniformly in x. In other words, given ϵ there exists $C > 0$ such that if $y > C$ or $y < -C$ and $(x, y) \in U$ then $|u(x, y)| < \epsilon$. Then $u = 0$ on U.

6. Let

$$u(x, y) = \operatorname{Re}\frac{i+z}{i-z} \quad \text{for } z \neq i \qquad \text{and} \qquad u(0, 1) = 0.$$

Show that u harmonic on the unit disc, is 0 on the unit circle, and is continuous on the closed unit disc except at the point $z = i$. This gives a counterexample to the uniqueness when u is not bounded.

7. Find an analytic function whose real part is the given function.
   (a) $u(x, y) = 3x^2y - y^3$ (b) $x - xy$
   (c) $\dfrac{y}{x^2 + y^2}$ (d) $\log\sqrt{x^2 + y^2}$
   (e) $\dfrac{y}{(x - t)^2 + y^2}$ where t is some real number.

8. Let $f(z) = \log z$. If $z = re^{i\theta}$, then

$$f(z) = \log r + i\theta,$$

so the real parts and imaginary parts are given by

$$u = \log r \qquad \text{and} \qquad v = \theta.$$

Draw the level curves $u = \text{constant}$ and $v = \text{constant}$. Observe that they intersect orthogonally.

9. Let V be the open set obtained by deleting the segment $[0, 1]$ from the right half plane, as shown on the figure. In other words V consists of all complex numbers $x + iy$ with $x > 0$, with the exception of the numbers $0 < x \leqq 1$.
   (a) What is the image of V under the map $z \mapsto z^2$.
   (b) What is the image of V under the map $z \mapsto z^2 - 1$?
   (c) Find an isomorphism of V with the right half plane, and then with the upper half plane. [*Hint*: Consider the function $z \mapsto \sqrt{z^2 - 1}$.]

10. Let U be the open set discussed at the end of the section, obtained by deleting the vertical segment of points $(0, y)$ with $0 \leqq y \leqq 1$ from the upper

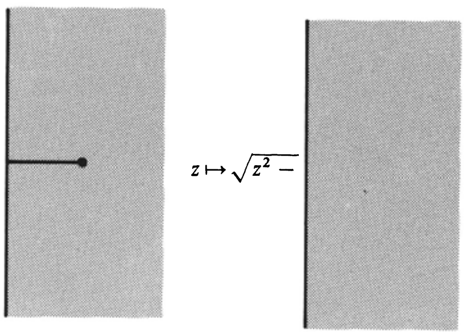


Figure 7

half plane. Find an analytic isomorphism

$$f: U \to H.$$

[*Hint*: Rotate the picture by 90° and use Exercise 9.]

11. Let φ be a complex harmonic function on a connected open set U. Suppose that φ^2 is also harmonic. Show that φ or $\bar{\varphi}$ is holomorphic.

12. **Green's Theorem** in calculus states: *Let $p = p(x, y)$ and $q = q(x, y)$ be C^1 functions on the closure of a bounded open set U whose boundary consists of a finite number of C^1 curves oriented so that U lies to the left of each one of these curves. Let C be this boundary. Then*

$$\int_C p\,dx + q\,dy = \iint_U \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) dy\,dx.$$

Suppose that f is analytic on U and on its boundary. Show that Green's theorem implies Cauchy's theorem for the boundary, i.e. show that

$$\int_C f = 0.$$

13. Let U be an open set and let $z_0 \in U$. The **Green's function for U originating at** z_0 is a real function g defined on the closure $\bar{U}$ of U, continuous except at z_0, and satisfying the following conditions:

GR 1. $g(z) = \log|z - z_0| + \psi(z)$, where ψ is harmonic on U.

GR 2. g vanishes on the boundary of U.

(a) Prove that a Green's function is uniquely determined if U is bounded.
(b) Let U be simply connected, with smooth boundary. Let

$$f: U \to D$$

be an analytic isomorphism of U with the unit disc such that $f(z_0) = 0$. Let

$$g(z) = \text{Re} \log f(z).$$

Show that g is a Green's function for U. You may assume that f extends to a continuous function from the boundary of U to the boundary of D.

VIII, §2. EXAMPLES

We shall give examples, some of which are formulated in physical terms. Let U be an open set whose boundary is a smooth curve C. We shall assume throughout that any physical function mentioned is harmonic.

In physics, functions like temperature, potential functions, are harmonic. We shall also assume that the uniqueness of a harmonic function on U with prescribed boundary value holds if the boundary value is assumed bounded and continuous except at a finite number of points.

Our examples are constructed for special open sets, which are simply connected. In general any such set is analytically isomorphic to the disc, or preferably to the upper half plane. Let

$$f: U \to H$$

be such an isomorphism. In practice, it is clear how f behaves at the boundary of U, and how it maps this boundary on the boundary of H, i.e. on the real axis. To construct a harmonic function on U with prescribed boundary values, it therefore suffices to construct a function φ on H, and then take the composite $\varphi \circ f$ (see Exercise 3 of the preceding section). In practice, there always exist nice explicit formulas giving the isomorphism f.

Example. We wish to describe the temperature in the upper half plane if the temperature is fixed with value 0 on the positive real axis, and fixed with value 20 on the negative real axis. As mentioned, temperature $v(z)$ is assumed to be harmonic. We recall that we can define

$$\log z = r(z) + i\theta(z)$$

for z in any simply connected region, in particular for

$$0 \leqq \theta(z) \leqq \pi, \qquad r(z) > 0$$

omitting the origin. We have

$$\theta(z) = 0 \text{ if } z \text{ is on the positive real axis,}$$

$$\theta(z) = \pi \text{ if } z \text{ is on the negative real axis.}$$

The function θ (sometimes denoted by arg) is the imaginary part of an analytic function, and hence is harmonic. The desired temperature is therefore obtainable as an appropriate constant multiple of $\theta(z)$, namely

$$v(z) = \frac{20}{\pi}\theta(z) = \frac{20}{\pi}\arg z.$$

In terms of x, y we can also write

$$\theta(z) = \arctan y/x.$$

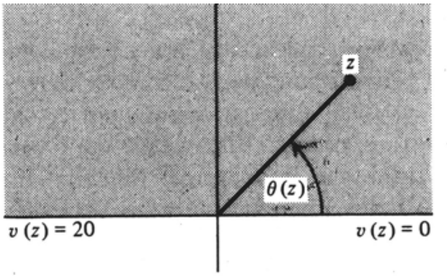


Figure 8

Example. If the temperature on the first quadrant has value 0 on the positive real axis, and 20 on the positive imaginary axis, give a formula for the temperature on the whole first quadrant.

Again we seek a harmonic function having the desired boundary values. We reduce the problem to the preceding example by using an analytic isomorphism between the first quadrant and the upper half plane, namely

$$z \mapsto z^2.$$

Therefore the solution of the problem in the present instance is given by

$$T(z) = \frac{20}{\pi} \arg z^2 = \frac{40}{\pi} \arctan y/x$$

if $z = x + iy$, and z lies in the first quadrant.

Example. We assume that you have worked Exercise 5. Let A be the upper semidisc. We wish to find a harmonic function φ on A which has value 20 on the positive real axis bounding the semidisc, and value 0 on the negative real axis bounding the semidisc. Furthermore, we ask that $\partial\varphi/\partial n = 0$ on the semicircle.

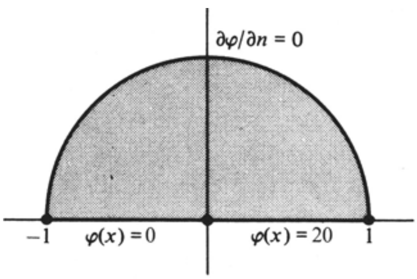


Figure 9

The analytic function $\log z$ maps the semidisc on the horizontal strip as shown on Fig. 10. We may therefore solve the problem with some function v on the strip, and take $\varphi(z) = v(\log z)$ as the solution on the semidisc.

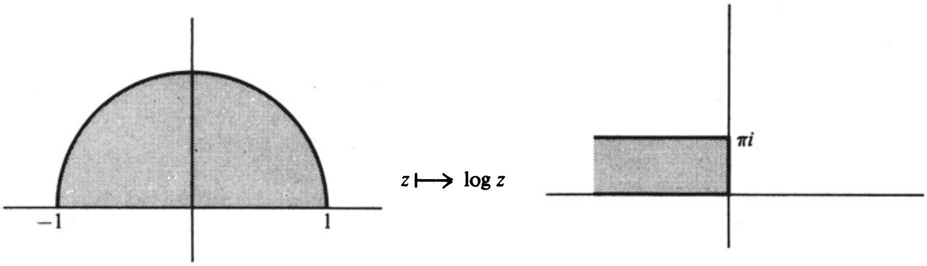


Figure 10

The semicircle is mapped on the vertical segment bounding the strip, and the condition $\partial\varphi/\partial n = 0$ on this segment means that if we view v as a function of two variables (x, y) then $\partial v/\partial x = 0$. Thus v is a function of y alone, and it must have the value 20 on the negative real axis, value 0 at any point $x + \pi i$. Such a function is

$$v(x, y) = 20 - \frac{20}{\pi} y.$$

Consequently,

$$\varphi(z) = 20 - \frac{20}{\pi} \theta(z),$$

where $0 \leqq \theta(z) \leqq \pi$.

Remark. The condition $\partial\varphi/\partial n = 0$ along a curve is usually interpreted physically as meaning that the curve is insulated, if the harmonic function is interpreted as temperature.

VIII, §2. EXERCISES

1. Find a harmonic function on the upper half plane with value 1 on the positive real axis and value -1 on the negative real axis.

2. Find a harmonic function on the indicated region, with the boundary values as shown.

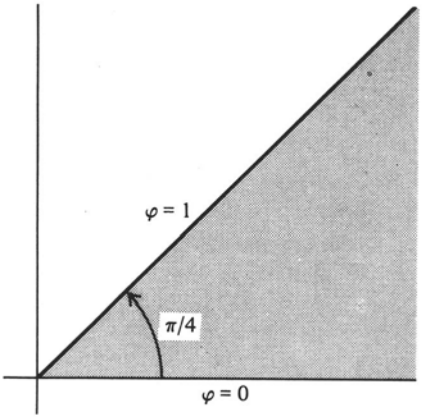


Figure 11

3. Find the temperature on a semicircular plate of radius 1, as shown on the figure, with the boundary values as shown. Value 0 on the semicircle, value 1 on one segment, value 0 on the other segment.

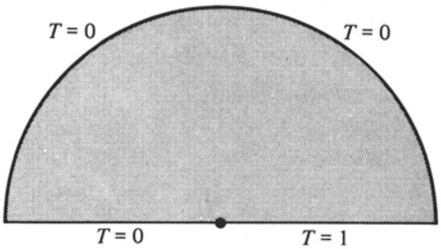


Figure 12

4. Find a harmonic function on the unit disc which has the boundary value 0 on the lower semicircle and boundary value 1 on the upper semicircle.

In the next exercise, recall that a function $\varphi: U \to \mathbf{R}$ is said to be of class C^1 if its partial derivatives $D_1\varphi$ and $D_2\varphi$ exist and are continuous. Let V be another open set. A mapping

$$f: V \to \mathbf{R}^2$$

where $f(x, y) = (u(x, y), v(x, y))$ is said to be of class C^1 if the two coordinate functions u, v are of class C^1.

If $\eta: [a, b] \to V$ is a curve in V, then we may form the composite curve $f \circ \eta$ such that

$$(f \circ \eta)(t) = f(\eta(t)).$$

Then $\gamma = f \circ \eta$ is a curve in U. Its coordinates are

$$f(\eta(t)) = (u(\eta(t)), v(\eta(t))).$$

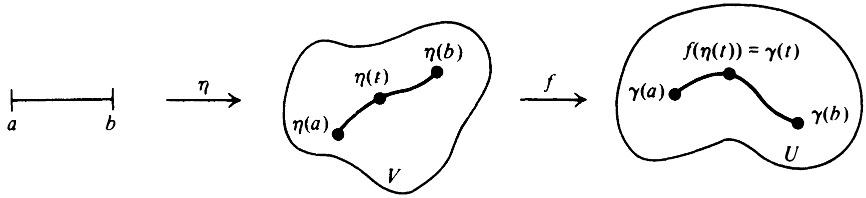


Figure 13

5. Let $\gamma : [a, b] \to U \subset \mathbf{R}^2$ be a smooth curve. Let

$$\gamma(t) = (\gamma_1(t), \gamma_2(t))$$

be the expression of γ in terms of its coordinates. The **tangent vector** is given by the derivative $\gamma'(t) = (\gamma_1'(t), \gamma_2'(t))$. We define

$$N_\gamma(t) = N(t) = (\gamma_2'(t), -\gamma_1'(t))$$

to be the **normal vector**. We shall write vectors vertically. The **normal derivative** of a smooth function φ on U along the curve γ is by definition

$$D_{N_\gamma \varphi} = (\text{grad } \varphi) \cdot N_\gamma = \varphi'(\gamma) N_\gamma,$$

using the derivative $\varphi'(\gamma)$, which at a value of t is a linear map

$$\varphi'(\gamma(t)) : \mathbf{R}^2 \to \mathbf{R}^2.$$

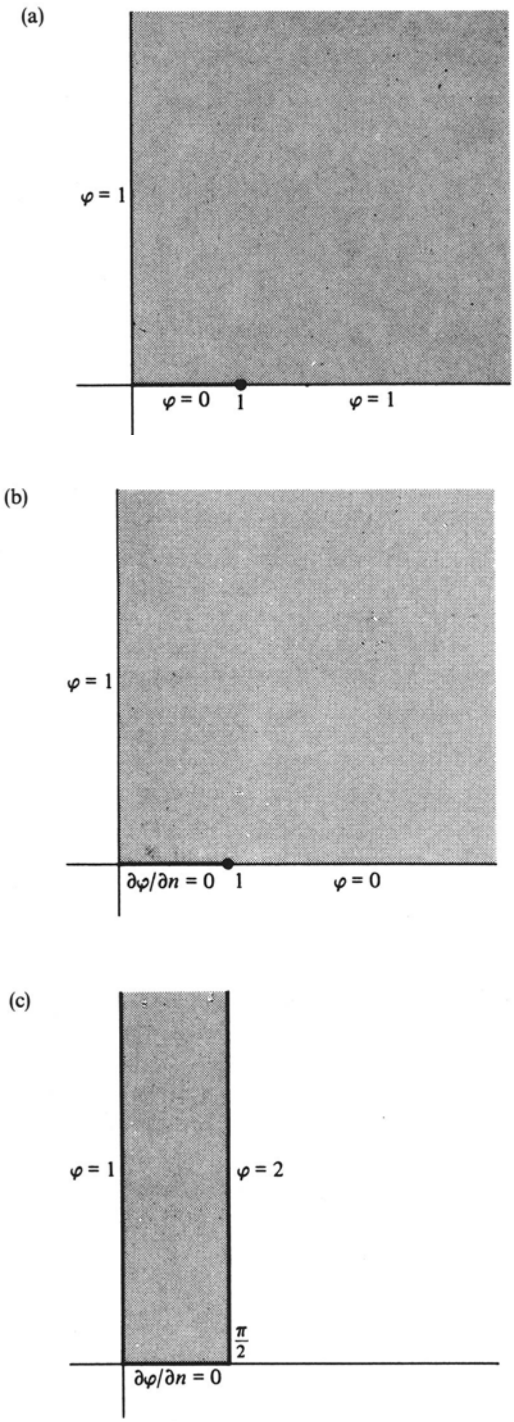
(a) Prove that the condition $D_{N_\gamma} \varphi = 0$ remains true under a change of parametrization of the interval of definition of γ.

(b) Let $f : V \to U$ be analytic, $\eta : [a, b] \to V$ and $\gamma = f \circ \eta$ as above. Show that the Cauchy–Riemann equations imply

$$N_{f \circ \eta} = f'(\eta) N_\eta = \begin{pmatrix} (\partial_1 f_1)(\eta) & (\partial_2 f_1)(\eta) \\ (\partial_1 f_2)(\eta) & (\partial_2 f_2)(\eta) \end{pmatrix} \begin{pmatrix} \eta_2' \\ -\eta_1' \end{pmatrix}.$$

Check the chain rule $D_{N(\eta)}(\varphi \circ f) = \varphi'(f \circ \eta) f'(\eta) N_\eta$. Conclude that if $D_{N(f \circ \eta)} \varphi = 0$, then $D_{N(\eta)} \varphi = 0$.

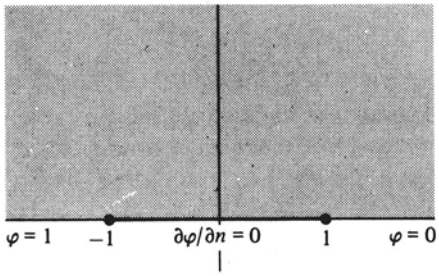
6. Find a harmonic function φ on the indicated regions, with the indicated boundary values. (Recall what sin z does to a vertical strip.)



(a)
$\varphi = 1$
$\varphi = 0$
1
$\varphi = 1$
(b)
$\varphi = 1$
$\partial\varphi/\partial n = 0$
1
$\varphi = 0$
(c)
$\varphi = 1$
$\varphi = 2$
$\frac{\pi}{2}$
$\partial\varphi/\partial n = 0$

Figure 14

(d)



(e)

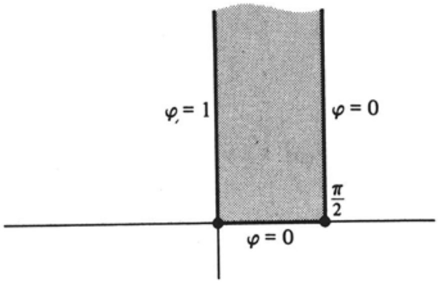


Figure 14 (*continued*)

VIII, §3. BASIC PROPERTIES OF HARMONIC FUNCTIONS

In §1 we saw that the real part of an analytic function is harmonic. Here we prove the converse. Recall that by definition, simply connected implies connected.

Theorem 3.1. *Let U be a simply connected open set. Let u be harmonic on U. Then there exists an analytic function f on U such that $u = \operatorname{Re} f$. The difference of two such functions is a pure imaginary constant.*

Proof. Let

$$h = 2\frac{\partial u}{\partial z} = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}.$$

Then h has continuous partials of first order. Furthermore h is analytic, because

$$\frac{\partial h}{\partial \bar{z}} = 2\frac{\partial}{\partial \bar{z}}\frac{\partial}{\partial z}u = \tfrac{1}{2}\Delta u = 0.$$

Since U is assumed simply connected, by Theorem 6.1 of Chapter III, h has a primitive f on U, so $f'(z) = h(z)$ for all $z \in U$. Let $u_1 = \operatorname{Re}(f)$ be the real part of f. Then

$$h(z) = f'(z) = 2\frac{\partial u_1}{\partial z} = \frac{\partial u_1}{\partial x} - i\frac{\partial u_1}{\partial y},$$

so u and u_1 have the same partial derivatives. It follows that there is a constant C such that $u = u_1 + C$. [*Proof*: Let z_0 be any point of U, and let $\gamma: [a, b] \to U$ be a curve joining z_0 with a point z in U. Let $g = u - u_1$. Then the partial derivatives of g are 0. By the chain rule we have

$$\frac{d}{dt}g(\gamma(t)) = \frac{\partial g}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial t} = 0.$$

Hence $g(\gamma(t))$ is constant, so $g(z_0) = g(z)$. This is true for all points $z \in U$, whence g is constant, as desired.] Subtracting the constant C from f yields the desired analytic function having the given real part u, and proves the existence.

Uniqueness is based on the following more general lemma.

Lemma 3.2. *Let f, g be analytic functions on a connected open set U. Suppose f, g have the same real part. Then $f = g + iK$ for some real constant K.*

Proof. Considering $f - g$, it suffices to prove that if the real part of an analytic function f on U is 0 then the function is a pure imaginary constant. But this is immediate from the open mapping Theorem 6.2 of Chapter II, because f cannot map an open set on a straight line, hence f is constant because U is connected by Theorem 1.2 of Chapter III. This constant is pure imaginary since the real part of f is 0. This proves the lemma, and also the theorem.

Remark. The condition that U is simply connected in Theorem 3.1 is needed in general; it is not sufficient that U is connected. Indeed, $\log r$ is a harmonic function on the open set obtained by deleting the origin from C, but there is no analytic function on that open set whose real part is $\log r$.

Remark. We consider composites of functions. In Exercise 3 of §1, you already considered a composite of a harmonic function with a holomorphic function. It was then assumed that the harmonic function was the real part of an analytic function, but now Theorem 3.1 tells us that this is always the case locally, that is in a neighborhood of a point. We then have the following properties, which are immediate, but which we record for possible future use.

Let $f\colon U \to V$ be an analytic function, and let g be a harmonic function on V. Then $g \circ f$ is harmonic on U.

Let $f\colon U \to V$ be analytic. Let $\bar{V}$ be the complex conjugate of V, that is the set of all points $\bar{z}$ with $z \in V$. Let g be harmonic on $\bar{V}$. Then the function

$$z \mapsto g\left(\overline{f(z)}\right)$$

is harmonic.

Prove these two statements as simple exercises.

Theorem 3.3 (Mean Value Theorem). *Let u be a harmonic function on an open set U. Let $z_0 \in U$, and let $r > 0$ be a number such that the closed disc of radius r centered at z_0 is contained in U. Then*

$$u(z_0) = \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{i\theta})\, d\theta.$$

Proof. There is a number $r_1 > r$ such that the disc of radius r_1 centered at z_0 is contained in U. Any $r_1 > r$ and close to r will do. By Theorem 3.1, there is an analytic function f on the disc of radius r_1 such that $u = \operatorname{Re} f$. By Cauchy's theorem,

$$f(z_0) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{\zeta - z_0}\, d\zeta,$$

where C is the circle of radius r centered at z_0.

We parametrize the circle by $\zeta = z_0 + re^{i\theta}$, so $d\zeta = ire^{i\theta}\, d\theta$. The integral then gives

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{i\theta})\, d\theta.$$

If we write $f = u + iv$, then the desired relation falls out for the real part u. This concludes the proof.

Theorem 3.4.

(a) *Let u be harmonic on a connected open set U. Suppose that u has a maximum at a point z_0 in U. Then u is constant.*

(b) *Let U be a connected open set and let $\bar{U}$ be its closure. Let u be a continuous function on $\bar{U}$, harmonic on U. If u is not constant on U, then a maximum of u on $\bar{U}$ occurs on the boundary of U in $\bar{U}$.*

Proof. First we prove that u is constant in a disc centered at z_0. By Theorem 3.1 there is an analytic function f on such a disc such that

Re $f = u$. Then $e^{f(z)}$ is analytic, and

$$|e^{f(z)}| = e^{u(z)}.$$

Since the exponential function is strictly increasing, it follows that a maximum for u is also a maximum for e^u, and hence also a maximum of $|e^f|$. By the maximum modulus principle for analytic functions, it follows that e^f is constant on the disc. Then e^u is constant, and finally u is constant, thus proving the theorem locally.

We now extend the theorem to an arbitrary connected open set. Let S be the set of points z in U such that u is constant in a neighborhood of z with value $u(z_0)$. Then S contains z_0, and S is open. By Theorem 1.6 of Chapter III, it will suffice to prove that S is closed in U. So let z_1 be a point in the closure of S, and z_1 contained in U. Since u is continuous, it follows that $u(z_1) = u(z_0)$ because points of S can be found arbitrarily close to z_1. Then u also has a maximum at z_1, and by the first part of the proof is constant in a neighborhood of z_1, which proves that $z_1 \in S$, and concludes the proof of part (a). Part (b) is an immediate consequence, thus concluding the proof of the theorem.

As an application of the maximum modulus theorem for harmonic functions, we obtain a convexity property of the maximum modulus of an analytic function.

Theorem 3.5 (Hadamard Three-Circle Theorem). *Let f be holomorphic on a closed annulus $0 < r_1 < |z| < r_2$. Let*

$$s = \frac{\log r_1 - \log r}{\log r_2 - \log r_1}.$$

Let $M(r) = M_f(r) = \|f\|_r = \max|f(z)|$ for $|z| = r$. Then

$$\log M(r) \leqq (1 - s)\log M(r_1) + s \log M(r_2).$$

Proof. Let α be a real number. The function $\alpha \log|z| + \log|f(z)|$ is harmonic outside the zeros of f. Near the zeros of f the above function has values which are large negative. Hence by the maximum modulus principle this function has its maximum on the boundary of the annulus, specifically on the two circles $|z| = r_1$ and $|z| = r_2$. Therefore

$$\alpha \log|z| + \log|f(z)| \leqq \max(\alpha \log r_1 + \log M(r_1), \alpha \log r_2 + \log M(r_2))$$

for all z in the annulus. In particular, we get the inequality

$$\alpha \log r + \log M(r) \leqq \max(\alpha \log r_1 + \log M(r_1), \alpha \log r_2 + \log M(r_2)).$$

Now let α be such that the two values inside the parentheses on the right are equal, that is

$$\alpha = \frac{\log M(r_2) - \log M(r_1)}{\log r_1 - \log r_2}.$$

Then from the previous inequality, we get

$$\log M(r) \leqq \alpha \log r_1 + \log M(r_1) - \alpha \log r,$$

which upon substituting the value for α gives the result stated in the theorem.

Remark. Another proof of the Hadamard theorem will be given later as a consequence of the Phragmen–Lindelöf theorem, giving another point of view and other convexity theorems.

We shall now give further properties of harmonic functions on an annulus. The first result is a mean value theorem which is actually an immediate consequence of Theorem 3.7, but we want to illustrate a method of proof giving a nice application of Fourier series and techniques of partial differential operators.

Theorem 3.6. *Let u be harmonic in an annulus $0 < r_1 < r < r_2$. Then there exist constants a, b such that*

$$\int_0^{2\pi} u(r, \theta)\frac{d\theta}{2\pi} = a \log r + b.$$

Proof. We shall use elementary properties of Fourier series with which readers are likely to be acquainted. For each integer n let

$$u_n(r) = \int_0^{2\pi} u(r, \theta)e^{-in\theta}\frac{d\theta}{2\pi}.$$

Thus $u_n(r)$ is the n-th Fourier coefficient of the function $\theta \mapsto u(r, \theta)$ for a fixed value of r. Since u is infinitely differentiable, one can differentiate under the integral sign, so u_n is infinitely differentiable. We let

$$u(r, \theta) = \sum_{-\infty}^{\infty} u_n(r)e^{in\theta}$$

be the Fourier series. Because u is C^∞, one can differentiate the Fourier series term by term, and the Fourier coefficients are uniquely determined.

Recall that the Laplace operator is given in polar coordinates for $r \neq 0$

by the formula

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}.$$

Differentiating the Fourier series with this operator, we note that the 0-th Fourier coefficient $u_0(r)$ must satisfy the partial differential equation

$$\Delta u_0 = 0,$$

or multiplying by r, must satisfy the differential equation

$$ru_0''(r) + u_0'(r) = 0.$$

This amounts to

$$u_0''/u_0' = -1/r,$$

which can be integrated to give

$$\log u_0' = -\log r + c_1 = \log(1/r) + c_1$$

with some constant c_1. Therefore there is some constant a such that

$$u_0' = a/r,$$

whence $u_0(r) = a \log r + b$ for some constant b. But integrating the Fourier series term by term, we obtain

$$\int_0^{2\pi} e^{in\theta}\, d\theta = 0 \qquad \text{if} \quad n \neq 0.$$

The integral of $u_0(r)$ with respect to θ gives $u_0(r)$, which concludes the proof of the theorem.

Next we consider harmonic functions on an annulus, which includes the special case of the punctured plane $\mathbf{C}^* = \mathbf{C} - \{0\}$ and the punctured disc $D^* = D - \{0\}$.

Theorem 3.7. *Let U be an annulus $0 \leqq r_1 < |z| < r_2$ (with r_2 possibly equal to ∞). Let u be harmonic on U. Then there exists a real constant a and an analytic function g on U such that*

$$u - a \log r = \mathrm{Re}(g).$$

Proof. We consider the half annuli as illustrated on Figure 15.

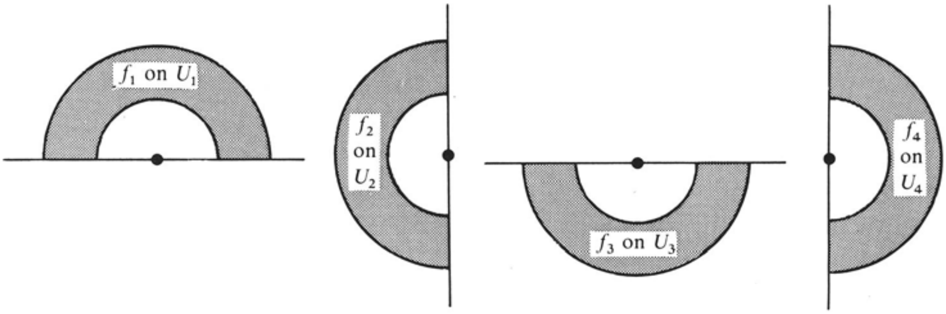


Figure 15

By Theorem 3.1, there exists an analytic function f_1 on the upper half annulus such that $\mathrm{Re}(f_1) = u$, and there exists an analytic function f_2 on the left half annulus U_2 such that $\mathrm{Re}(f_2) = u$. Then f_1 and f_2 have the same real part on the upper left quarter annulus, so $f_1 - f_2$ is constant on this quarter annulus by Lemma 3.2. After subtracting a constant from f_2 we may assume that $f_1 = f_2$ on the upper left quarter annulus. Similarly, we obtain f_3 on the lower half annulus such that $f_3 = f_2$ on the lower left quarter annulus and $\mathrm{Re}(f_3) = u$; and we obtain f_4 on the right half annulus with $\mathrm{Re}(f_4) = u$ and $f_4 = f_3$ on the lower half annulus. But in the final step, there is a constant Ki that we cannot get rid of such that

$$f_4 = f_1 + Ki \quad \text{on the upper quarter annulus.}$$

The condition $\mathrm{Re}(f_4) = u$ implies that K is real.

Now consider the functions

$$f_j(z) - \frac{K}{2\pi}\log_j(z) \quad (j = 1, \dots, 4) \text{ on the } j\text{-th half disc,}$$

where $\log_j(z)$ is given by

$$\log_j(z) = \log|z| + i\theta_j(z)$$

on the j-th half annulus and the angle θ_j on the j-th half annulus satisfies

$$0 < \theta_1 < \pi, \qquad \frac{\pi}{2} < \theta_2 < \frac{3\pi}{2}, \qquad \pi < \theta_3 < 2\pi, \qquad \frac{3\pi}{2} < \theta_4 < 2\pi + \frac{\pi}{2}.$$

Then $\log_4(z) = \log_1(z) + 2\pi i$ on the upper fourth quadrant. Hence

$$f_4(z) - \frac{K}{2\pi}\log_4(z) = f_1 - \frac{K}{2\pi}\log_1(z).$$

Therefore there is an analytic function g on u such that

$$g = f_j - (K/2\pi) \log_j$$

on the j-th half annulus. This concludes the proof of the theorem.

Remark 1. In the terminology of Chapter XI, we may say that the functions f_1, f_2, f_3, f_4 are **analytic continuations** of each other, and similarly for $f_j - (K/2\pi) \log_j$ for $j = 1, \ldots, 4$.

Remark 2. The result of Theorem 3.7 is similar to a result in real analysis, which we recall. On $\mathbf{R}^2$ from which the origin is deleted, we have a vector field

$$G(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}\right),$$

or in terms of differentials,

$$\frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy = d\theta$$

as the reader will verify by the chain rule. Now let F be an arbitrary C^1 vector field on $\mathbf{R}^2$ minus the origin, $F = (f, g)$ where f, g are C^1 functions. Assume that $\partial f/\partial y = \partial g/\partial x$ (the so-called integrability conditions). Let γ be the circle of radius 1 centered at the origin, and let

$$K = \int_\gamma F \cdot d\gamma = \int_0^{2\pi} F(\gamma(\theta)) \cdot \gamma'(\theta)\, d\theta.$$

Then there exists a function φ on $\mathbf{R}^2$ minus the origin such that

$$F = \operatorname{grad} \varphi + \frac{K}{2\pi} G.$$

Cf. Chapter 15, §4, Exercise 13, in my *Undergraduate Analysis* (Springer-Verlag, 1983).

Theorem 3.7 is of interest for its own sake: because it gives a first illustration of analytic continuation to whose general properties will be discussed later; and also because we shall use it in the next result, which proves for harmonic functions the analogue of the theorem concerning removable singularities for analytic functions.

Theorem 3.8. *Let u be a harmonic function on the punctured disc D^*. Assume that u is bounded. Then u extends to a harmonic function on D.*

Proof. By Theorem 3.7 there exists an analytic function g on D^* such that $u(z) = \mathrm{Re}(g(z)) + a \log r$. We then have the Laurent expansion

$$g(z) = \sum_{-\infty}^{\infty} a_n z^n,$$

and it suffices to show that there are no negative terms, i.e. $a_n = 0$ for n negative, for then g is bounded in a neighborhood of 0, so $a = 0$ because u is bounded, and we are done. Suppose there are only a finite number of negative terms, so g has a pole or order $N \geqq 1$ at the origin. Thus

$$g(z) = \frac{a_{-N}}{z^N} + \cdots + \frac{a_{-1}}{z} + \text{higher terms}$$

and $a_{-N} \neq 0$. Consider values of z such that $|z| = r$ approaches 0 and also such that $a_{-N}z^{-N}$ is real positive. Then

$$|\mathrm{Re}(g(z))| \geqq \frac{a_{-N}}{z^N} + O\left(\frac{1}{|z|^{N-1}}\right),$$

and so $u = \mathrm{Re}(g(z)) + a \log|z|$ cannot be bounded for such z. Hence g has an essential singularity at the origin. By the Casorati–Weierstrass theorem, g takes on values arbitrarily close to any given complex number on any disc of small radius. However, the equation

$$\mathrm{Re}(g(z)) = u(z) - a \log r$$

shows that the set of values will miss the points in some region with real part large negative or large positive depending on the sign of a, and even if $a = 0$. Hence we have shown that the Laurent expansion has only terms with $n \geqq 0$, thereby concluding the proof of the theorem.

Although we wanted to illustrate ideas of analytic continuation in the proof of Theorem 3.7, readers should also be aware of the following extension and a more powerful argument used to prove both results, as follows.

Theorem 3.9. *Let U be a simply connected open set in $\mathbf{C}$. Let $z_1, \ldots, z_n$ be distinct points in U, and let $U^* = U - \{z_1, \ldots, z_n\}$ be the open set obtained by deleting these points. Let u be a real harmonic function on U^*. Then there exist constants $a_1, \ldots, a_n$ and an analytic function f on*

U^* such that for all $z \in U^*$ we have

$$u(z) - \sum a_k \log|z - z_k| = \operatorname{Re} f(z).$$

Proof. On each open disc W contained in U^*, u is the real part of an analytic function A_W, uniquely determined up to an additive constant. We want to determine the obstruction for u to be the real part of an analytic function on U^*, and more precisely we want to show that there exist real constants a_k and an analytic function f on U^* such that

$$u(z) - \sum a_k \log|z - z_k| = \operatorname{Re} f(z).$$

We consider the functions A_W as above. For each W, the derivative A'_W is uniquely determined, and the collection of such functions $\{A'_W\}$ defines an analytic function A' on U^*. Let γ_k be a small circle around z_k, and let

$$a_k = \frac{1}{2\pi i}\int_{\gamma_k} A'(z)\, dz.$$

Let z_0 be a point in U^* and let γ_z be a piecewise C^1 path in U^* from z_0 to a point z in U^*. Let

$$g(z) = A'(z) - \sum_k a_k \frac{1}{z - z_k} \qquad \text{and} \qquad f(z) = \int_{\gamma_z} g(\zeta)\, d\zeta + u(z_0).$$

We claim that this last integral is independent of the path γ_z, and gives the desired function. To show independence of the path, it suffices to show that for any closed path γ in U^* we have $\int_\gamma g(\zeta)\, d\zeta = 0$. By Theorem 2.4 of Chapter IV, there are integers m_k such that $\gamma \sim \sum m_j \gamma_j$. Then

$$\int_\gamma g(\zeta)\, d\zeta = \sum_j m_j \int_{\gamma_j} g(\zeta)\, d\zeta = \sum_j m_j \int_{\gamma_j}\left[A'(\zeta) - \sum_k \frac{a_k}{\zeta - z_k}\right] d\zeta = 0.$$

This proves that the definition of $f(z)$ is independent of the path γ_z, so f is an analytic function on U^*.

Next we claim that a_k is real. Writing $dz = dx + i\, dy$ and

$$A'(z) = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}$$

we get

$$\int A'(z)\, dz = \int\left(\frac{\partial u}{\partial x}\, dx + \frac{\partial u}{\partial y}\, dy\right) + i\int\left(\frac{\partial u}{\partial x}\, dy - \frac{\partial u}{\partial y}\, dx\right).$$

The first integral over the circle is 0 because it is $\int du$ over the circle, i.e. it is the integral of grad u, which has a potential function u. The second integral is pure imaginary, and dividing by $2\pi i$ yields a real number, as desired.

Finally we want to prove that $\operatorname{Re} f(z) = u(z) - \sum a_k \log|z - z_k|$. Let W be a disc in U^* containing z_0. We prove the above relation first for $z \in W$. We can take the path γ_z to be contained in W. We then write the integral defining $f(z)$ as a sum of two integrals, to get

$$\begin{aligned}\operatorname{Re} f(z) &= \operatorname{Re} \int_{\gamma_z} A'_W(\zeta)\, d\zeta + u(z_0) - \sum_k \operatorname{Re} a_k \log(z - z_k) \\ &= u(z) - \sum a_k \log|z - z_k|.\end{aligned}$$

Finally we use the following lemma to conclude the proof.

Lemma 3.10. *Let U be a connected open set. Let u be harmonic on U, and f analytic on U. If $u = \operatorname{Re}(f)$ on some open disc contained in U, then $u = \operatorname{Re}(f)$ on U.*

Proof. Let V be the union of all open subsets of U where $u = \operatorname{Re}(f)$. Then V is not empty and is open. We need only show that V is closed in U. Let $\{z_n\}$ be a sequence of points in V converging to a point $w \in U$. Let D be a small open disc centered at w and contained in U, such that $u = \operatorname{Re}(g)$ for some analytic function g on D. Then D contains infinitely many z_n, and f, g have the same real part in a neighborhood of such z_n. Hence $f - g$ is a pure imaginary constant on such a neighborhood, whence on D. Therefore $u = \operatorname{Re}(f)$ on D, so $D \subset V$, $w \in V$, and V is closed in U, which proves the lemma.

Remark 1. We have encountered two similar situations with applications of Theorem 2.4 of Chapter IV, namely holomorphic functions and harmonic functions. In Appendix 6, we shall meet a third similar situation with locally integrable vector fields.

Remark 2. One could delete disjoint discs from U instead of points. The conclusion and proof of the theorem are valid. With this more general assumption, the result then includes Theorem 3.7 as a special case.

VIII, §3. EXERCISES

1. The **Gauss theorem** (a variation of Green's theorem) can be stated as follows. *Let γ be a closed piecewise C^1 curve in an open set U, and suppose γ has an interior contained in U. Let F be a C^1 vector field on U. Let n be the unit*

normal vector on γ. Then

$$\int_\gamma F \cdot n = \iint_{\mathrm{Int}(\gamma)} (\operatorname{div} F)\, dy\, dx.$$

Using the Gauss theorem, prove the following. Let u be a C^2 function on U, harmonic on the interior $\mathrm{Int}(\gamma)$. Then

$$\int_\gamma D_n u = 0.$$

Here $D_n u$ is the normal derivative $(\operatorname{grad} u) \cdot \mathbf{n}$, as in Exercise 5 of §2.

Subharmonic Functions

Define a real function φ to be **subharmonic** if φ is of class C^2 (i.e. has continuous partial derivatives up to order 2) and

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} \geqq 0.$$

The next exercise gives examples of subharmonic functions.

2. (a) Let u be real harmonic. Show that u^2 is subharmonic.
   (b) Let u be real harmonic, $u = u(x, y)$. Show that

   $$(\operatorname{grad} u)^2 = (\operatorname{grad} u) \cdot (\operatorname{grad} u)$$

   is subharmonic.
   (c) Show that the function $u(x, y) = x^2 + y^2 - 1$ is subharmonic.
   (d) Let u_1, u_2 be subharmonic, and c_1, c_2 positive numbers. Show that $c_1 u_1 + c_2 u_2$ is subharmonic.

3. Let φ be subharmonic on an open set containing a closed disc of radius r_1 centered at a point a. For $r < r_1$ let

   $$h(r) = \int_0^{2\pi} \varphi(a + re^{i\theta}) \frac{d\theta}{2\pi}.$$

   Show that $h(r)$ is increasing as a function of r. [*Hint*: Let $u(r, \theta) = \varphi(a + re^{i\theta})$. Then

   $$r \frac{d}{dr}\big(rh'(r)\big) = \int_0^{2\pi} r \frac{\partial}{\partial r}\left(r \frac{\partial u}{\partial r}\right) \frac{d\theta}{2\pi}.$$

   Use the expression for Δ in polar coordinates, and the fact that the integral of $\partial^2 u/\partial \theta^2$ is 0 to show that $rh'(r)$ is weakly increasing. Since $rh'(r) = 0$ for $r = 0$, it follows that $rh'(r) \geqq 0$, so $h'(r) \geqq 0$.]

4. Using Exercise 3, or any other way, prove the inequality

$$\varphi(a) \leqq \int_0^{2\pi} \varphi(a + re^{i\theta}) \frac{d\theta}{2\pi} \qquad \text{for every } r.$$

5. Suppose that φ is defined on an open set U and is subharmonic on U. Prove the maximum principle, that no point $a \in U$ can be a strict maximum for φ, i.e. that for every disc of radius r centered at a with r sufficiently small, we have

$$\varphi(a) \leqq \max \varphi(z) \qquad \text{for} \quad |z - a| = r.$$

6. Let φ be subharmonic on an open set U. Assume that the closure $\bar{U}$ is compact, and that φ extends to a continuous function on $\bar{U}$. Show that a maximum for φ occurs on the boundary.

7. Let U be a bounded open set. Let u, v be continuous functions on $\bar{U}$ such that u is harmonic on U, v is subharmonic on U, and $u = v$ on the boundary of U. Show that $v \leqq u$ on U. Thus a subharmonic function lies below the harmonic function having the same boundary value, whence its name.

Remarks. We gave a definition of subharmonic functions which would exhibit a number of properties rapidly, emphasizing the effect of the Laplace operator. Actually, in some of the most important applications, our definition is too strong, and one defines a function u to be subharmonic if it is upper semicontinuous, allowing $-\infty$ as a value (with $-\infty < c$ for all real numbers c), and if u satisfies the maximum principle locally in the neighborhood of every point. For a wide class of connected open sets U, not necessarily simply connected, one proves the existence of a harmonic function on U having given boundary value (satisfying suitable integrability conditions) by taking the sup of the subharmonic functions having this boundary value. Taking the sup of two functions does not preserve differentiability but it preserves continuity at a point, so just for that reason (among others), one has more flexibility in dealing with subharmonic functions rather than harmonic functions. For a systematic treatment using this approach, see [Ah 66] and [Fi 83], especially Chapter 1, §3 and §4.

The exercises on subharmonic functions will also be found worked out in [StW 71].

VIII, §4. THE POISSON FORMULA

In this section, we express an analytic function in terms of its real part. We first prove the formula in a rather ad hoc way. After Theorem 4.2, we shall make some comments explaining more about the structure of the formula. Some readers might prefer to look at these comments first, or simultaneously. We start with an identity for $z = x + iy = re^{i\varphi}$:

$$\text{(1)} \qquad \operatorname{Re} \frac{1+z}{1-z} = \frac{1-r^2}{1-2x+r^2} = \frac{1-r^2}{1-2r\cos\varphi + r^2}.$$

Define the **Poisson kernel** $P_{R,r}$ for $0 \leqq r < R$ by

$$P_{R,r}(\theta) = \frac{1}{2\pi} \frac{R^2 - r^2}{R^2 - 2Rr\cos\theta + r^2}.$$

This is a real periodic function of θ. From (1) one shows at once that

$$P_{R,r}(\theta) = \frac{1}{2\pi} \operatorname{Re} \frac{Re^{i\theta} + r}{Re^{i\theta} - r},$$

$$\text{(2)} \qquad P_{R,r}(\varphi - \theta) = \operatorname{Re} \frac{Re^{i\theta} + re^{i\varphi}}{Re^{i\theta} - re^{i\varphi}} = \frac{R^2 - r^2}{R^2 - 2rR\cos(\theta - \varphi) + r^2}.$$

Theorem 4.1. *Let f be holomorphic on the closed disc $\bar{D}_R$. Let $z \in D_R$. Then*

$$f(z) = \int_0^{2\pi} f(Re^{i\theta}) \operatorname{Re} \frac{Re^{i\theta} + z}{Re^{i\theta} - z} \frac{d\theta}{2\pi}.$$

Proof. Write $z = re^{i\varphi}$. Let C_R denote the circle of radius R, parametrized by $\zeta = Re^{i\theta}$, $d\zeta = i\,Re^{i\theta}\,d\theta$. Then by Cauchy's theorem,

$$f(z) = \frac{1}{2\pi i} \int_{C_R} \frac{f(\zeta)}{\zeta - z} d\zeta = \int_0^{2\pi} f(Re^{i\theta}) \frac{Re^{i\theta}}{Re^{i\theta} - re^{i\varphi}} \frac{d\theta}{2\pi}.$$

On the other hand, let $w = R^2/\bar{z} = (R^2/r)e^{i\varphi}$. Then $\zeta \mapsto f(\zeta)/(\zeta - w)$ is holomorphic on $\bar{D}_R$ so

$$0 = \frac{1}{2\pi i} \int_{C_R} \frac{f(\zeta)}{\zeta - w} d\zeta = \int_0^{2\pi} f(Re^{i\theta}) \frac{re^{i\theta}}{re^{i\theta} - Re^{i\varphi}} \frac{d\theta}{2\pi}.$$

Subtract and collect terms. The desired identity comes directly from (2).

Let us decompose f into its real and imaginary part, so

$$f = u + iv \qquad \text{with} \quad u = \operatorname{Re} f.$$

Then the integral expression of the theorem gives separately an integral expression for both u and v. For the real part, we thus get

$$u(z) = \int_0^{2\pi} u(Re^{i\theta}) \operatorname{Re} \frac{Re^{i\theta} + z}{Re^{i\theta} - z} \frac{d\theta}{2\pi}.$$

In particular, for f itself, we get:

Theorem 4.2. *Let f be holomorphic on the closed disc $\bar{D}_R$, then there is a real constant K such that for all $z \in D_R$ we have*

$$f(z) = \int_0^{2\pi} \operatorname{Re} f(Re^{i\theta}) \frac{Re^{i\theta} + z}{Re^{i\theta} - z} \frac{d\theta}{2\pi} + iK,$$

Proof. The right-hand side is analytic in z. One can see this either by applying Theorem 7.7 of Chapter III, or by differentiating under the integral sign using Theorem A3 of §6, which justifies such differentiation. By Theorem 4.1, the right-hand side and the left-hand side, namely f, have the same real part. As we saw in Lemma 3.2, this implies that the right-hand side and the left-hand side differ by a pure imaginary constant, as was to be shown.

Remark. In Theorem 3.1 we proved the existence of an analytic function having a given real part, say on a disc. Theorem 4.2 gives an explicit expression for this analytic function, useful to estimate both the function and its real part. Applications will be given in exercises.

The Poisson Integral as a Convolution

We shall now comment on the expression which occurs in the integral of Theorem 4.1, the Poisson kernel.

Given two periodic functions g, h one defines their **convolution** $g * h$ by the integral

$$(g * h)(\varphi) = \int_0^{2\pi} g(\theta)h(\varphi - \theta)\, d\theta = \int_0^{2\pi} g(\varphi - \theta)h(\theta)\, d\theta.$$

If g is an even function, that is $g(-\theta) = g(\theta)$ for all θ, then we can write this integral in the form

$$(g * h)(\varphi) = \int_0^{2\pi} g(\theta - \varphi)h(\theta)\, d\theta.$$

Then the integral of Theorem 4.1 is seen to be a convolution, namely

$$f(z) = f(re^{i\varphi}) = (P_{R,r} * h)(\varphi),$$

where $h(\theta) = f(Re^{i\theta})$. This type of convolution integral and more of its properties will appear again in the next section.

Remarks. Suppose we are given a continuous function g on the unit circle. We can define its **Cauchy transform**

$$h(z) = h_g(z) = \frac{1}{2\pi i}\int_C \frac{g(\zeta)}{\zeta - z}\, d\zeta,$$

so that h is holomorphic on the unit disc. Cases of all types can occur when h does not extend to a continuous function on C, when h does extend to a continuous function on C but this continuous function is different from g, and when finally this continuous function is g itself. We now give some examples.

Examples. Let $g(\zeta) = \zeta^n$ when n is an integer $\geqq 0$. Then

$$h(z) = z^n$$

by Cauchy's theorem. On the other hand, if n is a positive integer, and $g(\zeta) = \zeta^{-n}$, then from the expansion

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

we see that $h = 0$. More generally, we can form series

$$g(\zeta) = \sum_{-\infty}^{\infty} a_n\zeta^n,$$

subject to suitable convergence to obtain more general functions. If g is an arbitrary continuous function on the circle, with $\zeta = e^{i\theta}$, then g has a Fourier series as above, and the elementary theory of Fourier series implies that $\sum |a_n|^2$ converges. Because of the Schwarz inequality, one can integrate term by term, and the above example with negative powers of ζ shows that under the Cauchy transform, all the negative terms disappear, and the corresponding holomorphic function h is given by the series

$$h(z) = \sum_{n=0}^{\infty} a_n z^n.$$

If negative terms were present in the Fourier expansion of g, then certainly a continuous extension of h to the circle is not equal to g.

On the other hand, if we pick z outside the unit circle, then under the Cauchy transform, the positive terms of the Fourier expansion will vanish rather than the negative terms.

The Cauchy transform actually defines two holomorphic functions: one on the unit disc, and one outside the unit disc. When f is holomorphic on the closed unit disc, Theorem 4.1 and 4.2 show how they patch together to give rise to an integral expression for f in terms of what could be called the **Poisson transform** (the integral in Theorem 4.2), which will be reconsidered from another point of view in the next section.

We now give an example of a continuous function on the unit circle whose Cauchy transform cannot be extended by continuity, say to the point 1. We let

$$g(e^{i\theta}) = \frac{1}{\log(1/\theta)} \qquad \text{for} \quad 0 < \theta \leqq \pi/4$$

We extend this function of θ linearly between $\pi/4$ and π so that at $\theta = \pi$ the function is 0; and we extend the function by periodicity for other values of θ. We leave it to the reader to verify that as $z \to 1$, z real, the Cauchy transform tends to infinity.

VIII, §4. EXERCISES

1. Give another proof for Theorem 4.1 as follows. First by Cauchy's theorem,

$$f(0) = \frac{1}{2\pi i} \int_{C_R} \frac{f(\zeta)}{\zeta} d\zeta.$$

Let g be the automorphism of the disc which interchanges 0 and z. Apply the above formula to the function $f \circ g$ instead of f, and change variables in the integral, with $w = g(\zeta)$, $\zeta = g^{-1}(w)$.

2. Define

$$P_{R,r}(\theta) = \frac{1}{2\pi} \frac{R^2 - r^2}{R^2 - 2Rr\cos\theta + r^2}$$

for $0 \leqq r < R$.

Prove the inequalities

$$\frac{R-r}{R+r} \leqq 2\pi P_{R,r}(\theta-\varphi) \leqq \frac{R+r}{R-r}$$

for $0 \leqq r < R$.

3. Let f be analytic on the closed disc $\bar{D}(\alpha, R)$ and let $u = \operatorname{Re}(f)$. Assume that $u \geqq 0$. Show that for $0 \leqq r < R$ we have

$$\frac{R-r}{R+r}u(\alpha) \leqq u(\alpha + re^{i\theta}) \leqq \frac{R+r}{R-r}u(\alpha).$$

After you have read the next section, you will see that this inequality holds also if $u \geqq 0$ is harmonic on the disc, with a continuous extension to the closed disc $\bar{D}(\alpha, R)$.

4. Let $\{u_n\}$ be a sequence of harmonic functions on the open disc. If it converges uniformly on compact subsets of the disc, then the limit is harmonic.

VIII, §5. CONSTRUCTION OF HARMONIC FUNCTIONS

Let U be a simply connected open set with smooth boundary. By the Riemann mapping theorem, there is an analytic isomorphism of U with the unit disc, extending to a continuous isomorphism at the boundary. To construct a harmonic function on U with prescribed boundary value, it suffices therefore to do so for the disc.

In this case, we use the method of Dirac sequences, or rather Dirac families. We recall what that means. We shall deal with periodic functions of period 2π in the sequel, so we make that assumption from the beginning. By a **Dirac sequence** we shall mean a sequence of functions $\{K_n\}$ of a real variable, periodic of period 2π, real valued, satisfying the following properties.

DIR 1. *We have $K_n(t) \geqq 0$ for all n and all x.*

DIR 2. *Each K_n is continuous, and*

$$\int_0^{2\pi} K_n(t)\,dt = 1.$$

DIR 3. *Given ϵ and δ, there exists N such that if $n \geqq N$, then*

$$\int_{-\pi}^{-\delta} K_n + \int_{\delta}^{\pi} K_n < \epsilon.$$

Condition **DIR 2** means that the area under the curve $y = K_n(t)$ is equal to 1. Condition **DIR 3** means that this area is concentrated near 0 if n is taken sufficiently large. Thus a family $\{K_n\}$ as above looks like Fig. 16.

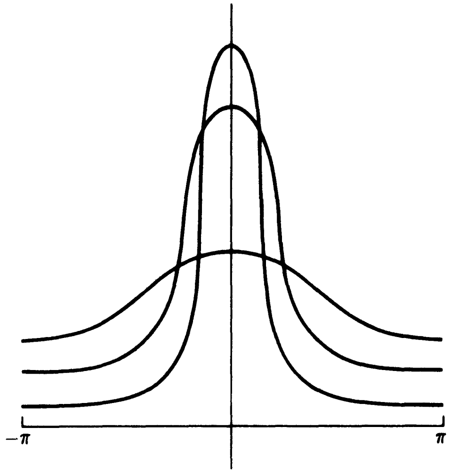


Figure 16

The functions $\{K_n\}$ have a peak near 0. In the applications, it is also true that the functions K_n are even, that is, $K_n(-x) = K_n(x)$, but we won't need this.

If f is any periodic function, we define the **convolution** with K_n to be

$$f_n(x) = K_n * f(x) = \int_{-\pi}^{\pi} f(t)K_n(x - t)\,dt.$$

Theorem 5.1. *Let f be continuous periodic. Then the sequence $\{K_n * f\}$ converges to f uniformly.*

Proof. Changing variables, we have

$$f_n(x) = \int_{-\pi}^{\pi} f(x - t)K_n(t)\,dt.$$

On the other hand, by **DIR 2**,

$$f(x) = f(x)\int_{-\pi}^{\pi} K_n(t)\,dt = \int_{-\pi}^{\pi} f(x)K_n(t)\,dt.$$

Hence

$$f_n(x) - f(x) = \int_{-\pi}^{\pi} [f(x - t) - f(x)]K_n(t)\,dt.$$

By the compactness of the circle, and the uniform continuity of f, we conclude that given ϵ there is δ such that whenever $|t| < \delta$ we have

$$|f(x-t) - f(x)| < \epsilon$$

for all x. Let B be a bound for f. Then we select N such that if $n \geqq N$,

$$\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi} K_n < \frac{\epsilon}{2B}.$$

We have

$$|f_n(x) - f(x)| \leqq \int_{-\pi}^{-\delta} + \int_{-\delta}^{\delta} + \int_{\delta}^{\pi} |f(x-t) - f(x)| K_n(t)\, dt.$$

To estimate the first and third integral, we use the given bound B for f so that $|f(x-t) - f(x)| \leqq 2B$. We obtain

$$\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi} |f(x-t) - f(x)| K_n(t)\, dt \leqq 2B \int_{-\pi}^{-\delta} + \int_{\delta}^{\pi} K_n(t)\, dt < \epsilon.$$

For the integral in the middle, we have the estimate

$$\int_{-\delta}^{\delta} |f(x-t) - f(x)| K_n(t)\, dt \leqq \int_{-\delta}^{\delta} \epsilon K_n \leqq \int_{-\pi}^{\pi} \epsilon K_n \leqq \epsilon.$$

This proves our theorem.

We leave it as an exercise to prove that if K_n is of class C^1, then

$$\frac{d}{dx}(K_n * f)(x) = \left(\frac{dK_n}{dx}\right) * f.$$

This is merely differentiating under the integral sign.

We shall work with polar coordinates r, θ. It is an exercise to see that the Laplace operator can be put in polar coordinates by

$$\boxed{\Delta = \left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}.}$$

It was convenient to formulate the general Dirac property for sequences, but we shall work here with families, indexed by r with $0 < r < 1$ and r

tending to 1, rather than n tending to infinity. We define the **Poisson kernel** as

$$P_r(\theta) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} r^{|k|} e^{ik\theta}.$$

The series is absolutely convergent, and uniformly so, dominated by a geometric series, if r stays away from 1. Simple trigonometric identities show that

$$(*) \qquad P_r(\theta) = \frac{1}{2\pi} \frac{1 - r^2}{1 - 2r\cos\theta + r^2}.$$

The smallest value of the denominator occurs when $\cos\theta = 1$, and we therefore see that

DIR 1. $P_r(\theta) \geqq 0$ *for all* r, θ.

Integrating the series term by term yields

DIR 2.
$$\int_{-\pi}^{\pi} P_r(\theta)\, d\theta = 1.$$

This follows immediately from the values:

$$\int_{-\pi}^{\pi} e^{ik\theta}\, d\theta = \begin{cases} 0 & \text{if } k \neq 0, \\ 2\pi & \text{if } k = 0. \end{cases}$$

Finally, we have the third condition:

DIR 3. *Given* ϵ *and* δ *there exists* r_0, $0 < r_0 < 1$, *such that if*

$$r_0 < r < 1,$$

then

$$\int_{-\pi}^{-\delta} P_r + \int_{\delta}^{\pi} P_r < \varepsilon.$$

This is easily seen as follows. Consider for definiteness the interval $[\delta, \pi]$. Then on that interval,

$$1 - 2r\cos\theta + r^2 \geqq 1 - 2\cos\delta + r^2.$$

As $r \to 1$, the right side approaches $2 - 2\cos\delta > 0$, so there exists $b > 0$ and r_1 such that if $r_1 < r < 1$ then

$$1 - 2r\cos\theta + r^2 \geqq b.$$

Hence

$$\int_\delta^\pi P_r\,d\theta \leqq \int_\delta^\pi \frac{1-r^2}{b}\,d\theta \leqq \frac{\pi}{b}(1-r^2).$$

Since $1 - r^2 \to 0$ as $r \to 1$, the desired estimate of **DIR 3** follows.

Thus we view $\{P_r\}$ as a Dirac family, with $r \to 1$. We let $f_r = P_r * f$ and write $f(r, \theta) = f_r(\theta)$. Then $f_r(\theta)$ is a function on the *open* disc. Theorem 5.1 yields:

$$f_r(\theta) \to f(\theta) \qquad \text{uniformly as} \quad r \to 1.$$

Theorem 5.2. *Let f be a real valued continuous function, periodic of period 2π. Then there exists a function u, continuous on the closed disc and harmonic on the open disc, such that $u = f$ on the circle, in other words $u(1, \theta) = f(\theta)$. This function is uniquely determined, and*

$$u(r, \theta) = f_r(\theta) = P_r * f(\theta).$$

Proof. The Laplace operator in polar coordinates can be applied to $P_r(\theta)$, differentiating the series term by term, which is obviously allowable. If you do this, you will find that

$$\left[r^2\left(\frac{\partial^2}{\partial r^2}\right) + r\left(\frac{\partial}{\partial r}\right) + \frac{\partial^2}{\partial \theta^2}\right] P_r(\theta) = 0.$$

Thus $\Delta P = 0$, where P denotes the function of two variables,

$$P(r, \theta) = P_r(\theta).$$

Differentiating under the integral sign, we then obtain

$$\Delta\big((P_r * f)(\theta)\big) = \big(\Delta P_r(\theta)\big) * f = 0.$$

We view the original periodic function f as a boundary value on the circle. The function

$$u(r, \theta) = f_r(\theta) = P_r * f(\theta)$$

is defined by convolution for $0 \leqq r < 1$, and by continuity for $r = 1$ by Theorem 5.1. This yields the existence of a harmonic function u having the prescribed value f on the circle. Uniqueness was proved in Theorem 1.3.

Remark. Theorem 5.2 gives an alternative proof of Theorem 4.1.

In many applications, e.g. physical applications, but even theoretical ones, it is not convenient to assume that the boundary value function is continuous. One should allow for at least a finite number of discontinuities, although still assuming that the function is bounded. In that case, an analysis of the proof shows that as much as one would expect of the theorem remains true, i.e. the reader will verify that the proof yields:

Theorem 5.3. *Let f be a bounded function on the reals, piecewise continuous, periodic of period 2π. Let S be a compact set where f is continuous. Then the sequence $\{K_n * f\}$ converges uniformly to f on S. The function*

$$u(r, \theta) = f_r(\theta) = P_r * f(\theta)$$

is harmonic on the open disc $0 \leqq r < 1$ and $0 \leqq \theta \leqq 2\pi$.

Not only can we prove the existence of a harmonic function having a prescribed boundary value as in Theorem 5.2, but we can also prove directly the existence of an *analytic function* having this boundary value for its real part by using the Poisson integral formula as the definition. Nothing is asserted about the imaginary part, and the example given at the end of §4 shows that the analytic function need not itself extend by continuity to the boundary.

Theorem 5.4. *Let u be continuous on the closed unit disc $\bar{D}$, and harmonic on the disc D. Then there exists an analytic function f on D such that $u = \operatorname{Re} f$, and two such functions differ by a pure imaginary constant. In fact,*

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

where C is the unit circle and K is a real constant.

Proof. The function f defined by the above integral is analytic on D by Theorem A3 of §6. We have to identify its real part with u. But the integrand is merely another expression for the convolution of the Poisson kernel with u. Indeed, the reader will easily verify that if $z = re^{i\theta}$ is the polar expression for z, then

$$P_r(\theta - t) = \frac{1}{2\pi} \operatorname{Re} \frac{e^{it} + z}{e^{it} - z} = \frac{1}{2\pi} \frac{1 - r^2}{1 - 2r\cos(\theta - t) + r^2}.$$

The unit circle is parametrized by $\zeta = e^{it}$, $d\zeta = e^{it}i\,dt$, and so the expression for the real part of f can also be written

$$\operatorname{Re} f = \frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{Re}\frac{e^{it}+z}{e^{it}-z}u(e^{it})\,dt,$$

which is the convolution integral. Applying Theorem 5.1 shows that the real part $\operatorname{Re} f$ on D extends continuously to the boundary and that its boundary value is precisely u. The uniqueness of harmonic functions with given boundary value shows that the real part of f is also equal to u on the interior, as was to be shown.

The integral expression of Theorem 5.4 gives a bound for f in terms of its real part and $f(0)$. Such a bound can be obtained in a simpler manner just using the maximum modulus principle, and we shall give this other proof in §3 of Chapter XII.

We shall now give an application of the construction of a harmonic function with given boundary value.

Theorem 5.5. *Let u be continuous on an open set U. Suppose that u satisfies the mean value property locally at every point of U, that is for $z_0 \in U$ and r sufficiently small,*

$$u(z_0) = \int_0^{2\pi} u(z_0 + re^{i\theta})\frac{d\theta}{2\pi}.$$

Then u is harmonic on U.

Proof. We first prove that u satisfies the maximum principle locally. Suppose $u(z_0) \geqq u(z_0 + re^{i\theta})$ for all r with $0 \leqq r \leqq r_0$. Then u is locally constant at z_0. Indeed, suppose that $u(z_1) < u(z_0)$ for some point z_1 on the circle of radius r. Then $u(z) \leqq u(z_0) - \varepsilon$, for some $\varepsilon > 0$ and all z sufficiently close to z_1 on the circle. It then follows that the above integral is $< u(z_0)$, a contradiction. This proves the local maximum principle, and similarly the local minimum principle.

Now to prove that u is harmonic, let $z_0 \in U$, and pick r_0 sufficiently small so that the mean value property holds with $0 \leqq r \leqq r_0$. Let u_1 be the harmonic function on the disc $D(z_0, r_0)$ having the given boundary value u on the circle of radius r_0 around z_0, as guaranteed by Theorem 5.4. Then $u - u_1$ has boundary value 0, and satisfies the mean value property at every point of $D(z_0, r_0)$. Hence $u - u_1$ has both its maximum and minimum on the boundary, so $u - u_1 = 0$ and $u = u_1$, thus proving the theorem.

VIII, §5. EXERCISES

One can also consider Dirac sequences or families over the whole real line. We use a notation which will fit a specific application. For each $y > 0$ suppose given a continuous function P_y on the real line, satisfying the following conditions:

DIR 1. *$P_y(t) \geqq 0$ for all y, and all real t.*

DIR 2.
$$\int_{-\infty}^{\infty} P_y(t)\,dt = 1.$$

DIR 3. *Given ϵ, δ there exists $y_0 > 0$ such that if $0 < y < y_0$, then*
$$\int_{-\infty}^{-\delta} + \int_{\delta}^{\infty} P_y(t)\,dt < \varepsilon.$$

We call $\{P_y\}$ a **Dirac family** again, for $y \to 0$. Prove:

1. Let f be continuous on $\mathbf{R}$, and bounded. Define the convolution $P_y * f$ by
$$P_y * f(x) = \int_{-\infty}^{\infty} P_y(x - t) f(t)\,dt.$$
Prove that $P_y * f(x)$ converges to $f(x)$ as $y \to 0$ for each x where f is continuous.
The proof should also apply to the case when f is bounded, and continuous except at a finite number of points, etc.

2. Let
$$P_y(t) = \frac{1}{\pi}\frac{y}{t^2 + y^2} \qquad \text{for} \quad y > 0.$$
Prove that $\{P_y\}$ is a Dirac family. It is called the Poisson family for the upper half plane, and it is classical.

3. Define for all real x and $y > 0$:
$$F(x, y) = P_y * f(x).$$
Prove that F is harmonic. In fact, show that the Laplace operator
$$\left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2$$
applied to
$$\frac{y}{(t - x)^2 + y^2}$$
yields 0.
You will have to differentiate under an integral sign, with the integral being taken over the real line. You can handle this in two ways.
(i) Work formally and assume everything is OK.
(ii) Justify all the steps. In this case, you have to use a lemma like Theorem A4 of §6.

The above procedure shows how to construct a harmonic function on the upper half plane, with given boundary value, just as was done for the disc in the text.

4. Let u be a bounded continuous function on the closure of the upper half plane (i.e. on the upper half plane and on the real line). Assume also that u is harmonic on the upper half plane, and that there are constants $c > 0$ and $K > 0$ such that

$$|u(t)| \leqq K \frac{1}{|t|^c} \qquad \text{for all } |t| \text{ sufficiently large.}$$

Using the Dirac family of the preceding exercise, prove that there exists an analytic function f on the upper half plane whose real part is u. [*Hint*: Recall the integral formula of Exercise 23 of Chapter VI, §1).]

Remark. If u is not bounded, then the conclusion need not hold, as shown by the function $u(x, y) = y$, which has boundary value 0 on the real line.

Let u be a continuous function on an open set U. We say that u satisfies the **circle mean value property** at a point $z_0 \in U$ if

$$u(z_0) = \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{i\theta})\, d\theta$$

for all $r > 0$ sufficiently small (so that in particular the disc $\bar{D}(z_0, r)$ is contained in U). We say that u satisfies the **disc mean value property** at a point $z_0 \in U$ if

$$u(z_0) = \frac{1}{\pi r^2} \iint_{\bar{D}(z_0, r)} u\, dx\, dy,$$

for all $r > 0$ sufficiently small. We say that the function satisfies the **mean value property** (either one) on U if it satisfies this mean value property at every point of U. By Theorem 3.3 and Theorem 5.5 we know that u is harmonic if and only if u satisfies the circle mean value property.

5. Prove that u is harmonic if and only if u satisfies the disc mean value property on U.

6. Let H^+ be the upper half plane. For $z \in H^+$ define the function

$$h_z(\zeta) = \frac{1}{2\pi i}\left(\frac{1}{\zeta - z} - \frac{1}{\zeta - \bar{z}}\right) \qquad \text{for} \quad \zeta \in H^+.$$

Then h_z is analytic on H^+ except for a simple pole at z. Let f be an analytic function on $H^+ \cup \mathbf{R}$ (i.e. on the closure of the upper half plane, meaning on an open set containing this closure). Suppose that f is bounded on $H^+ \cup \mathbf{R}$. Prove that

$$\int_{-\infty}^{\infty} f(t) h_z(t)\, dt = f(z).$$

This is the analogue of Theorem 4.1 for the upper half plane. [*Hint*: Integrate over the standard region from the calculus of residues, namely over the interval $[-R, R]$ and over the semicircle of radius R.]

7. In Exercise 6, consider the case $z = i$. Let

$$w = \frac{z-i}{z+i} \qquad \text{so} \quad z = -i\frac{w+1}{w-1}$$

be the standard isomorphisms between the upper half plane and the unit disc D. Show that

$$\left(\frac{1}{z-i} - \frac{1}{z+i}\right) dz = \frac{dw}{w}.$$

In light of Exercise 1, this shows that the kernel function in Exercise 2 corresponds to the Poisson kernel under the isomorphism between H^+ and D.

8. Let $x = r\cos\theta$, $y = r\sin\theta$ be the formulas for the polar coordinates. Let

$$f(x, y) = f(r\cos\theta, r\sin\theta) = g(r, \theta).$$

Show that

$$\boxed{\frac{\partial^2 g}{\partial r^2} + \frac{1}{r}\frac{\partial g}{\partial r} + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.}$$

For the proof, start with the formulas

$$\frac{\partial g}{\partial r} = (D_1 f)\cos\theta + (D_2 f)\sin\theta \qquad \text{and} \qquad \frac{\partial g}{\partial \theta} = -(D_1 f)r\sin\theta + (D_2 f)r\cos\theta,$$

and take further derivatives with respect to r and with respect to θ, using the rule for derivative of a product, together with the chain rule. Then add the expression you obtain to form the left-hand side of the relation you are supposed to prove. There should be enough cancellation on the right-hand side to prove the desired relation.

9. (a) For $t > 0$, let

$$K(t, x) = K_t(x) = \frac{1}{\sqrt{4\pi t}}e^{-x^2/4t}.$$

Prove that $\{K_t\}$ for $t \to 0$ is a Dirac family indexed by t, and $t \to 0$ instead of $n \to \infty$. One calls K the **heat kernel**.

(b) Let $D = (\partial/\partial x)^2 - \partial/\partial t$. Then D is called the **heat operator** (just as we defined a Laplace operator Δ). Show that $DK = 0$. (This is the analogue of the statement that $\Delta P = 0$ if P is the Poisson kernel.)

(c) Let f be a piecewise continuous bounded function on $\mathbf{R}$. Let $F(t, x) = (K_t * f)(x)$. Show that $DF = 0$, i.e. F satisfies the heat equation.

VIII, §6. APPENDIX. DIFFERENTIATING UNDER THE INTEGRAL SIGN

In this section, we used theorems of basic real analysis. We shall give here proofs in special cases which are sufficient for the applications we make here. We do this for the convenience of the reader. See Chapter XV, §1 for a version specific to complex situations. We first deal with continuity, then with differentiability.

Theorem A1. *Let f be a continuous function of two variables, $f = f(x, y)$ with x, y in finite intervals, $x \in [a, b]$ and $y \in [c, d]$. Let*

$$g(x) = \int_c^d f(t, x)\, dt.$$

Then g is continuous.

Proof. Since the rectangle $[a, b] \times [c, d]$ is compact, f is uniformly continuous on this rectangle. For x, $x' \in [a, b]$, we have

$$|g(x) - g(x')| \leqq \int_c^d |f(t, x) - f(t', x)|\, dt.$$

By uniform continuity, given ϵ there exists δ such that if $|x - x'| < \delta$ then $|f(t, x) - f(t', x)| < \epsilon$ for all $y \in [c, d]$. Hence

$$|g(x) - g(x')| \leqq \epsilon(d - c)$$

which concludes the proof.

Next we deal with an infinite interval, but still consider continuity.

Theorem A2. *Let f be a continuous function of two variables (t, x) defined for $t \geqq a$ and x in some compact set of numbers S. Assume that the integral*

$$\int_a^\infty f(t, x)\, dt = \lim_{B \to \infty} \int_a^B f(t, x)\, dt$$

converges uniformly for $x \in S$. Let

$$g(x) = \int_a^\infty f(t, x)\, dt.$$

Then g is continuous.

Proof. For given $x \in S$ we have

$$\begin{aligned} g(x+h) - g(x) &= \int_a^\infty f(t, x+h)\, dt - \int_a^\infty f(t, x)\, dt \\ &= \int_a^\infty \big(f(t,x+h) - f(t,x)\big)\, dt. \end{aligned}$$

Given ϵ, select B such that for all $y \in S$ we have

$$\left| \int_B^\infty f(t, y)\, dt \right| < \epsilon.$$

Then

$$\begin{aligned} |g(x+h) - g(x)| \leqq & \left| \int_a^B \big(f(t, x+h) - f(t, x)\big)\, dt \right| \\ & + \left| \int_B^\infty f(t, x+h)\, dt \right| + \left| \int_B^\infty f(t, x)\, dt \right|. \end{aligned}$$

We know that f is uniformly continuous on the compact set $[a, B] \times S$. Hence there exists δ such that whenever $|h| < \delta$ we have

$$|f(t, x+h) - f(t, x)| \leqq \epsilon/B.$$

The first integral on the right is then estimated by $B\epsilon/B = \epsilon$. The other two are estimated each by ϵ, so we have a 3ϵ-proof for the theorem.

We shall now prove special cases of the theorem concerning differentiation under the integral sign which are sufficient for our applications. They may be called the **absolutely convergent** cases.

For a function f of two variables, we let $D_1 f$ and $D_2 f$ be the partial derivatives of f with respect to the first and second variable respectively.

Theorem A3. *Let $f(t, x)$ be defined for $a \leqq t \leqq b$ and $x \in U$ (either an interval for the real variable case, or an open set in* $\mathbf{C}$ *for the case of complex x). Suppose f and $D_2 f$ defined and continuous. Let*

$$g(x) = \int_a^b f(t, x)\, dt.$$

Then g is differentiable, and

$$g'(x) = \int_a^b D_2 f(t,x)\, dt.$$

Proof. We have by linearity,

$$\frac{g(x+h)-g(x)}{h} - \int_a^b D_2 f(t, x)\, dt = \int_a^b \left[\frac{f(t, x+h)-f(t, x)}{h} - D_2 f(t, x)\right] dt.$$

By the mean value inequality (see below),

$$\left|\frac{f(t, x+h)-f(t, x)}{h} - D_2 f(t, x)\right| \leqq \sup_w |D_2 f(t, w) - D_2 f(t, x)|$$

where the sup is taken for all w on the segment from x to $x+h$. Let $S(h)$ be a closed interval or closed disc of radius $|h|$ centered at x, and contained in U. By the uniform continuity of $D_2 f$ on compact sets, the right side approaches 0 as $h \to 0$, uniformly for $t \in [a, b]$. The conclusion of the theorem follows at once.

Theorem A4. *Let f be a function of two variables (t, x) defined for $t \geqq a$ and x either in a finite interval $[c, d]$, or on some open set U in* $\mathbf{C}$. *Assume that $D_2 f$ exists, and that both f and $D_2 f$ are continuous. Assume that there are functions $\varphi(t)$ and $\psi(t)$ which are $\geqq 0$, such that*

$$|f(t, x)| \leqq \varphi(t) \qquad \textit{and} \qquad |D_2 f(t, x)| \leqq \psi(t),$$

for all t, x, and such that the integrals

$$\int_a^\infty \varphi(t)\, dt \qquad \textit{and} \qquad \int_a^\infty \psi(t)\, dt$$

converge. Let

$$g(x) = \int_a^\infty f(t, x)\, dt.$$

Then g is differentiable, and

$$Dg(x) = \int_a^\infty D_2 f(t, x)\, dt.$$

Proof. Let $S = [c, d]$, or $S =$ closed disc centered at x and contained in U. We have

$$\left|\frac{g(x+h)-g(x)}{h} - \int_a^\infty D_2 f(t, x)\, dt\right|$$
$$\leqq \int_a^\infty \left|\frac{f(t, x+h)-f(t, x)}{h} - D_2 f(t, x)\right| dt.$$

Select B so large that

$$\int_B^\infty \psi(t)\,dt < \epsilon.$$

Then we estimate our expression by

$$\int_a^\infty = \int_a^B + \int_B^\infty .$$

Since $D_2 f$ is uniformly continuous on $[a, B] \times S$ we can find δ such that whenever $|h| < \delta$,

$$\sup_w |D_2 f(t, w) - D_2 f(t, x)| < \frac{\epsilon}{B}$$

where the sup is taken for w on the segment between x and $x + h$. The integral between a and B is then bounded by ϵ. The integral between B and ∞ is bounded by 2ϵ because

$$\sup_w |D_2 f(t, w) - D_2 f(t, x)| \leqq 2\psi(t).$$

This proves our theorem.

Remark. In the above proof, we used the mean value theorem in its manifestation as a mean value inequality. We recall briefly how it is proved.

For a continuously differentiable function f on some open U, a point $x \in U$ such that the interval between x and $x + h$ is contained in U, we have the exact relation

$$f(x + h) - f(x) = \int_0^1 \frac{d}{dt} f(x + th)\,dt,$$

and hence the estimate

$$\left| \int_0^1 f'(x + th)\,dt \right| \leqq |h| \ \text{sup}\, |f'(x + th)|.$$

This gives

MVT 1. $$|f(x + h) - f(x)| \leqq |h| \sup_w |f'(w)|,$$

the sup being taken over all w on the segment. Apply this estimate to the function g such that $g(z) = f(z) - f'(x)z$. Then we get

MVT 2. $$|f(x+h) - f(x) - f'(x)h| \leqq |h| \sup_w |f'(w) - f'(x)|.$$

After dividing by h, this is precisely the inequality we used in the previous proofs.

PART TWO

Geometric Function Theory

This part contains three chapters which give a very strong geometric flavor to complex analysis. These chapters are logically independent of the third part, and continue some ideas concerning analytic isomorphisms and harmonic functions. Thus these three chapters form one possible natural continuation of Chapters VII and VIII.

On the other hand, the maximum modulus theorem and its applications give rise to another direction, taken up in Part Three, concerning the rate of growth of analytic functions.

The ordering of the book fits the breakdown of courses into semesters, not the mathematical ideas. For instance, the analytic continuation along curves could be done immediately after the definition of the logarithm in Chapter I, §6. This continuation relies on the same idea of homotopy and curves close together. Similarly, the Riemann mapping theorem could be done immediately after Chapter VII, as giving the general result in the background of the specific mappings for simple concrete open sets isomorphic to the disc. However, students may prefer to get an overview of the simpler cases first, and get to the general theorems afterward. Unfortunately, a book has to be projected in a totally ordered way on the page axis. Students should reorder the material according to their taste and individual ways of comprehension.

CHAPTER IX

Schwarz Reflection

Let f be an analytic function on an open set U, and let V be an open set. We shall give various criteria when f can be extended to an analytic function on $U \cup V$. The process of extending f in this way is called **analytic continuation**. If U, V are connected, and have in common an infinite set of points which have a point of accumulation in $U \cap V$, then an analytic continuation of f to $U \cup V$ is uniquely determined. Indeed, if g is analytic on V and $g = f$ on $U \cap V$, then g is the only such function by Theorem 1.2 of Chapter III.

In this chapter, we deal with the analytic continuation across a boundary consisting of a real analytic arc, which will be defined precisely. We shall deal with the continuation of harmonic functions as well as analytic functions, following the pattern established in the last chapter.

IX, §1. SCHWARZ REFLECTION (BY COMPLEX CONJUGATION)

First, we set some notation which will remain in force in this section.

Let U^+ be a connected open set in the upper half plane, and suppose that the boundary of U^+ contains an open interval I of real numbers. Let U^- be the reflection of U^+ across the real axis (i.e. the set of $\bar{z}$ with $z \in U^+$), and as in Figure 1, *let*

$$U = U^+ \cup I \cup U^-.$$

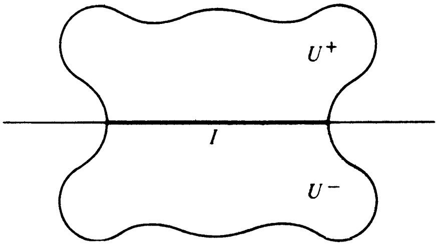


Figure 1

Theorem 1.1. *Suppose throughout this section that U is open.*

(i) *If f is a function on U, analytic on U^+ and U^-, and continuous on I, then f is analytic on U.*
(ii) *If f is a function on $U^+ \cup I$, analytic on U^+ and continuous on I, and f is real valued on I, then f has a unique analytic continuation F on U, and F satisfies*

$$F(z) = \overline{f(\bar{z})}.$$

Proof. We reduce (ii) to (i) if we define F by the above formula. From the hypothesis that f is real valued on I, it is clear that F is continuous at all points of I, whence on U. Furthermore, from the power series expansion of f at some point z_0 in the upper half plane, it is immediate from the formula that F is analytic on U^-. There remains to prove (i).

We consider values of z near I, and especially near some point of I. Such values lie inside a rectangle, as shown on Fig. 2(a). This rectangle has a boundary $C = C^+ + C^-$, oriented as shown.

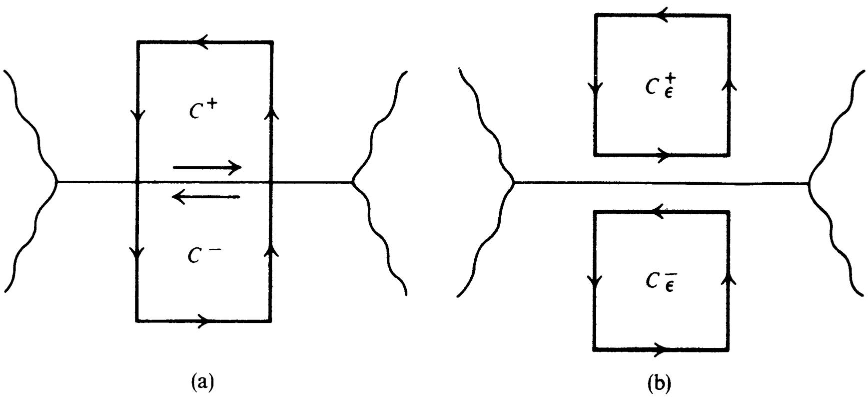


Figure 2

We define

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

for z inside the rectangle. Then for z not on I, we have

$$g(z) = \frac{1}{2\pi i}\int_{C^+} \frac{f(\zeta)}{\zeta - z}\,d\zeta + \frac{1}{2\pi i}\int_{C^-} \frac{f(\zeta)}{\zeta - z}\,d\zeta.$$

Let C_ϵ^+ be the rectangle as shown on Fig. 2(b). Then for z inside C_ϵ^+, Cauchy's formula gives

$$f(z) = \frac{1}{2\pi i}\int_{C_\epsilon^+} \frac{f(\zeta)}{\zeta - z}\,d\zeta.$$

An easy continuity argument shows that taking the limit as $\epsilon \to 0$, we get in fact

$$f(z) = \frac{1}{2\pi i}\int_{C^+} \frac{f(\zeta)}{\zeta - z}\,d\zeta,$$

for z inside C_ϵ^+, and hence inside C^+. On the other hand, a similar argument combined with Cauchy's theorem shows that

$$\int_{C^-} \frac{f(\zeta)}{\zeta - z}\,d\zeta = 0.$$

Hence $g(z) = f(z)$ if z is in U^+ near I. By continuity, we also obtain that $g(z) = f(z)$ if z is on I. By symmetry, the same arguments would show that $g(z) = f(z)$ if z is in U^-, and z is near I. This proves that $g(z) = f(z)$ for z near I in U, and hence that g is the analytic continuation of f in U, as was to be shown.

Next we prove the analogue for harmonic functions.

Theorem 1.2. *Let v be a continuous function on $U^+ \cup I$, harmonic on U^+, and equal to 0 on I. Then v extends to a harmonic function on $U^+ \cup I \cup U^-$.*

Proof. Define $v(\bar{z}) = -v(z)$. Then v is harmonic on U^-, because the property of being harmonic can be verified locally, and on a small disc centered at a point z_0 in U^+, v is the imaginary part of an analytic function f, so $-v$ is the imaginary part of $\bar{f}$, whence the definition of v on U^- gives a harmonic function. From the hypothesis that $v = 0$ on I, we conclude that v is continuous on I, so v is continuous on all of U.

There remains only to show that v is harmonic at points of I. Let $x_1 \in I$, and let $D(x_1, r_1)$ be a small disc centered at x_1. Since v is harmonic on U^+ and U^-, it satisfies the mean value property at points of U^+ and U^-. Since $v(x_1) = 0$, and $v(\bar{z}) = -v(z)$, it follows at once that v also satisfies the mean value property at x_1. Hence by Theorem 5.5 of Chapter VIII, it follows that v is harmonic in a neighborhood of x_1. This proves the theorem.

Peter Jones showed me another elegant argument which avoids the classical use of Theorem 5.5 and runs as follows. Let D be a disc centered at a point of a real interval I. Let v be a continuous function on the closure of the upper half disc D^+. We suppose that v is harmonic on D^+, and $v = 0$ on the real segment. Define the function v on D^- by

$$v(z) = -v(\bar{z}).$$

Then v is harmonic on D^-, and since $v = 0$ on I, it follows that v is defined continuously on the closure of D. We have to prove that v is harmonic on D. By the existence and uniqueness theorem, there exists a unique harmonic function v_1 on D having continuous boundary value equal to v on the boundary circle. We want to show $v_1 = 0$ on the real segment. Define v_2 on the closed disc by

$$v_2(z) = -v_1(\bar{z}).$$

Then v_2 is continuous on the closure of D, and harmonic on D. But v_2 also has the same boundary value as v_1. Therefore $v_2 = v_1$. From the above equation, we conclude that

$$v_1(x) = -v_1(x) \qquad \text{for } x \text{ on the real segment,}$$

whence $v_1(x) = 0$ for x on the real segment. Finally v_1 has the same boundary value as v itself on the closed upper half disc, and is therefore equal to v on this upper half disc. This concludes the proof.

Remark. Theorem 1.1 is also an immediate consequence of Theorem 1.2, and readers should be aware of both approaches, via Cauchy's theorem and via harmonic functions.

Next we consider cases when the analytic function may be an isomorphism. The next result will be applied in Theorem 2.5 to complement the Riemann mapping theorem in the next chapter. We let H^+ be the upper half plane and H^- the lower half plane.

Proposition 1.3. *Let f be an analytic function on U, real valued on I. Assume in addition that f gives an isomorphism of U^+ with $f(U^+) \subset H^+$ and an isomorphism of U^- with $f(U^-) \subset H^-$. Then f is an isomorphism of U with $f(U)$.*

Proof. We first show that f is a local isomorphism at each point of I. After making translations, we may assume without loss of generality that the point of I is 0, and that $f(0) = 0$. Let

$$f(z) = cz^m + \text{higher terms}, \qquad \text{with} \quad c \neq 0 \text{ and } m \geqq 1$$

be the power series expansion of f. Since $f(I) \subset \mathbf{R}$, we must have c real. Since by hypothesis $f(U^+) \subset H^+$, we must have $c > 0$, as one sees by looking at values on $z = re^{i\theta}$ with $r > 0$ and θ small. Then $m = 1$, for otherwise, there is some θ with $0 < \theta < \pi$ such that $\pi < m\theta < 2\pi$, contradicting the hypothesis that $f(U^+) \subset H^+$. Hence $f'(z) \neq 0$ for all $z \in I$, and f is a local isomorphism at each point of I.

There remains to prove that f is injective on I. Suppose $f(z_1) = f(z_2)$ with $z_1 \neq z_2$ on I. Let $z \in U^+$ be close to z_2. Then $f(z)$ is close to $f(z_1)$ and lies in H^+. But f gives an isomorphism of some disc centered at z_1 with a neighborhood of $f(z_1)$, and by hypothesis, f induces an isomorphism of the upper half of this disc with its image in H^+, so $f(z) = f(z')$ for some z' in the upper half of the disc, contradicting the hypothesis that f is injective on U^+. This concludes the proof of the proposition. Figure 3 illustrates the argument.

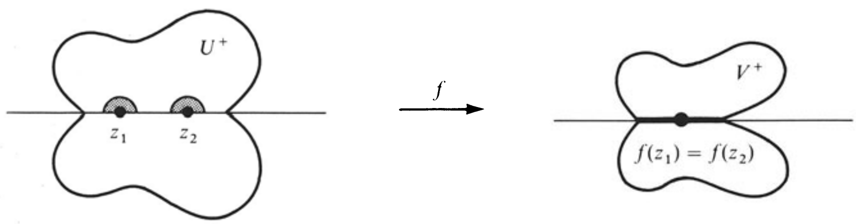


Figure 3

IX, §2. REFLECTION ACROSS ANALYTIC ARCS

The theorems of the preceding section apply to more general situations which are analytically isomorphic to those of the theorems. More precisely, let V be an open set in the complex numbers, and suppose that V

is the disjoint union

$$V = V^+ \cup \gamma \cup V^-,$$

where V^+ is open, V^- is open, and γ is some curve. We suppose that there exists an analytic isomorphism

$$\psi : U \to V$$

such that

$$\psi(U^+) = V^+, \qquad \psi(I) = \gamma, \qquad \psi(U^-) = V^-.$$

The notation $U = U^+ \cup I \cup U^-$ is the same as at the beginning of §1.

Theorem 2.1.

(i) *Given a function g on V which is analytic on V^+ and V^-, and continuous on γ, then g is analytic on V.*

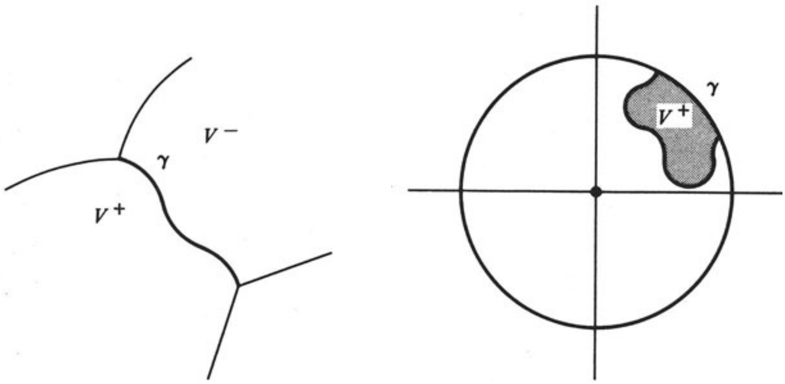
(ii) *If g is an analytic function on V^+ which extends to a continuous function on $V^+ \cup \gamma$, and is real valued on γ, then g extends to an analytic function on V.*

Proof. Obvious, using successively parts (i) and (ii) of the theorem, applied to the function $f = g \circ \psi$.

A standard example of this situation occurs when V^+ is an open set on one side of an arc of the unit circle, as on Fig. 4(b). It is then useful to remember that the map

$$z \mapsto 1/z$$

interchanges the inside and outside of the unit circle.



(a) (b)

Figure 4

We can also perform an analytic continuation in the more general situation, as in part (ii) of the theorem. More generally, let us define a curve

$$\gamma: [a, b] \to \mathbf{C}$$

to be **real analytic** if for each point t_0 in $[a, b]$ there exists a convergent power series expansion

$$\gamma(t) = \sum a_n(t - t_0)^n$$

for all t sufficiently close to t_0. Using these power series, we see that γ extends to an analytic map of some open neighborhood of $[a, b]$. We shall say that γ is a **proper analytic arc** if γ is injective, and if $\gamma'(t) \neq 0$ for all $t \in [a, b]$. We leave it as an exercise to prove:

If γ is a proper analytic arc, then there exists an open neighborhood W of $[a, b]$ such that γ extends to an analytic isomorphism of W.

Let $\gamma: [a, b] \to \mathbf{C}$ be a proper analytic arc, which is contained in the boundary of an open set U (in other words, the image of γ is contained in the boundary of U). We shall say that U **lies on one side of** γ if there exists an extension γ_W of γ to some open neighborhood W of $[a, b]$ as above, such that $\gamma_W^{-1}(U)$ lies either in the upper half plane, or in the lower half plane and is an open neighborhood of $[a, b]$ in that half plane. As usual, $\gamma_W^{-1}(U)$ is the set of points $z \in W$ such that $\gamma_W(z) \in U$.

Let f be an analytic function on an open set U. Let γ be a proper analytic arc which is contained in the boundary of U, and such that U lies on one side of γ. We say that f has an **analytic continuation across** γ if there exists an open neighborhood W of γ (without its end points) such that f has an analytic continuation to $U \cup W$.

Theorem 2.2. *Let f be analytic on an open set U. Let γ be a proper analytic arc which is contained in the boundary of U, and such that U lies on one side of γ. Assume that f extends to a continuous function on $U \cup \gamma$ (i.e. $U \cup$ Image γ), and that $f(\gamma)$ is contained in a proper analytic arc η, such that $f(U)$ lies on one side of η. Then f has an analytic continuation across γ.*

Proof. There exist analytic isomorphisms

$$\varphi: W_1 \to \text{neighborhood of } \gamma,$$

$$\psi: W_2 \to \text{neighborhood of } \eta,$$

where W_1, W_2 are open sets as illustrated on Fig. 4, neighborhoods of real intervals I_1 and I_2, respectively. These open sets are selected sufficiently small that U and $f(U)$ lie on one side of γ, η, respectively.

Define

$$g = \psi^{-1} \circ f \circ \varphi.$$

Then g is defined and analytic on one side of the interval I_1, and is also continuous on I_1 (without its end points). Furthermore g is real valued. Hence g has an analytic continuation by Theorem 1.1 to the other side of I_1. The function

$$\psi \circ g \circ \varphi^{-1}$$

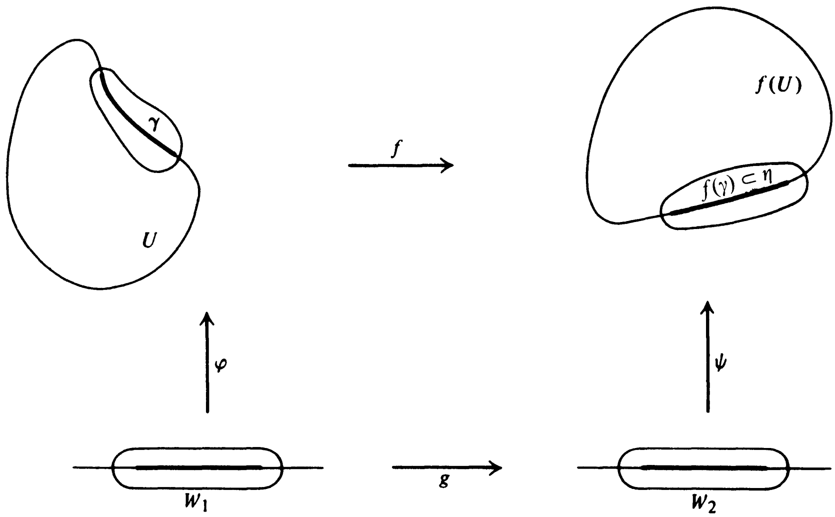


Figure 5

then defines an analytic continuation of f on some open neighborhood of γ (without its end points), as desired.

The analytic continuation of f in Theorem 2.2 is also often called the **reflection** of f across γ.

In exactly the same way, we can extend Theorem 1.2 to a harmonic function. More precisely:

Theorem 2.3. *Let v be harmonic on an open set U. Let γ be a proper analytic arc which is contained in the boundary of U, and such that U lies on one side of γ. Assume that v extends to a continuous function on $U \cup \gamma$, taking the value 0 on γ. Then v extends to a harmonic function on some open set containing $U \cup \gamma$.*

We shall now give some applications. Let U be an open set, and let $\{z_n\}$ be a sequence in U. We shall say that this sequence **approaches the**

boundary of U if given a compact subset K of U there exists n_0 such that $z_n \notin K$ for all $n \geqq n_0$. This is the most convenient definition for the moment. Readers can convince themselves that this coincides with the naive notion of the distance between z_n and the boundary approaching 0. For arguments showing the equivalence between the two notions, see Lemma 2.2 of Chapter X.

For the next result, we need another definition. Let $f: U \to V$ be a continuous map (which in practice will be analytic). We define f to be a **proper map** if for every compact set $K' \subset V$ the inverse image $f^{-1}(K')$ is compact. For instance, if f is an analytic isomorphism, then f is proper. Furthermore, let m be a positive integer. The map of $\mathbf{C} \to \mathbf{C}$ given by $z \mapsto z^m$ is proper.

Lemma 2.4. *Let $f: U \to V$ be a proper analytic map. If $\{z_n\}$ is a sequence in U approaching the boundary of U, then $\{f(z_n)\}$ approaches the boundary of V.*

Proof. Given K' compact in V, let $K = f^{-1}(K')$. There is some n_0 such that for $n \geqq n_0$ we have $z_n \notin K$, so $f(z_n) \notin K'$ as desired.

Example 1. Let $f: U \to D$ be an isomorphism of an open set with the unit disc. If $\{z_n\}$ is a sequence in U approaching the boundary of U, then $|f(z_n)| \to 1$ as $n \to \infty$.

Example 2. Let $f: U \to H$ be an isomorphism with the upper half plane. Let $\{z_n\}$ be a sequence in U approaching the boundary of U. Then $\{f(z_n)\}$ approaches the boundary of H, meaning that the values $f(z_n)$ either come close to the real numbers of become arbitrary large in absolute value, i.e. $|f(z_n)| \to \infty$.

Theorem 2.5. *Let $f: U \to D$ be an analytic isomorphism. Let γ be a proper analytic arc contained in the boundary of U and such that U lies on one side of γ. Then f extends to an analytic isomorphism on $U \cup \gamma$.*

Proof. In this proof we see the usefulness of dealing with harmonic functions globally. Let $v(z) = \log|f(z)|$. By Lemma 2.4, it follows that

$$\lim_{z \to \partial U} v(z) = 0.$$

Therefore v extends to a continuous function on $\overline{U}$ taking the value 0 on γ. By Theorem 2.3, v extends to a harmonic function on an open set containing $U \cup \gamma$. But locally, in a neighborhood of a point of γ, v is the real part of an analytic function g, which extends analytically across γ. Choosing the imaginary constant suitably makes g an analytic continuation of $\log f(z)$ in a neighborhood of the point. Then e^g gives the analytic continuation of f to $U \cup \gamma$. By Proposition 1.3 (in its isomorphic form corresponding to the analytic arc γ), we conclude that the analytic extension of f is injective on a sufficiently small open set containing $U \cup \gamma$, thus concluding the proof.

IX, §2. EXERCISES

1. Let C be an arc of unit circle $|z| = 1$, and let U be an open set inside the circle, having that arc as a piece of its boundary. If f is analytic on U if f maps U into the upper half plane, f is continuous on C, and takes real values on C, show that f can be continued across C by the relation

$$f(z) = \overline{f(1/\bar{z})}.$$

2. Suppose, on the other hand, that instead of taking real values on C, f takes on values on the unit circle, that is,

$$|f(z)| = 1 \qquad \text{for} \quad z \text{ on } C.$$

Show that the analytic continuation of f across C is now given by

$$f(z) = 1/\overline{f(1/\bar{z})}.$$

3. Let f be a function which is continuous on the closed unit disc and analytic on the open disc. Assume that $|f(z)| = 1$ whenever $|z| = 1$. Show that the function f can be extended to a meromorphic function, with at most a finite number of poles in the whole plane.

4. Let f be a meromorphic function on the open unit disc and assume that f has a continuous extension to the boundary circle. Assume also that f has only a finite number of poles in the unit disc, and that $|f(z)| = 1$ whenever $|z| = 1$. Prove that f is a rational function.

5. Work out the exercise left for you in the text, that is:

Let W be an open neighborhood of a real interval $[a, b]$. Let g be analytic on W, and assume that $g'(t) \neq 0$ for all $t \in [a, b]$, and g is injective on $[a, b]$. Then there exists an open subset W_0 of W containing $[a, b]$ such that g is an analytic isomorphism of W_0 with its image.

[*Hint*: First, by compactness, show that there is some neighborhood of $[a, b]$ on which g' does not vanish, and so g is a local isomorphism at each point of this neighborhood. Let $\{W_n\}$ be a sequence of open sets shrinking to $[a, b]$, for instance the set of points at distance $< 1/n$ from $[a, b]$. Suppose g is not injective on each W_n. Let $z_n \neq z_n'$ be two points in W_n such that $g(z_n) = g(z_n')$. The sequences $\{z_n\}$ and $\{z_n'\}$ have convergent subsequences, to points on $[a, b]$. If these limit points are distinct, this contradicts the injectivity of g on the real interval. If these limit points are equal, then for large n, the points are close to a point on the real interval, and this contradicts the fact that g is a local isomorphism at each point of the interval.]

6. Let U be an open connected set. Let f be analytic on U, and suppose f extends continuously to a proper analytic arc on the boundary of U, and this extension has value 0 on the arc. Show that $f = 0$ on U.

IX, §3. APPLICATION OF SCHWARZ REFLECTION

We begin with a relevant example of an analytic isomorphism involving circles. Let C_1, C_2 be two circles which intersect in two points z_1, z_2 and are perpendicular to each other at these points, as shown on Fig. 6. We suppose that C_2 does not go through the center of C_1. Let T be inversion through C_1. By Theorem 5.2, Chapter VII, T maps C_2 on another circle, and preserves orthogonality. Since every point of C_1 is fixed under reflection, and since given two points on C_1 there is exactly one circle passing through z_1 and z_2 and perpendicular to C_1, it follows that T maps C_2 onto itself. A point α on C_2 which lies outside C_1 is mapped on a point α' again on C_2 but inside C_1. Thus T interchanges the two arcs of C_2 lying outside and inside C_1, respectively.

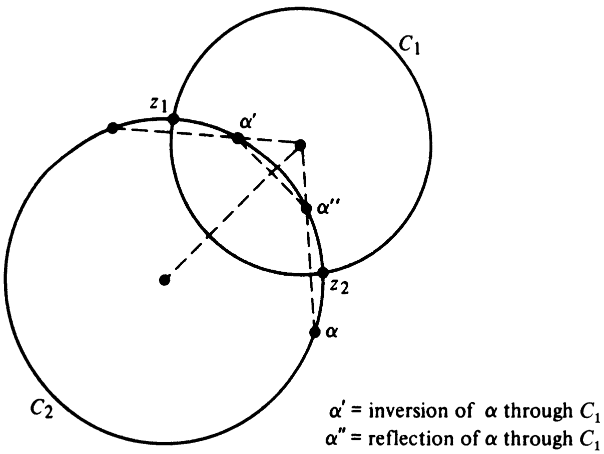


Figure 6

For the next example, it is convenient to make a definition. Let U be the open set bounded by three circular arcs perpendicular to the unit circle, as on Fig. 7. We shall call U a **triangle**. We suppose that the circular sides of the triangle do not pass through the center.

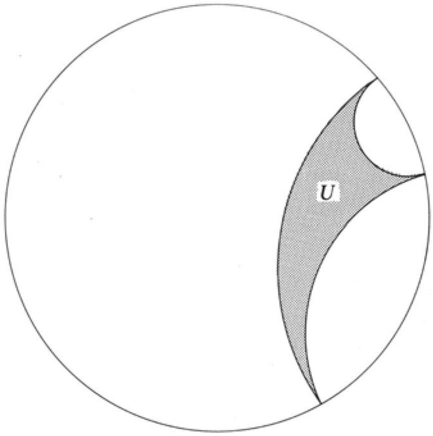


Figure 7

The remark implies that the reflection of U across any one of its "sides" is again a triangle, whose sides are circular arcs perpendicular to the unit circle.

Take three equidistant points on the unit circle, and join them with circular arcs perpendicular to the unit circle at these points. The region

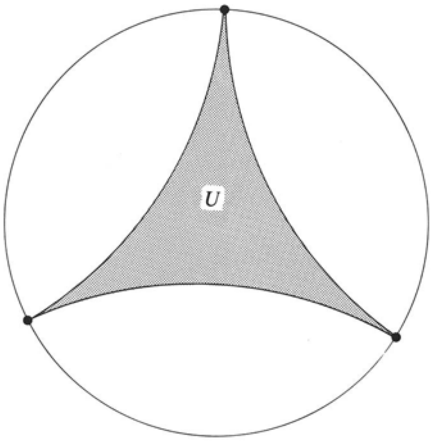


Figure 8

bounded by these three arcs is what we have called a triangle U, as shown on Fig. 8.

If we invert U across any one of its sides, we obtain another triangle U'. Inverting successively such triangles, we obtain a figure such as is illustrated in Fig. 9.

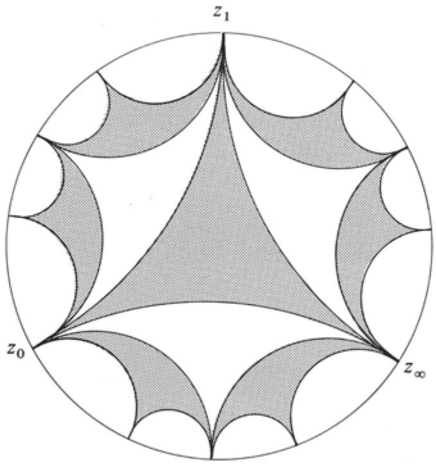


Figure 9

Let

$$g: U \to H \qquad \text{and} \qquad h: H \to U$$

by inverse analytic isomorphisms between the triangle U and the upper half plane. Such isomorphisms can always be found by the Riemann mapping theorem. Furthermore, we can always select h such that the three vertices z_0, z_1, z_∞ are mapped on 0, 1, ∞, respectively (why?). We proved in Theorem 2.5 that g can be extended continuously to the boundary of U, which consists of three analytic arcs. It follows that g maps the three arcs on the intervals

$$[-\infty, 0], \qquad [0, 1], \qquad [1, \infty]$$

respectively, and in particular, g is real valued on each arc.

By the Schwarz reflection principle, one may continue g analytically across each arc. It is not difficult to show that the union of all reflections of the original triangle repeatedly as above is equal to the whole unit disc but we omit the proof. Granting this we obtain a surjective analytic map

$$g: D \to \mathbf{C} - \{0, 1\}.$$

The inverse image of every point $z \in \mathbf{C}$, $z \neq 0, 1$ consists of a discrete set of points in the disc. Readers acquainted with elementary topology will realize that we have constructed the universal covering space of $\mathbf{C} - \{0, 1\}$.

In the chapter on analytic continuation, we shall give a further application of the above construction to prove Picard's theorem.

CHAPTER X

The Riemann Mapping Theorem

In this chapter we give the general proof of the Riemann mapping theorem. We also prove a general result about the boundary behavior.

X, §1. STATEMENT OF THE THEOREM

The **Riemann mapping theorem** asserts:

Let U be a simply connected open set which is not the whole plane. Then U is analytically isomorphic to the disc of radius 1. More precisely, given $z_0 \in U$, there exists an analytic isomorphism

$$f\colon U \to D(0, 1)$$

of U with the unit disc, such that $f(z_0) = 0$. Such an isomorphism is uniquely determined up to a rotation, i.e. multiplication by $e^{i\theta}$ for some real θ, and is therefore uniquely determined by the additional condition

$$f'(z_0) > 0.$$

We have seen in Chapter VII, §2 that the only analytic automorphisms of the disc are given by the mappings

$$w \mapsto e^{i\varphi} \frac{w - \alpha}{1 - \bar{\alpha}w},$$

where $|\alpha| < 1$ and φ is real. This makes the uniqueness statement obvious: If f, g are two analytic isomorphisms of U and $D(0, 1)$ satisfying the prescribed condition, then $f \circ g^{-1}$ is an analytic automorphism of the disc leaving the origin fixed and having positive derivative at the origin. It is then clear from the above formula that $f \circ g^{-1} = \text{id}$.

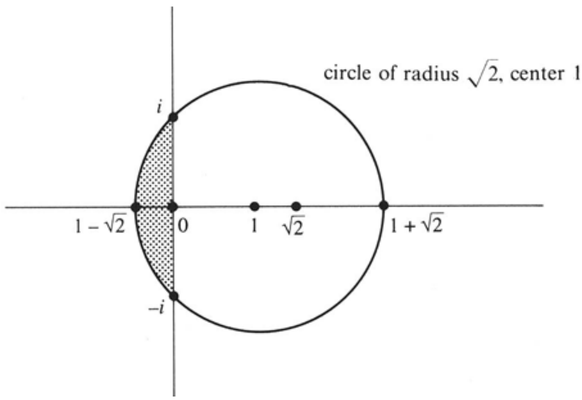
There are two main ingredients in the proof of the theorem. First, we shall prove that there exist injective analytic maps of U into the disc, mapping a given point z_0 to the origin. Furthermore, the family of such maps is relatively compact in the sense that given a sequence in the family, there exists a subsequence which converges uniformly on every compact subset of U.

Second, we consider the derivative $f'(z_0)$, which is bounded in absolute value. We pick a sequence $\{f_n\}$ in our family such that $|f_n'(z_0)|$ converges to the supremum of all values $|f'(z_0)|$ for f ranging in the family. We then prove that there is some element f in the family such that $|f'(z_0)|$ is actually equal to this supremum, which is therefore actually a maximum, and we prove that such a mapping f gives the desired isomorphism from U to D. We reduce the proof to Theorem 1.2 of Chapter VII, namely the Schwarz lemma for the derivative.

X, §1. EXERCISES

1. Let U be a simply connected open set. Let z_1, z_2 be two points of U. Prove that there exists a holomorphic automorphism f of U such that $f(z_1) = z_2$. (Distinguish the cases when $U = \mathbf{C}$ and $U \neq \mathbf{C}$.)

2. Let $f(z) = 2z/(1 - z^2)$. Show that f gives an isomorphism of the shaded region with a half disc. Describe the effect of f on the boundary.



What is the effect of f on the reflection of the region across the y-axis?

X, §2. COMPACT SETS IN FUNCTION SPACES

Let U be an open set. We denote by $\mathrm{Hol}(U)$ the space of holomorphic functions on U. A subset Φ of $\mathrm{Hol}(U)$ will be called **relatively compact** if every sequence in Φ has a subsequence which converges uniformly on every compact subset of U, not necessarily to an element of Φ itself. (*Note*: Instead of relatively compact, one sometimes calls such subsets **normal**, or **normal families**. The word relatively compact fits general notions of metric spaces somewhat better.) Recall that a subset S of complex numbers is called **relatively compact** if its closure is compact (closed and bounded). Such a subset S is relatively compact if and only if every sequence in S has a convergent subsequence. (The subsequence is allowed to converge to a point not in S by definition.)

A subset Φ of $\mathrm{Hol}(U)$ is said to be **uniformly bounded on compact sets in** U if for each compact set K in U there exists a positive number $B(K)$ such that

$$|f(z)| \leqq B(K) \qquad \text{for all} \quad f \in \Phi,\ z \in K.$$

A subset Φ of $\mathrm{Hol}(U)$ is said to be **equicontinuous on a compact set** K if, given ϵ, there exists δ such that if $z, z' \in K$ and $|z - z'| < \delta$, then

$$|f(z) - f(z')| < \epsilon \qquad \text{for all} \quad f \in \Phi.$$

Ascoli's theorem from real analysis states that an equicontinuous family of continuous functions on a compact set is relatively compact. We shall actually reprove it in the context of the next theorem.

Theorem 2.1. *Let $\Phi \subset \mathrm{Hol}(U)$, and assume that Φ is uniformly bounded on compact sets in U. Then Φ is equicontinuous on each compact set, and is relatively compact.*

Proof. After proving the equicontinuity, we shall use a diagonal procedure to find the convergent subsequence.

Let K be compact and contained in U. Let $3r$ be the distance from K to the complement of U. Let $z, z' \in K$ and let C be the circle centered at z' of radius $2r$. Suppose that $|z - z'| < r$. We have

$$\frac{1}{\zeta - z} - \frac{1}{\zeta - z'} = \frac{z - z'}{(\zeta - z)(\zeta - z')},$$

whence by Cauchy's formula,

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

Therefore

$$|f(z) - f(z')| < 2|z - z'| \, \|f\|_{K(2r)} \frac{1}{r},$$

where the sup norm of f is taken on the compact set $K(2r)$ consisting of those $z \in U$ such that $\operatorname{dist}(z, K) \leqq 2r$. This proves the equicontinuity of the family on K.

Given a sequence $\{f_n\}$ in Φ, we now prove that there exists a subsequence which converges uniformly on K by the standard proof for Ascoli's theorem. Let $\{z_j\}$ be a countable dense subset of K. For each j, the sequence $\{f_n(z_j)\}$ is bounded. There exists a subsequence $\{f_{n,1}\}$ such that

$$\{f_{n,1}(z_1)\}$$

converges. There exists a subsequence $\{f_{n,2}\}$ of $\{f_{n,1}\}$ such that

$$\{f_{n,2}(z_2)\}$$

converges. Proceeding in this manner we get subsequences $f_{n,j}$ such that

$$\{f_{n,j}(z_1)\}, \ldots, \{f_{n,j}(z_j)\}$$

converge. Then the diagonal subsequence $\{f_{n,n}\}$ is such that

$$\{f_{n,n}(z_j)\}$$

converges for each j.

In fact, we now prove that $\{f_{n,n}\}$ converges uniformly on K. Given ϵ, let δ be as in the definition of equicontinuity. Then for some k, the compact set K is contained in the union of discs

$$K \subset D(z_1, \delta) \cup \cdots \cup D(z_k, \delta).$$

Select N such that if $m, n > N$, then

$$|f_{n,n}(z_j) - f_{m,m}(z_j)| < \varepsilon \qquad \text{for} \quad j = 1, \ldots, k.$$

Let $z \in K$. Then $z \in D(z_i, \delta)$ for some i, and we get

$$\begin{aligned} |f_{n,n}(z) - f_{m,m}(z)| \leqq |f_{n,n}(z) - f_{n,n}(z_i)| &+ |f_{n,n}(z_i) - f_{m,m}(z_i)| \\ &+ |f_{m,m}(z_i) - f_{m,m}(z)|. \end{aligned}$$

The first and third term are $< \epsilon$ by the definition of equicontinuity. The middle term is $< \epsilon$ by what was just proved. We have therefore obtained a subsequence of the original sequence which converges uniformly on K.

We now perform another diagonal procedure.

Lemma 2.2. *There exists a sequence of compact sets K_s $(s = 1, 2, \ldots)$ such that K_s is contained in the interior of K_{s+1} and such that the union of all K_s is U.*

Proof. Let $\bar{D}_s$ be the closed disc of radius s, let $\bar{U}$ be the closure of U, and let

$$K_s = \text{set of points } z \in \bar{U} \cap \bar{D}_s \text{ such that } \operatorname{dist}(z, \text{boundary } U) \geqq 1/s.$$

Then K_s is contained in the interior of K_{s+1}. For instance, K_s is contained in the open set of elements $z \in U \cap D_{s+1}$ such that

$$\operatorname{dist}(z, \text{boundary } U) > \frac{1}{s+1}.$$

It is clear that the union of all K_s is equal to U. It then follows that any compact set K is contained in some K_s because the union of these open sets covers U, and a finite number of them covers K.

Let $\{f_n\}$ be the original sequence in Φ. There exists a subsequence $\{f_{n,1}\}$ which converges uniformly on K_1, then a subsequence $\{f_{n,2}\}$ which converges uniformly on K_2, and so forth. The diagonal sequence

$$\{f_{n,n}\}$$

converges uniformly on each K_s, whence on every compact set, and the theorem is proved.

X, §2. EXERCISES

Let U be an open set, and let $\{K_s\}$ $(s = 1, 2, \ldots)$ be a sequence of compact subsets of U such that K_s is contained in the interior of K_{s+1} for all s, and the union of the sets K_s is U. For f holomorphic on U, define

$$\sigma_s(f) = \min(1, \|f\|_s),$$

where $\|f\|_s$ is the sup norm of f on K_s. Define

$$\sigma(f) = \sum_{s=1}^{\infty} \frac{1}{2^s} \sigma_s(f).$$

1. Prove that σ satisfies the triangle inequality on $\operatorname{Hol}(U)$, and defines a metric on $\operatorname{Hol}(U)$.

2. Prove that a sequence $\{f_n\}$ in Hol(U) converges uniformly on every compact subset of U if and only if it converges for the metric σ.

3. Prove that Hol(U) is complete under the metric σ.

4. Prove that the map $f \mapsto f'$ is a continuous map of Hol(U) into itself, for the metric σ.

5. Show that a subset of Hol(U) is relatively compact in the sense defined in the text if and only if it is relatively compact with respect to the metric σ in the usual sense, namely its closure is compact.

Using these exercises, you may then combine the fact that Φ is equicontinuous in Theorem 2.1, with the usual statement of Ascoli's theorem, *without reproving the latter ad hoc*, to conclude the proof of Theorem 2.1.

6. Let Φ be the family of all analytic functions

$$f(z) = z + a_2 z^2 + a_3 z^3 + \cdots$$

on the open unit disc, such that $|a_n| \leqq n$ for each n. Show that Φ is relatively compact.

7. Let $\{f_n\}$ be a sequence of analytic functions on U, uniformly bounded. Assume that for each $z \in U$ the sequence $\{f_n(z)\}$ converges. Show that $\{f_n\}$ converges uniformly on compact subsets of U.

X, §3. PROOF OF THE RIEMANN MAPPING THEOREM

The theorem will be proved by considering an appropriate family of mappings, and maximizing the derivatives at one point. We first make a reduction which makes things technically simpler later. We let D be the unit disc $D(0, 1)$.

Let U be a simply connected open set $\neq \mathbf{C}$, and let $z_0 \in U$. We consider the family of all holomorphic functions $f\colon U \to D$ such that $f(z_0) = 0$ and such that f is injective. We shall prove that this family is not empty, and it is uniformly bounded. We shall then prove that there exists an element in the family for which $|f'(z_0)|$ is maximal in the family, and that this element gives the desired isomorphism of U and D. Since the only automorphisms of D leaving the origin fixed are rotations, we may then rotate such f so that $f'(z_0)$ is real and positive, thus determining f uniquely. We now carry out the program. We recall:

Lemma 3.1. *Let U be an open connected set. Let $f\colon U \to \mathbf{C}$ be analytic and injective. Then $f'(z) \neq 0$ for all $z \in U$, and f is an analytic isomorphism of U and its image.*

This is merely Theorem 6.4 of Chapter II.

Lemma 3.2. *Let U be a connected open set, and let $\{f_n\}$ be a sequence of injective analytic maps of U into $\mathbf{C}$ which converges uniformly on every compact subset of U. Then the limit function f is either constant or injective.*

Proof. Suppose f is not injective, so there exist two points $z_1 \neq z_2$ in U such that

$$f(z_1) = f(z_2) = \alpha.$$

Let $g_n = f_n - f_n(z_1)$. Then $\{g_n\}$ is a sequence which converges uniformly to

$$g = f - f(z_1).$$

Since f_n is assumed injective, it follows that g_n has no zero on U except at z_1. Suppose that g is not identically zero, and hence not identically zero near z_2 since we assumed that U is connected. Then z_2 is an isolated zero of g. Let γ be a small circle centered at z_2. Then $|g(z)|$ has a lower bound $\neq 0$ on γ (which is compact), and $\{1/g_n\}$ converges uniformly to $1/g$ on γ. By the formula of Chapter VI, §1 we know that

$$0 = \operatorname{ord}_{z_2} g_n = \frac{1}{2\pi i} \int_\gamma \frac{g_n'(z)}{g_n(z)}\, dz.$$

Taking the limit as $n \to \infty$ by Theorem 2.4 of Chapter III, we conclude that $\operatorname{ord}_{z_2} g = 0$ also, a contradiction which proves the lemma.

We now come to the main proof. We first make a reduction. We can always find some isomorphism of U with an open subset of the disc. To see this, we use the assumption that there exists some point $\alpha \in \mathbf{C}$ and $\alpha \notin U$. Since U is simply connected, there exists a determination

$$g(z) = \log(z - \alpha)$$

for $z \in U$ which is analytic on U. This function g is injective, for

$$g(z_1) = g(z_2) \quad \Rightarrow \quad e^{g(z_1)} = e^{g(z_2)} \quad \Rightarrow \quad z_1 - \alpha = z_2 - \alpha,$$

whence $z_1 = z_2$. If g takes on some value $g(z_0)$, then

$$g(z) \neq g(z_0) + 2\pi i$$

for all $z \in U$, as one sees again by exponentiating. We claim that there exists a disc around $g(z_0) + 2\pi i$ such that g takes on no value in this disc. Otherwise, there is a sequence $w_n \in U$ such that $g(w_n)$ approaches $g(z_0) + 2\pi i$, and exponentiating shows that w_n approaches z_0, so $g(w_n)$

approaches $g(z_0)$, a contradiction. Hence the function

$$\frac{1}{g(z) - g(z_0) - 2\pi i}$$

is bounded on U, and is analytic injective. By a translation and multiplication by a small positive real number, we may then obtain a function f such that $f(z_0) = 0$ and $|f(z)| < 1$ for all $z \in U$. This function f is injective, and so an isomorphism of U onto an open subset of the disc.

To prove the Riemann mapping theorem, we may now assume without loss of generality that U is an open subset of D and contains the origin. Let Φ be the family of all injective analytic maps

$$f: U \to D$$

such that $f(0) = 0$. This family is not empty, as it contains the identity. Furthermore, the absolute values $|f'(0)|$ for $f \in \Phi$ are bounded. This is obvious from Cauchy's formula

$$2\pi i f'(0) = \int \frac{f(\zeta)}{\zeta^2}\, d\zeta$$

and the uniform boundedness of the functions f, on some small closed disc centered at the origin. The integral is taken along the circle bounding the disc.

Let λ be the least upper bound of $|f'(0)|$ for $f \in \Phi$. Let $\{f_n\}$ be a sequence in Φ such that $|f_n'(0)| \to \lambda$. Picking a subsequence if necessary, Theorem 2.1 implies that we can find a limit function f such that

$$|f'(0)| = \lambda,$$

and $|f(z)| \leqq 1$ for all $z \in U$. Lemma 3.2 tells us that f is injective, and the maximum modulus principle tells us that in fact,

$$|f(z)| < 1 \qquad \text{for all} \quad z \in U.$$

Therefore $f \in \Phi$, and $|f'(0)|$ is maximal in the family Φ. The next result concludes the proof of the Riemann mapping theorem.

Theorem 3.3. *Let $f \in \Phi$ be such that $|f'(0)| \geqq |g'(0)|$ for all $g \in \Phi$. Then f is an analytic isomorphism of U with the disc.*

Proof. The Schwarz lemma for the derivative, Theorem 1.2 of Chapter VII, provides an essential case of the present theorem, and we reduce the proof to that case. All we have to prove is that f is surjective. Suppose not. Let $\alpha \in D$ be outside the image of f. Let T be an automorphism of the disc such that $T(\alpha) = 0$ (for instance, $T(z) = (\alpha - z)/(1 - \bar{\alpha}z)$, but the particular shape is irrelevant here). Then $T \circ f$ is an isomorphism of U onto an open subset of D which does not contain 0, and is simply connected since U is simply connected. We can therefore define a square root function on $T(f(U))$, for instance we can define

$$\sqrt{T(f(z))} = \exp(\tfrac{1}{2} \log T(f(z))) \qquad \text{for} \quad z \in U.$$

Note that the map

$$z \mapsto \sqrt{T(f(z))} \qquad \text{for} \quad z \in U$$

is injective, because if z_1, z_2 are two elements of U at which the map takes the same value, then $Tf(z_1) = Tf(z_2)$, whence $f(z_1) = f(z_2)$ because T is injective, and $z_1 = z_2$ because f is injective. Furthermore $\sqrt{T(f(U))}$ is contained in D, and does not contain 0.

Let R be an automorphism of D which sends $\sqrt{T(f(0))}$ to 0. Then

$$g \colon z \mapsto R(\sqrt{Tf(z)})$$

is an injective map of U into D such that $g(0) = 0$, so $g \in \Phi$. It suffices to prove that $|g'(0)| > |f'(0)|$ to finish the proof. But if we let S be the square function, and we let

$$\varphi(w) = T^{-1}(S(R^{-1}(w))),$$

then

$$f(z) = \varphi(g(z)).$$

Furthermore, $\varphi \colon D \to D$ is a map of D into itself, such that $\varphi(0) = 0$, and φ is not injective because of the square function S. By Theorem 1.2 of Chapter VII (the complement to the Schwarz lemma), we know that $|\varphi'(0)| < 1$. But

$$f'(0) = \varphi'(0)g'(0),$$

so $|g'(0)| > |f'(0)|$, contradicting our assumption that $|f'(0)|$ is maximal. This concludes the proof.

X, §4. BEHAVIOR AT THE BOUNDARY

In Chapter IX, Theorem 2.5, we have seen that the Riemann mapping function extends analytically to a proper analytic arc on the boundary. This result is sufficient for all the applications I know of, but it may be of interest for its own sake to consider the more general situation, so I include a proof of the continuous extension to more general boundaries, basically due to Lindelöf and Koebe. This section should normally be omitted from a course, but provides further reading in analytic techniques.

We investigate the extent to which an isomorphism

$$f: U \to D$$

of U with the disc can be extended by continuity to the boundary of U. There is a standard picture which shows the type of difficulty (impossibility) which can happen if the boundary is too complicated. In Fig. 1, the open set U consists of the interior of the rectangle, from which vertical segments as shown are deleted. These segments have their base point at $1/n$ for $n = 1, 2, \ldots$. It can be shown that for such an open set U there is no way to extend the Riemann mapping function by continuity to the origin, which in some sense is "inaccessible".

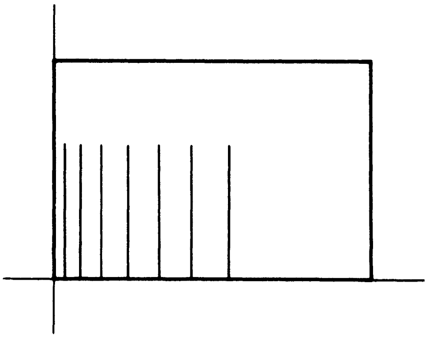


Figure 1

On the other hand, we shall now prove that the mapping can so be extended under more general conditions than before.

Throughout this section, the word **curve** will mean **continuous** curve. No further smoothness is needed, and in fact it is useful to have the flexibility of continuity to work with

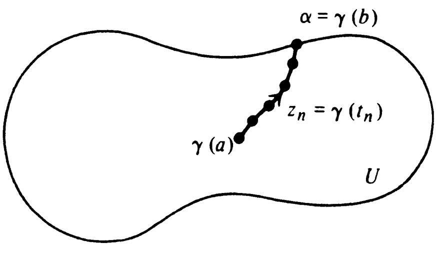


Figure 2

Let U be an open set, and α a boundary point. We say that α is **accessible** if given a sequence of points $z_n \in U$ such that $\lim z_n = \alpha$, there exists a continuous curve

$$\gamma\colon [a, b] \to \mathbf{C}$$

such that $\gamma(b) = \alpha$, $\gamma(t) \in U$ for all $t \neq b$, and there exist $t_n \in [a, b]$ such that $\gamma(t_n) = z_n$ (in other words, the curve passes through the given sequence), and $a < t_1 < t_2 < \cdots$, $\lim t_n = b$. Note that the curve lies entirely in U except for its end point.

Theorem 4.1. *Let U be simply connected and bounded, and let*

$$f\colon U \to D$$

be an isomorphism with the disc. If α is an accessible boundary point of U, then

$$\lim_{z \to \alpha} f(z)$$

exists for $z \in U$, and lies on the unit circle.

Proof. Suppose not. Then there exists a sequence $\{z_n\}$ in U tending to α, but $\{f(z_n)\}$ has no limit. We find a curve γ as in the definition of accessibility.

Lemma 4.2.

$$\lim_{t \to b} |f(\gamma(t))| = 1.$$

Proof. Suppose not. Given ϵ there exists a sequence of increasing numbers s_n such that $|f(\gamma(s_n))| \leqq 1 - \epsilon$, and taking a subsequence if necessary, we may assume $f(\gamma(s_n))$ converges to some w with $|w| \leqq 1 - \epsilon$. Let

$$g\colon D \to U$$

be the inverse function to f. Then

$$\gamma(s_n) = g(f(\gamma(s_n)))$$

and $\gamma(s_n) \to g(w)$, so $\gamma(s_n)$ cannot tend to the boundary of U, a contradiction which proves the lemma.

Since $\{f(z_n)\}$ has no limit, and since the closed disc $\bar{D}$ is compact, there exist two subsequences of $\{z_n\}$, which we will denote by $\{z'_n\}$ and $\{z''_n\}$, such that $f(z'_n)$ tends to a point w' on the unit circle (by the lemma), and $f(z''_n)$ tends to a point $w'' \neq w'$ on the unit circle. Let $z'_n = \gamma(t'_n)$ and $z''_n = \gamma(t''_n)$. Then $t'_n \to b$ and $t''_n \to b$ by assumption. Say

$t_n' \leqq t_n''$. By the uniform continuity of γ, the image $\gamma([t_n', t_n''])$ is uniformly close to $\gamma(b) = \alpha$ for large n. Let γ_n be the restriction of γ to $[t_n', t_n'']$. The image $f(\gamma_n)$ is a curve in D, joining $f(z_n') = w_n'$ to $f(z_n'') = w_n''$, and $f(\gamma_n)$ tends to the unit circle uniformly by the lemma. The picture is as on Fig. 3.

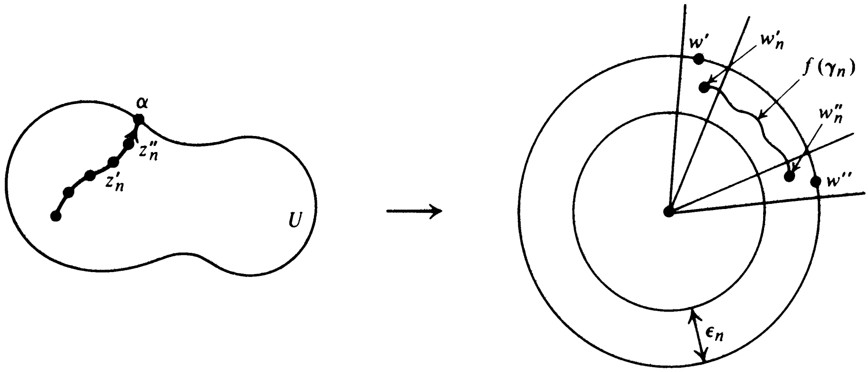


Figure 3

Let us draw rays from the origin to points on the circle close to w', and also two rays from the origin to points on the circle close to w'', as shown on the figure. We take n sufficiently large. Then for infinitely many n, the curves $f(\gamma_n)$ will lie in the same smaller sector. Passing to a subsequence if necessary, we may assume that for all n, the curves $f(\gamma_n)$ lie in the same sector.

As before, we let $g: D \to U$ be the inverse function to f, and we let

$$h = g - \alpha,$$

so $h(w) = g(w) - \alpha$ for $|w| < 1$. Then $h(f(\gamma_n)) \to 0$ as $n \to \infty$. We wish to apply the maximum modulus, but we have to make an auxiliary construction of another function which takes on small values on a curve surrounding the whole circle of radius $1 - \epsilon_n$. This is done as follows. We state a general lemma, due to Lindelöf and Koebe, according to Bieberbach.

Lemma 4.3. *Let h be analytic on the unit disc D, and bounded. Let w', w'' be two distinct points on the circle. Let $\{w_n'\}$ and $\{w_n''\}$ be sequences in the unit disc D converging to w' and w'', respectively, and let ψ_n be a curve joining w_n' with w_n'' such that ψ_n lies in the annulus*

$$1 - \epsilon_n < |w| < 1,$$

and $\epsilon_n \to 0$ as $n \to \infty$. Assume that $r_n > 0$ is a sequence tending to 0 *such that*

$$|h(w)| < r_n \quad \textit{for} \quad w \textit{ on } \psi_n.$$

Then h *is identically* 0.

Proof. Dividing $h(w)$ by some power of w if necessary, we may assume without loss of generality that $h(0) \neq 0$. Without loss of generality, after a rotation, we may assume that

$$w'' = \overline{w'},$$

so w' and w'' are symmetric about the horizontal axis. We pick a large integer M and let L' be the ray having angle $2\pi/2M$ with the x-axis. Let L'' be the ray having angle $-2\pi/2M$ with the axis, as shown on Fig. 4.

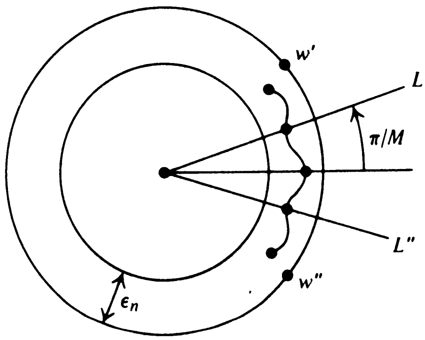


Figure 4

We let ψ_n be defined on an interval $[a_n, b_n]$. Let u_n be the largest value of the parameter such that $\psi_n(u_n)$ is on L', and let v_n be the smallest value of the parameter $> u_n$ such that $\psi_n(v_n)$ lies on the x-axis. Then ψ_n restricted to $[u_n, v_n]$ is a curve inside the sector lying between the x-axis and L', and connecting the point $\psi_n(u_n)$ with the point $\psi_n(v_n)$. If we reflect this curve across the x-axis, then we obtain a curve which we denote by $\overline{\psi}_n$, joining $\psi_n(v_n)$ with $\overline{\psi_n(u_n)}$. We let σ_n be the join of these two curves, so that σ_n is symmetric about the x-axis, and joins $\psi_n(u_n)$ with $\overline{\psi_n(u_n)}$, passing through $\psi_n(v_n)$, as shown on Fig. 5.

Then the beginning and end points of σ_n lie at the same distance from the origin. (This is what we wanted to achieve to make the next step valid.) Let T be rotation by the angle $2\pi/M$. If we rotate σ_n by T iterated M times, i.e. take

$$\sigma_n,\ T\sigma_n,\ \ldots, T^{M-1}\sigma_n,$$

then we obtain a closed curve, lying inside the annulus $1 - \epsilon_n < |w| < 1$.

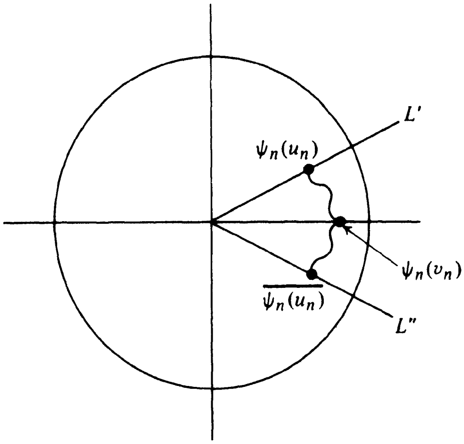


Figure 5

Finally, define the function

$$h^*(w) = h(w)\overline{h(\bar{w})},$$

and the function

$$G(w) = h^*(w)h^*(Tw) \ldots h^*(T^{M-1}w).$$

Let B be a bound for $|h(w)|$, $w \in D$. Each factor in the definition of G is bounded by B^2. For any w on the above closed curve, some rotation $T^k w$ lies on σ_n, and then we have

$$|h^*(w)| \leqq r_n B.$$

Therefore G is bounded on the closed curve by

$$|G(w)| \leqq r_n B^{2M-1}.$$

For each ray L from the origin, let w_L be the point of L closest to the origin, and lying on the closed curve. Let W be the union of all segments $[0, w_L]$ open on the right, for all rays L. Then W is open, and the boundary of W consists of points of the closed curve. By the maximum modulus principle, it follows that

$$|G(0)| \leqq \max|G(w)|,$$

where the max is taken for w on the closed curve. Letting n tend to infinity, we see that $G(0) = 0$, whence $h(0) = 0$, a contradiction which proves the lemma, and therefore also the theorem.

Theorem 4.4. *Let U be bounded. Let $f: U \to D$ be an isomorphism with the disc, and let α_1, α_2 be two distinct boundary points of U which are accessible. Suppose f extended to α_1 and α_2 by continuity. Then*

$$f(\alpha_1) \neq f(\alpha_2).$$

Proof. We suppose $f(\alpha_1) = f(\alpha_2)$. After multiplying f by a suitable constant, we may assume $f(\alpha_i) = -1$. Let

$$g: D \to U$$

again be the inverse function of f. Let γ_1, γ_2 be the curves defined on an interval $[a, b]$ such that their end points are α_1, α_2, respectively, and $\gamma_i(t) \in U$ for $i = 1, 2$ and $t \in [a, b]$, $t \neq b$. There exists a number c with $a < c < b$ such that

$$|\gamma_1(t) - \gamma_2(t)| > \tfrac{1}{2}|\alpha_1 - \alpha_2|, \qquad \text{if} \quad c < t < b,$$

and there exists δ such that

$$f(\gamma_1([a, c])) \qquad \text{and} \qquad f(\gamma_2([a, c]))$$

do not intersect the disc $D(-1, \delta)$ as shown on Fig. 6. Let

$$A(\delta) = D \cap D(-1, \delta).$$

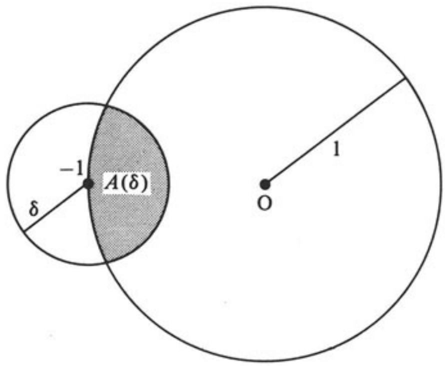


Figure 6

Then $A(\delta)$ is described in polar coordinates by

$$0 \leqq r \leqq \delta \qquad \text{and} \qquad -\varphi(r) \leqq \theta \leqq \varphi(r)$$

with an appropriate function $\varphi(r)$. Note that $\varphi(r) < \pi/2$. We have:

$$\begin{aligned}\text{Area } g(A(\delta)) &= \iint_{A(\delta)} |g'(z)|^2\, dy\, dx \\ &= \int_0^\delta \int_{-\varphi(r)}^{\varphi(r)} |g'(-1 + re^{i\theta})|^2 r\, d\theta\, dr.\end{aligned}$$

For each $r < \delta$ let w_1, w_2 be on $f(\gamma_1)$, $f(\gamma_2)$, respectively, such that

$$|w_i + 1| = r, \qquad i = 1, 2,$$

and

$$|g(w_1) - g(w_2)| > \tfrac{1}{2}|\alpha_1 - \alpha_2|.$$

Then

$$g(w_1) - g(w_2) = \int_{w_1}^{w_2} g',$$

where the integral is taken over the circular arc from w_1 to w_2 in D, with center -1. For $0 < r < \delta$ we get

$$\tfrac{1}{2}|\alpha_1 - \alpha_2| < \int_{-\varphi(r)}^{\varphi(r)} |g'(-1 + re^{i\theta})| r\, d\theta,$$

whence by the Schwarz inequality, we find

$$\frac{|\alpha_1 - \alpha_2|^2}{4\pi r} \leqq r \int_{-\varphi(r)}^{\varphi(r)} |g'(-1 + re^{i\theta})|^2\, d\theta.$$

We integrate both sides with respect to r from 0 to δ. The right-hand side is bounded, and the left-hand side is infinite unless $\alpha_1 = \alpha_2$. This proves the theorem.

The technique for the above proof is classical. It can also be used to prove the continuity of the mapping function at the boundary. Cf. for instance Hurwitz–Courant, Part III, Chapter 6, §4.

CHAPTER XI

Analytic Continuation Along Curves

In this chapter we give further means to extend the domain of definition of an analytic function. We shall apply Theorem 1.2 of Chapter III in the following context. Suppose we are given an analytic function f of an open connected set U. Let V be open and connected, and suppose that $U \cap V$ is not empty, so is open. We ask whether there exists an analytic function g on V such that $f = g$ on $U \cap V$, or only such that $f(z) = g(z)$ for all z in some set of points of $U \cap V$ which is not discrete. The above-mentioned Theorem 1.2 shows that such a function g if it exists is uniquely determined. One calls such a function g a **direct analytic continuation** of f, and we also say that (g, V) is a **direct analytic continuation** of (f, U). We use the word "direct" because later we shall deal with analytic continuation along a curve and it is useful to have an adjective to distinguish the two notions. For simplicity, however, one usually omits the word "direct" if no confusion can result from this omission. If a direct analytic continuation exists as above, then there is is unique analytic function h on $U \cup V$ such that $h = f$ on U and $h = g$ on V.

In the case of Schwarz reflection, the two sets have a common boundary curve. We studied this situation in Chapter IX.

In a final application, we shall put the Riemann mapping theorem together with Schwarz reflection to prove Picard's theorem, which follows conceptually in a few lines, once this more extensive machinery is available.

XI, §1. CONTINUATION ALONG A CURVE

Let D_0 be a disc centered at a point z_0. Let γ be a path whose beginning point is z_0 and whose end point is w, and say γ is defined on the interval $[a, b]$. In this section, smoothness of the path plays no role, and "path" will mean continuous path. Let

$$a = a_0 \leqq a_1 \leqq a_2 \leqq \cdots \leqq a_{n+1} = b$$

be a partition of the interval. Let D_i be a disc containing $\gamma(a_i)$ as shown on Fig. 1. *Instead of discs, we could also take convex open sets.* The intersection of two convex sets is also convex. Actually, what we shall need precisely is that the intersection of a finite number of the sets $D_0, \ldots, D_n$ is connected if it is not empty.

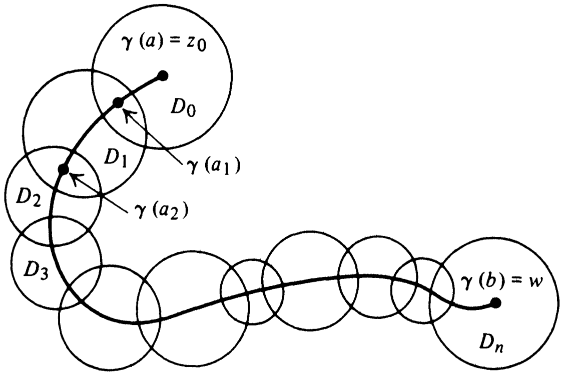


Figure 1

We shall say that the sequence

$$\{D_0, D_1, \ldots, D_n\}$$

is **connected by the curve along the partition** if the image $\gamma([a_i, a_{i+1}])$ is contained in D_i. Then the intersection $D_i \cap D_{i+1}$ contains $\gamma(a_{i+1})$.

Let f_0 be analytic on D_0. By an **analytic continuation of** (f_0, D_0) **along a connected sequence** $[D_0, \ldots, D_n]$ we shall mean a sequence of pairs

$$(f_0, D_0), (f_1, D_1), \ldots, (f_n, D_n)$$

such that (f_{i+1}, D_{i+1}) is a direct analytic continuation of (f_i, D_i) for $i = 0, \ldots, n-1$. This definition appears to depend on the choice of partition and the choice of the connected sequence $\{D_0, \ldots, D_n\}$. We shall prove in the next theorem that it does not depend on these choices. Thus we shall obtain a well-defined analytic function in a neighborhood of the end point of the path, which is called the **analytic continuation of** (f_0, D_0) **along the path** γ. As a matter of notation, we may also denote this function by f_γ.

Theorem 1.1. *Let* $(g_0, E_0), \ldots, (g_m, E_m)$ *be another analytic continuation of* (g_0, E_0) *along a connected sequence* $\{E_0, \ldots, E_m\}$ *with respect to a* *partition of the path γ. If $f_0 = g_0$ in some neighborhood of z_0, then $g_m = f_n$ in some neighborhood of $\gamma(b)$, so (g_m, E_m) is a direct analytic continuation of (f_n, D_n).*

Proof. The proof that the analytic continuation does not depend on the choices of partition and discs or convex sets will be similar to that used in Chapter III, §4, when we dealt with the integral along a path. The proof here is equally easy and straightforward.

Suppose first that the partition is fixed, and let

$$(g_0, E_0), \ldots, (g_n, E_n)$$

be an analytic continuation along another connected sequence

$$\{E_0, \ldots, E_n\}.$$

Suppose $g_0 = f_0$ in a neighborhood of z_0, which means that (g_0, E_0) is a direct analytic continuation of (f_0, D_0). Since $D_0 \cap E_0$ is connected it follows that $f_0 = g_0$ on the whole set $D_0 \cap E_0$, which also contains z_1. By hypothesis, $f_1 = f_0$ on $D_0 \cap D_1$ and $g_1 = g_0$ on $E_0 \cap E_1$. Hence $f_1 = g_1$ on $D_0 \cap E_0 \cap D_1 \cap E_1$ which contains z_1. Since $D_1 \cap E_1$ is connected, it follows that $f_1 = g_1$ on $D_1 \cap E_1$. Thus (g_1, E_1) is a direct analytic continuation of (f_1, D_1). We can now proceed by induction to see that (g_n, E_n) is a direct analytic continuation of (f_n, D_n), thus concluding the proof in this case.

Next let us consider a change in the partition. Any two partitions have a common refinement. To show the independence of the partition it suffices to do so when we insert one point in the given partition, say we insert c in the interval $[a_k, a_{k+1}]$ for some k. On one hand, we take the connected sequence

$$\{D_0, \ldots, D_k, D_k, \ldots, D_n\},$$

where D_k is repeated twice, so that $\gamma([a_k, c]) \subset D_k$ and $\gamma([c, a_{k+1}]) \subset D_k$. Then

$$(f_0, D_0), \ldots, (f_k, D_k), (f_k, D_k), \ldots, (f_n, D_n)$$

is an analytic continuation of (f_0, D_0) along this connected sequence. On the other hand, suppose that

$$(g_0, E_0), \ldots, (g_k, E_k), (g_k^*, E_k^*), \ldots, (g_n, E_n)$$

is an analytic continuation of (g_0, E_0) along another connected sequence $\{E_0, \ldots, E_k, E_k^*, \ldots, E_n\}$ with respect to the new partition, and $g_0 = f_0$ in some neighborhood of z_0. By the first part of the proof, we know that (g_k, E_k) is a direct analytic continuation of (f_k, D_k). By hypothesis, g_k^* is equal to g_k on $E_k \cap E_k^*$, $f_k = g_k$ on $D_k \cap E_k$, so

$$g_k^* = f_k \quad \text{on} \quad D_k \cap E_k \cap E_k^*, \qquad \text{which contains} \quad z_k^* = \gamma(c).$$

Therefore (g_k^*, E_k^*) is a direct analytic continuation of (f_k, D_k). Again we can apply the first part of the proof to the second piece of the path which is defined on the interval $[a_{k+1}, a_{n+1}]$, with respect to the partition $[a_{k+1}, a_{k+2}, \ldots, a_{n+1}]$ to conclude the proof of the theorem.

Example. Let us start with the function $\log z$ defined by the usual power series on the disc D_0 which is centered at 1 and has radius < 1 but > 0. Let the path be the circle of radius 1 oriented counterclockwise as usual. If we continue $\log z$ along this path, and let (g, D) be its continuation, then near the point 1 it is easy to show that

$$g(z) = \log z + 2\pi i.$$

Thus g differs from f_0 by a constant, and is not equal to f_0 near $z_0 = 1$.

Example. Let

$$f(z) = e^{(1/2)\log z},$$

where f is defined in a neighborhood of 1 by the principal value for the log. [We could write $f(z) = \sqrt{z}$ in a loose way, but the square root sign has the usual indeterminacy, so it is meaningless to use the expression $\sqrt{z}$ for a function unless it is defined more precisely.] We may take the analytic continuation of f along the unit circle. The analytic continuation of $\log z$ along the circle has the value

$$\log z + 2\pi i$$

near 1. Hence the analytic continuation of f along this circle has the value

$$g(z) = e^{(1/2)(\log z + 2\pi i)} = -e^{(1/2)\log z} = -f(z)$$

for z near 1. This is the other solution to the equation

$$f(z)^2 = z$$

near 1. If we continue g analytically once more around the circle, then we obtain $f(z)$, the original function.

Remark. In some texts, one finds a picture of the "Riemann surface" on which "$f(z) = \sqrt{z}$" is defined, and that picture represents two sheets crossing themselves, and looking as if there was some sort of singularity at the origin. **It should be emphasized that the picture is totally and irretrievably misleading**. The proper model for the domain of definition of $f(z)$ is obtained by introducing another plane, so that the correspondence between $f(z)$ and z is represented by the map

$$\mathbf{C} \to \mathbf{C}$$

given by $\zeta \mapsto \zeta^2$. The association $\zeta \mapsto \zeta^2$ gives a double covering of $\mathbf{C}$ by $\mathbf{C}$ at all points except the origin (it is the function already discussed in Chapter I, §3). On every simply connected open set U not containing 0, one has the inverse function $\zeta = z^{1/2}$, $z \in U$.

More generally, let P be a polynomial in two variables, $P = P(T_1, T_2)$, not identically zero. A solution (analytic) $f(z)$ of the equation

$$P(f(z), z) = 0$$

for z in some open set is called an **algebraic function**. It can be shown that if we delete a finite number of points from the plane, then such a solution f can be continued analytically along every path. Furthermore, we have the following theorem.

Theorem 1.2. *Let $P(T_1, T_2)$ be a polynomial in two variables. Let γ be a curve with beginning point z_0 and end point w. Let f be analytic at z_0, and suppose that f has an analytic continuation f_γ along the curve γ. If*

$$P(f(z), z) = 0 \qquad \text{for} \quad z \text{ near } z_0,$$

then

$$P(f_\gamma(z), z) = 0 \qquad \text{for} \quad z \text{ near } w.$$

Proof. This is obvious, because the relation holds in each successive disc $D_0, D_1, \ldots, D_m$ used to carry out the analytic continuation.

Theorem 1.3 (Monodromy Theorem). *Let U be a connected open set. Let f be analytic at a point z_0 of U, and let γ, η be two paths from z_0 to a point w of U. Assume:*

(i) *γ is homotopic to η in U.*
(ii) *f can be continued analytically along any path in U.*

Let f_γ, f_η be the analytic continuations of f along γ and η, respectively. Then f_γ and f_η are equal in some neighborhood of w.

Proof. The proof follows the ideas of Lemma 4.3 and Theorem 5.1 of Chapter III. Let

$$\psi: [a, b] \times [c, d] \to U$$

be a homotopy so that if we put $\gamma_u(t) = \psi(t, u)$, then $\gamma = \gamma_c$ and $\eta = \gamma_d$.

Then we have:

Lemma 1.4. *For each $u \in [c, d]$, if u' is sufficiently close to u, then a continuation of f along γ_u is equal to a continuation of f along $\gamma_{u'}$ in some neighborhood of w.*

Proof. Given a continuation along a connected sequence of discs or convex open sets

$$(f_0, D_0), \ldots, (f_n, D_n)$$

along γ_u, connected by the curve along the partition, it is immediately verified that if u' is sufficiently close to u, then this is also a continuation along $\gamma_{u'}$ (by the uniform continuity of the homotopy ψ). In Theorem 1.1, we have seen that the analytic continuation along a curve does not depend on the choice of partition and $D_0, \ldots, D_n$. This proves the lemma.

To finish the proof, define two paths, α, β from z_0 to w to be **f-equivalent** if f_α and f_β are equal in some neighborhood of w. Let S be the set of points $u \in [c, d]$ such that γ_c and γ_u are f-equivalent. We claim that $S = [c, d]$, which implies the theorem. Note that $c \in S$ so S is not empty. By Lemma 1.4, S is open in $[c, d]$. We show that S is closed in $[c, d]$. Let $u \in [c, d]$ be in the closure of S. Then there exist points $u' \in S$ close to u, so again by Lemma 1.4, γ_u is f-equivalent to $\gamma_{u'}$, so γ_u is f-equivalent to γ_c. Since $[c, d]$ is connected, it follows that $S = [c, d]$, as was to be shown.

Remark. The proof of Theorems 5.1 and 5.2 in Chapter III could also have been formulated in terms of an equivalence, namely that the integrals of f along the two paths are equal. Lemmas 4.3 and 4.4 of Chapter III play the role of Lemma 1.4 used here. We presented both variations of the argument to illustrate different intuitions of the situation.

Theorem 1.5. *Let U be a simply connected open set. Let $z_0 \in U$, and let f be analytic at z_0. Assume that f can be continued along any path from z_0 to any point in U. Let γ_z be a path from z_0 to a point z in U, and let $f_{\gamma_z}(z)$ be the analytic continuation of f along this path. Then $f_{\gamma_z}(z)$ is independent of the path from z_0 to z, and the association $z \mapsto f_{\gamma_z}(z)$ defines an analytic function on U.*

Proof. Suppose we have shown the independence from the path. Let z_1 be some point in U and let z be a variable point in a disc centered at z_1. Then the analytic continuation of f from z_0 to z may be first taken from z_0 to z_1 along some path, and then we take the continuation on the disc centered at z_1, which shows the analyticity. So it remains to deal with the independence from the path, which is a consequence of the following lemma, combined with Theorem 1.3.

Lemma 1.6. *Let U be simply connected. Let γ, η be two paths in U from a point z_0 to a point z_1. Then there is a homotopy in U between the two paths, leaving the end points z_0, z_1 fixed.*

Proof. The arguments are routine. Cf. my *Undergraduate Analysis*, Second Edition, Springer-Verlag, 1997, Chapter XVI, §6, especially Theorem 6.4 and Proposition 6.6. These arguments do not involve complex analysis but merely juggling with homotopies which require being written down, or at least being clearly shown on pictures. We let readers look the matter up in the above reference, or work it out for themselves as an exercise.

We now return to the analytic continuation, and the use of discs or other sets as in Lemma 1.4. Although for the general theory one thinks of $D_0, \ldots, D_n$ as small discs, because the open set U may have a complicated shape and the curve may wind around a lot, in practice the set U may be much simpler.

Example. Let U be obtained from $\mathbf{C}$ by deleting the origin. Let f be an analytic function on the upper half plane. We want to continue the function along all curves in U. Then we might choose only four sets E_0, E_1, E_2, E_3, which are the upper half plane, left half plane, lower half plane, and right half plane, respectively. The intersection of two successive sets is a quarter plane.

Example. Let g be an analytic function on a connected open set U. We do not assume that U is simply connected. Let $z_0 \in U$. Let f_0 be a primitive of g in some disc containing z_0, which we may choose to be

$$f_0(z) = \int_{z_0}^{z} g(\zeta)\, d\zeta,$$

where the integral is taken along any path from z_0 to z inside the disc. Then f_0 can be analytically continued along any path in U, essentially by integration along the path. For instance let γ be a path, and let z_1 be a point on γ. Say

$$\gamma\colon [a, b] \to U, \qquad \text{with} \quad \gamma(a) = z_0,$$

and $z_1 = \gamma(c)$ with $a < c \leqq b$. Then we can define f_1 in a neighborhood of z_1 by taking the integral of g along the path from z_0 to z_1, that is along the restriction of γ to the interval $[a, c]$, and then from z_1 to z along any path contained in a small disc D containing z_1, as illustrated in Fig. 2.

In Fig. 2, we have drawn U with a shaded hole, so U is not simply connected. The monodromy theorem applies to this situation.

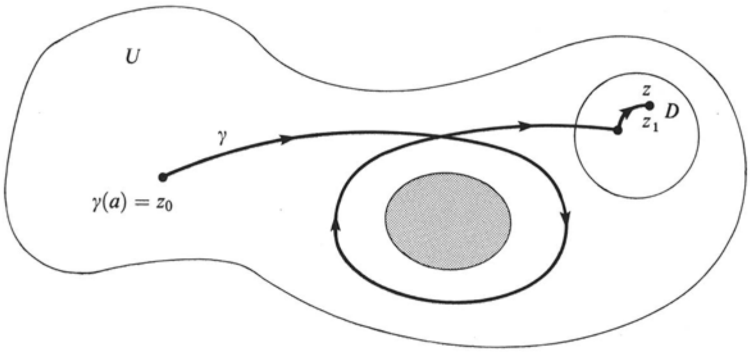


Figure 2

A concrete case can be given as follows. Let $R(z)$ be a rational function, that is a quotient of polynomials. We can factor the numerator and denominator into linear factors, say

$$R(z) = \frac{(z - a_1) \cdots (z - a_n)}{(z - b_1) \cdots (z - b_m)}.$$

Then R has poles at $b_1, \ldots, b_m$ (which we do not assume all distinct). Let U be the plane from which the numbers $b_1, \ldots, b_m$ are deleted. Then U is connected but not simply connected if $m \geqq 1$. We can define the integral

$$f(z) = \int_{z_0, \gamma}^{z} R(\zeta)\, d\zeta,$$

along a path from z_0 to z. This integral depends on the path, and in fact depends on the winding numbers of the path around the points $b_1, \ldots, b_m$.

Example. The preceding example can be generalized, we do not really need g to be analytic on U; we need only something weaker. For instance, let U be the plane from which the three points 1, 2, 3 have been deleted. Then for any simply connected open set V in U, there is a function

$$[(z-1)(z-2)(z-3)]^{1/2},$$

i.e. a square root of $(z-1)(z-2)(z-3)$, and any two such square roots on V differ by a sign. Starting from a point z_0 in U, let

$$g_0(z) = \frac{1}{[(z-1)(z-2)(z-3)]^{1/2}},$$

where the square root is one of the possible determinations analytic in a neighborhood of z_0. Then we can form the integral

$$\int_{z_0}^{z} g_0(\zeta)\, d\zeta,$$

along any path in U from z_0 to z. The integral will depend on the path, but the mondromy theorem applies to this situation as well.

We end this section with an important special case of the monodromy theorem, which is the one to be used in §3.

Theorem 1.7. *Let U be a simply connected open set. Let f be analytic at a point z_0 and assume that f can be continued along every path from z_0 to every point in U. Then the analytic continuation of f along a path from z_0 to a point w is independent of the path, and defines an analytic function on all of U.*

Proof. This is a special case of Theorem 1.3, because two paths from z_0 to w are homotopic.

XI, §1. EXERCISES

1. Let f be analytic in the neighborhood of a point z_0. Let k be a positive integer, and let $P(T_1, \dots, T_k)$ be a polynomial in k variables. Assume that

$$P(f, Df, \dots, D^k f) = 0,$$

where $D = d/dz$. If f can be continued along a path γ, show that

$$P(f_\gamma, Df_\gamma, \dots, D^k f_\gamma) = 0.$$

2. **(Weierstrass).** Prove that the function

$$f(z) = \sum z^{n!}$$

cannot be analytically continued to any open set strictly larger than the unit disc. [*Hint*: If z tends to 1 on the real axis, the series clearly becomes infinite. Rotate z by a k-th root of unity for positive integers k to see that the function becomes infinite on a dense set of points on the unit circle.]

3. Let U be a connected open set and let u be a harmonic function on U. *Let D* be a disc contained in U and let f be an analytic function on D such that $u = \operatorname{Re}(f)$. Show that f can be analytically continued along every path in U.

4. Let U be a simply connected open set in $\mathbf{C}$ and let u be a real harmonic function on U. Reprove that there exists an analytic function f on U such that $u = \operatorname{Re}(f)$ by showing that if (f_0, D_0) is analytic on a disc and $\operatorname{Re}(f_0) = u$

on the disc, then (f_0, D_0) can be continued along every curve in U. [*Note*: The result was proved as Theorem 3.1 of Chapter VIII.]

XI, §2. THE DILOGARITHM

For $|z| < 1$ we define the function

$$f(z) = -\frac{\log(1-z)}{z} = \sum_{n=1}^{\infty} \frac{z^{n-1}}{n}.$$

This function is holomorphic at 0. Let now $z \neq 0$ and let γ be a path in $\mathbf{C} - \{0, 1\}$ except for the beginning point at 0. We then have an analytic continuation $f_\gamma(z)$, which depends on γ. Since the continuation of $\log(1-z)$ picks up a constant of integration after integrating around loops around 1, it follows that the analytic continuation of f may have a pole at 0, with residue an integral multiple of $2\pi i$. Instead of integrating along paths starting at 0, we can use a slightly arbitrary device of picking some point, say 1/2, and defining f by analytic continuation from 1/2 to z, along paths γ from 1/2 to z in $\mathbf{C} - \{0, 1\}$.

The **dilogarithm** is defined for $|z| < 1$ by the integral and power series

$$L_2(z) = \int_0^z f(\zeta)\, d\zeta = \sum_{n=1}^{\infty} \frac{z^n}{n^2},$$

and is defined by analytic continuation in general, so we get a function $L_{2,\gamma}(z)$ for each path γ as above. We may write

$$L_{2,\gamma}(z) = -\int_{0,\gamma}^{z} \log_\gamma (1-\zeta) \frac{d\zeta}{\zeta}.$$

Let U be the simply connected set obtained by deleting the set of real numbers $\geqq 1$ from $\mathbf{C}$. Then for $z \in U$, the function $f_\gamma(z)$ is independent of the path γ in U, and may still be denoted by f. It is analytic.

However, we are interested in the analytic continuation $L_{2,\gamma}$ in general, and whatever problem arises has to do with the argument, which we now work out.

Theorem 2.1. *For $z \neq 0, 1$ the function*

$$z \mapsto D_\gamma(z) = \operatorname{Im} L_{2,\gamma}(z) + \arg_\gamma(1-z) \log|z|$$

is independent of the path γ in $\mathbf{C} - \{0, 1\}$.

Sketch of Proof. We proceed in steps.

First, if γ is homotopic to η in $\mathbf{C} - \{0, 1\}$ then $D_\gamma = D_\eta$.

Indeed, we have separately that $\arg_\gamma(1 - z) = \arg_\eta(1 - z)$ and

$$L_{2,\gamma}(z) = L_{2,\eta}(z)$$

because we can reduce these equalities to the case when the paths are close together, and to a local question in a small disc as usual.

Second, if η differs from γ by a loop winding once around 1, *then*

$$\arg_\eta(1 - z) = \arg_\gamma(1 - z) + 2\pi.$$

This is immediate.

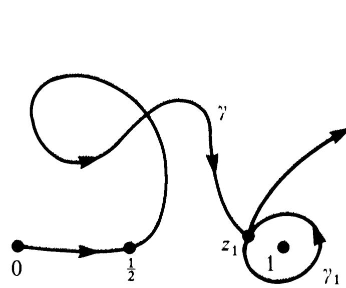


Figure 3

Third, suppose that η is homotopic to a path as shown in Fig. 3, so again differs from γ by a loop winding once around 1. *We prove that* $D_\eta = D_\gamma$ by a computation of the integral. We may assume that γ_1 is contained in a small disc centered at 1. Then using the *principal branch* of $(\log(1 - \zeta))/\zeta$ we get:

$$\begin{aligned}
\int_{\gamma_1} \frac{\log(1 - \zeta)}{\zeta}\, d\zeta &= \int_{\gamma_1} \frac{\log(1 - \zeta)}{1 - (1 - \zeta)}\, d\zeta \\
&= \int_{\gamma_1} \log(1 - \zeta) \sum_{n=0}^{\infty} (1 - \zeta)^n\, d\zeta \\
[\text{by parts}] &= -\sum_{n=0}^{\infty} \frac{(1 - z)^{n+1}}{n + 1} \log(1 - z)\Bigg|_{z_1,\gamma_1}^{z_1} + \underbrace{\int_{\gamma_1} \sum_{n=0}^{\infty} \frac{(1 - \zeta)^n}{n + 1}\, d\zeta}_{=0} \\
&= \log(z)\log(1 - z)\Bigg|_{z_1,\gamma_1}^{z_1} \\
&= (2\pi i)\log_{\mathrm{pr}}(z_1) \text{ (where } \log_{\mathrm{pr}} \text{ is the principal value of log).}
\end{aligned}$$

Two analytic continuation of $(\log(1-\zeta)/\zeta)$ in a neighborhood of 1 differ by $m2\pi i/\zeta$ for some integer m. Further, as we have seen, the argument of any branch of $\log(1-\zeta)$ changes by 2π after continuation along γ_1. Therefore

$$\begin{aligned}\int_\eta \frac{\log(1-\zeta)}{\zeta}\,d\zeta &= \int_\gamma \frac{\log(1-\zeta)}{\zeta}\,d\zeta + \int_{z_1,\gamma}^{z} (2\pi i)\frac{d\zeta}{\zeta} + (2\pi i)\log_{\mathrm{pr}}(z_1)\\ &= -L_{2,\gamma}(z) + (2\pi i)(\log|z| + m2\pi i) \qquad \text{with some } m \in \mathbf{Z}\\ &= -L_{2,\gamma}(z) + (2\pi i)\log|z| - (2\pi)^2 m.\end{aligned}$$

But the left-hand side of this equality is $-L_{2,\eta}(z)$, so the proof of Theorem 2.1 is concluded, since $(2\pi)^2 m$ is real.

The case when the two paths η and γ differ by a loop around 0 can be carried out similarly. We omit this. We also omit a complete proof that any path is homotopic to one which is obtained by a combination of the above two types. We merely wanted to give a non-trivial example of analytic continuation along a path.

The function $D(z)$ is called the **Bloch–Wigner function**. The dilogarithm occurs in classical analysis, see for instance [Co 35] and [Ro 07]. The function D, which does not depend on the path, was discovered by David Wigner, and used extensively by Spencer Bloch and others in algebraic geometry and number theory, in fancy contexts. See, for instance, [Bl 77], [Za 88], [Za 90], [Za 91].

BIBLIOGRAPHY

[Bl 77] S. Bloch, Applications of the dilogarithm function in algebraic K-theory and algebraic geometry, *Intl. Symp. on Algebraic Geometry*, Kyoto, 1977 pp. 103–114

[Co 35] H.S.M. Coxeter, The functions of Schlafli and Lobatschevsky, *Quart. J. Math.* **6** (1935) pp. 13–29

[La 82] S. Lang, Units and class groups in number theory and algebraic geometry, Appendix, *Bull. Math. Soc.* **6** No. 3 (1982) pp. 307–312

[Le 81] L. Lewin, *Polylogarithms and associated functions*, North-Holland, Amsterdam, 1981

[Ro 07] L.J. Rogers, On function sums connected with the series $\sum(x^n/n^2)$, *Proc. London Math. Soc.*, **4-2** (1907) pp. 169–189

[Za 88] D. Zagier, The remarkable dilogarithm, *J. Math. Phys. Sci.* **22** (1988) pp. 131–145

[Za 90] D. Zagier, The Bloch–Wigner–Ramakrishnan polylogarithm function, *Math. Ann.* **286** (1990) pp. 613–624

[Za 91] D. Zagier, Polylogarithms, Dedekind zeta functions, and the algebraic K-theory of fields, in *Arithmetic Algebraic Geometry* (eds. G. van der

Geer, F. Oort, J. Steenbrink). *Progr. in Math.* **89**, Birkhauser, Boston (1991) pp. 391–430

XI, §2. EXERCISES

1. We investigate the analytic continuation of the dilogarithm for the curves illustrated in Fig. 4. Let $z_1 = 1/2$.

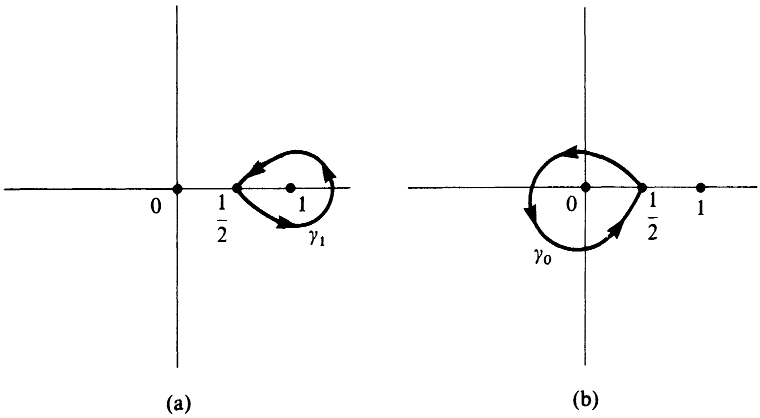


Figure 4

(a) Let γ_1 be a curve as shown on Fig. 4(a), circling 1 exactly once. How does the analytic continuation of L_2 along γ_1 differ from L_2 in a neighborhood of z_1?
(b) How does the analytic continuation of L_2 along the path γ_0 on Fig. 4(b) differ from L_2?
(c) If you continue L_2 first around γ_0 and then around γ_1, how does this continuation differ from continuing L_2 first around γ_1 and then around γ_0? [They won't be equal!]

2. Let D be the Bloch–Wigner function.
(a) Show that $D(1/z) = -D(z)$ for $z \neq 0$, and so D extends in a natural way to a continuous function on $\mathbf{C} \cup \{\infty\}$.
(b) Show that $D(z) = -D(1-z)$.

3. For $k \geqq 2$ define the **polylogarithm function**

$$L_k(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^k} \qquad \text{for } |z| < 1.$$

Show that for every positive integer N,

$$L_k(z^N) = N^{k-1} \sum_{\zeta^N = 1} L_k(\zeta z)$$

where the sum is taken over all N-th roots of unity ζ. [*Hint*: Observe that if ζ

is an N-th root of unity, $\zeta \neq 1$, then

$$1 + \zeta + \cdots + \zeta^{N-1} = \frac{1 - \zeta^N}{1 - \zeta} = 0.]$$

4. Prove the relation

$$\prod_{\zeta^N = 1} (1 - \zeta X) = 1 - X^N,$$

where the product is taken over all N-th roots of unity ζ.

For other examples of such relations, see Chapter XV, formula Γ **8** of §2, and the exercise of Chapter XV, §3.

XI, §3. APPLICATION TO PICARD'S THEOREM

We return to the Schwarz reflection across circular triangles in the disc as in Chapter IX, §3, and we follow the notation of that section. Thus we have inverse isomorphisms

$$g\colon U \to H \qquad \text{and} \qquad h\colon H \to U,$$

where U is the triangle. We could continue g analytically by reflection, but vice versa, one may continue h analytically across each interval, by reflection across the sides of the triangle. If h_γ denotes the analytic continuation of h along any path γ not passing through 0, 1, then for any z on the real line, $z \neq 0$, 1 we have $h_\gamma(z)$ on the side of some iterated reflection of the original triangle. For any complex number $z \neq 0$, 1 the value $h_\gamma(z)$ lies inside the unit disc D.

Picard's Theorem. *Let f be an entire function whose values omit at least two complex numbers. Then f is constant.*

Proof. After composing f with a linear map, we may assume that the omitted values are 0, 1. Without loss of generality, we may assume that there is some z_0 such that $f(z_0)$ lies in the upper half plane. (Otherwise, we would proceed in a similar manner relative to the lower half plane.) The analytic function $h(f(z))$ for z near z_0 maps a neighborhood of z_0 inside U. We may continue this function $h \circ f$ analytically along any path in $\mathbf{C}$, because f may be so continued, and the image of any path in $\mathbf{C}$ under f is a path which does not contain 0 or 1. Since $\mathbf{C}$ is simply connected, the analytic continuation of $h \circ f$ to $\mathbf{C}$ is then well defined on all of $\mathbf{C}$, and its values lie in the unit disc, so are bounded. By Liouville's theorem, we conclude that $h \circ f$ is constant, whence f is constant, which proves the theorem.

Remark. The crucial step in the proof was the use of the functions g and h. In classical times, they did not like functions defined essentially in an "abstract nonsense" manner. The situation of the triangle is so concrete that one might prefer to exhibit the desired functions explicitly. This can be done, and was done in Picard's original proof. He used the standard "modular function" denoted classically by λ. Its properties and definition follow easily from the theory of elliptic functions. Cf. for instance my book on *Elliptic Functions*, [La 73] Chapter 18, §5. It is usually easier to deal with the upper half plane rather than the unit disc in such concrete situations, because the upper half plane is the natural domain of definition of the classical modular functions.

PART THREE

Various Analytic Topics

This final part deals with various topics which are not all logically interrelated. For the most part, these topics involve analytic estimates, and the study of entire or meromorphic functions, and their rates of growth such as Jensen's formula. We give examples of special functions, for instance, elliptic functions, the gamma function, and the zeta function. I recommend the classical British texts for other special functions (e.g. Bessel functions, hypergeometric functions), including Copson, Titchmarsh, and Whittaker–Watson, listed in the Bibliography at the end of this book.

The final chapter on the prime number theorem can be used for supplementary reading, partly because of the intrinsic interest of the theorem itself, and partly because it shows how complex analysis connects with number theory. In particular, the proof of the prime number theory involves a very striking application and example of the calculus of residues.

CHAPTER XII

Applications of the Maximum Modulus Principle and Jensen's Formula

We return to the maximum principle in a systematic way, and give several ways to apply it, in various contexts.

We begin the chapter by some estimates of an analytic function in terms of its zeros. Then we give the more precise exact relationship of Jensen's formula. This formula has a remarkable application to an old proof by Borel of Picard's theorem that an entire function omits at most one value. This application has been overlooked for many decades. I think it should be emphasized in light of the renewed importance of extensions of Jensen's formula by Nevanlinna.

The remainder of the chapter consists of varied applications, somewhat independent of each other and in various contexts.

One of the most striking applications omitted from standard courses, is to the problem of transcendence: Given some analytic function, describe those points z such that $f(z)$ is an integer, or a rational number, or an algebraic number. This type of question first arose at the end of the nineteenth century, among analysts in Weierstrass' school (Stäckel, Strauss). The question was again raised by Pólya, and Gelfond recognized the connection with the transcendence problems. See the books on transcendence by, for instance, Baker, Gelfond, Lang, Schneider, in the Bibliography. In order not to assume too much for the reader, we shall limit ourselves in this chapter to the case when the function takes on certain values in the rational numbers. The reader who knows about algebraic numbers will immediately see how to extend the proof to that case.

XII, §1. JENSEN'S FORMULA

The first result gives an inequality involving the zeros of a function and its maximum modulus.

Theorem 1.1. *Let f be holomorphic on the closed disc of radius R, and assume that $f(0) \neq 0$. Let the zeros of f in the open disc be ordered by increasing absolute value,*

$$z_1, z_2, \ldots, z_N,$$

each zero being repeated according to its multiplicity. Then

$$|f(0)| \leqq \frac{\|f\|_R}{R^N} |z_1 z_2 \cdots z_N|.$$

Proof. Let

$$g(z) = \prod_{n=1}^{N} \frac{R(z_n - z)}{R^2 - \bar{z}_n z} \quad \text{and} \quad F(z) = \frac{f(z)}{g(z)}.$$

Then the function F is holomorphic on the closed disc of radius R, and

$$|F(z)| = |f(z)| \qquad \text{when} \quad |z| = R.$$

Hence the maximum modulus principle implies that

$$|F(z)| \leqq \|f\|_R \qquad \text{for} \quad |z| \leqq R,$$

and the theorem follows by putting $z = 0$.

The theorem is known as **Jensen's inequality**. Let $\nu(r)$ be the number of zeros of f in the closed disc $|z| \leqq r$. Then

$$\sum_{n=1}^{N} \log(R/|z_n|) = \sum_{n=1}^{N} \int_{|z_n|}^{R} \frac{dx}{x} = \int_0^R \frac{\nu(x)}{x} dx.$$

Hence we obtain another formulation of **Jensen's inequality**,

$$\boxed{\int_0^R \frac{\nu(x)}{x} dx \leqq \log\|f\|_R - \log|f(0)|.}$$

In many applications, the estimate given by Jensen's inequality suffices, but it is of interest to get the exact relation which exists between the zeros of f and the mean value on the circle.

We begin with preliminary remarks. Let us assume that f is meromorphic on the closed disc $\bar{D}_R$ (centered at the origin). Then f has only a finite number of zeros and poles on this disc. For $a \in \bar{D}_R$ we abbreviate

$$n_f(a) = \operatorname{ord}_a(f).$$

Observe that $n_f(a) = 0$ except for the finite number of possibilities that f has a zero of pole at a. Thus we may form the sum

$$\sum_{\substack{a \in D_R \\ a \neq 0}} n_f(a) \log \frac{R}{|a|},$$

where the sum is taken over all elements $a \in D_R$ and $a \neq 0$. The sum is really a finite sum, since only a finite number of terms are $\neq 0$.

We omitted the term with $a = 0$ because we cannot divide by zero, so we have to deal with this term separately. For this purpose, we consider the power series expansion of f at 0, and we let c_f be its leading coefficient:

$$f(z) = c_f z^m + \text{higher terms}.$$

We assume f is not constant, and so $c_f \neq 0$. Thus f has order m at 0, and we have $n_f(0) = m$. If f has no zero or pole at 0, then $c_f = f(0)$.

Even though f may have zeros or poles on the circle of radius R, we shall prove in a subsequent lemma that the improper integral

$$\int_0^{2\pi} \log|f(Re^{i\theta})|\, d\theta$$

converges. We may then state the main theorem of this section.

Theorem 1.2 (Jensen's Formula). *Let f be meromorphic and not constant on the closed disc $\bar{D}_R$. Then*

$$\int_0^{2\pi} \log|f(Re^{i\theta})| \frac{d\theta}{2\pi} + \sum_{\substack{a \in D_R \\ a \neq 0}} n_f(a) \log \frac{|a|}{R} + n_f(0) \log \frac{1}{R} = \log|c_f|.$$

In particular, if f has no zero or pole at 0, *then the right side is equal to the constant* $\log|f(0)|$.

Proof. Suppose first that f has no zeros or poles in the closed disc $|z| \leqq R$. Then $\log f(z)$ is analytic on this disc, and

$$\log f(0) = \frac{1}{2\pi i} \int_{|z|=R} \frac{\log f(z)}{z}\, dz = \int_0^{2\pi} \log f(Re^{i\theta}) \frac{d\theta}{2\pi}.$$

Taking the real part proves the theorem in this case.

Next, suppose $f(z) = z$. Then $\log|f(Re^{i\theta})| = \log R$ and $n_f(0) = 1$, so that the left side of Jensen's formula is equal to 0. But $c_f = 1$, so the right side of Jensen's formula is 0 also, and the formula is true.

Let α be a complex number $\neq 0$. We shall now prove Jensen's formula for the function

$$f(z) = z - \alpha, \qquad \text{where} \quad \alpha = ae^{i\varphi},$$

with $0 < a \leqq R$. In this case, $|f(0)| = |\alpha| = a$, and the formula to be proved is

$$\log R = \frac{1}{2\pi}\int_0^{2\pi} \log|Re^{i\theta} - ae^{i\varphi}|\,d\theta.$$

It will therefore suffice to prove:

Lemma 1.3. *If* $0 < a \leqq R$, *then* $\displaystyle\int_0^{2\pi} \log\left|e^{i\theta} - \frac{a}{R}e^{i\varphi}\right| d\theta = 0.$

Proof. Suppose first $a < R$. Then the function

$$\frac{\log\left(1 - \dfrac{a}{R}z\right)}{z}$$

is analytic for $|z| \leqq 1$, and from this it is immediate that the desired integral is 0. Next suppose $a = R$, so we have to prove that

$$(*) \qquad \int_0^{2\pi} \log|1 - e^{i\theta}|\,d\theta = 0.$$

To see this we note that the left-hand side is the real part of the complex integral

$$\int_C \log(1 - z)\frac{dz}{iz},$$

taken over the circle of radius 1, $z = e^{i\theta}$, $dz = ie^{i\theta}$. Let $\gamma(\epsilon)$ be the contour as shown on Fig. 1.

For $\operatorname{Re} z < 1$ the function

$$\frac{\log(1 - z)}{z}$$

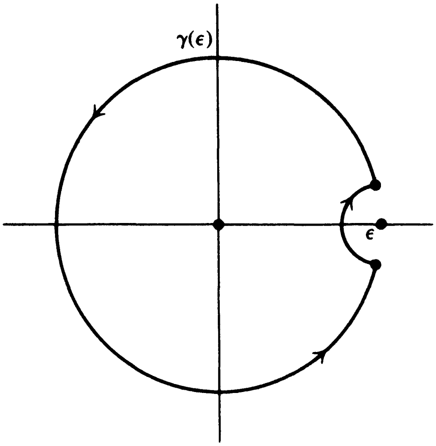


Figure 1

is holomorphic, and so the integral is equal to 0:

$$\int_{\gamma(\epsilon)} \frac{\log(1-z)}{z} dz = 0.$$

The integral over the small indented circular arc of radius ϵ is bounded in absolute value by a constant times $\epsilon \log \epsilon$, which tends to 0 as ϵ tends to 0. This proves (*), settling the present case.

We have now proved Jensen's formula in special cases, and the general case can be deduced as follows. Let

$$h(z) = f(z) \prod_{a \in \bar{D}_r} (z-a)^{-n_f(a)}.$$

We have multiplied f by a rational function which cancels all the zeros and poles, so h has no zero and pole on $\bar{D}_R$. We can also write

$$f(z) = h(z) \prod_{a \in \bar{D}_R} (z-a)^{n_f(a)}.$$

Thus we can go from h to f in steps by multiplication or division with factors of the special type that we considered previously, and for which we proved Jensen's formula. Therefore, to conclude the proof, it will suffice to prove:

If Jensen's formula is valid for a function f and for a function g, then it is valid for fg and for f/g.

This is easily proved as follows. Let $\mathrm{LS}(f)$ be the left side of Jensen's formula for the function f, and let $\mathrm{RS}(f)$ be the right side, so $\mathrm{RS}(f) = \log|c_f|$. We note that for two meromorphic functions f, g we have:

$$\mathrm{LS}(fg) = \mathrm{LS}(f) + \mathrm{LS}(g) \qquad \text{and} \qquad \mathrm{RS}(fg) = \mathrm{RS}(f) + \mathrm{RS}(g).$$

This is immediately verified for each term entering in Jensen's formula. For example, we have

$$\log|fg| = \log|f| + \log|g| \qquad \text{and similarly for the integral;}$$

$$n_{fg}(a) = n_f(a) + n_g(a);$$

$$c_{fg} = c_f c_g \qquad \text{so that} \quad \log|c_{fg}| = \log|c_f| + \log|c_g|.$$

We also have at once that $\mathrm{LS}(f^{-1}) = -\mathrm{LS}(f)$ and $\mathrm{RS}(f^{-1}) = -\mathrm{RS}(f)$, because this relationship for the inverse of a function is also true for each term. From this formalism, it follows that if Jensen's formula is valid for $f(z)$ and for $(z - a)$, then it is valid for $f(z)(z - a)$ and for $f(z)/(z - a)$. Starting with $h(z)$ we may then multiply and divide successively by factors of degree 1 to reach f. This concludes the proof of Theorem 1.2.

XII, §1. EXERCISES

1. Let f be analytic on the closed unit disc and assume $|f(z)| \leqq 1$ for all z in this set. Suppose also that $f(1/2) = f(i/2) = 0$. Prove that $|f(0)| \leqq 1/4$.

2. Let f be analytic on a disc $\bar{D}(z_0, R)$, and suppose f has at least n zeros in a disc $D(z_0, r)$ with $r < R$ (counting multiplicities). Assume $f(z_0) \neq 0$. Show that

$$\left(\frac{R}{r}\right)^n \leqq \|f\|_R/|f(z_0)|.$$

3. Let f be an entire function. Write $z = x + iy$ as usual. Assume that for every pair of real numbers $x_0 < x_1$ there is a positive integer M such that $f(x + iy) = O(y^M)$ for $y \to \infty$, uniformly for $x_0 \leqq x \leqq x_1$. The implied constant in the estimate depends on x_0, x_1 and f. Let $a_1 < a_2$ be real numbers. Assume that $1/f$ is bounded on $\mathrm{Re}(z) = a_2$. For $T > 0$, let $N_f(T)$ be the number of zeros of f in the box

$$a_1 \leqq x \leqq a_2 \qquad \text{and} \qquad T \leqq y \leqq T + 1.$$

Prove that $N_f(T) = O(\log T)$ for $T \to \infty$. [*Hint*: Use an estimate as in Exercise 2 applied to a pair of circles centered at $a_2 + iy$ and of constant radius.]

Remark. The estimate of Exercise 3 is used routinely in analytic number theory to estimate the number of zeros of a zeta function in a vertical strip.

Next we develop extensions of the Poisson formula. We first set some notation. For $a \in D_R$, define

$$G_R(z, a) = G_{R,a}(z) = \frac{R^2 - \bar{a}z}{R(z-a)}.$$

Then $G_{R,a}$ has precisely one pole on $\bar{D}_R$ and no zeros. We have

$$|G_{R,a}(z)| = 1 \qquad \text{for} \quad |z| = R.$$

4. Apply the Poisson formula of Chapter VIII, §4 to prove the following theorem.

Poisson–Jensen Formula. *Let f be meromorphic on $\bar{D}_R$. Let U be a simply connected open subset of D_R not containing the zeros or poles of f. Then there is a real constant K such that for z in this open set, we have*

$$\log f(z) = \int_0^{2\pi} \log|f(Re^{i\theta})| \frac{Re^{i\theta} + z}{Re^{i\theta} - z} \frac{d\theta}{2\pi} - \sum_{a \in D_R} (\operatorname{ord}_a f) \log G_R(z, a) + iK.$$

[For the proof, assume first that f has no zeros and poles on the circle C_R. Let

$$h(z) = f(z) \prod G_R(z, a)^{\operatorname{ord}_a f}$$

and apply Poisson to $\log h$. Then take care of the zeros and poles on C_R in the same way as in the Jensen formula.]

5. Let f be meromorphic. Define $n_f^+(0) = \max\left(0, n_f(0)\right)$, and:

$$N_f(\infty, R) = \sum_{\substack{a \in D_R \\ f(a)=\infty \\ a \neq 0}} -(\operatorname{ord}_a f) \log \frac{R}{|a|} + n_{1/f}^+(0) \log R$$

$$N_f(0, R) = \sum_{\substack{a \in D_R \\ f(a)=0 \\ a \neq 0}} (\operatorname{ord}_a f) \log \frac{R}{|a|} + n_f^+(0) \log R.$$

Show that Jensen's formula can be written in the form

$$\int_0^{2\pi} \log|f(Re^{i\theta})| \frac{d\theta}{2\pi} + N_f(\infty, R) - N_f(0, R) = \log|c_f|.$$

6. Let α be a positive real number. Define $\log^+(\alpha) = \max(0, \log \alpha)$.
   (a) Show that $\log \alpha = \log^+(\alpha) - \log^+(1/\alpha)$.
   (b) Let f be meromorphic on $\bar{D}_R$. For $r < R$ define

$$m_f(r) = \int_0^{2\pi} \log^+|f(re^{i\theta})| \frac{d\theta}{2\pi} \quad \text{and} \quad G_{R,f}^{\infty}(z) = \prod_{\substack{a \in D_R \\ f(a)=\infty}} G_R(z, a)^{-\operatorname{ord}_a f}.$$

Let $G = G^{\infty}_{R,f}$. Show that

$$m_G(r) = N_f(\infty, R) - N_f(\infty, r).$$

(c) Following Nevanlinna, define the **height function**

$$T_f(r) = m_f(r) + N_f(\infty, r).$$

Deduce Nevanlinna's formulation of Jensen's formula:

$$T_{1/f}(r) = T_f(r) - \log|c_f|.$$

XII, §2. THE PICARD–BOREL THEOREM

Picard's theorem states that an entire function which omits two distinct complex numbers is constant. We shall give a proof of this theorem due to Borel and dating back to 1887. For Picard's original proof, see Chapter XI, §3. This section will not be used later and may be omitted by uninterested readers. We begin by setting some notation.

Let α be a positive real number. We define

$$\log^+(\alpha) = \max(0, \log \alpha).$$

If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are positive, then

$$\log^+(\alpha_1\alpha_2) \leqq \log^+(\alpha_1) + \log^+(\alpha_2)$$

and

$$\log^+(\alpha_1 + \cdots + \alpha_n) \leqq \sum \log^+(\alpha_i) + \log n.$$

The first inequality is immediate, and for the second, the stronger inequality holds:

$$\log^+(\alpha_1 + \cdots + \alpha_n) \leqq \max \log^+ \alpha_i + \log n.$$

For the proof, we note that

$$\log^+(\alpha_1 + \cdots + \alpha_n) \leqq \log^+(n \cdot \max \alpha_i) \leqq \log^+ \max \alpha_i + \log n.$$

Lemma 2.1. *Let $b \in \mathbf{C}$. Then*

$$\int_0^{2\pi} \log|b - e^{i\theta}| \frac{d\theta}{2\pi} = \log^+|b|.$$

Proof. If $|b| > 1$ then $\log|b - z|$ for $|z| < 1 + \varepsilon$ is harmonic, and $\log^+|b| = \log|b|$, so the formula is true by the mean value property for harmonic functions. If $|b| < 1$, then

$$\int_0^{2\pi} \log|b - e^{i\theta}| \frac{d\theta}{2\pi} = \int_0^{2\pi} \log|be^{-i\theta} - 1| \frac{d\theta}{2\pi}$$
$$= \log|-1| = 0 = \log^+|b|.$$

If $|b| = 1$, then the lemma has already been proved in Lemma 1.3.

Let f be a meromorphic function. We define the **proximity function**

$$m_f(r) = \int_0^{2\pi} \log^+|f(re^{i\theta})| \frac{d\theta}{2\pi}.$$

The integral is of course an improper integral, which converges. The situation is similar to that discussed in Jensen's formula. Then the previous inequalities for $\log^+$ immediately yield:

If f, g are meromorphic functions, then

$$m_{fg} \leqq m_f + m_g.$$

If $f_1, \ldots, f_n$ are meromorphic functions, then

$$m_{f_1 + \cdots + f_n} \leqq m_{f_1} + \cdots + m_{f_n} + \log n.$$

We note that

$$\log^+ \alpha - \log^+ 1/\alpha = \log \alpha \qquad \text{and} \qquad |\log \alpha| = \log^+ \alpha + \log^+ 1/\alpha.$$

Example. Suppose h is entire without zeros. Then

$$m_h = m_{1/h} + \log|h(0)|.$$

Indeed, from the definition and the preceding remark, we get

$$m_h(r) - m_{1/h}(r) = \int_0^{2\pi} \log|h(re^{i\theta})| \frac{d\theta}{2\pi} = \log|h(0)|$$

by the mean value property of a harmonic function, Theorem 3.2 of Chapter VIII, §3. This proves the formula.

We shall also need to count zeros of a function, and how many times a function takes a given value. Suppose f is an entire function. We had defined $v_f(r)$ to be the number of zeros of f in the closed disc of radius r.

It is now convenient to count the zeros in the open disc, so we define:

$n_f(0, r) =$ number of zeros of f in the open disc of radius r, counted with multiplicities.

If $f(0) \neq 0$ we define

$$N_f(0, r) = \int_0^r n_f(0, t) \frac{dt}{t} = \sum_{\substack{a \in D_r \\ a \neq 0}} (\mathrm{ord}_a f) \log \left| \frac{r}{a} \right|,$$

$$N_f(b, r) = N_{f-b}(0, r) \qquad \text{if} \quad b \in \mathbf{C} \text{ is such that } f(0) - b \neq 0.$$

Proposition 2.2 (Cartan). *Let f be an entire function. Then*

$$m_f(r) = \int_0^{2\pi} N_f(e^{i\theta}, r) \frac{d\theta}{2\pi} + \log^+ |f(0)|.$$

In particular, m_f is an increasing function of r.

Proof. For each θ such that $f(0) \neq e^{i\theta}$ we apply Jensen's formula to the function $f(z) - e^{i\theta}$, to get

$$N_f(e^{i\theta}, r) + \log|f(0) - e^{i\theta}| = \int_0^{2\pi} \log|f(re^{i\varphi}) - e^{i\theta}| \frac{d\varphi}{2\pi}.$$

Then we integrate each side with respect to θ, and use Lemma 2.1 to prove the formula. That m_f is increasing then follows because for each θ, the function

$$r \mapsto N_f(e^{i\theta}, r)$$

is increasing. This proves the proposition.

Let f be an entire function. We define

$$M_f(R) = \log\|f\|_R \qquad \text{where} \qquad \|f\|_R = \sup_{|z|=R} |f(z)|.$$

Thus M_f is the log of the maximum modulus on the circle of radius R. It is obvious that

$$m_f \leqq \max(M_f, 0).$$

Conversely:

Theorem 2.3. *Let f be an entire function. Then for $r < R$ we have*

$$M_f(r) \leqq \frac{R+r}{R-r} m_f(R),$$

and in particular, $M_f(r) \leqq 3m_f(2r)$.

Proof. We shall use the theorem only when f has no zeros, and the proof in this case is slightly shorter. Hence we shall give the proof only in this case. The key step is to show that for $r < R$ we have the inequality

$$M_f(r) \leqq \frac{R+r}{R-r} m_f(R).$$

Since we assume that f is entire without zeros, we can define $\log f(z)$ as an entire function of z as in Chapter III, §6. We then apply Poisson's formula to the real part of $\log f(z)$, taking $z = re^{i\varphi}$ and $\zeta = Re^{i\theta}$, to get

$$\log|f(z)| \leqq \int_0^{2\pi} \log|f(Re^{i\theta})| \operatorname{Re}\frac{\zeta + z}{\zeta - z}\frac{d\theta}{2\pi}.$$

But the max and min of $\operatorname{Re}\big((\zeta + z)/(\zeta - z)\big)$ occurs when $\cos(\theta - \varphi) = 1$ or -1, so

$$0 < \frac{R-r}{R+r} \leqq \operatorname{Re}\frac{\zeta+z}{\zeta-z} \leqq \frac{R+r}{R-r}$$

and we get

$$\log|f(z)| \leqq \int_0^{2\pi} \log^+|f(Re^{i\theta})| \frac{R+r}{R-r}\frac{d\theta}{2\pi} = \frac{R+r}{R-r} m_f(R).$$

Now take the sup for $|z| = r$ and let $R = 2r$ to conclude the proof.

Corollary 2.4. *Let f be an entire function. If m_f is bounded for $r \to \infty$ then f is constant. If there exists a constant k such that*

$$m_f(R_j) \leqq k \log R_j$$

for a sequence of numbers $R_j \to \infty$ then f is a polynomial of degree $\leqq k$.

Proof. The first assertion follows from Theorem 2.3 and Liouville's theorem that a bounded entire function is constant. The second assertion is essentially Exercise 5 of Chapter V, §1, but we give the short proof. Select A large positive, and let $R = Ar$. By Cauchy's theorem, if $f =$

$\sum a_n z^n$ then $|a_n| \leqq \|f\|_R/R^n$. By Theorem 2.3, we find for $R = R_j = Ar_j$:

$$\log|a_n| \leqq M_f(R) - n \log R \leqq \frac{A+1}{A-1} k \log(Ar) - n \log(Ar).$$

If $n > k$, then for A sufficiently large, we also have $n > k(A+1)/(A-1)$. Letting $R = R_j \to \infty$ we find that $a_n = 0$. Taking A sufficiently large shows that $a_n = 0$ if $n > k$, thus proving the corollary.

We now turn to the main application.

Theorem 2.5. *Let h be an entire function without zeros. Then*

$$m_{h'/h}(r) \ll \log r + \log m_h(r)$$

for all r lying outside a set of finite measure, and $r \to \infty$.

The proof will require some lemmas. Recall that a set has **finite measure** if it can be covered by a union of intervals $I_1, I_2, \ldots$ such that the sum of the lengths is finite, i.e.

$$\sum_{n=1}^{\infty} \text{length}(I_n) < \infty.$$

Lemma 2.6. *Let h be an entire function without zeros. Let $1 \leqq r < R$. Then*

$$m_{h'/h}(r) \leqq \log^+ R + 2 \log^+ \frac{1}{R-r} + 2 \log^+ m_h(R) + \text{a constant}.$$

Proof. Note that $\log h$ is defined as an entire function, and we have the Poisson representation

$$\log h(z) = \int_0^{2\pi} \log|h(Re^{i\theta})| \frac{Re^{i\theta} + z}{Re^{i\theta} - z} \frac{d\theta}{2\pi} + iK$$

for some constant K. Differentiating with respect to z with $|z| < R$ we obtain

$$h'/h(z) = \int_0^{2\pi} \log|h(Re^{i\theta})| \frac{2\, Re^{i\theta}}{(Re^{i\theta} - z)^2} \frac{d\theta}{2\pi}.$$

Therefore we find the estimate

$$|h'/h(z)| \leqq \frac{2R}{(R-r)^2} [m_h(R) + m_{1/h}(R)].$$

Taking $\log^+$ and integrating, we obtain the bound

$$m_{h'/h}(r) \leqq \log^+ R + 2\log^+ \frac{1}{R-r} + \log^+ m_h(R) + \log^+ m_{1/h}(R) + \log 2.$$

But as we saw at the beginning of the section, $m_{1/h} = m_h - \log|h(0)|$, so the lemma follows, with the constant $\log 2 - \log|h(0)|$.

The lemma tells us that $m_{h'/h}(r)$ is very small compared to m_h, provided we can find R slightly bigger than r such that $m_h(R)$ is not much bigger than $m_h(r)$. The next lemma shows that we can find arbitrarily large values of such numbers r.

Lemma 2.7. *Let S be a continuous, non-constant, increasing function defined for $r > 0$. Then*

$$S\left(r + \frac{1}{S(r)}\right) < 2S(r)$$

for all $r > 0$ except for r lying in a set of finite measure.

Proof. Let E be the exceptional set where the stated inequality is false, that is $S(r + 1/S(r)) \geqq 2S(r)$. Suppose there is some $r_1 \in E$, $S(r_1) \neq 0$. Let

$$r_2 = \inf\left\{r \in E \text{ such that } r \geqq r_1 + \frac{1}{S(r_1)}\right\}.$$

Then $r_2 \in E$ and $r_2 \geqq r_1 + 1/S(r_1)$. Let

$$r_3 = \inf\left\{r \in E \text{ such that } r \geqq r_2 + \frac{1}{S(r_2)}\right\}.$$

Then $r_3 \in E$ and $r_3 \geqq r_2 + 1/S(r_2)$. We continue in this way to get the sequence $\{r_n\}$.

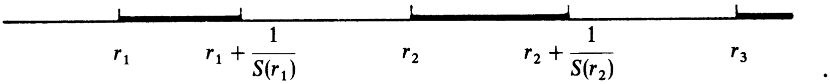


Figure 2

Since S is monotone, by construction we find

$$S(r_{n+1}) \geqq S\left(r_n + \frac{1}{S(r_n)}\right) \geqq 2S(r_n) \geqq 2^n S(r_1).$$

Hence

$$\sum \frac{1}{S(r_n)} \leqq \frac{2}{S(r_1)}.$$

If the sequence $\{r_n\}$ is infinite, then it cannot be bounded, for otherwise one sees immediately that S would become infinite before the least upper bound, which does not happen since S is continuous. Then E is covered by the union of the intervals

$$\bigcup \left[r_n, r_n + \frac{1}{S(r_n)} \right]$$

which has measure $\leqq 2/S(r_1)$, thereby proving the lemma.

We apply the lemma to the function

$$S(r) = \log^+ m_h(r),$$

and we take

$$R = r + \frac{1}{\log^+ m_h(r)}$$

in Lemma 2.6. This concludes the proof of Theorem 2.5. (See Exercise 1.)

We now come to Picard's theorem.

Theorem 2.8. *Let f be an entire function. Let a, b be two distinct complex numbers such that $f(z) \neq a, b$ for all z. Then f is constant.*

Proof. Let

$$L(w) = \frac{w - a}{b - a}.$$

Then L carries a, b to $0, 1$. Thus after replacing f by $L \circ f$, we may assume without loss of generality that $a = 0$ and $b = 1$. Thus we let h be an entire function which has no zeros and such that $1 - h$ also has no zeros, and we have to prove that h is constant. Letting $h_1 = h$ and $h_2 = 1 - h$, we see that it suffices to prove:

Theorem 2.9. *Let h_1, h_2 be entire functions without zeros. If*

$$h_1 + h_2 = 1$$

then h_1, h_2 are constant.

Differentiating the given relation, we obtain two linear equations

$$h_1 + h_2 = 1,$$
$$h_1' + h_2' = 0.$$

Write $h_1' = (h_1'/h_1)h_1$ and $h_2' = (h_2'/h_2)h_2$. Solving for h_1, say, we get

$$h_1 = \frac{h_2'/h_2}{h_2'/h_2 - h_1'/h_1},$$

unless the denominator is 0, in which case $(h_2/h_1)'/(h_2/h_1) = 0$, and then h_2/h_1 is constant, and it follows at once that both h_1, h_2 are constant. We then apply the inequalities for m_{fg} and m_{f+g} at the beginning of this section, and we find by Theorem 2.5 that

$$m_{h_1}(r) \ll \log r + \log m_{h_1}(r) + \log m_{h_2}(r)$$

for all r outside a set of finite measure. We have a similar inequality for m_{h_2}. Then we get that

$$m_{h_1}(r) + m_{h_2}(r) \ll \log r + \log m_{h_1}(r) + \log m_{h_2}(r).$$

It follows that

$$m_{h_1}(r) \text{ and } m_{h_2}(r) = O(\log r)$$

for arbitrarily large values of r. By Corollary 2.4, we conclude that h_1, h_2 are polynomials, which are constant since they have no zeros. This concludes Borel's proof of Picard's theorem.

XII, §2. EXERCISES

1. Let h be an entire function without zeros. Show that m_h is continuous. [This is essentially trivial, by the uniform continuity of continuous functions on compact sets.]

2. Let f be a meromorphic function. Show that m_f is continuous. This is less trivial, but still easy. Let $z_1, \ldots, z_s$ be the poles of f on a circle of radius r. Let $z_j = re^{i\theta_j}$, and let $I(\theta_j, \delta)$ be the open interval or radius δ centered at θ_j. Let I be the union of these intervals. Then for $\theta \notin I$, $f(te^{i\theta})$ converges uniformly to $f(re^{i\theta})$ as $t \to r$, so as $t \to r$,

$$\int_{\theta \notin I} \log^+|f(te^{i\theta})| \frac{d\theta}{2\pi} \to \int_{\theta \notin I} \log^+|f(re^{i\theta})| \frac{d\theta}{2\pi}.$$

Given ε, there exists δ such that if $\theta \in I(\theta_j, \delta)$ for some j and $|t - r| < \delta$, then

$$\log^+|f(re^{i\theta})| = \log|f(re^{i\theta})| \qquad \text{and} \qquad \log^+|f(te^{i\theta})| = \log|f(te^{i\theta})|$$

because $|f(z)|$ is large when z is near z_j. But for z near z_j,

$$f(z) = (z - z_j)^{-e} g(z),$$

where $e > 0$ and $|g(z)|$ is bounded away from 0. Hence for z near z_j,

$$\log|f(z)| = -e \log|z - z_j| + \text{bounded function}.$$

Then

$$\int_{\theta \in I(\theta_j, \delta)} \log|f(re^{i\theta})| \frac{d\theta}{2\pi} \quad \text{and} \quad \int_{\theta \in I(\theta_j, \delta)} \log|f(te^{i\theta})| \frac{d\theta}{2\pi}$$

are both small, because they essentially amount to

$$\int_{-\delta}^{\delta} \log|\theta| \, d\theta,$$

up to a bounded factor. Put in the details of this part of the argument.]

3. Let f be meromorphic. Let a be a complex number. Show that

$$T_f(r) = T_{f-a}(r) + O_a(1),$$

where $|O_a(1)| \leqq \log^+|a| + \log 2$.

So far you will have proved the first main theorem of Nevanlinna theory. For a development of the theory, see Nevanlinna's book [Ne 53]. You can also look up the self-contained Chapters VI and VIII of [La 87].

The next sections deal with separate topics of more specialized interest. First we show how to estimate an analytic function by its real part. The method continues the spirit of the beginning of §1, *using the maximum modulus principle. Next, we show that in the estimate of a function, one can use the fact that derivatives are small instead of the function having actual zeros to get bounds on the function. Such a technique is useful in contexts which deepen the result stated in* §4 *concerning rational values of entire or meromorphic functions. Then we go on to applications of the maximal modulus principle to regions other than discs, namely strips. These topics are logically independent of each other and can be covered in any order that suits the reader's tastes or needs. Readers may also wish to skip the next sections and go immediately to the next chapter.*

XII, §3. BOUNDS BY THE REAL PART, BOREL–CARATHÉODORY THEOREM

We shall now give a simple proof that an analytic function is essentially bounded by its real part. As usual, we use the technique of two circles. If u is a real function, we let $\sup_R u = \sup u(z)$ for $|z| = R$.

Theorem 3.1 (Borel–Carathéodory). *Let f be holomorphic on a closed disc of radius R, centered at the origin. Let $\|f\|_r = \max|f(z)|$ for*

$|z| = r < R$. *Then*

$$\|f\|_r \leqq \frac{2r}{R-r} \sup_R \operatorname{Re} f + \frac{R+r}{R-r}|f(0)|.$$

Proof. Let $A = \sup_R \operatorname{Re} f$. Assume first that $f(0) = 0$. Then $A \geqq 0$ (why?). Let

$$g(z) = \frac{f(z)}{z(2A - f(z))}.$$

Then g is holomorphic for $|z| \leqq R$. Furthermore, if $|z| = R$, then

$$|2A - f(z)| \geqq |f(z)|.$$

Hence $\|g\|_R \leqq 1/R$. By the maximum modulus principle, we have the inequality $\|g\|_r \leqq \|g\|_R$, and hence, if $|w| = r$, we get

$$\frac{|f(w)|}{r|2A - f(w)|} \leqq \frac{1}{R},$$

whence

$$|f(w)| \leqq \frac{r}{R}(2A + |f(w)|),$$

and therefore

$$\|f\|_r \leqq \frac{2r}{R-r}A,$$

which proves the lemma in this case.

In general, we apply the preceding estimate to the function

$$h(z) = f(z) - f(0).$$

Then

$$\sup_R \operatorname{Re} h \leqq \sup_R \operatorname{Re} f + |f(0)|,$$

and if $|w| = r$, we get

$$|f(w) - f(0)| \leqq \frac{2r}{R-r}[A + |f(0)|],$$

whence

$$|f(w)| \leqq \frac{2r}{R-r}[A + |f(0)|] + |f(0)|$$

thereby proving the theorem.

In Chapter VIII we gave a more precise description of the relationship which exists between an analytic function and its real part, by expressing the function in terms of an appropriate integral involving only the real part and the Poisson kernel function. The analytic function is essentially uniquely determined by its real part, except for adding a constant, which explains the occurrence of the term $f(0)$ in the above estimate. However, the proof here with the maximum modulus principle is so simple that we found it worthwhile including it anyway. Besides, we give an estimate in terms of $\sup_R \operatorname{Re} f$, not just $\|\operatorname{Re} f\|_R$. This is significant in the next corollaries.

Corollary 3.2. *Let h be an entire function. Let $\rho > 0$. Assume that there exists a number $C > 0$ such that for all sufficiently large R we have*

$$\sup_R \operatorname{Re} h \leqq CR^\rho.$$

Then h is a polynomial of degree $\leqq \rho$.

Proof. In the Borel–Carathéodory theorem, use $R = 2r$. Then we have $\|h\|_r \ll r^\rho$ for $r \to \infty$. Let $h(z) = \sum a_n z^n$. By Cauchy's formula we have $|a_n| \leqq \|h\|_R/R^n$ for all R, so $a_n = 0$ if $n > \rho$, as desired.

Corollary 3.3 (Hadamard). *Let f be an entire function with no zeros. Assume that there is a constant $C \geqq 1$ such that $\|f\|_R \leqq C^{R^\rho}$ for all R sufficiently large. Then $f(z) = e^{h(z)}$ where h is a polynomial of degree $\leqq \rho$.*

Proof. By Chapter III, §6 we can define an analytic function $\log f(z) = h(z)$ such that $e^{h(z)} = f(z)$. The assumption implies that Re h satisfies the hypotheses of Corollary 3.2, whence h is a polynomial, as desired.

XII, §4. THE USE OF THREE CIRCLES AND THE EFFECT OF SMALL DERIVATIVES

We begin by illustrating the use of three circles in estimating the growth of a function in terms of the number of zeros.

We study systematically the situation when an analytic function has zeros in a certain region. This has the effect of making the function small throughout the region. We want quantitative results showing how small in terms of the number of zeros, possibly counted with their multiplicities. If the function has zeros at points $z_1, \ldots, z_n$ with multiplicities $k_1, \ldots, k_n$, then

$$\frac{f(z)}{(z - z_1)^{k_1} \cdots (z - z_n)^{k_n}}$$

is again analytic, and the maximum modulus principle can be applied to it. On a disc of radius R centered at the origin such that the disc contains all the points $z_1, \ldots, z_n$, suppose that R is large with respect to $|z_1|, \ldots, |z_n|$, say $|z_j| \leqq R/2$. Then for $|z| = R$ we get

$$|z - z_j| \geqq R/2.$$

If we estimate the above quotient on the circle of radius R (which gives an estimate for this quotient inside by the maximum modulus principle) we obtain an estimate of the form

$$2^{k_1 + \cdots + k_n} \|f\|_R / R^{k_1 + \cdots + k_n}.$$

Specific situations then compare $\|f\|_R$ with the power of R in the denominator to give whatever result is sought. In the applications most often the multiplicities are equal to the same integer. We state this formally.

Theorem 4.1. *Let f be holomorphic on the closed disc of radius R. Let $z_1, \ldots, z_N$ be points inside the disc where f has zeros of multiplicities $\geqq M$, and assume that these points lie in the disc of radius R_1. Assume*

$$R_1 \leqq R/2.$$

Let $R_1 \leqq R_2 \leqq R$. Then on the circle of radius R_2 we have the estimate

$$\|f\|_{R_2} \leqq \frac{\|f\|_R 4^{MN}}{e^{MN \log(R/R_2)}}.$$

Proof. Let $|w| = R_2$. We estimate the function

$$\frac{f(z)}{[(z - z_1) \cdots (z - z_N)]^M} [(w - z_1) \cdots (w - z_N)]^M$$

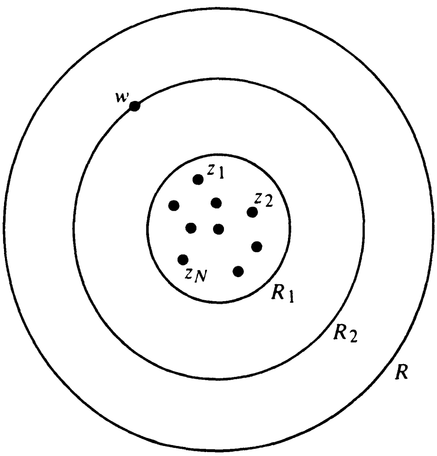


Figure 3

on the circle of radius R. This function has precisely the value $f(w)$ at $z = w$. The estimate $|w - z_j| \leqq 2R_2$ is trivial, and the theorem follows at once.

If a function does not have zeros at certain points, but has small derivatives, then it is still true that the function is small in a region not too far away from these points. A quantitative estimate can be given, with a main term which is the same as if the function had zeros, and an error term, measured in terms of the derivatives.

Let $z_1, \ldots, z_N$ be distinct points in the open disc of radius R, and let

$$Q(z) = [(z - z_1) \cdots (z - z_N)]^M.$$

Let f be holomorphic on the closed disc of radius R. Let Γ be the circle of radius R, and let Γ_j be a circle around z_j not containing z_k for $k \neq j$, and contained in the interior of Γ. Then we have for z not equal to any z_j:

Hermite Interpolation Formula

$$\frac{f(z)}{Q(z)} = \frac{1}{2\pi i} \int_\Gamma \frac{f(\zeta)}{Q(\zeta)} \frac{d\zeta}{(\zeta - z)}$$
$$- \frac{1}{2\pi i} \sum_{j=1}^{N} \sum_{m=0}^{M-1} \frac{D^m f(z_j)}{m!} \int_{\Gamma_j} \frac{(\zeta - z_j)^m}{Q(\zeta)} \frac{d\zeta}{(\zeta - z)}.$$

This formula, due to Hermite, is a direct consequence of the residue formula. We consider the integral

$$\int_\Gamma \frac{f(\zeta)}{Q(\zeta)} \frac{1}{(\zeta - z)} d\zeta.$$

The function has a simple pole at $\zeta = z$ with residue $f(z)/Q(z)$. This gives the contribution on the left-hand side of the formula. The integral is also equal to the sum of integrals taken over small circles around the points $z_1, \ldots, z_N, z$. To find the residue at z_j, we expand $f(\zeta)$ at z_j, say

$$f(\zeta) = a_0 + a_1(\zeta - z_j) + \cdots + a_M(\zeta - z_j)^M + \cdots.$$

Looking at the quotient by $Q(\zeta)$ immediately determines the residues at z_j in terms of coefficients of the expression, which are such that

$$a_m = D^m f(z_j)/m!.$$

The formula then drops out.

It is easy to estimate $|f(z)|$. Multiplying by $Q(z)$ introduces the quotients

$$\frac{Q(z)}{Q(\zeta)} = \prod_{j=1}^{N} \left[\frac{(z - z_j)}{(\zeta - z_j)}\right]^M,$$

which are trivially estimated. The denominator is small according as the radius of Γ_j is small. In applications, one tries to take the Γ_j of not too small radius, and this depends on the minimum distance between the points $z, z_1, \ldots, z_N$. It is a priori clear that if the points are close together, then the information that the function has small derivatives at these points is to a large extent redundant. This information is stronger the wider apart the points are.

Making these estimates, the following result drops out.

Theorem 4.2. *Let f be holomorphic on the closed disc of radius R. Let $z_1, \ldots, z_N$ be distinct points in the disc of radius R_1. Assume*

$$2R_1 < R_2 \quad \text{and} \quad 2R_2 < R.$$

Let σ be the minimum of 1, and the distance between any pair of distinct points among $z_1, \ldots, z_N$. Then

$$\|f\|_{R_2} \leqq \frac{\|f\|_R C^{MN}}{(R/R_2)^{MN}} + (CR_2/\sigma)^{MN} \max_{m,j} \frac{|D^m f(z_j)|}{m!},$$

where C is an absolute constant.

An estimate for the derivative of f can then be obtained from Cauchy's formula

$$\frac{D^k f(z)}{k!} = \frac{1}{2\pi i} \int_{|\zeta|=R_2} \frac{f(\zeta)}{(\zeta - z)^{k+1}} d\zeta.$$

from which we see that such a derivative is estimated by a similar expression, multiplied by

$$2^k/R_2^k.$$

We may summarize the estimate of Theorem 4.2 by saying that the first term is exactly the same as would arise if f had zeros at the points $z_1, \ldots, z_N$, and the second term is a correcting factor describing the extent to which those points differ from actual zeros. In practice, the derivatives of f are very small at these points, which thus do not differ too much from zeros.

XII, §5. ENTIRE FUNCTIONS WITH RATIONAL VALUES

We shall give a theorem showing how the set of points where an entire function takes on rational values is limited. The method follows a classical method of Gelfond and Schneider in proving the transcendence of α^β (when α, β are algebraic $\neq 0, 1$ and β is irrational). It develops ideas of Schneider, who had partially axiomatized the situation, but in a manner which makes the theorem also applicable to a wider class of functions.

We recall that an analytic function is said to be entire if it is analytic on all of $\mathbf{C}$. Let $\rho > 0$. We shall say that f has **strict order** $\leqq \rho$ if there exists a number $C > 1$ such that for all sufficiently large R we have

$$|f(z)| \leqq C^{R^\rho} \qquad \text{whenever} \quad |z| \leqq R.$$

Two functions f, g are called **algebraically independent** if for any polynomial function

$$\sum a_{ij} f^i g^j = 0$$

with complex coefficients a_{ij}, we must have $a_{ij} = 0$ for all i, j.

We denote the rational numbers by $\mathbf{Q}$, as usual. If f, g take on rational values at certain points, we shall construct an auxiliary function $\sum b_{ij} f^i g^j$ which has zeros of high order at these points, and then estimate this latter function.

Theorem 5.1. *Let $f_1, \ldots, f_n$ be entire functions of strict order $\leqq \rho$. Assume that at least two of these functions are algebraically independent. Assume that the derivative $d/dz = D$ maps the ring $\mathbf{Q}[f_1, \ldots, f_n]$ into itself, i.e. for each j there is a polynomial P_j with rational coefficients such that*

$$Df_j = P_j(f_1, \ldots, f_n).$$

Let $w_1, \ldots, w_N$ be distinct complex numbers such that

$$f_j(w_i) \in \mathbf{Q}$$

for $j = 1, \ldots, n$ and $i = 1, \ldots, N$. Then $N \leqq 4\rho$.

The most classical application of the theorem is to the ring of functions

$$\mathbf{Q}[z, e^z]$$

which is certainly mapped into itself by the derivative $d/dz = D$. The theorem then implies that e^w cannot be rational for any integer $w \neq 0$. For otherwise,

$$e^w, e^{2w}, \ldots, e^{Nw}$$

would be rational, so we would obtain infinitely many numbers at which the exponential takes on rational values, as well as the function z, which contradicts the theorem. The same argument works in the more general case when dealing with algebraic numbers, to show that e^w is not algebraic when w is algebraic $\neq 0$. Cf. [La 62], Theorem 3, and [La 66], Chapter III, Theorem 1.

Before giving the main arguments proving the theorem, we state some lemmas. The first, due to Siegel, has to do with integral solutions of linear homogeneous equations.

Lemma 5.2. *Let*

$$a_{11}x_1 + \cdots + a_{1n}x_n = 0,$$
$$\cdots$$
$$a_{r1}x_1 + \cdots + a_{rn}x_n = 0$$

be a system of linear equations with integer coefficients a_{ij}, and $n > r$. Let $A \geqq 1$ be a number such that $|a_{ij}| \leqq A$ for all i, j. Then there exists an integral, non-trival solution with

$$|x_j| \leqq 2(2nA)^{r/(n-r)}.$$

Proof. We view our system of linear equations as a linear equation $L(X) = 0$, where L is a linear map, $L\colon \mathbf{Z}^{(n)} \to \mathbf{Z}^{(r)}$, determined by the matrix of coefficients. If B is a positive number, we denote by $\mathbf{Z}^{(n)}(B)$ the set of vectors X in $\mathbf{Z}^{(n)}$ such that $|X| \leqq B$ (where $|X|$ is the maximum of the absolute values of the coefficients of X). Then L maps $\mathbf{Z}^{(n)}(B)$ into $\mathbf{Z}^{(r)}(nBA)$. The number of elements in $\mathbf{Z}^{(n)}(B)$ is $\geqq B^n$ and $\leqq (2B+1)^n$. We seek a value of B such that there will be two distinct elements X, Y in $\mathbf{Z}^{(n)}(B)$ having the same image, $L(X) = L(Y)$. For this, it will suffice that $B^n > (2nBA + 1)^r$, and thus it will suffice that $B = (3nA)^{r/(n-r)}$. We take $X - Y$ as the solution of our problem.

The next lemma has to do with estimates of derivatives. By the **size** of a polynomial with rational coefficients we shall mean the maximum of the absolute values of the coefficients. A **denominator** for a set of rational numbers will be any positive integer whose product with every element of the set is an integer. We define in a similar way a denominator for a polynomial with rational coefficients. We abbreviate denominator by "den".

Let

$$P(T_1, \ldots, T_n) = \sum \alpha_{i_1 \ldots i_n} T_1^{i_1} \cdots T_n^{i_n}$$

be a polynomial with complex coefficients, and let

$$Q(T_1, \ldots, T_n) = \sum \beta_{i_1 \ldots i_n} T_1^{i_1} \cdots T_n^{i_n}$$

be a polynomial with real coefficients $\geqq 0$. We say that Q **dominates** P and write $P \prec Q$, if $|\alpha_{(i)}| \leqq \beta_{(i)}$ for all $(i) = (i_1, \ldots, i_n)$. It is then immediately verified that the relation of domination is preserved under addition, multiplication, and taking partial derivatives with respect to the variables $T_1, \ldots, T_n$. Thus if $P \prec Q$, then $D_i P \prec D_i Q$, where $D_i = \partial/\partial T_i$.

Lemma 5.3. *Let $f_1, \ldots, f_n$ be functions such that the derivative $D = d/dz$ maps the ring $\mathbf{Q}[f_1, \ldots, f_n]$ into itself. There exists a number C_1 having the following property. If $Q(T_1, \ldots, T_n)$ is a polynomial with rational coefficients, of total degree $\leqq r$, then*

$$D^m(Q(f_1, \ldots, f_n)) = Q_m(f_1, \ldots, f_n)$$

where $Q_m \in \mathbf{Q}[T_1, \ldots, T_n]$ is a polynomial satisfying:

(i) $\deg Q_m \leqq r + md$ *with d defined below.*
(ii) $\operatorname{size} Q_m \leqq (\operatorname{size} Q) m!\, C_1^{m+r}$.
(iii) *There exists a denominator for the coefficients of Q_m bounded by* $\operatorname{den}(Q) C_1^{m+r}$.

Proof. Let $P_j(T_1, \ldots, T_n)$ be a polynomial such that

$$Df_j = P_j(f_1, \ldots, f_n).$$

Let d be the maximum of the degrees of $P_1, \ldots, P_n$. There exists a "differentiation" $\bar{D}$ on the polynomial ring $\mathbf{Q}[T_1, \ldots, T_n]$ such that

$$\bar{D}T_j = P_j(T_1, \ldots, T_n),$$

and for any polynomial P we have

$$\bar{D}(P(T_1, \ldots, T_n)) = \sum_{j=1}^{n} (D_j P)(T_1, \ldots, T_n) P_j(T_1, \ldots, T_n).$$

This is just obtained by the usual chain rule for differentiation, and

$$D_j = \partial/\partial T_j$$

is the usual partial derivative. But the polynomial Q is dominated by

$$Q \prec \operatorname{size}(Q)(1 + T_1 + \cdots + T_n)^r$$

and each polynomial P_i is dominated by $\operatorname{size}(P_i)(1 + T_1 + \cdots + T_n)^d$. Thus for some constant C_2 we have

$$\bar{D}Q \prec \operatorname{size}(Q) C_2 r (1 + T_1 + \cdots + T_n)^{r+d}.$$

Proceeding inductively, we see that $\bar{D}^k Q$ is dominated by

$$\bar{D}^k Q \prec \text{size}(Q) C_3^k r(r+d) \cdots (r+kd)(1 + T_1 + \cdots + T_n)^{r+kd}.$$

Since

$$\begin{aligned} r(r+d) \cdots (r+kd) &\leqq dr(dr+d) \cdots (dr+kd) \\ &\leqq d^{k+1} r(r+1) \cdots (r+k), \end{aligned}$$

this product is estimated by

$$d^{k+1} \frac{(r+k)!}{r!\,k!} rk! \leqq C_4^{r+k} k!.$$

This proves the lemma.

We apply the lemma when we want to evaluate a derivative

$$D^k f(w)$$

at some point w, where $f = Q(f_1, \ldots, f_n)$ is a polynomial in the functions $f_1, \ldots, f_n$. Then all we have to do is plug in $f_1(w), \ldots, f_n(w)$ in $Q_k(T_1, \ldots, T_n)$ to obtain

$$D^k f(w) = Q_k(f_1(w), \ldots, f_n(w)).$$

If w is regarded as fixed, this gives us an estimate for $D^k f(w)$ as in (ii) and (iii) of the theorem, whenever $f_1(w), \ldots, f_n(w)$ are rational numbers. Thus the previous discussion tells us how fast the denominators and absolute values of a derivative

$$D^k f(w)$$

grow when w is a point such that $f_1(w), \ldots, f_n(w)$ are rational numbers.

We now come to the main part of the proof of the theorem. Let f, g be two functions among $f_1, \ldots, f_n$ which are algebraically independent. Let L be a positive integer divisible by $2N$. We shall let L tend to infinity at the end of the proof.

Let

$$F = \sum_{i,j=1}^{L} b_{ij} f^i g^j$$

have integer coefficients, and let $L^2 = 2MN$. We wish to select the coefficients b_{ij} not all 0 such that

$$D^m F(w_\nu) = 0$$

for $m = 0, \ldots, M - 1$ and $v = 1, \ldots, N$. This amounts to solving a system of linear equations

$$\sum_{i,j=1}^{L} b_{ij} D^m(f^i g^j)(w_v) = 0,$$

and by hypothesis $D^m(f^i g^j)(w_v)$ is a rational number for each v. We treat the b_{ij} as unknowns, and wish to apply Siegel's lemma. We have:

Number of unknowns $= L^2$,

Number of equations $= MN$.

Then our choice of L related to M is such that

$$\frac{\#\text{equations}}{\#\text{unknowns} - \#\text{equations}} = 1.$$

We multiply the equations by a common denominator for the coefficients. Using the estimate of Lemma 5.3, and Siegel's lemma, we can take the b_{ij} to be integers, whose size is bounded by

$$\text{size } b_{ij} \ll M!\, C_2^{M+L} \ll M^M C_2^{M+L}$$

for $M \to \infty$.

Since f, g are algebraically independent, the function F is not identically zero. Let s be the smallest integer such that all derivatives of F up to order $s - 1$ vanish at all points $w_1, \ldots, w_N$, but such that D^sF does not vanish at one of the w_v, say w_1. Then $s \geqq M$. We let

$$\alpha = D^sF(w_1).$$

Then $\alpha \neq 0$ is a rational number, and by Lemma 5.3 it has a denominator which is $\ll C_1^s$ for $s \to \infty$. Let c be this denominator. Then $c\alpha$ is an integer, and its absolute value is therefore $\geqq 1$.

We shall obtain an upper bound for $|D^sF(w_1)|$ by the technique of Theorem 4.1. We have

$$D^sF(w_1) = s!\, \frac{F(z)}{(z - w_1)^s}\bigg|_{z=w_1}$$

We estimate the function

$$H(z) = s!\, \frac{F(z)}{[(z - w_1) \cdots (z - w_N)]^s} \prod_{v \neq 1} (w_1 - w_v)^s$$

on the circle of radius $R = s^{1/2\rho}$. By the maximum modulus principle, we have

$$|D^s F(w_1)| = |H(w_1)| \leqq \|H\|_R \leqq s^s C^{Ns} \|F\|_R / R^{Ns}$$

for a suitable constant C. Using the estimate for the coefficients of F, and the order of growth of the functions $f_1, \ldots, f_n$, together with the fact that $L \leqq M \leqq s$, we obtain

$$\|F\|_R \leqq \frac{C_3^s s^s C_4^{R^\rho L}}{R^{Ns}} \leqq \frac{C_3^s s^s C_4^{sN}}{e^{Ns(\log s)/2\rho}}.$$

Hence

$$1 \leqq |c\alpha| = |cD^s F(w_1)| \leqq \frac{s^{2s} C_5^{Ns}}{e^{Ns(\log s)/2\rho}}.$$

Taking logs yields

$$\frac{Ns \log s}{2\rho} \leqq 2s \log s + C_6 Ns.$$

We divide by $s \log s$, and let $L \to \infty$ at the beginning of the proof, so $s \to \infty$. The inequality

$$N \leqq 4\rho$$

drops out, thereby proving the theorem.

XII, §6. THE PHRAGMEN–LINDELÖF AND HADAMARD THEOREMS

We write a complex number in the form

$$s = \sigma + it$$

with real σ, t.

We shall use the O notation as follows. Let f, g be functions defined on a set S, and g real positive. We write

$$f(z) = O(g(z)) \qquad \text{or} \qquad |f(z)| \ll g(z) \qquad \text{for} \quad |z| \to \infty$$

if there is a constant C such that $|f(z)| \leqq Cg(z)$ for $|z|$ sufficiently large. When the context makes it clear, we omit the reference that $|z| \to \infty$.

Consider a strip of complex numbers s such that the real part σ lies in some finite interval $[\sigma_1, \sigma_2]$. We are interested in conditions under which f is bounded in the strip. Suppose that $|f| \leqq 1$ on the sides of the strip. It turns out that if f has some bound on its order of growth inside the strip, then $|f| \leqq 1$ on the whole strip. The double exponential $f(s) = e^{e^{is}}$ on the strip $[-\pi/2, \pi/2]$ gives an example when f is not bounded on the interior. However, one has:

Theorem 6.1 (Phragmen–Lindelöf). *Let f be continuous on the strip*

$$-\pi/2 \leqq \sigma \leqq \pi/2$$

and holomophic on the interior. Suppose $|f| \leqq 1$ on the sides of the strip, and suppose there exists $0 < \alpha < 1$ and $C > 0$ such that

$$|f(s)| \leqq \exp(Ce^{\alpha|s|})$$

for s in the strip with $|s| \to \infty$. Then $|f| \leqq 1$ in the strip.

Proof. Let $\alpha < \beta < 1$. For each $\epsilon > 0$, define

$$g_\epsilon(s) = f(s)e^{-2\epsilon \cos(\beta s)}.$$

We have $2\,\mathrm{Re}(\cos \beta s) = (e^{\beta t} + e^{-\beta t})\cos \beta\sigma$. Since $0 < \beta < 1$, it follows that there exists $c > 0$ such that $c \leqq \cos \beta\sigma$ for all $\sigma \in [-\pi/2, \pi/2]$. Hence

$$|g_\epsilon(s)| \to 0 \qquad \text{as } |s| \to \infty, \quad s \text{ in the strip.}$$

Since it is immediate that g_ϵ is bounded by 1 on the sides of the strip, it follows from the maximum modulus principle applied to truncated rectangles of the strip that $|g_\epsilon| \leqq 1$ on the whole strip. Then we conclude that

$$|f(s)| \leqq e^{2\epsilon \,\mathrm{Re} \cos \beta s} \qquad \text{for all } s \text{ in the strip.}$$

We now fix some finite rectangle, and take s in this rectangle. The previous inequality holds for every $\epsilon > 0$, and the values $|\cos \beta s|$ are bounded for s in the rectangle. This implies that $|f(s)| \leqq 1$ for s in the rectangle, and concludes the proof of the theorem.

From Theorem 6.1 we shall obtain several variations. First, we say that a function f on a strip $\sigma_1 \leqq \sigma \leqq \sigma_2$ is of **finite order** on the strip if there exists $\lambda > 0$ such that

$$\log|f(s)| \ll |s|^\lambda \qquad \text{for } |s| \to \infty, \ s \text{ in the strip.}$$

Theorem 6.2 (Phragmen–Lindelöf, second version). *Let f be continuous on a strip $\mathrm{Re}(s) \in [\sigma_1, \sigma_2]$ and holomophic on the interior. Suppose $|f| \leqq 1$ on the sides of the strip, and f is of finite order on the strip. Then $|f| \leqq 1$ on the strip.*

Proof. By a linear change of variables $s = aw + b = \varphi(w)$ we see that the given strip corresponds to a strip as in Theorem 6.1. It is immediately verified that $f \circ \varphi$ satisfies the hypothesis of Theorem 6.1, so $|f \circ \varphi| \leqq 1$, and finally $|f| \leq 1$, which proves the corollary.

Remark. The bound of 1 on the sides is a convenient normalization. If instead $|f| \leqq B$ on the sides of the strip for some $B > 0$, then by considering f/B instead of f one concludes that $|f| \leqq B$ inside the strip.

The Phragmen–Lindelöf can be applied to give estimates for a function even when the function is not bounded in the strip. For instance, we obtain the following statement as an immediate consequence.

Corollary 6.3. *Suppose that f is of finite order in the strip, and that there is some positive integer M such that*

$$f(\sigma_1 + it) = O(|t|^M) \qquad \text{for } |t| \to \infty$$

and similarly for σ_2 instead of σ_1, that is $|f(s)|$ is bounded polynomially on the sides of the strip. Then for s in the strip, we have

$$f(s) = O(|s|^M) \qquad \text{for } |s| \to \infty.$$

Proof. Let s_0 be some point away from the strip. Then the function

$$g(s) = f(s)/(s - s_0)^M.$$

is bounded in the strip and we can apply the Phragment–Lindelöf theorem to conclude the proof.

Example. Define the **zeta function**

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}.$$

It is easily shown that the series converges absolutely for $\operatorname{Re}(s) > 1$, and uniformly on every compact subset of this right half-plane, thus defining a holomorphic function. It is also easily shown that

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1} \qquad \text{(called the \textbf{Euler product})},$$

where the product is taken over all the prime numbers. In Chapter XIII you will learn about the gamma function $\Gamma(s)$. Define

$$f(s) = s(1 - s)\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s).$$

It can be shown that there is an extension of f to an entire function, satisfying the functional equation

$$f(s) = f(1 - s).$$

For $\operatorname{Re}(s) \geqq 2$, say, the Euler product is uniformly bounded. The functional equation gives a bound

$$f(\sigma + it) = O(|t|^M) \qquad \text{for } |t| \to \infty$$

for each $\sigma \leqq -2$. The Phragmen–Lindelöf theorem then gives a bound in the middle.

In the Phragmen–Lindelöf theorem we were interested in the crude asymptotic behavior for large t. In the next theorem, we want a more refined behavior, and so we must assume that the function is holomorphic and bounded in a whole strip.

Theorem 6.4 (First Convexity Theorem). *Let $s = \sigma + it$. Let f be holomorphic and bounded on the strip $a \leqq \sigma \leqq b$. For each σ let*

$$M_f(\sigma) = M(\sigma) = \sup_t |f(\sigma + it)|.$$

Then $\log M(\sigma)$ *is a convex function of σ.*

Proof. The statement is defined to be the inequality

$$\log M(\sigma) \leqq \frac{b - \sigma}{b - a} \log M(a) + \frac{\sigma - a}{b - a} \log M(b)$$

Expressed in multiplicative notation, this is equivalent with

$$M(\sigma)^{b-a} \leqq M(a)^{b-\sigma} M(b)^{\sigma-a}.$$

The case when $M(a) = M(b) = 1$ is simply a special case of the Phragmen–Lindelöf theorem.

In general, let

$$h(s) = M(a)^{(b-s)/(b-a)} M(b)^{(s-a)/(b-a)}.$$

Then h is entire, has no zeros, and $1/h$ is bounded on the strip. We have

$$|h(a + it)| = M(a) \qquad \text{and} \qquad |h(b + it)| = M(b)$$

for all t. Consequently,

$$M_{f/h}(a) = M_{f/h}(b) = 1.$$

The first part of the theorem implies that $|f/h| \leqq 1$, whence $|f| \leqq |h|$, thus proving our theorem.

Corollary 6.5 (Hadamard Three Circle Theorem). *Let $f(z)$ be holomorphic on the annulus $\alpha \leqq |z| \leqq \beta$, centered at the origin. Let*

$$M(r) = \sup_{|z|=r} |f(z)|.$$

Then $\log M(r)$ *is a convex function of* $\log r$. *In other words,*

$$\log(\beta/\alpha) \log M(r) \leqq \log(\beta/r) \log M(\alpha) + \log(r/\alpha) \log M(\beta).$$

Proof. Let $f^*(s) = f(e^s)$. Then f^* is holomorphic and bounded on the strip $a \leqq \sigma \leqq b$, where $e^a = \alpha$ and $e^b = \beta$. We simply apply the theorem, to get the corollary.

In the next corollary, we analyze a growth exponent. Let f be holomorphic in the neighborhood of a vertical line $\sigma + it$, with fixed σ, and suppose that

$$f(\sigma + it) \ll |t|^\gamma$$

for some positive number γ. The inf of all such γ can be called the **growth exponent** of f, and will be denoted by $\psi(\sigma)$. Thus

$$f(\sigma + it) \ll |t|^{\psi(\sigma)+\epsilon}$$

for every $\epsilon > 0$, and $\psi(\sigma)$ is the least exponent which makes this inequality true.

Theorem 6.6 (Second Convexity Theorem). *Let f be holomorphic in the strip $a \leqq \sigma \leqq b$. For each σ assume that $f(\sigma + it)$ grows at most like a power of $|t|$, and let $\psi(\sigma)$ be the least number $\geqq 0$ for which*

$$f(\sigma + it) \ll |t|^{\psi(\sigma)+\epsilon}$$

for every $\epsilon > 0$. Assume also that f is of finite order in the strip. Then $\psi(\sigma)$ is convex as a function of σ, and in particular is continuous on $[a, b]$.

Proof. The corollary of the Phragmen–Lindelöf theorem shows that there is a uniform M such that $f(\sigma + it) \ll |t|^M$ in the strip. Let $L_\epsilon(s)$ be the formula for the straight line segment between $\psi(a) + \epsilon$ and $\psi(b) + \epsilon$; in other words, let

$$L_\epsilon(s) = \frac{b - s}{b - a}[\psi(a) + \epsilon] + \frac{s - a}{b - a}[\psi(b) + \epsilon].$$

The function

$$f(s)(-is)^{-L_\epsilon(s)}$$

is then immediately seen to be bounded in the strip, and our theorem follows, since we get $\psi(\sigma) \leqq L_\epsilon(\sigma)$ for each σ in the strip, and every $\epsilon > 0$.

Example. We have already mentioned the zeta function. It is a standard fact that it has a pole of order 1 at $s = 1$, and is otherwise holomorphic. Hence let us put

$$f(s) = (s - 1)\zeta(s).$$

Then f is an entire function. We then have the corresponding $\psi(\sigma)$. The **Riemann hypothesis** states that all the zeros of $\zeta(s)$ (or $f(s)$) in the strip $0 < \sigma < 1$ lie on the line $\mathrm{Re}(s) = 1/2$. It can be shown fairly easily that this hypothesis is equivalent to the property that the graph of ψ is as shown on Fig. 4, in other words, $\psi(\sigma) = 0$ for $\sigma \geqq 1/2$. Cf. my *Algebraic Number Theory*, Springer-Verlag, Chapter XIII, §5.

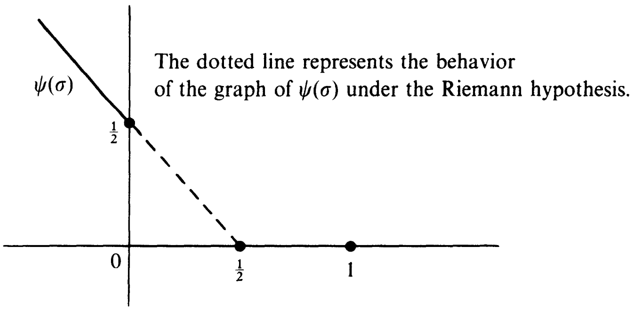


Figure 4

XII, §6. EXERCISES

Phragmen–Lindelöf for Sectors

1. Let U be the right half plane ($\mathrm{Re}\, z > 0$). Let f be continuous on the closure of U and analytic on U. Assume that there are constants $C > 0$ and $\alpha < 1$ such that

$$|f(z)| \leqq Ce^{|z|^\alpha}$$

for all z in U. Assume that f is bounded by 1 on the imaginary axis. Prove that f is bounded by 1 on U. Show that the assertion is not true if $\alpha = 1$.

2. More generally, let U be the open sector between two rays from the origin. Let f be continuous on the closure of U (i.e. the sector and the rays), and analytic on U. Assume that there are constants $C > 0$ and α such that

$$|f(z)| \leqq Ce^{|z|^{\alpha}}$$

for all $z \in U$. If π/β is the angle of the sector, assume that $0 < \alpha < \beta$. If f is bounded by 1 on the rays, prove that f is bounded by 1 on U.

3. Consider again a finite strip $\sigma_1 \leqq \sigma \leqq \sigma_2$. Suppose that f is holomorphic on the strip, $|f(s)| \to 0$ as $|s| \to \infty$ with s in the strip, and $|f(s)| \leqq 1$ on the sides of the strip. Prove that $|f(s)| \leqq 1$ in the strip.
(**Remark.** The bound of 1 is used only for normalization purposes. If f is bounded by some constant B, then dividing f by B reduces the problem to the case when the bound is 1.)

4. Let f be holomorphic on the disc D_R of radius R. For $0 \leqq r < R$ let

$$I(r) = \frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\theta})|^2 \, d\theta.$$

Let $f = \sum a_n z^n$ be the power series for f.
(a) Show that
$$I(r) = \sum |a_n|^2 r^{2n}.$$

(b) $I(r)$ is an increasing function of r.
(c) $|f(0)|^2 \leqq I(r) \leqq \|f\|_r^2$.
(d) $\log I(r)$ is a convex function of $\log r$, assuming that f is not the zero function. [*Hint*: Put $s = \log r$,

$$J(s) = I(e^s).$$

Show that $(\log J)'' = \dfrac{J''J - (J')^2}{J^2}$. Use the Schwarz inequality to show that

$$J''J - (J')^2 \geqq 0.]$$

CHAPTER XIII

Entire and Meromorphic Functions

A function is said to be **entire** if it is analytic on all of **C**. It is said to be **meromorphic** if it is analytic except for isolated singularities which are poles. In this chapter we describe such functions more closely. We develop a multiplicative theory for entire functions, giving factorizations for them in terms of their zeros, just as a polynomial factors into linear factors determined by its zeros. We develop an additive theory for meromorphic functions, in terms of their principal part (polar part) at the poles.

Examples of classical functions illustrating the general theory of Weierstrass products and Mittag-Leffler expansions will be given in the three subsequent chapters: elliptic functions, gamma functions, and zeta functions. For other classical functions (including hypergeometric functions), see, for instance, Copson [Co 35].

XIII, §1. INFINITE PRODUCTS

Let $\{u_n\}$ $(n = 1, 2, \ldots)$ be a sequence of complex numbers $\neq 0$. We say that the infinite product

$$\prod_{n=1}^{\infty} u_n$$

converges absolutely if $\lim u_n = 1$ and if the series

$$\sum_{n=1}^{\infty} \log u_n$$

converges absolutely, i.e. $\sum |\log u_n|$ converges. For a finite number of n

we take any determination of $\log u_n$, but if n is sufficiently large, then u_n can be written in the form $u_n = 1 - \alpha_n$ where $|\alpha_n| < 1$, and then we let $\log u_n = \log(1 - \alpha_n)$ be defined by the usual series for the logarithm. Under this condition of absolute convergence, it follows that the series

$$\sum \log u_n = \sum \log(1 - \alpha_n)$$

converges, so the partial sums

$$\sum_{n=1}^{N} \log u_n$$

have a limit. Since the exponential function is continuous, we can exponentiate these partial sums (which are the log of the partial products) and we see that

$$\prod_{n=1}^{\infty} u_n = \lim_{N\to\infty} \prod_{n=1}^{N} u_n$$

exists.

Lemma 1.1. *Let $\{\alpha_n\}$ be a sequence of complex numbers $\alpha_n \neq 1$ for all n. Suppose that*

$$\sum |\alpha_n|$$

converges. Then

$$\prod_{n=1}^{\infty} (1 - \alpha_n)$$

converges absolutely.

Proof. For all but a finite number of n, we have $|\alpha_n| < \frac{1}{2}$, so

$$\log(1 - \alpha_n)$$

is defined by the usual series, and for some constant C,

$$|\log(1 - \alpha_n)| \leqq C|\alpha_n|.$$

Hence the product converges absolutely by the definition and our assumption on the convergence of $\sum |\alpha_n|$.

The lemma reduces the study of convergence of an infinite product to the study of convergence of a series, which is more easily manageable.

In the applications, we shall consider a product

$$\prod_{n=1}^{\infty} (1 - h_n(z)),$$

where $h_n(z)$ is a function $\neq 1$ for all z in a certain set K, and such that we have a bound

$$|h_n(z)| \leqq |\alpha_n|$$

for all but a finite number of n, and all z in K. Taking again the partial sums of the logarithms,

$$\sum_{n=1}^{N} \log(1 - h_n(z)),$$

we can compare this with the sums of the lemma to see that these sums converge absolutely and uniformly on K. Hence the product taken with the functions $h_n(z)$ converges absolutely and uniformly for z in K.

It is useful to formulate a lemma on the logarithmic derivative of an infinite product, which will apply to those products considered above.

Lemma 1.2. *Let $\{f_n\}$ be a sequence of analytic functions on an open set U. Let $f_n(z) = 1 + h_n(z)$, and assume that the series*

$$\sum h_n(z)$$

converges uniformly and absolutely on U. Let K be a compact subset of U not containing any of the zeros of the functions f_n for all n. Then the product $\prod f_n$ converges to an analytic function f on U, for $z \in K$ we have

$$f'/f(z) = \sum_{n=1}^{\infty} f_n'/f_n(z),$$

and the convergence is absolute and uniform on K.

Proof. By covering K with a finite number of discs of sufficiently small radius, using the compactness, we may assume that K is a closed disc. Write

$$f(z) = \prod_{n=1}^{N-1} f_n(z) \prod_{n=N}^{\infty} f_n(z),$$

where N is picked so large that $\sum_{n=N}^{\infty} |h_n(z)| < 1$. Then the series

$$\sum_{n=N}^{\infty} \log f_n(z) = \sum_{n=N}^{\infty} \log(1 + h_n(z))$$

converges uniformly and absolutely, to define a determination of

$$\log G(z), \qquad \text{where} \quad G(z) = \prod_{n=N}^{\infty} f_n(z).$$

Then

$$f'/f(z) = \sum_{n=1}^{N} f_n'/f_n(z) + G'/G(z).$$

But we can differentiate the series for $\log G(z)$ term by term, whence the expression for $f'/f(z)$ follows. If the compact set is away from the zeros of f_n for all n, then the convergence is clearly uniform, as desired.

XIII, §1. EXERCISES

1. Let $0 < |\alpha| < 1$ and let $|z| \leqq r < 1$. Prove the inequality

$$\left|\frac{\alpha + |\alpha| z}{(1 - \bar{\alpha} z)\alpha}\right| \leqq \frac{1 + r}{1 - r}.$$

2. **(Blaschke Products).** Let $\{\alpha_n\}$ be a sequence in the unit disc D such that $\alpha_n \neq 0$ for all n, and

$$\sum_{n=1}^{\infty} (1 - |\alpha_n|)$$

converges. Show that the product

$$f(z) = \prod_{n=1}^{\infty} \frac{\alpha_n - z}{1 - \bar{\alpha}_n z} \frac{|\alpha_n|}{\alpha_n}$$

converges uniformly for $|z| \leqq r < 1$, and defines a holomorphic function on the unit disc having precisely the zeros α_n and no other zeros. Show that $|f(z)| \leqq 1$,

3. Let $\alpha_n = 1 - 1/n^2$ in the preceding exercise. Prove that

$$\lim_{x \to 1} f(x) = 0 \qquad \text{if} \quad 0 < x < 1.$$

In fact, prove the estimate for $\alpha_{n-1} < x < \alpha_n$:

$$|f(x)| < \prod_{k=1}^{n-1} \frac{x - \alpha_k}{1 - \alpha_k x} < \prod_{k=1}^{n-1} \frac{\alpha_n - \alpha_k}{1 - \alpha_k} < 2e^{-n/3}.$$

4. Prove that there exists a bounded analytic function f on the unit disc for which each point of the unit circle is a singularity.

5. **(q-Products).** Let $z = x + iy$ be a complex variable, and let $\tau = u + iv$ with u, v real, $v > 0$ be a variable in the upper half-plane H. We define

$$q_\tau = e^{2\pi i\tau} \qquad \text{and} \qquad q_z = e^{2\pi iz}.$$

Consider the infinite product

$$(1 - q_z) \prod_{n=1}^{\infty} (1 - q_\tau^n q_z)(1 - q_\tau^n / q_z).$$

(a) Prove that the infinite product is absolutely convergent.

(b) Prove that for fixed τ, the infinite product defines a holomorphic function of z, with zeros at the points

$$m + n\tau, \quad m, n \text{ integers.}$$

We define the **second Bernoulli polynomial**

$$\mathbf{B}_2(y) = y^2 - y + \tfrac{1}{6}.$$

Define the **Néron–Green function**

$$\lambda(z, \tau) = \lambda(x, y, \tau) = -\log\left| q_\tau^{\mathbf{B}_2(y/v)/2}(1 - q_z) \prod_{n=1}^{\infty} (1 - q_\tau^n q_z)(1 - q_\tau^n/q_z) \right|.$$

(c) Prove that for fixed τ, the function $z \mapsto \lambda(z, \tau)$ is periodic with periods $1, \tau$.

XIII, §2. WEIERSTRASS PRODUCTS

Let f, g be entire functions with the same zeros, at which they have the same multiplicities. Then f/g is an entire function without zeros. We first analyze this case.

Theorem 2.1. *Let f be an entire function without zeros. Then there exists an entire function h such that*

$$f(z) = e^{h(z)}.$$

Proof. Since $\mathbf{C}$ is simply connected, this is merely a restatement of the result of Chapter III, §6 where we defined the logarithm $\log f(z)$ for any function f which has no zeros.

We see that if f, g are two functions with the same zeros and same multiplicites, then

$$f(z) = g(z)e^{h(z)}$$

for some entire function $h(z)$. Conversely, if $h(z)$ is entire, then $g(z)e^{h(z)}$ has the same zeros as g, counted with their multiplicities.

We next try to give a standard form for a function with prescribed zeros. Suppose we order these zeros by increasing absolute value, so let $z_1, z_2, \ldots$ be a sequence of complex numbers $\neq 0$, satisfying

$$|z_1| \leqq |z_2| \leqq \cdots.$$

Assume that $|z_n| \to \infty$ as $n \to \infty$. If we try to define the function by the product

$$\prod_{n=1}^{\infty}\left(1-\frac{z}{z_n}\right),$$

then we realize immediately that this product may not converge, and so we have to insert a convergence factor. We do not want this factor to introduce new zeros, so we make it an exponential. We want it to be as simple as possible, so we make it the exponential of a polynomial, whose degree will depend on the sequence of z_n. Thus we are led to consider factors of the form

$$E_n(z) = (1-z)e^{z+z^2/2+\cdots+z^{n-1}/(n-1)}.$$

The polynomial in the exponent is exactly what is needed to cancel the first n terms in the series for the log, so that

$$\log E_n(z) = \log(1-z) + z + \frac{z^2}{2} + \cdots + \frac{z^{n-1}}{n-1}$$

$$= \sum_{k=n}^{\infty} \frac{-z^k}{k}.$$

Lemma 2.2. *If* $|z| \leqq 1/2$, *then*

$$|\log E_n(z)| \leqq 2|z|^n.$$

Proof.

$$|\log E_n(z)| \leqq \frac{|z|^n}{n}\sum_{k=0}^{\infty}\frac{1}{2^k} \leqq 2|z|^n.$$

Given the sequence $\{z_n\}$, we pick integers k_n such that the series

$$\sum_{n=1}^{\infty}\left(\frac{R}{|z_n|}\right)^{k_n}$$

converges for all positive real R. Since $|z_n| \to \infty$ we can find such k_n, for instance $k_n = n$, so that $(R/|z_n|)^n \leqq 1/2^n$ for $n \geqq n_0(R)$. We let

$$P_n(z) = z + \frac{z^2}{2} + \cdots + \frac{z^{k_n-1}}{k_n-1}$$

and

$$E_n(z, z_n) = \left(1-\frac{z}{z_n}\right)e^{P_n(z/z_n)}.$$

Note that

$$P_n\left(\frac{z}{z_n}\right) = \frac{z}{z_n} + \frac{1}{2}\left(\frac{z}{z_n}\right)^2 + \cdots + \frac{1}{k_n - 1}\left(\frac{z}{z_n}\right)^{k_n - 1}.$$

Theorem 2.3. *Given the sequences $\{z_n\}$, $\{k_n\}$, $\{P_n\}$ as above; if the series*

$$\sum_{n=1}^{\infty}\left(\frac{R}{|z_n|}\right)^{k_n}$$

converges for all positive real R (which is the case if $k_n = n$), then the product

$$\prod_{n=1}^{\infty} E_n(z, z_n) = \prod_{n=1}^{\infty}\left(1 - \frac{z}{z_n}\right)e^{P_n(z/z_n)}$$

converges uniformly and absolutely on every disc $|z| \leqq R$, and defines an entire function with zeros at the points of the sequence $\{z_n\}$, and no other zeros.

Proof. Fix R. Let N be such that

$$|z_N| \leqq 2R < |z_{N+1}|.$$

Then for $|z| \leqq R$ and $n > N$ we have $|z/z_n| \leqq 1/2$, and hence

$$|\log E_n(z, z_n)| \leqq 2\left(\frac{R}{|z_n|}\right)^{k_n}.$$

Therefore the series

$$\sum_{N+1}^{\infty} \log E_n(z, z_n)$$

converges absolutely and uniformly when $|z| \leqq R$, thereby implying the absolute and uniform convergence of the exponentiated product.

The limiting function obviously has the sequence $\{z_n\}$ as zeros, with the multiplicity equal to the number of times that z_n is repeated in the sequence. We still have to show that the limiting function has no other zeros. We fix some radius R and consider only $|z| \leqq R$. Given ϵ there exists N_0 such that if $N \geqq N_0$, then

$$\left|\log \prod_{n=N_0}^{N} E_n(z, z_n)\right| = \left|\sum_{n=N_0}^{N} \log E_n(z, z_n)\right| < \epsilon,$$

by the absolute uniform convergence of the log sequence proved previously. Hence the product

$$\prod_{N_0}^{N} E_n(z, z_n)$$

is close to 1. But

$$f(z) = \prod_{n=1}^{N_0-1} E_n(z, z_n) \lim_{N\to\infty} \prod_{N_0}^{N} E_n(z, z_n).$$

The first product on the right has the appropriate zeros in the disc $|z| \leqq R$. The limit of the second product on the right is close to 1, and hence has no zero. This proves the theorem.

The sequence $\{z_n\}$ was assumed such that $z_n \neq 0$. Of course an entire function may have a zero at 0, and to take this into account, we have to form

$$\boxed{z^m \prod_{n=1}^{\infty} \left(1 - \frac{z}{z_n}\right) e^{P_n(z/z_n)}.}$$

This function has the same zeros as the product in Theorem 2.3, with a zero of order m at the origin in addition.

In most of the examples, one can pick $k_n - 1$ equal to a fixed integer, which in all the applications we shall find in this book is equal to 1 or 2. The chapter on elliptic functions gives examples of order 2. We now give an example of order 1.

Example 2.4. We claim that

$$\frac{\pi^2}{\sin^2 \pi z} = \sum_{n\in \mathbf{Z}} \frac{1}{(z-n)^2},$$

$$\pi \cot \pi z = \pi \frac{\cos \pi z}{\sin \pi z} = \frac{1}{z} + \sum_{n\neq 0} \left(\frac{1}{z-n} + \frac{1}{n}\right),$$

$$\sin \pi z = \pi z \prod_{n\neq 0} \left(1 - \frac{z}{n}\right) e^{z/n} = \pi z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right).$$

Proof. Consider the difference

$$h(z) = \frac{1}{z} + \sum_{n\neq 0} \left(\frac{1}{z-n} + \frac{1}{n}\right) - \pi \frac{\cos \pi z}{\sin \pi z}. \tag{1}$$

The function $\pi \cot \pi z$ has simple poles at the integers, with residue 1, and so does the sum on the right-hand side. Hence the difference on the right-hand side representing the function h is an entire function, which we must prove is equal to 0. We take the derivative

$$h'(z) = \sum_{n \in \mathbf{Z}} \frac{-1}{(z-n)^2} + \frac{\pi^2}{\sin^2 \pi z}. \tag{2}$$

It is clear that h' is periodic of period 1 (but it would not be so immediate that the function h itself is periodic). Next we prove that h' is constant. By Liouville's theorem, it suffices to prove that h' is bounded. But for $|y| \geqq 2$ (say), putting $z = x + iy$, it is immediate that the series

$$\sum_{n \in \mathbf{Z}} \frac{1}{|x + iy - n|^2} \qquad \text{with} \quad 0 \leqq x \leqq 1$$

is bounded, and in fact the series approaches 0 as $|y| \to \infty$. Furthermore, substituting $z = x + iy$ directly in the expression

$$\sin \pi z = \frac{e^{i\pi z} - e^{-i\pi z}}{2i}$$

shows that $1/\sin^2 \pi z$ is also bounded in the cut-off strips $0 \leqq x \leqq 1$, $|y| \geqq 2$, and approaches 0 as $|y| \to \infty$ (like $e^{-2\pi|y|}$). Since h' is entire, it follows that h' is continuous also for $|y| \leqq 2$, whence h' is bounded in the entire strip, and is therefore bounded on $\mathbf{C}$ by periodicity. Hence h' is constant by Liouville's theorem. Furthermore, this constant is 0 because as we have seen, $h'(x + iy) \to 0$ as $|y| \to \infty$.

Having proved $h' = 0$ we now conclude that h itself is constant. Immediately from the definition (1), one sees that $h(0) = 0$, whence $h = 0$ and therefore the formula for $\cot \pi z$ has been proved.

Finally we consider $\sin \pi z$. We let $z_n = n$ where n ranges over the integers $\neq 0$ (possibly negative). Let $k_n = 2$, $k_n - 1 = 1$. The series

$$\sum_{n=1}^{\infty} (R/n)^2$$

converges for all R. Hence Theorem 2.3 implies that the function

$$f(z) = \pi z \prod_{n \neq 0} \left(1 - \frac{z}{n}\right) e^{z/n}$$

is an entire function with zeros of order 1 at the integers. *We show that it is equal to* $\sin \pi z$. The Weierstrass product has the property stated in Lemma 1.2, and therefore taking the logarithmic derivative term by term yields

$$f'/f(z) = \frac{1}{z} + \sum_{n \neq 0} \left(\frac{1}{z-n} + \frac{1}{n}\right).$$

But

$$\cot \pi z = \pi \frac{\cos \pi z}{\sin \pi z} = \frac{d(\sin \pi z)/dz}{\sin \pi z}.$$

Since the logarithmic derivatives of $\sin \pi z$ and $f(z)$ are equal, we now conclude that there exists a constant C such that

$$f(z) = C \sin \pi z.$$

We divide both sides by πz and let z tend to 0. We then see that $C = 1$, thereby proving the Weierstrass product expression for $\sin \pi z$.

Remark. Of course, we gave two version of the Weierstrass product for $\sin \pi z$. In the second version, the product is taken for $n = 1, 2, \ldots$ ranging over the positive integers. As a simple exercise, justify the fact that we could combine the terms with n positive and the term with $-n$, so that the exponential factor cancels in this second version.

XIII, §2. EXERCISES

1. Let f be an entire function, and n a positive integer. Show that there is an entire function g such that $g^n = f$ if and only if the orders of the zeros of f are divisible by n.

2. Prove that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

[*Hint*: Use the constant term of the Laurent expansion of $\pi^2/\sin^2 \pi z$ at $z = 0$.]

3. More generally, show:

(a) $\pi z \cot \pi z = 1 - 2 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} z^{2m}/n^{2m}$.

(b) Define the **Bernoulli numbers** B_k by the series

$$\frac{t}{e^t - 1} = 1 - \frac{t}{2} + \sum_{k=2}^{\infty} B_k \frac{t^k}{k!}.$$

Setting $t = 2i\pi z$ and comparing coefficients, prove:

If k is an even positive integer, then

$$2\zeta(k) = -\frac{B_k}{k!}(2\pi i)^k.$$

If $k = 2$, you recover the computation of Exercise 2.

4. In the terminology of algebra, the set E of entire functions is a ring, and in fact a subring of the ring of all functions; namely E is closed under addition and multiplication, and contains the function 1. By an **ideal** J, we mean a subset of E such that if $f, g \in J$ then $f + g \in J$, and if $h \in E$ then $hf \in J$. In other words, J is closed under multiplication elements of E, and under addition. If there exists functions $f_1, \ldots, f_r \in J$ such that all elements of J can be expressed in the form $A_1 f_1 + \cdots + A_r f_r$ with $A_i \in E$, then we call $f_1, \ldots, f_r$ **generators** of J, and we say that J is **finitely generated**. Give an example of an ideal of E which is not finitely generated.

   For finitely generated ideals, see the exercises of §4.

XIII, §3. FUNCTIONS OF FINITE ORDER

It is useful to have a simple criterion when the integers k_n in the Weierstrass product can all be taken equal to a fixed integer. Let ρ be a positive real number.

An entire function f is said to be of **order** $\leqq \rho$ if, given $\epsilon > 0$, there exists a constant C (depending on ϵ) such that

$$\|f\|_R \leqq C^{R^{\rho+\epsilon}} \qquad \text{for all } R \text{ sufficiently large.}$$

or equivalently,

$$\log \|f\|_R \ll R^{\rho+\epsilon} \qquad \text{for} \quad R \to \infty.$$

Observe that to verify such an asymptotic inequality, it suffices to verify it when R ranges through positive integers N, because we have

$$(N+1)^{\rho+\epsilon} \ll N^{\rho+\epsilon} \qquad \text{for } N \text{ sufficiently large.}$$

A function is said to be of **strict order** $\leqq \rho$ if the same estimate holds without the ϵ, namely

$$\|f\|_R \leqq C^{R^{\rho}} \qquad \text{for } R \text{ sufficiently large.}$$

The function is said to be of **order** ρ if ρ is the greatest lower bound of those positive numbers which make the above inequality valid, and similarly for the definition of the exact **strict order**.

Example. The function e^z has strict order 1 because $|e^z| = e^x \leqq e^{|z|}$.

Further examples will be given by constructing functions of order ρ using Weierstrass products, as follows.

Let $\rho > 0$. Let k be the smallest integer $> \rho$. Let $\{z_n\}$ be a sequence of complex numbers $\neq 0$, ordered by increasing absolute value, and such that

$$\sum \frac{1}{|z_n|^{\rho+\epsilon}}$$

converges for every ϵ. As before let

$$P_k(z) = P(z) = z + \frac{z^2}{2} + \cdots + \frac{z^{k-1}}{k-1}.$$

We call

$$\boxed{E^{(k)}(z, \{z_n\}) = E(z) = \prod \left(1 - \frac{z}{z_n}\right) e^{P(z/z_n)} = \prod E_k(z, z_n)}$$

the **canonical product** determined by the sequence $\{z_n\}$ and the number ρ, or canonical product for short. Lemma 2.2 shows that it defines an entire function.

Theorem 3.1. *The above canonical product is an entire function of order $\leqq \rho$.*

Proof. Let $\epsilon > 0$ be such that $\rho + \epsilon < k$ and let $\lambda = \rho + \epsilon$. There exists a constant C such that

$$|E_k(z)| = |(1 - z)e^{P(z)}| \leqq C^{|z|^\lambda}.$$

This is true for $|z| \leqq \frac{1}{2}$ by Lemma 2.2, and it is even more obvious for $|z| \geqq 1$ and $\frac{1}{2} \leqq |z| \leqq 1$. Then

$$|E^{(k)}(z, \{z_n\})| \leqq \prod C^{|z/z_n|^\lambda} \leqq C^{|z|^\lambda \Sigma 1/|z_n|^\lambda},$$

thus proving the theorem.

The next results essentially prove the converse of what we just did, by showing that all entire functions of order $\leqq \rho$ are essentially Weierstrass products of the above type, up to an exponential factor and a power of z.

Theorem 3.2. *Let f be an entire function of strict order $\leqq \rho$. Let $v_f(R)$ be the number of zeros of f in the disc of radius R. Then*

$$v_f(R) \ll R^\rho.$$

Proof. Dividing f by a power of z if necessary, we may assume without loss of generality that f does not vanish at the origin. The estimate is an immediate consequence of Jensen's inequality, but corresponds to a coarse form of it, which can be proved ad hoc. Indeed, let $z_1, \ldots, z_n$ be the zeros in the circle of radius R, and let

$$g(z) = \frac{f(z)}{(z - z_1)\cdots(z - z_n)} z_1 \cdots z_n.$$

Then g is entire, $g(0) = \pm f(0)$, and we apply the maximum modulus principle to the disc of radius $3R$, say. Then the inequality

$$|f(0)| \leqq \|f\|_{3R}/2^n$$

falls out, thereby proving Theorem 3.2 directly.

Theorem 3.3. *Let f have strict order $\leqq \rho$, and let $\{z_n\}$ be the sequence of zeros $\neq 0$ of f, repeated with their multiplicities, and ordered by increasing absolute value. For every $\delta > 0$ the series*

$$\sum \frac{1}{|z_n|^{\rho+\delta}}$$

converges

Proof. We sum by parts with a positive integer $R \to \infty$:

$$\sum_{|z_n| \leqq R} \frac{1}{|z_n|^{\rho+\delta}} \ll \sum_{r=1}^{R} \frac{\nu(r+1) - \nu(r)}{r^{\rho+\delta}}$$

$$\ll \frac{\nu(R)}{R^{\rho+\delta}} + \sum_{r=1}^{R} \frac{\nu(r)}{r^{\rho+\delta+1}} + \text{constant}.$$

Each quotient $\nu(r)/r^{\rho}$ is bounded, so the first term is bounded, and the sum is bounded by $\sum 1/r^{1+\delta}$ which converges. This proves the theorem.

Theorem 3.4 (Minimum Modulus Theorem). *Let f be an entire function of order $\leqq \rho$. Let $z_1, z_2, \ldots$ be its sequence of zeros, repeated according to their multiplicities. Let $s > \rho$. Let U be the complement of the closed discs of radius $1/|z_n|^s$ centered at z_n, for $|z_n| > 1$. Then there exists $r_0(\epsilon, f)$ such that for $z \in U$, $|z| = r > r_0(\epsilon, f)$ we have*

$$|f(z)| > e^{-r^{\rho+\epsilon}}, \qquad \text{i.e.} \quad \log|f(z)| > -r^{\rho+\epsilon}.$$

Proof. We first prove the minimum modulus theorem for the canonical product. Let $|z| = r$. We write

$$E(z) = \prod_{|z_n|<2r} \left(1 - \frac{z}{z_n}\right) \prod_{|z_n|<2r} e^{P(z/z_n)} \prod_{|z_n|\geqq 2r} E_k(z, z_n).$$

The third product over $|z_n| \geqq 2r$ is seen to have absolute value $> C^{-r^{\rho+\epsilon}}$ by the usual arguments. Indeed, we have the analogous estimate from below instead of Lemma 2.2, namely:

If $|w| \leqq 1/2$, *then*

$$\log|E_k(w)| \geqq -\frac{2}{k}|w|^k.$$

This is seen at once from the expansion

$$-\log E_k(w) = \sum_{m=k}^{\infty} \frac{w^m}{m} = \frac{w^k}{k} + \text{higher terms},$$

and the higher terms can be estimated by the geometric series. We put $w = z/z_n$ to obtain

$$\sum_{|z_n|\geqq 2r} \log|E_k(z, z_n)| \geqq -\frac{2}{k} \sum_{|z_n|\geqq 2r} \left|\frac{z}{z_n}\right|^k \geqq -\frac{2}{k} \sum_{|z_n|\geqq 2r} \left|\frac{z}{z_n}\right|^{\rho+\epsilon}.$$

Since $\sum 1/|z_n|^{\rho+\epsilon}$ converges, we obtain the desired estimate

$$\log|\text{third product}| \geqq -C_1 r^{\rho+\epsilon} \qquad \text{for some constant } C_1 > 0.$$

Consider next the second product. There is a constant $C_2 > 0$ such that

$$-|P(z/z_n)| \geqq -C_2 \sum_{j=0}^{k-1} |z/z_n|^j.$$

If $|z_n| \geqq r$, then the term $|z/z_n|^j$ is bounded and there are

$$\leqq \nu(2r) \ll r^{\rho+\epsilon}$$

such terms. For $|z_n| < r$ we can replace j by its highest possible value $k - 1$. Then we obtain

$$\log|\text{second product}| \geqq -C_2 \sum_{|z_n|<r} \left|\frac{z}{z_n}\right|^{k-1} - O(1).$$

$$\geqq -C_2 r^{k-1} \sum_{|z_n|<r} \left|\frac{1}{z_n}\right|^{k-1} - O(1).$$

Using summation by parts, we find

$$\begin{aligned}\sum_{1 \leqq |z_n| < r} \left|\frac{1}{z_n}\right|^{k-1} &\ll \sum_{m \leqq r} \frac{\nu(m+1) - \nu(m)}{m^{k-1}} \\ &\ll \frac{\nu(r)}{r^{k-1}} + \sum_{m \leqq r} \frac{\nu(m)}{m^k} + \text{constant} \\ &\ll \frac{r^{\rho+\epsilon}}{r^{k-1}} + r^{\rho+\epsilon-k+1} + O(1).\end{aligned}$$

Combining this with the lower bound for log|second product| proves that

$$\log|\text{second product}| \gg -r^{\rho+\epsilon},$$

as desired.

In the first product we have

$$\left|1 - \frac{z}{z_n}\right| = \frac{|z - z_n|}{|z_n|} \geqq \frac{1}{|z_n|^{s+1}} \geqq \frac{1}{(2r)^{s+1}}.$$

Hence the first product satisfies the lower bound

$$\begin{aligned}\sum_{|z_n|<2r} \log\left|1 - \frac{z}{z_n}\right| &\geqq \nu(2r) - (s+1)\log(2r) \\ &\geqq -C_3 r^{\rho+\epsilon}\log(2r) \quad \text{(by Theorem 3.2).}\end{aligned}$$

which concludes the proof of the minimum modulus theorem for canonical products.

Before proving the full minimum modulus theorem, we give the most important application of minimum modulus for Weierstrass products, which now allow us to characterize entire functions of finite order.

Theorem 3.5 (Hadamard). *Let f be an entire function of order ρ, and let $\{z_n\}$ be the sequence of its zeros $\neq 0$. Let k be the smallest integer $> \rho$. Let $P = P_k$. Then*

$$f(z) = e^{h(z)} z^m \prod \left(1 - \frac{z}{z_n}\right) e^{P(z/z_n)},$$

where m is the order of f at 0, and h is a polynomial of degree $\leqq \rho$.

Proof. The series $\sum 1/|z_n|^s$ converges for $s > \rho$. Hence for every r sufficiently large, there exists R with $r \leqq R \leqq 2r$ such that for all n the circle of radius R does not intersect the disc of radius $1/|z_n|^s$ centered at z_n. By the minimum modulus theorem, we get a lower bound for the canonical product $E(z)$ on the circle of radius R, which shows that the quotient $f(z)/E(z)z^m$ is entire of order $\leqq \rho$. We can apply Corollary 3.3 of Chapter XII to conclude the proof.

The full minimum modulus theorem for f is now obvious since the exponential term $e^{h(z)}$ obviously satisfies the desired lower bound.

XIII, §3. EXERCISES

1. Let f, g be entire of order ρ. Show that fg is entire of order $\leqq \rho$, and $f+g$ is entire of order $\leqq \rho$.
2. Let f, g be entire of order $\leqq \rho$, and suppose f/g is entire. Show that f/g is entire of order $\leqq \rho$.

XIII, §4. MEROMORPHIC FUNCTIONS, MITTAG-LEFFLER THEOREM

Suppose f has a pole at z_0, with the power series expansion

$$f(z) = \frac{a_{-m}}{(z-z_0)^m} + \cdots + a_0 + a_1(z-z_0) + \cdots.$$

We call

$$\mathrm{Pr}(f, z_0) = \frac{a_{-m}}{(z-z_0)^m} + \cdots + \frac{a_{-1}}{(z-z_0)} = P\left(\frac{1}{z-z_0}\right)$$

the **principal part** of f at z_0.

We shall consider the additive analogue to Section §2, which is to construct a meromorphic function having given principal parts at a sequence of points $\{z_n\}$ which is merely assumed to be discrete.

Theorem 4.1 (Mittag-Leffler). *Let $\{z_n\}$ be a sequence of distinct complex numbers such that $|z_n| \to \infty$. Let $\{P_n\}$ be polynomials without constant term. Then there exists a meromorphic function f whose only poles are at $\{z_n\}$ with principal part $P_n(1/(z-z_n))$. The most general function of this kind can be written in the form*

$$f(z) = \sum_n \left[P_n\left(\frac{1}{z-z_n}\right) - Q_n(z)\right] + \varphi(z),$$

where Q_n is some polynomial, and φ is entire. The series converges absolutely and uniformly on any compact set not containing the poles.

Proof. Since a principal part at 0 can always be added *a posteriori*, we assume without loss of generality that $z_n \neq 0$ for all n. We expand

$$P_n\left(\frac{1}{z - z_n}\right)$$

in a power series of z/z_n at the origin. This power series is a linear combination of power series arising from a single term

$$\frac{1}{(z - z_n)^k} = \frac{(-1)^k}{z_n^k\left(1 - \dfrac{z}{z_n}\right)^k} = \frac{1}{z_n^k}\sum b_j\left(\frac{z}{z_n}\right)^j$$

with coefficients b_j related to binomial coefficients (depending on k). In particular, since the radius of convergence of $1/(1 - T)^k$ is 1, we have the estimate

$$|b_j| \ll (1 + \epsilon)^j \qquad \text{for } j \to \infty,$$

with the implied constant depending on k. Let d_n be a positive integer. It follows from these estimates that if we let $Q_n(T)$ be the polynomial of degree $\leqq d_n - 1$ in the power series expansion of $P_n(1/(z - z_n))$ at 0, then there exists a constant B_n (depending only on n) such that

$$\left|P_n\left(\frac{1}{z - z_n}\right) - Q_n(z)\right| \leqq B_n \sum_{j=d_n}^{\infty} (1 + \epsilon)^j \left|\frac{z}{z_n}\right|^j.$$

Therefore we can pick d_n such that if $|z/z_n| \leqq 1/2$, then

$$\left|P_n\left(\frac{1}{z - z_n}\right) - Q_n(z)\right| \leqq \frac{1}{2^n}.$$

Then the series

$$\sum\left[P_n\left(\frac{1}{z - z_n}\right) - Q_n(z)\right]$$

converges absolutely and uniformly for z in any compact set not containing the z_n. In fact, given a radius R, let $R \leqq |z_N|$, and split the series,

$$\sum_{n=1}^{N}\left[P_n\left(\frac{1}{z - z_n}\right) - Q_n(z)\right] + \sum_{N+1}^{\infty}\left[P_n\left(\frac{1}{z - z_n}\right) - Q_n(z)\right].$$

The first part is a finite sum. If $|z| \leqq R/2$, then the second sum is dominated by $\sum 1/2^n$. The finite sum has the desired poles, and the infinite sum on the right has no poles in the disc of radius $R/2$. This is true for every R, thereby completing the proof of the theorem.

XIII, §4. EXERCISES

1. Let g be a meromorphic function on $\mathbf{C}$, with poles of order at most one, and integral residues. Show that there exists a meromorphic function f such that $f'/f = g$.

2. Given entire functions f, g without common zeros, prove that there exist entire functions A, B such that $Af + Bg = 1$. [*Hint*: By Mittag-Leffler, there exists a meromorphic function M whose principal parts occur only at the zeros of g, and such that the principal part $\mathrm{Pr}(M, z_n)$ at a zero z_n of g is the same as $\mathrm{Pr}(1/fg, z_n)$, so $M - 1/fg$ is holomorphic at z_n. Let $A = Mg$, and take it from there.]

3. Let f, g be entire functions.
   (a) Show that there exists an entire function h and entire functions f_1, g_1 such that $f = hf_1$, $g = hg_1$, and f_1, g_1 have no zeros in common.
   (b) Show that there exist entire functions A, B such that $Af + Bg = h$.

4. Let $f_1, \ldots, f_m$ be a finite number of entire functions, and let J be the set of all combinations $A_1 f_1 + \cdots + A_m f_m$, where A_i are entire functions. Show that there exists a single entire function f such that J consists of all multiples of f, that is, J consists of all entire functions Af, where A is entire. In the language of rings, this means that every finitely generated ideal in the ring of entire functions is principal.

5. Let $\{a_k\}$, $\{z_k\}$ be sequences of non-zero complex numbers, with $|z_k| \to \infty$ and $|z_k| \leqq |z_{k+1}|$ for all k. Let ρ be a real number >0 such that

$$\sum_{k=1}^{\infty} \frac{|a_k|}{|z_k|^{\rho}} < \infty.$$

   Define

$$A_n = \sum_{k=1}^{n} |a_k|.$$

   (a) Prove that $A_n = o(|z_n|^{\rho})$ for $n \to \infty$, meaning the $\lim A_n/|z_n|^{\rho} = 0$.
   (b) Let d be the smallest integer $\geqq \rho$. Let G_d be the polynomial

$$G_d(z) = \sum_{n=0}^{d-1} z^n.$$

   Define

$$F_k(z, z_k) = \frac{a_k}{z - z_k} + \frac{a_k}{z_k} G_d(z/z_k).$$

Prove that the series

$$F(z) = \sum_{k=1}^{\infty} F(z, z_k)$$

converges absolutely and uniformly on every compact set not containing any z_k.

(c) Let S be a subset of $\mathbf{C}$ at finite non-zero distance from all z_k, that is, there exists $c > 0$ such that $|z - z_k| \geqq c$ for all $z \in S$ and all k. Show that

$$F(z) = O(|z|^d) \qquad \text{for } z \in S, \quad |z| \to \infty.$$

(d) Let U be the complement of the union of all discs $D(z_k, \delta_k)$, centered at z_k, of radius $\delta_k = 1/|z_k|^d$. Show that

$$F(z) = O(|z|^{p+d}) \qquad \text{for } z \in U, \quad |z| \to \infty.$$

Note: For part (d), you will probably need part (a), but for (c), you won't.

CHAPTER XIV

Elliptic Functions

In this chapter we give the classical example of entire and meromorphic functions of order 2. The theory illustrates most of the theorems proved so far in the book. A self-contained "analytic" continuation of the topics discussed in this chapter can be found in Chapters 3, 4, and 18 of my book on *Elliptic Functions* [La 73].

XIV, §1. THE LIOUVILLE THEOREMS

Let ω_1, ω_2 be two complex numbers which are linearly independent over the real numbers. This means that there is no relation

$$a\omega_1 + b\omega_2 = 0 \qquad \text{with} \quad a, b \in \mathbf{R} \quad \text{not both } 0.$$

By the **lattice** generated by ω_1, ω_2 we mean the set of all complex numbers of the form

$$m\omega_1 + n\omega_2 \qquad \text{with} \quad m, n \in \mathbf{Z}.$$

Thus a lattice looks like the set of points of intersections in the lines of the following diagram. We shall use the notation $L = [\omega_1, \omega_2]$. Observe that if ω, $\omega' \in L$, then $\omega + \omega'$, and $n\omega \in L$ for all integers n. One may say that L is closed under addition and under multiplication by integers. Those who know the definition of a group will therefore see that a lattice is a subgroup of $\mathbf{C}$. Furthermore, since $\mathbf{C}$ has dimension 2

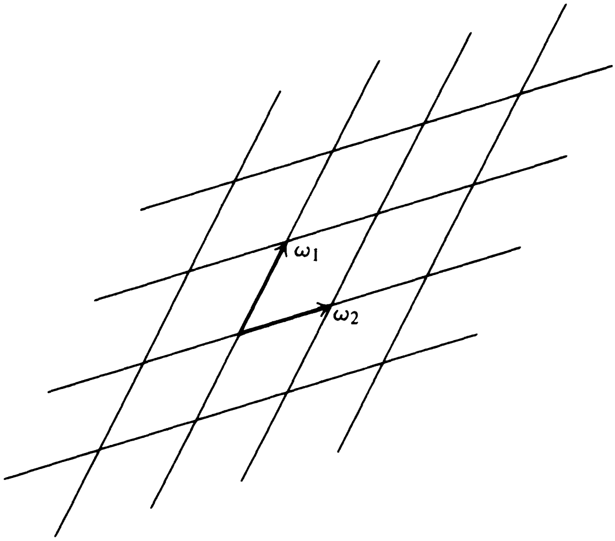


Figure 1

as a vector space over **R**, we see also that ω_1, ω_2 form a basis of **C** over **R**.

Let z, $w \in \mathbf{C}$. We define $z \equiv w \bmod L$ to mean that $z - w \in L$. We let the reader verify that the relation $z \equiv w \bmod L$, called **congruence modulo** L, is an equivalence relation, that is:

$z \equiv z \bmod L$;

if $z \equiv w$ and $w \equiv u$ then $z \equiv u$;

if $z \equiv w$ then $w \equiv z$.

The set of equivalence classes mod L is denoted by $\mathbf{C}/L$, which we read C **modulo** L. Since congruence is an equivalence relation, we may add congruence classes. Also observe that if $z \equiv w \bmod L$ and n is an integer, then $nz \equiv nw \bmod L$. If λ is an arbitrary complex number $\neq 0$, then in general it is of course not true that $\lambda L \subset L$, and even less true that $\lambda L = L$. However, the reader can verify that if $z \equiv w \bmod L$, then λL is also a lattice, and $\lambda z \equiv \lambda w \bmod \lambda L$.

If $L = [\omega_1, \omega_2]$ as above, and $\alpha \in \mathbf{C}$, we call the set consisting of all points

$$\alpha + t_1\omega_1 + t_2\omega_2, \qquad 0 \leqq t_i \leqq 1,$$

a **fundamental parallelogram** for the lattice (with respect to the given basis). We could also take the values $0 \leqq t_i < 1$ to define a fundamental parallelogram. Let P be the fundamental parallelogram defined in this latter fashion. Then one has:

Given $z \in \mathbf{C}$, there exists a unique element $z_0 \in P$ such that $z \equiv z_0 \bmod L$.

Proof. First let $\alpha = 0$. Write $z = a\omega_1 + b\omega_2$ with $a, b \in \mathbf{R}$. Let m be the largest integer $\leqq a$, and let n be the largest integer $\leqq b$. Let $t_1 = a - m$ and $t_2 = b - n$. Then

$$z = m\omega_1 + n\omega_2 + t_1\omega_1 + t_2\omega_2 \equiv t_1\omega_1 + t_2\omega_2 \bmod L.$$

Observe that $0 \leqq t_i < 1$, so the congruence class of $z \bmod L$ has a representative in the fundamental parallelogram P. On the other hand, suppose that

$$t_1\omega_1 + t_2\omega_2 \equiv s_1\omega_1 + s_2\omega_2 \qquad \text{with} \quad 0 \leqq s_i < 1.$$

Then $|s_i - t_i| < 1$, but $(s_1 - t_1)\omega_1 + (s_2 - t_2)\omega_2 \equiv 0 \bmod L$, so $s_i = t_i$ for $i = 1, 2$. Thus $z_0 = t_1\omega_1 + t_2\omega_2$ is the unique element of P which is $\equiv z \bmod L$.

If $\alpha \neq 0$, then we apply the above arguments to $z - \alpha$, to get both the existence and uniqueness of an element z_0 in a fundamental parallelogram which is $\equiv z \bmod L$.

Since we assumed that ω_1, ω_2 are not real scalar multiples of each other (because they are linearly independent over $\mathbf{R}$), it follows that ω_1/ω_2 or ω_2/ω_1 has a positive imaginary part. After changing the order in which we consider ω_1 and ω_2, we can achieve that the imaginary part of ω_1/ω_2 is positive.

Unless otherwise specified, we shall assume that $\mathrm{Im}(\omega_1/\omega_2) > 0$, i.e. that ω_1/ω_2 lies in the upper half plane $H = \{x + iy, y > 0\}$. An **elliptic function** f (with respect to L) is a meromorphic function on $\mathbf{C}$ which is L-periodic, that is,

$$f(z + \omega) = f(z)$$

for all $z \in \mathbf{C}$ and $\omega \in L$. Note that f is periodic if and only if

$$f(z + \omega_1) = f(z) = f(z + \omega_2).$$

An elliptic function which is entire (i.e. without poles) must be constant. Indeed, the function on a fundamental parallelogram is continuous and so bounded, and by periodicity it follows that the function is bounded on all of $\mathbf{C}$, whence constant by Liouville's theorem.

Theorem 1.1. *Let P be a fundamental parallelogram for L, and assume that the elliptic function f has no poles on its boundary ∂P. Then the sum of the residues of f in P is 0.*

Proof. We have

$$2\pi i \sum \operatorname{Res} f = \int_{\partial P} f(z)\, dz = 0,$$

this last equality being valid because of the periodicity, so the integrals on opposite sides cancel each other (Fig. 2).

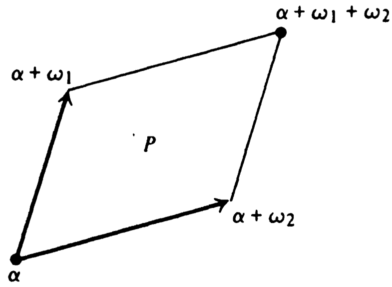


Figure 2

Corollary. *An elliptic function has at least two poles (counting multiplicities) in* $\mathbf{C}/L$.

Theorem 1.2. *Let P be a fundamental parallelogram, and assume that the elliptic function f has no zero or pole on its boundary. Let* $\{a_i\}$ *be the singular points (zeros and poles) of f inside P, and let f have order* m_i *at* a_i. *Then*

$$\sum m_i = 0.$$

Proof. Observe that f elliptic implies that f' and f'/f are elliptic. We then obtain

$$0 = \int_{\partial P} f'/f(z)\, dz = 2\pi\sqrt{-1} \sum \text{Residues} = 2\pi\sqrt{-1} \sum m_i,$$

thus proving our assertion.

Theorem 1.3. *Hypotheses being as in Theorem* 1.2, *we have*

$$\sum m_i a_i \equiv 0 \pmod{L}.$$

Proof. This time, we take the integral

$$\int_{\partial P} z \frac{f'(z)}{f(z)} dz = 2\pi\sqrt{-1} \sum m_i a_i,$$

because

$$\operatorname{res}_{a_i} z\frac{f'(z)}{f(z)} = m_i a_i.$$

On the other hand we compute the integral over the boundary of the parallelogram by taking it for two opposite sides at a time. One pair of such integrals is equal to

$$\int_{\alpha}^{\alpha+\omega_1} z\frac{f'(z)}{f(z)}dz - \int_{\alpha+\omega_2}^{\alpha+\omega_1+\omega_2} z\frac{f'(z)}{f(z)}dz.$$

We change variables in the second integral, letting $u = z - \omega_2$. Both integrals are then taken from $\alpha + \omega_1$, and after a cancellation, we get the value

$$-\omega_2 \int_{\alpha}^{\alpha+\omega_1} \frac{f'(u)}{f(u)}du = 2\pi\sqrt{-1}k\omega_2,$$

for some integer k. The integral over the opposite pair of sides is done in the same way, and our theorem is proved.

XIV, §2. THE WEIERSTRASS FUNCTION

We now prove the existence of elliptic functions by writing some analytic expression, namely the Weierstrass function

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in L^*}\left[\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right],$$

where the sum is taken over the set of all non-zero periods, denoted by L^*. We have to show that this series converges uniformly on compact sets not including the lattice points. For bounded z, staying away from the lattice points, the expression in the brackets has the order of magnitude of $1/|\omega|^3$. Hence it suffices to prove:

Lemma 2.1. *If* $\lambda > 2$, *then* $\displaystyle\sum_{\omega \in L^*} \frac{1}{|\omega|^\lambda}$ *converges.*

Proof. Let A_n be the annulus consisting of all $z \in \mathbf{C}$ such that

$$n - 1 \leqq |z| < n.$$

The partial sum for $|\omega| < N$ can be decomposed into a sum for $\omega \in A_n$, and then a sum for $1 \leqq n \leqq N$. We claim:

There exists a number $C > 0$ such that the number of lattice points in A_n is $\leqq Cn$.

Assuming this claim for a moment, we obtain

$$\sum_{|\omega| \leqq N} \frac{1}{|\omega|^\lambda} \ll \sum_1^\infty \frac{n}{n^\lambda} \ll \sum_1^\infty \frac{1}{n^{\lambda-1}}$$

which converges for $\lambda > 2$.

Now to prove the claim, let d be a positive integer such that the diameter of a fundamental parallelogram is $\leqq d$. Then A_n is contained in the annulus of complex numbers z such that

$$n - 1 - d \leqq |z| \leqq n + d.$$

By direct computation, the area of this annulus is $C_1 n$ for some constant C_1 (depending on d). Let k_n be the number of fundamental parallelograms of the lattice which intersect A_n. The number of lattice points in A_n is certainly bounded by k_n. But if P is a fundamental parallelogram, then

$$k_n(\text{area of } P) \leqq \text{area of the annulus}\{n - 1 - d \leqq |z| \leqq n + d\} \leqq C_1 n.$$

This proves that $k_n \leqq Cn$, with $C = C_1/(\text{area of } P)$, and thus proves the claim.

The series expression for $\wp$ shows that it is meromorphic, with a double pole at each lattice point, and no other pole. It is clear that $\wp$ is even, that is,

$$\wp(z) = \wp(-z)$$

(summing over the lattice points is the same as summing over their negatives). We get $\wp'$ by differentiating term by term,

$$\wp'(z) = -2 \sum_{\omega \in L} \frac{1}{(z - \omega)^3},$$

the sum being taken for all $\omega \in L$. Note that $\wp'$ is clearly periodic, and is odd, that is

$$\wp'(-z) = -\wp'(z).$$

From its periodicity, we conclude that there is a constant C such that

$$\wp(z + \omega_1) = \wp(z) + C.$$

Let $z = -\omega_1/2$ (not a pole of $\wp$). We get

$$\wp\left(\frac{\omega_1}{2}\right) = \wp\left(-\frac{\omega_1}{2}\right) + C,$$

and since $\wp$ is even, it follows that $C = 0$. Hence $\wp$ is itself periodic, something which we could not see immediately from its series expansion.

Theorem 2.2. *Let f be an elliptic function periodic with respect to L. Then f can be expressed as a rational function of $\wp$ and $\wp'$.*

Proof. If f is elliptic, we can write f as a sum of an even and an odd elliptic function, namely

$$f(z) = \frac{f(z) + f(-z)}{2} + \frac{f(z) - f(-z)}{2}.$$

If f is odd, then the product $f\wp'$ is even, so it will suffice to prove that if f is even, then f is a rational function of $\wp$.

Suppose that f is even and has a zero of order m at some point u. Then clearly f also has a zero of the same order at $-u$ because

$$f^{(k)}(u) = (-1)^k f^{(k)}(-u).$$

Similarly for poles.

If $u \equiv -u$ (mod L), then the above assertion holds in the strong sense, namely f has a zero (or pole) of even order at u.

Proof. First note that $u \equiv -u$ (mod L) is equivalent to

$$2u \equiv 0 \pmod{L}.$$

In $\mathbf{C}/L$ there are exactly four points with this property, represented by

$$0, \frac{\omega_1}{2}, \frac{\omega_2}{2}, \frac{\omega_1 + \omega_2}{2}$$

in a fundamental parallelogram. If f is even, then f' is odd, that is,

$$f'(u) = -f'(-u).$$

Since $u \equiv -u$ (mod L) and f' is periodic, it follows that $f'(u) = 0$, so that f has a zero of order at least 2 at u. If $u \not\equiv 0$ (mod L), then the above argument shows that the function

$$g(z) = \wp(z) - \wp(u)$$

has a zero of order at least 2 (hence exactly 2 by Theorem 1.2 and the fact that $\wp$ has only one pole of order 2 on $\mathbf{C} \bmod L$). Then f/g is even, elliptic, holomorphic at u. If $f(u)/g(u) \neq 0$, then $\operatorname{ord}_u f = 2$. If $f(u)/g(u) = 0$, then f/g again has a zero of order at least 2 at u and we can repeat the argument. If $u \equiv 0 \pmod{L}$ we use $g = 1/\wp$ and argue similarly, thus proving that f has a zero of even order at u.

Now let u_i $(i = 1, \dots, r)$ be a family of points containing one representative from each class $(u, -u) \pmod{L}$ where f has a zero or pole, other than the class of L itself. Let

$$\begin{aligned} m_i &= \operatorname{ord}_{u_i} f \qquad \text{if} \quad 2u_i \not\equiv 0 \pmod{L}, \\ m_i &= \tfrac{1}{2}\operatorname{ord}_{u_i} f \qquad \text{if} \quad 2u_i \equiv 0 \pmod{L}. \end{aligned}$$

Our previous remarks show that for $a \in \mathbf{C}$, $a \not\equiv 0 \pmod{L}$, the function $\wp(z) - \wp(a)$ has a zero of order 2 at a if and only if $2a \equiv 0 \pmod{L}$, and has distinct zeros of order 1 at a and $-a$ otherwise. Hence for all $z \not\equiv 0 \pmod{L}$ the function

$$\prod_{i=1}^{r} [\wp(z) - \wp(u_i)]^{m_i}$$

has the same order at z as f. This is also true at the origin because of Theorem 1.2 applied to f and the above product. The quotient of the above product by f is then an elliptic function without zero or pole, hence a constant, thereby proving Theorem 2.2.

Next, we obtain the power series development of $\wp$ and $\wp'$ at the origin, from which we shall get the algebraic relation holding between these two functions. We do this by brute force.

$$\begin{aligned} \wp(z) &= \frac{1}{z^2} + \sum_{\omega \in L^*} \left[\frac{1}{\omega^2}\left(1 + \frac{z}{\omega} + \left(\frac{z}{\omega}\right)^2 + \cdots\right)^2 - \frac{1}{\omega^2} \right] \\ &= \frac{1}{z^2} + \sum_{\omega \in L^*} \sum_{m=1}^{\infty} (m+1)\left(\frac{z}{\omega}\right)^m \frac{1}{\omega^2} \\ &= \frac{1}{z^2} + \sum_{m=1}^{\infty} c_m z^m, \end{aligned}$$

where

$$c_m = \sum_{\omega \neq 0} \frac{m+1}{\omega^{m+2}}.$$

Note that $c_m = 0$ if m is odd.

Using the notation

$$s_m(L) = s_m = \sum_{\omega \neq 0} \frac{1}{\omega^m}$$

we get the expansion

$$\boxed{\wp(z) = \frac{1}{z^2} + \sum_{n=1}^{\infty} (2n+1)s_{2n+2}(L)z^{2n},}$$

from which we write down the first few terms explicitly:

$$\boxed{\wp(z) = \frac{1}{z^2} + 3s_4z^2 + 5s_6z^4 + \cdots}$$

and differentiating term by term,

$$\boxed{\wp'(z) = \frac{-2}{z^3} + 6s_4z + 20s_6z^3 + \cdots.}$$

Theorem 2.3. *Let $g_2 = g_2(L) = 60s_4$ and $g_3 = g_3(L) = 140s_6$. Then*

$$\wp'^2 = 4\wp^3 - g_2\wp - g_3.$$

Proof. We expand out the function

$$\varphi(z) = \wp'(z)^2 - 4\wp(z)^3 + g_2\wp(z) + g_3$$

at the origin, paying attention only to the polar term and the constant term. This is easily done, and one sees that there is enough cancellation so that these terms are 0, in other words, $\varphi(z)$ is an elliptic function without poles, and with a zero at the origin. Hence φ is identically zero, thereby proving our theorem.

The preceding theorem shows that the points $(\wp(z), \wp'(z))$ lie on the curve defined by the equation

$$y^2 = 4x^3 - g_2x - g_3.$$

The cubic polynomial on the right-hand side has a discriminant given by

$$\Delta = g_2^3 - 27g_3^2.$$

We shall see in a moment that this discriminant does not vanish.

Let

$$e_i = \wp\left(\frac{\omega_i}{2}\right), \qquad i = 1, 2, 3,$$

where $L = [\omega_1, \omega_2]$ and $\omega_3 = \omega_1 + \omega_2$. Then the function

$$h(z) = \wp(z) - e_i$$

has a zero at $\omega_i/2$, which is of even order so that $\wp'(\omega_i/2) = 0$ for $i = 1, 2, 3$, by previous remarks. Comparing zeros and poles, we conclude that

$$\wp'(z)^2 = 4\big(\wp(z) - e_1\big)\big(\wp(z) - e_2\big)\big(\wp(z) - e_3\big).$$

Thus e_1, e_2, e_3 are the roots of $4x^3 - g_2 x - g_3$. Furthermore, $\wp$ takes on the value e_i with multiplicity 2 and has only one pole of order 2 mod L, so that $e_i \neq e_j$ for $i \neq j$. This means that the three roots of the cubic polynomial are distinct, and therefore

$$\Delta = g_2^3 - 27g_3^2 \neq 0.$$

XIV, §3. THE ADDITION THEOREM

Given complex numbers g_2, g_3 such that $g_2^3 - 27g_3^2 \neq 0$, one can ask whether there exists a lattice for which these are the invariants associated to the lattice as in the preceding section. The answer is yes. For the moment, we consider the case when g_2, g_3 are given as in the preceding section, that is, $g_2 = 60s_4$ and $g_3 = 140s_6$.

We have seen that the map

$$z \mapsto \big(\wp(z), \wp'(z)\big)$$

parametrizes points on the cubic curve defined by the equation

$$y^2 = 4x^3 - g_2 x - g_3.$$

If $z \notin L$ then the image of z under this map is a point of the curve, and if $z \in L$, then we can define its image to be a "point at infinity". If $z_1 \equiv z_2 \pmod{L}$ then z_1 and z_2 have the same image under this map.

Let

$$P_1 = \big(\wp(u_1), \wp'(u_1)\big) \qquad \text{and} \qquad P_2 = \big(\wp(u_2), \wp'(u_2)\big)$$

be two points on the curve. Let $u_3 = u_1 + u_2$. Let

$$P_3 = \big(\wp(u_3), \wp'(u_3)\big).$$

We also write

$$P_1 = (x_1, y_1), \qquad P_2 = (x_2, y_2), \qquad P_3 = (x_3, y_3).$$

Then we shall express x_3, y_3 as rational functions of (x_1, y_1) and (x_2, y_2). We shall see that P_3 is obtained by taking the line through P_1, P_2, intersecting it with the curve, and reflecting the point of intersection through the x-axis, as shown on Fig. 3.

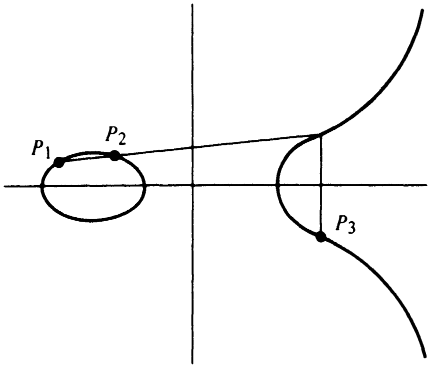


Figure 3

Select u_1, $u_2 \in \mathbf{C}$ and $\notin L$, and assume $u_1 \not\equiv u_2 \pmod{L}$. Let a, b be complex numbers such that

$$\wp'(u_1) = a\wp(u_1) + b,$$
$$\wp'(u_2) = a\wp(u_2) + b,$$

in other words, $y = ax + b$ is the line through $(\wp(u_1), \wp'(u_1))$ and $(\wp(u_2), \wp'(u_2))$. Then

$$\wp'(z) - (a\wp(z) + b)$$

has a pole of order 3 at 0, whence it has three zeros, counting multiplicities, and two of these are at u_1 and u_2. If, say, u_1 had multiplicity 2, then by Theorem 1.3 we would have

$$2u_1 + u_2 \equiv 0 \pmod{L}.$$

If we fix u_1, this can hold for only one value of $u_2 \bmod L$. Let us assume that we do not deal with this value. Then both u_1, u_2 have multiplicity 1, and the third zero lies at

$$u_3 \equiv -(u_1 + u_2) \pmod{L}$$

again by Theorem 1.3. So we also get

$$\wp'(u_3) = a\wp(u_3) + b.$$

The equation

$$4x^3 - g_2x - g_3 - (ax + b)^2 = 0$$

has three roots, counting multiplicities. They are $\wp(u_1)$, $\wp(u_2)$, $\wp(u_3)$, and the left-hand side factors as

$$4(x - \wp(u_1))(x - \wp(u_2))(x - \wp(u_3)).$$

Comparing the coefficient of x^2 yields

$$\wp(u_1) + \wp(u_2) + \wp(u_3) = \frac{a^2}{4}.$$

But from our original equations for a and b, we have

$$a(\wp(u_1) - \wp(u_2)) = \wp'(u_1) - \wp'(u_2).$$

Therefore from

$$\wp(u_3) = \wp(-(u_1 + u_2)) = \wp(u_1 + u_2)$$

we get

$$\boxed{\wp(u_1 + u_2) = -\wp(u_1) - \wp(u_2) + \frac{1}{4}\left(\frac{\wp'(u_1) - \wp'(u_2)}{\wp(u_1) - \wp(u_2)}\right)^2}$$

or in algebraic terms,

$$x_3 = -x_1 - x_2 + \frac{1}{4}\left(\frac{y_1 - y_2}{x_1 - x_2}\right)^2.$$

Fixing u_1, the above formula is true for all but a finite number of $u_2 \equiv u_1 \pmod{L}$, whence for all $u_2 \not\equiv u_1 \pmod{L}$ by analytic continuation.

For $u_1 \equiv u_2 \pmod{L}$ we take the limit as $u_1 \to u_2$ and get

$$\boxed{\wp(2u) = -2\wp(u) + \frac{1}{4}\left(\frac{\wp''(u)}{\wp'(u)}\right)^2.}$$

These give us the desired algebraic addition formulas. Note that the formulas involve only g_2, g_3 as coefficients in the rational functions.

XIV, §4. THE SIGMA AND ZETA FUNCTIONS

Both in number theory and analysis one factorizes elements into prime powers. In analysis, this means that a function gets factored into an infinite product corresponding to its zeros and poles.

In this chapter, we are concerned with the analytic expressions.

Our first task is to give a universal gadget allowing us to factorize an elliptic function, with a numerator and denominator which are entire functions, and are as periodic as possible.

One defines a **theta function** (on **C**) with respect to a lattice L, to be an entire function θ satisfying the condition

$$\theta(z+u) = \theta(z)e^{2\pi i[l(z,u)+c(u)]}, \qquad z \in \mathbf{C}, \quad u \in L,$$

where l is **C**-linear in z, **R**-linear in u, and $c(u)$ is some function depending only on u. We shall construct a theta function.

We write down the **Weierstrass sigma function**, which has zeros of order 1 at all lattice points, by the Weierstrass product

$$\sigma(z) = z \prod_{\omega \in L^*} \left(1 - \frac{z}{\omega}\right) e^{z/\omega + 1/2(z/\omega)^2}.$$

Here L^* means the lattice from which 0 is deleted, i.e. we are taking the product over the non-zero periods. We note that σ also depends on L, and so we write $\sigma(z, L)$, which is homogeneous of degree 1, namely

$$\boxed{\sigma(\lambda z, \lambda L) = \lambda\sigma(z, L),} \qquad \lambda \in \mathbf{C}, \quad \lambda \neq 0.$$

Taking the logarithmic derivative formally yields the **Weierstrass zeta function**

$$\zeta(z, L) = \zeta(z) = \frac{\sigma'(z)}{\sigma(z)} = \frac{1}{z} + \sum_{\omega \in L^*} \left[\frac{1}{z-\omega} + \frac{1}{\omega} + \frac{z}{\omega^2}\right].$$

It is clear that the sum on the right converges absolutely and uniformly for z in a compact set not containing any lattice point, and hence integrating and exponentiating shows that the infinite product for $\sigma(z)$ also converges absolutely and uniformly in such a region. Differentiating $\zeta(z)$ term by term shows that

$$\zeta'(z) = -\wp(z) = -\frac{1}{z^2} - \sum_{\omega \in L^*} \left[\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right].$$

Also from the product and sum expressions, we see at once that *both*

σ and ζ are odd functions, that is

$$\sigma(-z) = -\sigma(z) \qquad \text{and} \qquad \zeta(-z) = -\zeta(z).$$

The series defining $\zeta(z, L)$ shows that it is homogeneous of degree -1, that is,

$$\boxed{\zeta(\lambda z, \lambda L) = \frac{1}{\lambda}\zeta(z, L).}$$

Differentiating the function $\zeta(z + \omega) - \zeta(z)$ for any $\omega \in L$ yields 0 because the $\wp$-function is periodic. Hence there is a constant $\eta(\omega)$ (sometimes written η_ω) such that

$$\zeta(z + \omega) = \zeta(z) + \eta(\omega).$$

It is clear that $\eta(\omega)$ is **Z**-linear in ω. If $L = [\omega_1, \omega_2]$, then one uses the notation

$$\eta(\omega_1) = \eta_1 \qquad \text{and} \qquad \eta(\omega_2) = \eta_2.$$

As with ζ, the form $\eta(\omega)$ satisfies the homogeneity relation of degree -1, as one verifies directly from the similar relation for ζ. Observe that the lattice should strictly be in the notation, so that in full, the relations should read:

$$\boxed{\begin{aligned} &\zeta(z + \omega, L) = \zeta(z, L) + \eta(\omega, L), \\ &\eta(\lambda\omega, \lambda L) = \frac{1}{\lambda}\eta(\omega, L). \end{aligned}}$$

Theorem 4.1. *The function σ is a theta function, and in fact*

$$\frac{\sigma(z + \omega)}{\sigma(z)} = \psi(\omega)e^{\eta(\omega)(z+\omega/2)},$$

where

$$\begin{aligned} \psi(\omega) &= 1 \qquad \textit{if} \quad \omega/2 \in L, \\ \psi(\omega) &= -1 \qquad \textit{if} \quad \omega/2 \notin L. \end{aligned}$$

Proof. We have

$$\frac{d}{dz}\log\frac{\sigma(z + \omega)}{\sigma(z)} = \eta(\omega).$$

Hence

$$\log \frac{\sigma(z+\omega)}{\sigma(z)} = \eta(\omega)z + c(\omega),$$

whence exponentiating yields

$$\sigma(z+\omega) = \sigma(z)e^{\eta(\omega)z+c(\omega)},$$

which shows that σ is a theta function. We write the quotient as in the statement of the theorem, thereby defining $\psi(\omega)$, and it is then easy to determine $\psi(\omega)$ as follows.

Suppose that $\omega/2$ is not a period. Set $z = -\omega/2$ in the above relation. We see at once that $\psi(\omega) = -1$ because σ is odd. On the other hand, consider

$$\frac{\sigma(z+2\omega)}{\sigma(z)} = \frac{\sigma(z+2\omega)}{\sigma(z+\omega)}\frac{\sigma(z+\omega)}{\sigma(z)}.$$

Using the functional equation twice and comparing the two sides, we see that $\psi(2\omega) = \psi(\omega)^2$. In particular, if $\omega/2 \in L$, then

$$\psi(\omega) = \psi(\omega/2)^2.$$

Dividing by 2 until we get some element of the lattice which is not equal to twice a period, we conclude at once that $\psi(\omega) = (-1)^{2n} = 1$.

The numbers η_1 and η_2 are called **basic quasi periods of** ζ.

Legendre Relation. *We have*

$$\eta_2\omega_1 - \eta_1\omega_2 = 2\pi i.$$

Proof. We integrate around a fundamental parallelogram P, just as we did for the $\wp$-function:

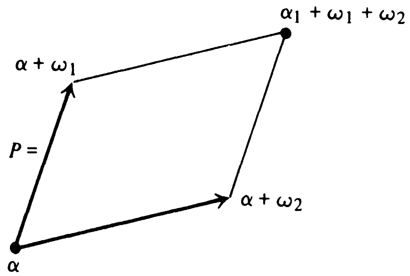


Figure 4

The integral is equal to

$$\int_{\partial P} \zeta(z)\, dz = 2\pi i \sum \text{residues of } \zeta$$
$$= 2\pi i$$

because ζ has residue 1 at 0 and no other pole in a fundamental parallelogram containing 0. On the other hand, using the quasi periodicity, the integrals over opposite sides combine to give

$$\eta_2 \omega_1 - \eta_1 \omega_2,$$

as desired.

Next, we show how the sigma function can be used to factorize elliptic functions. We know that the sum of the zeros and poles of an elliptic function must be congruent to zero modulo the lattice. Selecting suitable representatives of these zeros and poles, we can always make the sum equal to 0.

For any $a \in \mathbf{C}$ we have

$$\frac{\sigma(z + a + \omega)}{\sigma(z + a)} = \psi(\omega) e^{\eta(\omega)(z + \omega/2)} e^{\eta(\omega) a}.$$

Observe how the term $\eta(\omega)a$ occurs linearly in the exponent. It follows that if $\{a_i\}$, $\{b_i\}$ $(i = 1, \ldots, n)$ are families of complex numbers such that

$$\sum a_i = \sum b_i,$$

then the function

$$\frac{\prod \sigma(z - a_i)}{\prod \sigma(z - b_i)}$$

is periodic with respect to our lattice, and is therefore an elliptic function. Conversely, any elliptic function can be so factored into a numerator and denominator involving the sigma function. We write down explicitly the special case with the $\wp$-function.

Theorem 4.2. *For any $a \in \mathbf{C}$ not in L, we have*

$$\wp(z) - \wp(a) = -\frac{\sigma(z + a)\sigma(z - a)}{\sigma^2(z)\sigma^2(a)}.$$

Proof. The function $\wp(z) - \wp(a)$ has zeros at a and $-a$, and has a double pole at 0. Hence

$$\wp(z) - \wp(a) = C\frac{\sigma(z+a)\sigma(z-a)}{\sigma^2(z)}$$

for some constant C. Multiply by z^2 and let $z \to 0$. Then $\sigma^2(z)/z^2$ tends to 1 and $z^2\wp(z)$ tends to 1. Hence we get the value $C = -1/\sigma^2(a)$, thus proving our theorem.

CHAPTER XV

The Gamma and Zeta Functions

We now come to a situation where the natural way to define a function is not through a power series but through an integral depending on a parameter. We shall give a natural condition when we can differentiate under the integral sign, and we can then use Goursat's theorem to conclude that the holomorphic function so defined is analytic.

We shall be integrating over intervals. For concreteness let us assume that we integrate on $[0, \infty[$. A function f on this interval is said to be **absolutely integrable** if

$$\int_0^\infty |f(t)|\, dt$$

exists. If the function is continuous, the integral is of course defined as the limit

$$\lim_{B\to\infty} \int_0^B |f(t)|\, dt.$$

We shall also deal with integrals depending on a parameter. This means f is a function of two variables, $f(t, z)$, where z lies in some domain U in the complex numbers. The integral

$$\int_0^\infty f(t, z)\, dt = \lim_{B\to\infty} \int_0^B f(t, z)\, dt$$

is said to be **uniformly convergent** for $z \in U$ if, given ϵ, there exists B_0 such that if $B_0 < B_1 < B_2$, then

$$\left| \int_{B_1}^{B_2} f(t, z)\, dt \right| < \epsilon.$$

The integral is **absolutely and uniformly convergent** for $z \in U$ if this same condition holds with $f(t, z)$ replaced by the absolute value $|f(t, z)|$.

XV, §1. THE DIFFERENTIATION LEMMA

Lemma 1.1. *Let I be an interval of real numbers, possibly infinite. Let U be an open set of complex numbers. Let $f = f(t, z)$ be a continuous function on $I \times U$. Assume:*

(i) *For each compact subset K of U the integral*

$$\int_I f(t, z)\, dt$$

is uniformly convergent for $z \in K$.

(ii) *For each t the function $z \mapsto f(t, z)$ is analytic. Let*

$$F(z) = \int_I f(t, z)\, dt.$$

Then F is analytic on U, $D_2 f(t, z)$ satisfies the same hypotheses as f, and

$$F'(z) = \int_I D_2 f(t, z)\, dt.$$

Proof. Let $\{I_n\}$ be a sequence of finite closed intervals, increasing to I. Let D be a disc in the z-plane whose closure is contained in U. Let γ be the circle bounding D. Then for each z in D we have

$$f(t, z) = \frac{1}{2\pi i} \int_\gamma \frac{f(t, \zeta)}{\zeta - z} d\zeta,$$

so

$$F(z) = \frac{1}{2\pi i} \int_I \int_\gamma \frac{f(t, \zeta)}{\zeta - z} d\zeta\, dt.$$

If γ has radius R, center z_0, consider only z such that $|z - z_0| \leqq R/2$. Then

$$\left| \frac{1}{\zeta - z} \right| \leqq 2/R.$$

For each n we can define

$$F_n(z) = \frac{1}{2\pi i} \int_{I_n} \int_\gamma \frac{f(t, \zeta)}{\zeta - z} d\zeta\, dt.$$

In view of the restriction on z above, we may interchange the integrals and get

$$F_n(z) = \frac{1}{2\pi i}\int_\gamma \frac{1}{\zeta - z}\left[\int_{I_n} f(t, \zeta)\, dt\right] d\zeta.$$

Then F_n is analytic by Theorem 7.7 of Chapter III. By hypothesis, the integrals over I_n converge uniformly to the integral over I. Hence F is analytic, being the uniform limit of the functions F_n for $|z - z_0| \leqq R/2$. On the other hand, $F_n'(z)$ is obtained by differentiating under the integral sign in the usual way, and converges uniformly to $F'(z)$. However

$$F_n'(z) = \frac{1}{2\pi i}\int_{I_n} D_2 f(t, z)\, dt.$$

This proves the theorem.

Observe that the hypotheses under which the theorem is proved are slightly weaker than in the real case, because of the peculiar nature of complex differentiable functions, whose derivative can be expressed as an integral. For the differentiation lemma see the Appendix of Chapter VIII.

Example. Let f be a continuous function with compact support on the real numbers. (**Compact support** means that the function is equal to 0 outside a compact set.) Consider the integral

$$F(z) = \int_{-\infty}^{\infty} f(t)e^{itz}\, dt.$$

Let $y_0 < y_1$ be real numbers. The integrand for

$$y_0 < \operatorname{Im} z < y_1$$

is of the form

$$f(t)e^{itz} = f(t)e^{itx}e^{-ty}$$

and we have $|e^{itx}| = 1$, whereas e^{-ty} lies between e^{-ty_1} and e^{-ty_0}. Since f has compact support, the values of t for which $f(t) \neq 0$ are bounded. Hence e^{itz} is bounded uniformly for $y_0 < \operatorname{Im} z < y_1$. Thus the integral converges absolutely and uniformly for z satisfying these inequalities.

Differentiating under the integral sign yields the integrand

$$itf(t)e^{itz},$$

and again the function $itf(t)$ has compact support. Thus the same argument can be applied. In fact we have the uniform bound

$$|itf(t)e^{itz}| \leqq |t||f(t)|e^{ty_1}$$

which is independent of z in the given regions. Therefore we obtain

$$F'(z) = \int_{-\infty}^{\infty} itf(t)e^{itz}\,dt \qquad \text{for} \quad y_0 < \operatorname{Im} z < y_1.$$

As this is true for every choice of y_0, y_1 we conclude that in fact F is an entire function.

The similar result for continuity rather than differentiability from Theorem A2 of the Appendix of Chapter VIII can be used. We give here another version which will suffice in what follows.

Lemma 1.2 (Continuity Lemma). *Let I be an interval, U an open set in the complex numbers, and $f(t, z)$ a continuous function on $I \times U$. Assume that there exists a function φ on I which is absolutely integrable on I, and such that for all $z \in U$ we have*

$$|f(t, z)| \leqq \varphi(t).$$

Then the function F defined by

$$F(z) = \int_I f(t, z)\,dt$$

is continuous.

XV, §1. EXERCISES

1. For $\operatorname{Re}(z) > 0$, prove that

$$\log z = \int_0^{\infty} (e^{-t} - e^{-zt})\frac{dt}{t}.$$

[*Hint*: Show that the derivatives of both sides are equal.]

2. Let f be analytic on the closed unit disc. Let

$$I(z) = \int_0^1 \frac{f(t)}{t+z}\,dt.$$

Show that $I(z) + f(-z) \log z$ is analytic for z in some neighborhood of 0. [*Hint*: First consider z real positive, or if you wish, z with positive real part. Use the power series expansion $f(t) = \sum c_k t^k$, and write $t = t + z - z$. Collect terms. The part

$$\sum_{k=0}^{\infty} c_k(-z)^k \int_0^1 \frac{dt}{t+z}$$

will give rise to the log term.]

The Laplace Transform

3. Let f be a continuous function with compact support on the interval $[0, \infty[$. Show that the function Lf given by

$$Lf(z) = \int_0^{\infty} f(t)e^{-zt}\,dt$$

is entire.

4. Let f be a continuous function on $[0, \infty[$, and assume that there is constant $C > 1$ such that

$$|f(t)| \ll C^t \qquad \text{for} \quad t \to \infty,$$

i.e. there exist constants A, B such that $|f(t)| \leqq Ae^{Bt}$ for all t sufficiently large.

(a) Prove that the function

$$Lf(z) = \int_0^{\infty} f(t)e^{-zt}\,dt$$

is analytic in some half plane $\operatorname{Re} z \geqq \sigma$ for some real number σ. In fact, the integral converges absolutely for some σ. Either such σ have no lower bound, in which case Lf is entire, or the greatest lower bound σ_0 is called the **abscissa of convergence of the integral**, and the function Lf is analytic for $\operatorname{Re}(z) > \sigma_0$. The integral converges absolutely for

$$\operatorname{Re} z \geqq \sigma_0 + \epsilon,$$

for every $\epsilon > 0$.

The function Lf is called the **Laplace transform** of f.

(b) Assuming that f is of class C^1, prove by integrating by parts that

$$L(f')(z) = zLf(z) - f(0).$$

Find the Laplace transform of the following functions, and the abscissa of convergence of the integral defining the transform. In each case, a is a real number $\neq 0$.

5. $f(t) = e^{-at}$
6. $f(t) = \cos at$
7. $f(t) = \sin at$
8. $f(t) = (e^t + e^{-t})/2$

9. Suppose that f is periodic with period $a > 0$, that is $f(t + a) = f(t)$ for all $t \geqq 0$. Show that

$$Lf(z) = \frac{\int_0^a e^{-zt}f(t)\,dt}{1 - e^{-az}} \quad \text{for} \quad \operatorname{Re} z > 0.$$

XV, §2. THE GAMMA FUNCTION

There are two natural approaches to the gamma function. One is via the Weierstrass product, and the other is via a Mellin transform integral. Certain properties are clear from one definition but not from the other, and there is also the problem of proving that the two definitions give the same function. There are many variations for achieving all this. I select one of them, starting with the Weierstrass product and its consequences, which constitute the more algebraic properties of the gamma function.

Weierstrass Product

Let γ be the **Euler constant**, that is

$$\gamma = \lim_{n \to \infty} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n} - \log n\right).$$

By the general theory of Weierstrass products, there is an entire function $g(z)$ whose zeros are the negative integers and 0, having the Weierstrass product

$$g(z) = e^{\gamma z} z \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right) e^{-z/n}.$$

We define the **gamma function** to be $\Gamma(z) = 1/g(z)$, so that

Γ 1.
$$\boxed{1/\Gamma(z) = z e^{\gamma z} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right) e^{-z/n},}$$

Thus the gamma function has poles of order 1 at the negative integers. We record at once its logarithmic derivative

Γ 2.
$$-\Gamma'/\Gamma(z) = g'/g(z) = \frac{1}{z} + \gamma + \sum_{n=1}^{\infty} \left(\frac{1}{z + n} - \frac{1}{n}\right).$$

The Euler constant is the unique constant such that the Weierstrass *product for g satisfies the property*

$$g(z+1) = \frac{1}{z} g(z),$$

and therefore the gamma function satisfies the property

Γ 3.
$$\boxed{\Gamma(z+1) = z\Gamma(z).}$$

To prove this, let $g_1(z) = z^{-1}g(z)$, so that $g_1(0) = 1$. Taking logarithmic derivatives, we find immediately from **Γ 2** that

$$\begin{aligned} g'/g(z+1) &= \frac{1}{z+1} + \gamma + \sum_{n=1}^{\infty} \left(\frac{1}{z+n+1} - \frac{1}{n} \right) \\ &= g_1'/g_1(z). \end{aligned}$$

Indeed, the term with $1/z$ disappears in the Mittag–Leffler expansion of g_1'/g_1, and looking at the partial sums of this expansion, we find that they differ from the partial sums of $g'/g(z+1)$ by a term $1/(z+n+1)$ which tends to 0 as n tends to ∞. Therefore there is a constant C such that

$$g(z+1) = Cg_1(z).$$

Evaluating at $z = 0$ yields $g(1) = Cg_1(0) = C$, so

$$C = e^{\gamma} \prod_{n=1}^{\infty} \left(1 + \frac{1}{n}\right) e^{-1/n}.$$

Hence

$$\begin{aligned} \log C &= \gamma + \sum_{n=1}^{\infty} \left[\log\left(1 + \frac{1}{n}\right) - \frac{1}{n} \right] \\ &= \gamma + \sum_{n=1}^{\infty} \left[\log(n+1) - \log n - \frac{1}{n} \right] \\ &= 0, \end{aligned}$$

as one sees by looking at the partial sums and using the definition

$$\gamma = \lim_{n \to \infty} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n} - \log n \right).$$

Hence $C = 1$ and we are done. In addition we have also proved that for positive integers n we have

Γ 4. $\Gamma(n) = (n-1)!$ and $\Gamma(1) = 1.$

Next comes the identity

Γ 5.
$$\boxed{\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}.}$$

To prove this, we note that from the Weierstrass product and Example 2.4 of Chapter X,

$$\Gamma(z)\Gamma(-z) = -\frac{1}{z^2}\prod_{n=1}^{\infty}\left(1 - \frac{z^2}{n^2}\right)^{-1} = -\frac{\pi}{z \sin \pi z}.$$

Using **Γ 3** the formula **Γ 5** drops out.

Putting $z = 1/2$ and noting that $\Gamma(x)$ is positive when x is positive, we get $\Gamma(1/2)^2 = \pi$, and so

$$\Gamma(\tfrac{1}{2}) = \sqrt{\pi}.$$

From **Γ 5** we shall also obtain the residue of Γ at negative integers, namely:

Γ 6. *The residue of* Γ *at* $z = -n$ *for* $n = 0, 1, 2, \ldots$ *is* $(-1)^n/n!$.

To see this, we multiply both sides of **Γ** 5 by $(z+n)$ and use the addition formula for the sine, to find

$$(z+n)\Gamma(z) = \frac{\pi(z+n)}{\sin \pi(z+n)} \frac{(-1)^n}{\Gamma(1-z)}.$$

We can evaluate the right side at $z = -n$ to be $(-1)^n/n!$, which therefore gives the residue from the left side, as desired. Several other proofs will be given subsequently.

Next we compute an alternate version of the Weierstrass product. Define

$$g_n(z) = \frac{z(z+1)\cdots(z+n)}{n!\,n^z}.$$

From the definition of $g_n(z)$, we get directly

$$g_n(z) = z\prod_{k=1}^{n}\left(\left(1 + \frac{z}{k}\right)e^{-z/k}\right)e^{z(1+(1/2)+\cdots+(1/n)-\log n)}.$$

This product is a partial product for the Weierstrass product of $1/\Gamma$. Thus we obtain the limiting value for the gamma function, that $\Gamma(z) = \lim_{n\to\infty} 1/g_n(z)$, or in other words,

Γ7.
$$\boxed{\Gamma(z) = \lim_{n\to\infty} \frac{n!\,n^z}{z(z+1)\cdots(z+n)}}$$

uniformly for z in a compact set not containing 0 or a negative integer.

The Gauss Multiplication Formula (Distribution Relation)

Define the entire function

$$D(z) = \frac{\sqrt{2\pi}}{\Gamma(z)} = \sqrt{2\pi}g(z).$$

Then

Γ8.
$$\boxed{\prod_{j=0}^{N-1} D\left(z + \frac{j}{N}\right) = D(Nz)N^{Nz-1/2}.}$$

For example, if $N = 2$ we find the **duplication formula**

$$\Gamma(z)\Gamma(z + \tfrac{1}{2}) = \sqrt{\pi}2^{1-2z}\Gamma(2z),$$

which allows us to express values of the gamma function at half integers in terms of gamma values at integers.

To prove the general formula, the main idea is that the left side and right side have the same zeros and poles, so the quotient is an exponential function of order 1, so of type AB^z, and one has to determine A and B. Of course, there is a computational part to the proof, which we carry out. So we let

$$\frac{\prod_{j=0}^{N-1} \Gamma(z + j/N)}{\Gamma(Nz)} = AB^z = h(z), \quad \text{say.}$$

By using $\Gamma(z + 1) = z\Gamma(z)$ repeatedly, we find

$$h(z + 1) = \prod_{j=0}^{N-1} \frac{z + j/N}{Nz + j} h(z) = N^{-N}h(z).$$

But from the definition of h, we also have $h(z + 1) = Bh(z)$, so

$$B = N^{-N}.$$

There remains to determine A. We can evaluate $\Gamma(z)/\Gamma(Nz)$ at $z = 0$ to find N. Hence from the definition of $h(z) = AB^z$ at $z = 0$, we get

$$A = h(0) = N \prod_{j=1}^{N-1} \Gamma(j/N).$$

In particular, $A > 0$, and it suffices to compute A^2. By **Γ 4** we find:

$$\begin{aligned}\left(\frac{A}{N}\right)^2 &= \prod_{j=1}^{N-1} \Gamma(j/N)\Gamma(1 - j/N) \\ &= \pi^{N-1} \prod_{j=1}^{N-1} (\sin \pi j/N)^{-1} = (2\pi i)^{N-1} \prod_{j=1}^{N-1} (e^{i\pi j/N} - e^{-i\pi j/N})^{-1}.\end{aligned}$$

But

$$\prod_{j=1}^{N-1} (e^{i\pi j/N} - e^{-i\pi j/N}) = \prod_{j=1}^{N-1} e^{i\pi j/N} \prod_{j=1}^{N-1} (1 - e^{-2\pi i j/N}) = i^{N-1}N,$$

by using some simple algebraic identities. Indeed, we use

$$\sum_{j=1}^{N-1} j = (N - 1)N/2,$$

and we note that $e^{-2\pi i j/N}$ $(j = 1, \ldots, N - 1)$ ranges over all N-th roots of unity $\neq 1$. From the factorization

$$\prod_{\zeta^N=1} (X - \zeta) = X^N - 1$$

we find

$$\prod_{\substack{\zeta^N=1 \\ \zeta \neq 1}} (X - \zeta) = \frac{X^N - 1}{X - 1} = X^{N-1} + X^{N-2} + \cdots + 1.$$

Substituting 1 for X on the right-hand side yields N. Then we note that i^{N-1} cancels, so $(A/N)^2 = (2\pi)^{N-1}/N$. Then formula **Γ8** drops out.

Remark. We formulated **Γ8** with the normalized function D so that the powers of π disappeared, for three reasons. First, the number of factors $\sqrt{2\pi}$ which occur is precisely the number of factors in the product. Hence such factors might as well be incorporated from the start into the formula.

The second reason is that if one lets

$$\zeta(s, z) = \sum_{n=0}^{\infty} \frac{1}{(z+n)^s}$$

and if we let $\zeta'(s, z)$ denote the derivative with respect to s, then one has the **Lerch formula**

$$D(z) = \exp(-\zeta'(0, z)).$$

Thus the normalization with $D(z)$ gets rid of all extra constants in this Lerch formula. See Theorem 3.2 below.

The third reason is that, when z ranges over rational numbers, then the term $N^{Nz-1/2}$ is a pure fractional root of an integer, and contains no transcendental factor involving π. It is a conjecture of Rohrlich that the only multiplicative relations of values of the gamma function at rational numbers (up to algebraic factors) are those which follow formally from **Γ3**, **Γ5**, and the multiplication formula **Γ8**.

Thus we see that the normalization $D(z)$ instead of $\Gamma(z)$ is much nicer in several respects. It exhibits the structure of relations involving the gamma function in a much clearer way, than when a random constant appears to be floating around.

For another basic example of a distribution relation (or addition formula) see the exercise of §3, and also Exercises 3, 4 of Chapter XI, §2.

The (Other) Gauss Formula

The Gauss Formula which follows is really a corollary of **Γ2**.

Γ9. $$\boxed{\Gamma'/\Gamma(z) = \int_0^{\infty} \left(\frac{e^{-t}}{t} - \frac{e^{-zt}}{1 - e^{-t}} \right) dt} \qquad \textit{for} \quad \mathrm{Re}(z) > 0.$$

Proof. We start with the simple fact that

$$(1) \qquad \log z = \int_0^{\infty} (e^{-t} - e^{-zt}) \frac{dt}{t} \qquad \text{for} \quad \mathrm{Re}(z) > 0.$$

Indeed, both sides are analytic in z. We can differentiate under the integral sign, and we find $1/z$ for both sides. Evaluating each side at $z = 1$ we find 0. Hence both sides are equal.

Then directly from the logarithmic derivative **Γ 2** and using the definition of the Euler constant

$$\gamma = \lim_{N \to \infty} \left(1 + \frac{1}{2} + \cdots + \frac{1}{N} - \log N\right)$$

we obtain

$$\Gamma'/\Gamma(z) = \lim_{N \to \infty} \left(\log N - \sum_{n=0}^{N-1} \frac{1}{z+n}\right). \tag{2}$$

Using (1) and

$$\frac{1}{z+n} = \int_0^\infty e^{-(z+n)t}\, dt,$$

we see that a partial sum for $\Gamma'/\Gamma(z)$ is given by

$$\begin{aligned}
\log N - \sum_{n=0}^{N-1} \frac{1}{z+n} &= \int_0^\infty \left[e^{-t} - e^{-Nt} - te^{-zt} \sum_{n=0}^{N-1} e^{-nt}\right] \frac{dt}{t} \\
&= \int_0^\infty \left[e^{-t} - e^{-Nt} - te^{-zt}\left(\frac{1}{1-e^{-t}} - \frac{e^{-Nt}}{1-e^{-t}}\right)\right] \frac{dt}{t} \\
&= \int_0^\infty \left(\frac{e^{-t}}{t} - \frac{e^{-zt}}{1-e^{-t}}\right) dt + \int_0^\infty e^{-Nt}\left(\frac{e^{-zt}}{1-e^{-t}} - \frac{1}{t}\right) dt.
\end{aligned}$$

Since

$$\frac{e^{-zt}}{1-e^{-t}} - \frac{1}{t} = O(1) \qquad \text{for} \quad t \to 0,$$

it follows at once that we can take the limit as $N \to \infty$ under the integral sign, giving 0 as this limit, and concluding the proof of the Gauss formula.

Remark. Note that $1/t$ is the principal part of $e^{-zt}/(1 - e^{-t})$ in the sense of meromorphic functions in the variable t. Subtracting this principal part in the integrand makes the integral converge, and is known as **regularizing** the integral.

From **Γ9** one immediately obtains:

Γ10.
$$\gamma = \int_0^\infty \left(\frac{1}{1-e^{-t}} - \frac{1}{t}\right) e^{-t}\, dt.$$

Γ11.
$$\Gamma'/\Gamma(z) = -\gamma + \int_0^\infty \frac{e^{-t} - e^{-tz}}{1 - e^{-t}}\, dt.$$

The Mellin Transform

We now start all over again. We suppose nothing known about the gamma function, and define the **Euler integral**

Γ12.
$$\boxed{\Gamma(z) = \int_0^\infty e^{-t} t^z \frac{dt}{t}.}$$

Then the integral fits the hypotheses of Lemma 1.1, provided

$$0 < a \leqq \operatorname{Re} z \leqq b \quad \text{if} \quad 0 < a < b \text{ are real numbers.}$$

Therefore $\Gamma(z)$ is an analytic function in the right half plane. (Give the details of the proof that the above integral satisfies the hypotheses of the lemma, and write down explicitly what the derivative is.)

Remarks. An integral of the form

$$F(z) = \int_0^\infty \varphi(t) t^z \frac{dt}{t}$$

is called the **Mellin transform** of a function φ. We write dt/t because this expression is invariant under "multiplicative translations". This means: Let f be any function which is absolutely integrable on $0 < t < \infty$. Let a be a positive number. Then

$$\int_0^\infty f(at) \frac{dt}{t} = \int_0^\infty f(t) \frac{dt}{t}.$$

This is verified trivially by the change of variables formula of freshman calculus. Use will be made of this in the exercises. For example, replacing t by nt where n is a positive integer, we obtain

$$\frac{\Gamma(s)}{n^s} = \int_0^\infty e^{-nt} t^s \frac{dt}{t} \quad \text{for} \quad \operatorname{Re}(s) > 0.$$

Summing over n yields what is called the Riemann zeta function of the complex variable s. See Chapter XV, §4.

The Mellin integral for the gamma function converges only for $\operatorname{Re}(z) > 0$. We illustrate right away a technique which allows us to give its analytic continuation to the whole plane, in the simplest context. We write the integral in the form

$$\int_0^\infty = \int_0^1 + \int_1^\infty.$$

Now we observe that the integral from 1 to ∞ defines an entire function of z, say

$$H(z) = \int_1^\infty e^{-t}t^z \frac{dt}{t},$$

because the convergence problem for t near 0 has disappeared. For the other integral

$$P(z) = \int_0^1 e^{-t}t^z \frac{dt}{t},$$

we write down the power series expansion for e^{-t} at $t = 0$. We leave to the reader the justification that we can interchange the finite integral and the series for $\mathrm{Re}(z) > 0$. We perform the integration, which is completely elementary, and we then find

$$P(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\,(z+n)}.$$

This series converges uniformly for z in any compact subset of the plane not containing 0 or a negative integer. Then

$$\Gamma(z) = P(z) + H(z)$$

gives the Mittag-Leffler decomposition of the gamma function in terms of its principal parts, and gives one form for its meromorphic continuation to $\mathbf{C}$.

Another technique can also be applied. We integrate the Euler integral **Γ 12** by parts, with $u = e^{-t}$, $dv = t^{z-1}\,dt$, to get:

$$\Gamma(z) = \frac{1}{z}\int_0^\infty e^{-t}t^z\,dt.$$

In particular, the formula

$$\Gamma(z+1) = z\Gamma(z)$$

follows directly from the definition with the Euler integral. Inductively we obtain

$$\Gamma(z+n+1) = z(z+1)\cdots(z+n)\Gamma(z)$$

which can also be written

$$\Gamma(z) = \frac{1}{z(z+1)\cdots(z+n)} \int_0^\infty e^{-t} t^{z+n}\, dt.$$

Note that the integral in this last expression is holomorphic in z for $\mathrm{Re}(z) > -n-1$, so this expression gives a meromorphic continuation for the Euler integral arbitrarily far to the left as $n \to \infty$. Putting successively $z = 1$, $z = 2, \dots, z = n$ we can also compute the residues at the negative integers and 0 directly from these integral expressions.

After these remarks, we proceed to the proof that the Euler integral is equal to the Weierstrass product for $\Gamma(z)$. Integrating by parts, the reader will prove easily that for x real > 0,

$$(*) \qquad \int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1}\, dt = \frac{n^x n!}{x(x+1)\cdots(x+n)} = 1/g_n(z)$$

for every integer $n \geq 1$. (But first change variables, let $t/n = u$, $dt = n\, du$.) We let the reader prove that

$$0 \leqq e^{-t} - (1 - t/n)^n \leqq t^2 e^{-t}/n \qquad \text{for} \quad 0 \leqq t \leqq n.$$

This is done by using the usual technique that if $f(0) = g(0)$ and $f'(t) \leqq g'(t)$ for $t \in [0, n]$, then $f(t) \leqq g(t)$, and also the inequality $(1 + t/n)^n \leqq e^t$. Since $\int_0^\infty t^2 e^{-t} t^{x-1}\, dt$ converges, it follows that we can take the limit on the left of $(*)$, and obtain for x real > 0:

$$\lim_{n\to\infty} \int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1}\, dt = \int_0^\infty e^{-t} t^{x-1}\, dt.$$

Using **Γ 7** concludes the proof that the Euler integral gives the same value as the Weierstrass product, and therefore we have identified the definition of the gamma function by means of the Mellin transform in the right half plane with the definition by means of the Weierstrass product. Thus the Weierstrass product gives the analytic continuation of the Mellin transform to a meromorphic function in the whole plane.

The Stirling Formula

This formula gives an asymptotic development for the gamma function, and the best statement is the one giving an exact error term, as follows.

Γ13. $$\boxed{\log \Gamma(z) = \left(z - \frac{1}{2}\right)\log z - z + \tfrac{1}{2}\log(2\pi) - \int_0^\infty \frac{P_1(t)}{z+t}\,dt,}$$

where $P_1(t) = t - [t] - \frac{1}{2}$ is the sawtooth function, $[t]$ denoting the largest integer $\leqq t$. One takes the principal value for the log, deleting the negative real axis where the gamma function has its poles. The usefulness of the error term involving the integral of $P_1(t)$ is that it tends to 0 uniformly in every sector of complex numbers $z = re^{i\theta}$ such that

$$-\pi + \delta \leqq \theta \leqq \pi - \delta, \qquad 0 < \delta < \pi.$$

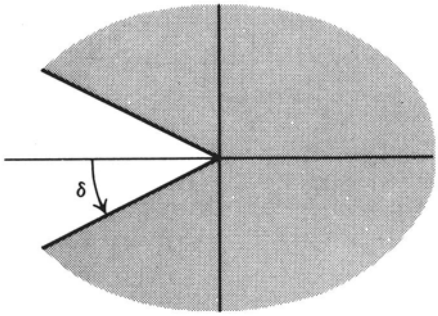


Figure 1

When $z = n$ is a positive integer, it is at the level of calculus to prove that

$$n! = n^n e^{-n}\sqrt{2\pi n}\, e^{\lambda/12n},$$

where $|\lambda| \leqq 1$. Since $n! = \Gamma(n+1)$, one sees that the asymptotic relation

$$n! \sim n^n e^{-n}\sqrt{2\pi n}$$

is a special case of the relation, valid for all $|z| \to \infty$:

Γ14. $$\boxed{\Gamma(z) \sim z^{z-1/2}e^{-z}\sqrt{2\pi},}$$

uniformly in the sector mentioned above. The twiddle sign $\sim$ means that the quotient of the left-hand side by the right-hand side approaches 1, for $|z| \to \infty$.

Proof of Stirling's Formula

We shall need a simple formula.

Euler Summation Formula. *Let f be any C^1 function of a real variable. Then*

$$\sum_{k=0}^{n} f(k) = \int_0^n f(t)\,dt + \tfrac{1}{2}(f(n) + f(0)) + \int_0^n P_1(t)f'(t)\,dt.$$

Proof. The sawtooth function looks like Fig. 2. Note that

$$P_1'(t) = 1$$

for t not an integer.

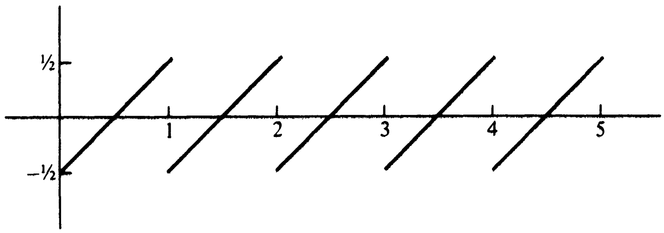


Figure 2

Integrating by parts with $u = P_1(t)$ and $dv = f'(t)\,dt$ yields

$$\int_{k-1}^{k} P_1(t)f'(t)\,dt = P_1(t)f(t)\Big|_{k-1}^{k} - \int_{k-1}^{k} f(t)\,dt$$
$$= \tfrac{1}{2}(f(k) + f(k-1)) - \int_{k-1}^{k} f(t)\,dt.$$

We take the sum from $k = 1$ to $k = n$. Adding the integral

$$\int_0^n f(t)\,dt \qquad \text{and} \qquad \tfrac{1}{2}(f(n) + f(0))$$

then yields the sum $\sum_{k=0}^{n} f(k)$ on the right side and proves the formula.

Before going further and applying the formula to the gamma function, we evaluate some constants. The first constant will not be used in the

Stirling formula, but it will be used in §3, and it gives a simple example for the formula, so we derive it now.

Lemma 2.1.

$$\frac{1}{2} + \int_0^\infty \frac{P_1(t)}{(1+t)^2}\,dt = \gamma.$$

Proof. We apply Euler's formula to the function $f(x) = 1/(1+x)$. Then the formula gives

$$1 + \frac{1}{2} + \cdots + \frac{1}{n+1} = \log(n+1) + \frac{1}{2}\left(\frac{1}{n+1} + 1\right) + \int_0^n P_1(t)\frac{1}{(1+t)^2}\,dt.$$

Since P_1 is bounded, the integral on the right converges absolutely. Hence after subtracting $\log(n+1)$ from both sides, the lemma drops out by the definition of γ.

Let

$$P_2(t) = \tfrac{1}{2}(t^2 - t) \qquad \text{for} \quad 0 \leqq t \leqq 1,$$

and extend $P_2(t)$ by periodicity to all of $\mathbf{R}$ (period 1). Then $P_2(n) = 0$ for all integers n, and P_2 is bounded. Furthermore, $P_2'(t) = P_1(t)$.

Lemma 2.2. *For z not on the negative real axis, we have*

$$\int_0^\infty \frac{P_1(t)}{z+t}\,dt = \int_0^\infty \frac{P_2(t)}{(z+t)^2}\,dt.$$

The integral is analytic in z in the open set U obtained by deleting the negative real axis from the plane, and one can differentiate under the integral sign on the right in the usual way.

Proof. We write

$$\int_0^\infty = \sum_{n=0}^\infty \int_n^{n+1}.$$

Integrating by parts on each interval $[n, n+1]$ gives the identity of the lemma. The integral involving P_2 is obviously absolutely convergent, and the differentiation lemma applies.

Lemma 2.3.

$$\lim_{y\to\infty} \int_0^\infty \frac{P_1(t)}{iy+t}\,dt = 0.$$

Proof. The limit is clear from Lemma 2.2.

We now apply the Euler formula to the functions

$$f(t) = \log(z + t), \qquad \text{and} \qquad f(t) = \log(1 + t)$$

and assume until further notice that z is real > 0. Then we have no difficulty dealing with the log and its properties from freshman calculus. Subtracting the expressions in Euler's formula for these two functions, and recalling that

$$\int \log x \, dx = x \log x - x,$$

we obtain

$$\begin{aligned}
\log \frac{z(z+1)\cdots(z+n)}{n!(n+1)} = {} & z \log(z+n) + n \log(z+n) - z \log z \\
& - (z+n) + z + \tfrac{1}{2}(\log(z+n) + \log z) \\
& - (n+1)\log(n+1) + (n+1) \\
& - 1 - \tfrac{1}{2}\log(n+1) \\
& + \text{the terms involving the integrals of } P_1(t).
\end{aligned}$$

None of this is so bad. We write

$$z \log(z+n) = z \log n\left(1 + \frac{z}{n}\right) = z \log n + z \log\left(1 + \frac{z}{n}\right).$$

The term $z \log n$ is just $\log n^z$, and we move it to the other side.

On the other hand, we note that $n + 1$ occurs in the denominator on the left, and we move $-\log(n+1)$ from the left-hand side to the right-hand side, changing signs. We also make as many cancellations as we can on the right-hand side. We end up with

$$-\left(n + \frac{1}{2}\right)\log\left(1 + \frac{1}{n}\right) + \left(n + \frac{1}{2}\right)\log\left(1 + \frac{z}{n}\right)$$

among other expressions. But from the Taylor expansion for large n (and fixed z) we know that

$$\log\left(1 + \frac{1}{n}\right) = \frac{1}{n} + O\left(\frac{1}{n^2}\right) \qquad \text{and} \qquad \log\left(1 + \frac{z}{n}\right) = \frac{z}{n} + O\left(\frac{1}{n^2}\right).$$

Therefore it is easy to take the limit as n tends to infinity, and we find by **Γ 7**,

$$(*) \qquad \log \Gamma(z) = \left(z - \frac{1}{2}\right)\log z - z + 1 + \int_0^\infty \frac{P_1(t)}{1+t}\,dt - \int_0^\infty \frac{P_1(t)}{z+t}\,dt.$$

All this is true for positive real z. By Lemma 2.2 the integral on the right is analytic in z for z in the open set U equal to the plane from which the negative real axis has been deleted. Since the other expressions

$$\log \Gamma(z), \qquad z \log z, \qquad \log z, \qquad z$$

are also analytic in this open set (which is simply connected) it follows that formula $(*)$ is valid for all z in this open set.

There remains to evaluate the constant.

Lemma 2.4.

$$1 + \int_0^\infty \frac{P_1(t)}{1+t}\,dt = \frac{1}{2}\log 2\pi.$$

Proof. From

$$\Gamma(z)\Gamma(-z) = \frac{-\pi}{z \cdot \sin \pi z}$$

we get

$$|\Gamma(iy)| = \sqrt{\frac{2\pi}{y(e^{\pi y} - e^{-\pi y})}}.$$

From $(*)$ we get

$$\begin{aligned}
1 + \int_0^\infty \frac{P_1(t)}{1+t}\,dt &= \mathrm{Re}\left\{\log \Gamma(iy) - \left(iy - \frac{1}{2}\right)\log(iy) + iy + \int_0^\infty \frac{P_1(t)}{iy+t}\,dt\right\} \\
&= \lim_{y\to\infty}\left\{\log|\Gamma(iy)| + \tfrac{1}{2}\log y + \frac{\pi y}{2}\right\} \quad \text{[by Lemma 2.3]} \\
&= \lim_{y\to\infty} \log\sqrt{\frac{2\pi y e^{\pi y}}{y(e^{\pi y} - e^{-\pi y})}} \\
&= \tfrac{1}{2}\log 2\pi.
\end{aligned}$$

This proves Stirling's formula.

Observe that differentiating under the integral sign involving P_2 yields a good error term for $\Gamma'/\Gamma(z)$, namely

Γ 15.
$$\boxed{\Gamma'/\Gamma(z) = \log z - \frac{1}{2z} + 2\int_0^\infty \frac{P_2(t)}{(z+t)^3}\,dt.}$$

Remark. To integrate by parts more than once, it is more useful to take $P_2(t) = \frac{1}{2}(t^2 - t + \frac{1}{6}) = \frac{1}{2}\mathbf{B}_2(t)$, where $\mathbf{B}_2$ is the second Bernoulli polynomial, and so forth.

XV, §2. EXERCISES

1. Prove that:
   (a) $\Gamma'/\Gamma(1) = -\gamma$.
   (b) $\Gamma'/\Gamma(\frac{1}{2}) = -\gamma - 2\log 2$.
   (c) $\Gamma'/\Gamma(2) = -\gamma + 1$.

2. Give the details for the proofs of formulas **Γ10** and **Γ11**.

3. Prove that $\displaystyle\int_0^\infty e^{-t}\log t\,dt = -\gamma$.

4. Show that
$$\int_0^1 \left(\frac{1}{e^t - 1} - \frac{1}{t}\right)dt + \int_1^\infty \frac{dt}{e^t - 1} = 0.$$

5. Let $a_1, \dots, a_r$ be distinct complex numbers, and let $m_1, \dots, m_r$ be integers. Suppose that
$$h(z) = \prod_{i=1}^r \Gamma(z + a_i)^{m_i}$$
is an entire function without zeros and poles.
   (a) Prove that there are constants A, B such that $h(z) = AB^Z$.
   (b) Assuming (a), prove that $m_i = 0$ for all i.

6. (a) Give an exact value for $\Gamma(1/2 - n)$ when n is a positive integer, and thus show that $\Gamma(1/2 - n) \to 0$ rapidly when $n \to \infty$. Thus the behavior at half the odd negative integers is quite opposite to the polar behavior at the negative integers themselves.
   (b) Show that $\Gamma(1/2 - n + it) \to 0$ uniformly for real t, as $n \to \infty$, n equal to a positive integer.

7. **Mellin Inversion Formula.** Show that for $x > 0$ we have
$$e^{-x} = \frac{1}{2\pi i}\int_{\sigma=\sigma_0} x^{-s}\Gamma(s)\,ds,$$

where $s = \sigma + it$, and the integral is taken on a vertical line with fixed real part $\sigma_0 > 0$, and $-\infty < t < \infty$. [*Hint*: What is the residue of $x^{-s}\Gamma(s)$ at $s = -n$?]

8. Define the alternate Laplace transform L^- by

$$L^-f(w) = \int_0^\infty f(t)e^{wt}\,dt.$$

(a) Let $f(t) = e^{-zt}$ for $t \geqq 0$. Show that

$$L^-f(w) = \frac{1}{z-w} \qquad \text{for} \quad \mathrm{Re}(w) < \mathrm{Re}(z).$$

(b) Let $f(t) = t^{s-1}e^{-zt}$ for $t \geqq 0$. Show that

$$L^-f(w) = \Gamma(w)(z-w)^{-s} \qquad \text{for} \quad \mathrm{Re}(w) < \mathrm{Re}(z).$$

Here $(z-w)^s$ is defined by taking $-\pi/2 < \arg(z-w) < \pi/2$.

9. Consider the gamma function in a vertical strip $x_1 \leqq \mathrm{Re}(z) \leqq x_2$. Let a be a complex number. Show that the function

$$z \mapsto \Gamma(z+a)/\Gamma(z) = h(z)$$

has polynomial growth in the strip (as distinguished from exponential growth). In other words, there exists $k > 0$ such that

$$|h(z)| = O(|z|^k) \qquad \text{for} \quad |z| \to \infty, \quad z \text{ in the strip.}$$

Let the **Paley–Wiener space** consist of those entire functions f for which there exists a positive number C having the following property. Given an integer $N > 0$, we have

$$|f(x+iy)| \ll \frac{C^{|x|}}{(1+|y|)^N},$$

where the implied constant in $\ll$ depends on f and N. We may say that f is at most of exponential growth with respect to x, and is rapidly decreasing, uniformly in every vertical strip of finite width.

10. If f is C^∞ (infinitely differentiable) on the open interval $]0,\infty[$, and has compact support, then its **Mellin transform** Mf defined by

$$Mf(z) = \int_0^\infty f(t)t^z\,\frac{dt}{t}$$

is in the Paley–Wiener space. [*Hint*: Integrate by parts.]

11. Let F be in the Paley-Wiener space. For any real x, define the function ${}^tM_xF$ by

$$ {}^tM_xF(t) = \int_{\mathrm{Re}\,z=x} F(z)t^z\,\frac{dz}{i}.$$

The integral is supposed to be taken on the vertical line $z = x + iy$, with fixed x, and $-\infty < y < \infty$. Show that ${}^tM_xF$ is independent of x, so can be written ${}^tMF$. [*Hint*: Use Cauchy's theorem.] Prove that ${}^tM_xF$ has compact support on $]0, \infty[$.

Remark. If you want to see these exercises worked out, cf. my book $SL_2(\mathbf{R})$, Chapter V, §3. The two maps M and ${}^tM$ are inverse to each other, but one needs the Fourier inversion formula to prove this.

12. Let a, b be real numbers > 0. Define the **K-Bessel function**

K 1. $$K_s(a, b) = \int_0^\infty e^{-(a^2t+b^2/t)}t^s \frac{dt}{t}.$$

Prove that K_s is an entire function of s. Prove that

K 2. $$K_s(a, b) = (b/a)^s K_s(ab),$$

where for $c > 0$ we define

K 3. $$K_s(c) = \int_0^\infty e^{-c(t+1/t)}t^s \frac{dt}{t}.$$

Prove that

K 4. $$K_s(c) = K_{-s}(c).$$

K 5. $$K_{1/2}(c) = \sqrt{\pi/c}\,e^{-2c}.$$

[*Hint*: Differentiate the integral for $\sqrt{x}K_{1/2}(x)$ under the integral sign.] Let $x_0 > 0$ and $\sigma_0 \leqq \sigma \leqq \sigma_1$. Show that there exists a number $C = C(x_0, \sigma_0, \sigma_1)$ such that if $x \geqq x_0$, then

K 6. $$K_\sigma(x) \leqq Ce^{-2x}.$$

Prove that

K 7. $$\int_{-\infty}^\infty \frac{1}{(u^2+1)^s} du = \sqrt{\pi}\frac{\Gamma(s - 1/2)}{\Gamma(s)}$$

for $\mathrm{Re}(s) > 1/2$. Also prove that

K 8. $$\Gamma(s)\int_{-\infty}^\infty \frac{e^{ixu}}{(u^2+1)^s} du = 2\sqrt{\pi}(x/2)^{s-1/2}K_{s-1/2}(x)$$

for $\mathrm{Re}(s) > 1/2$.
[To see this worked out, cf. [La 73], Chapter 20.]

XV, §3. THE LERCH FORMULA

We shall study zeta functions *per se* in the next section, but here we want to give a continuation of the technique used to derive Sterling's formula to obtain an important formula due to Lerch, concerning the gamma function. The next section does not depend on this one, and may be read independently.

Let $u > 0$. We introduce the **Hurwitz zeta function**

$$\zeta(s, u) = \sum_{n=0}^{\infty} \frac{1}{(n+u)^s}.$$

The series converges absolutely and defines an analytic function for $\operatorname{Re}(s) > 1$. In particular, setting $u = 1$ we obtain the Riemann zeta function

$$\zeta(s) = \zeta(s, 1) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Theorem 3.1. *For* $\operatorname{Re}(s) > -1$ *we have an analytic continuation of* $\zeta(s, u)$ *given by*

$$\zeta(s, u) = \frac{u^{1-s}}{s-1} + \frac{u^{-s}}{2} - s \int_0^{\infty} \frac{P_1(t)}{(t+u)^{s+1}}\, dt.$$

Proof. First, for $\operatorname{Re}(s) > 1$, we apply Euler's summation formula to the function

$$f(t) = \frac{1}{(t+u)^s}.$$

We write down this summation formula with a finite number of terms $\sum f(k)$, with $1 \leqq k \leqq n$, and then we let n tend to infinity. The integrals in the Euler formula converge uniformly, and the formula of Theorem 3.1 drops out, except that we still have $\operatorname{Re}(s) > 1$. However, now that we have the formula in this range for s, we note that the oscillating integral on the right, involving P_1 as in Lemma 2.2, is uniformly convergent for $\operatorname{Re}(s) > -1$, and therefore gives the analytic continuation of $\zeta(s, u)$ in this larger domain. Furthermore, by Lemma 1.1 we may differentiate under the integral sign, so we get a way of finding the derivative in this large domain.

For each fixed value of u, it is now easy to find the first few terms of the power series expansion of $\zeta(s, u)$ at the origin. We put

$$D(u) = \frac{\sqrt{2\pi}}{\Gamma(u)}$$

and we let $O(s^2)$ denote an analytic function divisible by s^2 near $s = 0$.

Theorem 3.2.

$$\zeta(s, u) = \tfrac{1}{2} - u - \big(\log D(u)\big)s + O(s^2).$$

Proof. We use the geometric series for $1/(s-1) = -1/(1-s)$ and use $u^{-s} = 1 - s\log u + O(s^2)$. The integral on the right of the formula in Theorem 3.1 is holomorphic at $s = 0$, and its value at $s = 0$ is obtained by substituting $s = 0$. Thus we find

$$\zeta(s, u) = -u(1 - s\log u)(1 + s) + \frac{1 - s\log u}{u} - s\int_0^\infty \frac{P_1(t)}{t + u}\,dt + O(s^2).$$

The formula stated in Theorem 3.2 is now a consequence of Sterling's formula **Γ 13.**

We denote by $\zeta'(s, u)$ the derivative with respect to s.

Corollary 3.3.

$$\zeta'(0) = -\frac{1}{2}\log 2\pi \qquad \text{and} \qquad \zeta(s) = \frac{1}{s-1} + \gamma + O(s-1).$$

Proof. For the first expression, put $u = 1$ in Theorem 3.2. For the second expression, write $s = s - 1 + 1$ in front of the integral of Theorem 3.1. Then the constant term γ drops out by using Lemma 2.1. This concludes the proof.

Corollary 3.4 (Lerch Formula).

$$\log D(u) = -\zeta'(0, u)$$

or completely in terms of the gamma function,

$$\log \Gamma(u) = \zeta'(0, u) - \zeta'(0).$$

Proof. Immediate from the power series expansion of Theorem 3.2, and the value of $\zeta'(0)$ in Corollary 3.3.

Remark. We could have replaced $u > 0$ throughout by a complex z with $\mathrm{Re}(z) > 0$. The formulas converge in this case. Furthermore, $\log \Gamma(z)$ is analytic for z in the plane from which the half line $(-\infty, 0]$ is deleted, and so from Theorem 3.3 we see that the Lerch formula holds if u is replaced by z on this open set. Defining the **regularized product**

$$D(z) = \frac{\sqrt{2\pi}}{\Gamma(z)},$$

we obtain the Lerch formula for the complex variable z, namely:

$$\boxed{\log D(z) = -\zeta'(0, z).}$$

This is the most classical and most elementary special case of a general formalism concerning regularized products and determinants, which arose in many contexts, including physics. The paper by Voros [Vo 87] gives a treatment, including several examples from classical partial differential equations. See especially Voros' Section 4, on zeta functions and functional determinants. For further results, see [JoL 93].

[Vo 87] A. Voros, Spectral functions, special functions and the Selberg zeta function, *Commun. Math. Phys.* **110** (1987) pp. 439–465

[JoL 93] J. Jorgenson and S. Lang, *Basic analysis of regularized series and products*, Lecture Notes in Mathematics **1564**, Springer-Verlag, 1993

XV, §3. EXERCISE

1. For each real number x, we let $\{x\}$ be the unique number such that $x - \{x\}$ is an integer and $0 < \{x\} \leqq 1$. Let N be a positive integer. Prove the **addition formula (distribution relation)**

$$N^{-s} \sum_{j=0}^{N-1} \zeta(s, \{x + j/N\}) = \zeta(s, \{Nx\}).$$

From this formula and Theorem 3.2, deduce another proof for the multiplication formula of the gamma function.

XV, §4. ZETA FUNCTIONS

In this section, we are concerned more with the zeta function, its analytic continuation to the whole plane, and the functional equation. We deal to a large extent with the **Hurwitz zeta function** introduced in the preceding section, namely for $u > 0$,

$$\zeta(s, u) = \sum_{n=0}^{\infty} \frac{1}{(n + u)^s}, \qquad \text{analytic for } \operatorname{Re}(s) > 1.$$

In particular, setting $u = 1$ we obtain the **Riemann zeta function**

$$\zeta(s) = \zeta(s, 1) = \sum_{n=1} \frac{1}{n^s}.$$

We shall now express the Riemann and Hurwitz zeta functions as Mellin transforms. We have

$$\Gamma(s) = \int_0^\infty e^{-t}t^s \frac{dt}{t} \qquad \text{for } \operatorname{Re}(s) > 0$$
$$= \int_0^\infty e^{-(n+u)t}(n+u)^s t^s \frac{dt}{t}$$

after making the multiplicative change of variables $t \mapsto (n+u)t$, which leaves the integral invariant. Therefore dividing both sides by $(n+u)^s$ and summing, we get for $\operatorname{Re}(s) > 1$:

$$\Gamma(s)\zeta(s, u) = \sum_{n=0}^\infty \frac{\Gamma(s)}{(n+u)^s} = \int_0^\infty \sum_{n=0}^\infty e^{-(n+u)t} t^s \frac{dt}{t}.$$

But

$$\sum_{n=0}^\infty e^{-(n+u)t} = \frac{e^{-ut}}{1-e^{-t}}.$$

Therefore, letting

$$G_u(z) = \frac{e^{-uz}}{1-e^{-z}},$$

we find

$$\Gamma(s)\zeta(s, u) = \int_0^\infty G_u(t)t^s \frac{dt}{t},$$

so the zeta function times the gamma function is the Mellin transform of G_u.

We let

$$F_u(z) = \frac{e^{uz}}{e^z - 1}.$$

Then

$$G_u(-z) = \frac{e^{uz}}{1-e^z} = -F_u(z).$$

We introduced these functions to get an analytic continuation of the zeta function to a meromorphic function on all of **C** by means of the **Hankel integral**

$$H_u(s) = \int_C F_u(z)z^s \frac{dz}{z},$$

where the integral is taken over the contour C as shown on Fig. 3, in which K_ϵ is a small circle of radius ϵ around the origin.

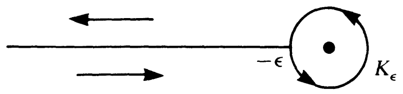


Figure 3

As usual, $z^s = e^{s \log z}$, and the log is taken as its principal value on the complex plane from which the negative real axis is deleted. Symbolically we may write

$$\int_C = \int_{-\infty}^{-\epsilon} + \int_{K_\epsilon} + \int_{-\epsilon}^{-\infty},$$

but the integrand in the first and last integral may not be the same, corresponding to the two values of z^s which differ by a constant.

The exponential decay of the integrand and Lemma 1.1 show at once that H_u is an entire function of s. We shall see that it gives the analytic continuation for $\zeta(s, u)$, and therefore also for the Riemann zeta function by putting $u = 1$.

Theorem 4.1.

$$H_u(s) = -(e^{i\pi s} - e^{-i\pi s})\Gamma(s)\zeta(s, u) = -2i \sin(\pi s)\Gamma(s)\zeta(s, u).$$

Proof. We change the variable, putting $z = -w$. Then writing $G = G_u$, $F = F_u$, we find

$$\begin{aligned} H_u(s) = e^{-i\pi s} &\int_{\infty}^{\epsilon} F(-w)e^{s \log w} \frac{dw}{w} \\ &+ \int_{-K_\epsilon} F(-w)e^{s \log(-w)} \frac{dw}{w} \\ &+ e^{i\pi s} \int_{\epsilon}^{\infty} F(-w)e^{s \log(w)} \frac{dw}{w}. \end{aligned}$$

Therefore taking the limit as $\epsilon \to 0$ (see the lemma below), we find

$$\begin{aligned} H_u(s) &= e^{-i\pi s} \int_0^\infty G(t)t^s \frac{dt}{t} - e^{i\pi s} \int_0^\infty G(t)t^s \frac{dt}{t} \\ &= -(e^{i\pi s} - e^{-i\pi s}) \int_0^\infty G(t)t^s \frac{dt}{t} \end{aligned}$$

which proves the theorem, except for proving the lemma.

Lemma 4.2. *If* $\mathrm{Re}(s) > 1$, *then*

$$\int_{-K_\epsilon} G(w) e^{s \log(-w)} \frac{dw}{w} \to 0 \quad \text{as} \quad \epsilon \to 0.$$

Proof. The length of K_ϵ and $|dw/w|$ have a product which is bounded. But putting $r = |z|$ and $\sigma = \mathrm{Re}(s)$, we have

$$e^{s \log z} \ll e^{\sigma \log r} = r^\sigma,$$

and

$$G(z) \ll 1/r \quad \text{for} \quad r \to 0.$$

This proves the lemma.

Remark. From Theorem 4.1, the fact that H_u is an entire function, and the fact that the zeros of $\sin \pi s$ at the negative integers and 0 cancel the poles of $\Gamma(s)$ at these integers, we see that $\zeta(s)$ is holomorphic at these integers. We shall determine where $\zeta(s)$ has a pole shortly. Before doing this, we get the values of ζ at the negative integers.

Theorem 4.3. *We have*

$$\zeta(s, u) = -\frac{1}{2\pi i} \Gamma(1 - s) H_u(s).$$

In particular, if n is a positive integer, then

$$\zeta(1 - n, u) = -\frac{1}{2\pi i} \Gamma(n) H_u(1 - n)$$

and

$$\frac{\zeta(1 - n, u)}{\Gamma(n)} = -\textit{residue of } F_u(z) z^{-n} \textit{ at } z = 0.$$

In particular, letting B_n denote the Bernoulli numbers, we find

$$\zeta(1 - n) = -\frac{1}{n} B_n \qquad \textit{except for} \quad \zeta(0) = B_1 = -\tfrac{1}{2}.$$

Proof. Observe that when $s = 1 - n$ in the Hankel integral, then the integrand is a meromorphic function, and so the integrals from $-\infty$ to $-\epsilon$ and from $-\epsilon$ to $-\infty$ cancel, leaving only the integral over K_ϵ. We can then apply Cauchy's formula to get the stated value. The assertion about the value of the zeta function at negative integers then comes immediately from the definition of the Bernoulli numbers in terms of the coefficients of a power series, namely

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!}$$

Next we shall obtain an expression for the Hankel function which will lead to the functional equation.

Theorem 4.4. *For* $\mathrm{Re}(s) < 0$ *we have*

$$\begin{aligned} H_u(s) &= (-2\pi)^s \sum_{n=1}^{\infty} \frac{e^{2\pi i u n} e^{i\pi s/2} - e^{-2\pi i u n} e^{-i\pi s/2}}{n^{1-s}} \\ &= -(2\pi)^s \sum_{n=1}^{\infty} \frac{2i \sin(2\pi u n + \pi s/2)}{n^{1-s}}. \end{aligned}$$

Proof. Let m be an integer $\geqq 2$, and let D_m be the path indicated on Fig. 4, consisting of the square and the portion of C inside the square, with the given orientation.

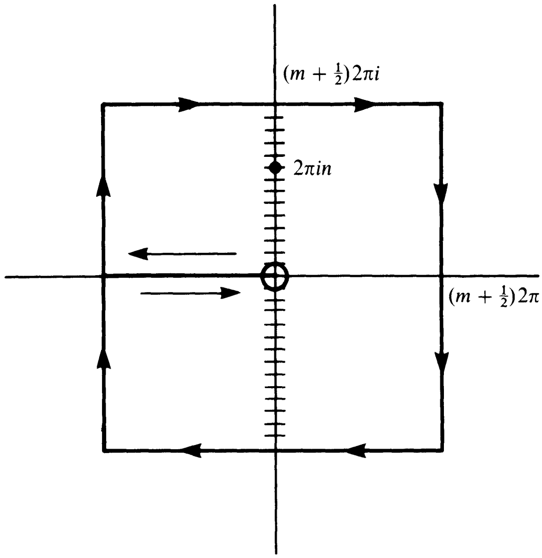


Figure 4

Thus the square crosses the axes at half integral multiples of $2\pi i$. Let

$$R_n = \text{residue of } F_u(z) z^{s-1} \text{ at } 2\pi i n, \quad n \neq 0, \quad -m \leqq n \leqq m.$$

Then

$$\int_{D_m} F(z) z^s \frac{dz}{z} = -2\pi i \sum_{\substack{n=-m \\ n \neq 0}}^{m} R_n.$$

For $n \geqq 1$ we have

$$R_n = \frac{1}{i} e^{2\pi i u n} (2\pi n)^s e^{i\pi s/2},$$

$$R_{-n} = \frac{1}{-i} e^{-2\pi i u n} (2\pi n)^{s-1} e^{-i\pi s/2}.$$

Note that $F(z)/z$ is bounded on the outside square. Hence if $\mathrm{Re}(s) < 0$, the Hankel integral over the outside square tends to 0 as $m \to \infty$. Hence

$$H_u(s) = \int_C F(z) z^s \frac{dz}{z} = \lim_{m \to \infty} \int_{D_m} F(z) z^s \frac{dz}{z}$$

and the theorem follows.

We apply Theorem 4.4 to the Riemann zeta function, taking $u = 1$. Then for $\mathrm{Re}(s) < 0$, we obtain

$$H_1(s) = -(2\pi)^s \sum_{n=1}^{\infty} \frac{e^{i\pi s/2} - e^{-i\pi s/2}}{n^{1-s}}$$

and therefore using Theorem 4.2:

Theorem 4.5.

$$\zeta(s) = (2\pi)^s \Gamma(1 - s) \frac{\sin(\pi s/2)}{\pi} \zeta(1 - s).$$

Observe that the formula of Theorem 4.5 was derived at first when $\mathrm{Re}(s) < 0$ so that the series $\sum 1/n^{1-s}$ converges absolutely. However, we know from Theorem 4.1 that $\zeta(s)$ is a meromorphic function of s, so that quite independently of the series representation in any region, the relation of Theorem 4.5 holds unrestrictedly for all s by analytic continuation. We may reformulate the functional equation relating $\zeta(s)$ and $\zeta(1 - s)$ in another more symmetric form as follows.

Theorem 4.6. *Let $\xi(s) = s(s - 1)\pi^{-s/2}\Gamma(s/2)\zeta(s)$. Then ξ is an entire function satisfying the functional equation*

$$\xi(s) = \xi(1 - s)$$

The zeta function itself is meromorphic, and holomorphic except for a simple pole at $s = 1$.

Proof. The term $s(s - 1)$ remains unchanged under the transformation $s \mapsto 1 - s$. That the other factor $\pi^{-s/2}\Gamma(s/2)\zeta(s)$ remains unchanged follows at once from Theorem 4.5, using the formulas

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s} \quad \text{and} \quad \Gamma(s)\Gamma(s+\tfrac{1}{2}) = 2\pi^{1/2}2^{-2s}\Gamma(2s).$$

We leave the actual easy computation to the reader. As to the fact that $\xi(s)$ is an entire function, and that $s = 1$ is the only pole of $\zeta(s)$, we argue as follows. From Theorem 4.1, from the fact that H_1 is entire, and that the zeros of $\sin(\pi s)\Gamma(s)$ occur only at the positive integers, we see that $\zeta(s)$ can have poles only at the positive integers. But the series $\sum 1/n^s$ shows that $\zeta(s)$ is holomorphic for $\mathrm{Re}(s) > 1$, so the only possible pole is at $s = 1$. From Theorem 4.5, we see that $s = 1$ is actually a pole of order 1, coming from the pole of $\Gamma(1 - s)$, since the other expressions on the right of the formula in Theorem 4.5 are analytic at $s = 1$. This concludes the proof of Theorem 4.6.

XV, §4. EXERCISES

1. (a) Show that $\zeta(s)$ has zeros of order 1 at the even negative integers.
   (b) Show that the only other zeros are such that $0 \leqq \mathrm{Re}(s) \leqq 1$.
   (c) Prove that the zeros of (b) actually have $\mathrm{Re}(s) = 1/2$. [You can ask the professor teaching the course for a hint on that one.]

2. Define $F(z) = \xi(\frac{1}{2} + iz)$. Prove that $F(z) = F(-z)$.

3. Let C be the contour as shown on Fig. 3. Thus the path consists of $]-\infty, -\epsilon]$, the circle which we denote by K_ϵ, and the path from $-\epsilon$ to $-\infty$. On the plane from which the negative real axis has been deleted, we take the principal value for the log, and for complex s,

$$z^{-s} = e^{-s \log z}.$$

   The integrals will involve z^s, and the two values for z^s in the first and third integral will differ by a constant.
   (a) Prove that the integral

$$\int_C e^z z^{-s}\, dz$$

   defines an entire function of s.
   (b) Prove that for $\mathrm{Re}(1 - s) > 0$ we have

$$\int_C e^z z^{-s}\, dz = 2i \sin \pi s \int_0^\infty e^{-u} u^{-s}\, du.$$

   (c) Show that

$$\frac{1}{\Gamma(s)} = \frac{1}{2\pi i} \int_C e^z z^{-s}\, dz.$$

   The contour integral gives another analytic continuation for $1/\Gamma(s)$ to the whole plane.

CHAPTER XVI

The Prime Number Theorem

At the turn of the century, Hadamard and de la Vallee Poussin independently gave a proof of the prime number theorem, exploiting the theory of entire functions which had been developed by Hadamard. Here we shall give D.J. Newman's proof, which is much shorter. I have also benefited from Korevaar's exposition. See:

D.J. NEWMAN, Simple analytic proof of the prime number theorem, *Amer. Math. Monthly* **87** (1980) pp. 693–696

J. KOREVAAR, On Newman's quick way to the prime number theorem, *Math. Intell.* **4**, No. 3 (1982) pp. 108–115

See also [BaN 97]. Newman's proof illustrates again several techniques of complex analysis: contour integration, absolutely convergent products in a context different from Weierstrass products, and various aspects of entire functions in a classical context. Thus this chapter gives interesting more advanced reading material, and displays the versatility of applications of complex analysis.

In Chapter XV we touched already on the zeta function, and gave a method to prove the functional equation and the analytic continuation to the whole plane by means of the Hankel integrals. Here we shall develop whatever we need of analysis from scratch, but we assume the unique factorization of integers into primes.

We recall the notation: if f, g are two functions of a variable x, defined for all x sufficiently large x, and g is positive, we write

$$f = O(g)$$

to mean that there exists a constant $C > 0$ such that $|f(x)| \leqq Cg(x)$ for all x sufficiently large. Thus a function $f = O(1)$ means that f is bounded for $x \geqq x_0$.

XVI, §1. BASIC ANALYTIC PROPERTIES OF THE ZETA FUNCTION

Let s be a complex variable. For $\mathrm{Re}(s) > 1$ the series

$$\sum_{n=1}^{\infty} \frac{1}{n^s}$$

converges absolutely, and uniformly for $\mathrm{Re}(s) \geqq 1 + \delta$, with any $\delta > 0$. One sees this by estimating

$$\left| \frac{1}{n^s} \right| \leqq \frac{1}{n^{1+\delta}}$$

and by using the integral test on the real series $\sum 1/n^{1+\delta}$, which has positive terms.

As you should know, a **prime number** is an integer $\geqq 2$ which is divisible only by itself and 1. Thus the prime numbers start with the sequence 2, 3, 5, 7, 11, 13, 17, 19, ...

Theorem 1.1. *The product*

$$\prod_p \left(1 - \frac{1}{p^s}\right)$$

converges absolutely for $\mathrm{Re}(s) > 1$, *and uniformly for* $\mathrm{Re}(s) \geqq 1 + \delta$ *with* $\delta > 0$, *and we have*

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}.$$

Proof. The convergence of the product is an immediate consequence of the definition given in Chapter XIII, §1 and the same estimate which gave the convergence of the series for the zeta function above. In the same region $\mathrm{Re}(s) \geqq 1 + \delta$, we can use the geometric series estimate to conclude that

$$\left(1 - \frac{1}{p^s}\right)^{-1} = 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \cdots = E_p(s), \quad \text{say.}$$

Using a basic fact from elementary number theory that every positive integer has unique factorization into primes, up to the order of the factors, we conclude that in the product of the terms $E_p(s)$ for all primes p, the expression $1/n^s$ will occur exactly once, thus giving the series for the zeta function in the region $\mathrm{Re}(s) > 1$. This concludes (Euler's) proof.

The product

$$\prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$$

is called the **Euler product**. The representation of the zeta function as such a product shows that $\zeta(s) \neq 0$ for $\mathrm{Re}(s) > 1$. We shall refine this statement to the line $\mathrm{Re}(s) = 1$ in Theorem 1.3.

In Exercise 5 of Chapter XV, §2, we gave one method to show how the zeta function extends to a meromorphic function on the whole plane. We do not use this exercise but reprove ad hoc what we need for the application to the prime number theorem in the next section.

Theorem 1.2. *The function*

$$\zeta(s) - \frac{1}{s-1}$$

extends to a holomorphic function on the region $\mathrm{Re}(s) > 0$.

Proof. For $\mathrm{Re}(s) > 1$, *we have*

$$\begin{aligned}\zeta(s) - \frac{1}{s-1} &= \sum_{n=1}^{\infty} \frac{1}{n^s} - \int_1^{\infty} \frac{1}{x^s}\,dx \\ &= \sum_{n=1}^{\infty} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s}\right) dx.\end{aligned}$$

We estimate each term in the sum by using the relations

$$f(b) - f(a) = \int_a^b f'(t)\,dt \qquad \text{and so} \qquad |f(b) - f(a)| \leqq \max_{a \leqq t \leqq b} |f'(t)|\,|b-a|.$$

Therefore each term is estimated as follows:

$$\begin{aligned}\left|\int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s}\right) dx\right| &\leqq \max_{n \leqq x \leqq n+1} \left|\frac{1}{n^s} - \frac{1}{x^s}\right| \leqq \max \left|\frac{s}{x^{s+1}}\right| \\ &\leqq \frac{|s|}{n^{\mathrm{Re}(s)+1}}.\end{aligned}$$

Thus the sum of the terms converges absolutely and uniformly for $\mathrm{Re}(s) > \delta$. This concludes the proof of the theorem.

We note that the zeta function has a certain symmetry about the x-axis, namely

$$\zeta(\bar{s}) = \overline{\zeta(s)}.$$

This is immediate from the Euler product and the series expansion of Theorem 1.2. It follows that if s_0 is a complex number where ζ has a zero of order m (which may be a pole, in which case m is negative), then the complex conjugate $\bar{s}_0$ is a complex number where ζ has a zero of the same order m.

We now define

$$\varphi(x) = \sum_{p \leqq x} \log p \qquad \text{and} \qquad \Phi(s) = \sum_p \frac{\log p}{p^s} \qquad \text{for} \quad \operatorname{Re}(s) > 1.$$

The sum defining $\Phi(s)$ converges uniformly and absolutely for

$$\operatorname{Re}(s) \geqq 1 + \delta,$$

by the same argument as for the sum defining the zeta function. We merely use the fact that given $\epsilon > 0$,

$$\log n \leqq n^\epsilon \qquad \text{for all} \quad n \geqq n_0(\epsilon).$$

Theorem 1.3. *The function Φ is meromorphic for* $\operatorname{Re}(s) > \frac{1}{2}$. *Furthermore, for* $\operatorname{Re}(s) \geqq 1$, *we have* $\zeta(s) \neq 0$ *and*

$$\Phi(s) - \frac{1}{s-1}$$

has no poles for $\operatorname{Re}(s) \geqq 1$.

Proof. For $\operatorname{Re}(s) > 1$, the Euler product shows that $\zeta(s) \neq 0$. By Chapter XIII, Lemma 1.2, we get

$$-\zeta'/\zeta(s) = \sum_p \frac{\log p}{p^s - 1}.$$

Using the geometric series we get the expansion

$$\begin{aligned}\frac{1}{p^s - 1} = \frac{1}{p^s}\frac{1}{1 - 1/p^s} &= \frac{1}{p^s}\left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots\right) \\ &= \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots\end{aligned}$$

so

$$(*) \qquad -\zeta'/\zeta(s) = \Phi(s) + \sum_p h_p(s) \qquad \text{where} \qquad |h_p(s)| \leqq C\frac{\log p}{|p^{2s}|}$$

for some constant C. But the series $\sum (\log n)/n^{2s}$ converges absolutely and uniformly for $\mathrm{Re}(s) \geqq \frac{1}{2} + \delta$, with $\delta > 0$, so Theorem 1.2 and (*) imply that Φ is meromorphic for $\mathrm{Re}(s) > \frac{1}{2}$, and has a pole at $s = 1$ and at the zeros of ζ, but no other poles in this region.

There remains only to prove that ζ has no zero on the line $\mathrm{Re}(s) = 1$. We follow Titchmarsh, *Theory of the Riemann zeta function*, III, 3.3 and 3.4. Put $s = \sigma + it$. Then

$$\zeta(s) = \exp \sum_p \sum_{m=1}^{\infty} \frac{1}{mp^{ms}},$$

so

$$|\zeta(s)| = \exp \sum_p \sum_{m=1}^{\infty} \frac{\cos(mt \log p)}{mp^{m\sigma}}.$$

It follows that

$$(1) \qquad \zeta^3(\sigma)|\zeta(\sigma + it)|^4|\zeta(\sigma + 2it)|$$

$$= \exp \sum_p \sum_{m=1}^{\infty} \frac{3 + 4\cos(mt \log p) + \cos(2mt \log p)}{mp^{m\sigma}}.$$

From the positivity

$$(2) \qquad 3 + 4\cos\theta + \cos 2\theta = 2(1 + \cos\theta)^2 \geqq 0,$$

it follows that

$$(3) \qquad \zeta^3(\sigma)|\zeta(\sigma + it)|^4|\zeta(\sigma + 2it)| \geqq 1 \qquad \text{for} \quad \sigma > 1.$$

Fix $t \neq 0$. Suppose $\zeta(1 + it) = 0$. Then for $\sigma > 1$, $\sigma \to 1$,

$$\zeta^3(\sigma) = O(1/(\sigma - 1)^3) \qquad \text{and} \qquad \zeta(\sigma + 2it) = O(1),$$

but $\zeta(\sigma + it) = O(\sigma - 1)$. Hence the left side of (3) is $O(\sigma - 1)$, contradiction. This proves Theorem 1.3.

It will be convenient to have an integral expression for Φ.

Proposition 1.4. *For* $\mathrm{Re}(s) > 1$ *we have*

$$\Phi(s) = s\int_1^{\infty} \frac{\varphi(x)}{x^{s+1}}\,dx.$$

Proof. To prove this, compute the integral on the right between successive prime numbers, where φ is constant. Then sum by parts. We leave the details as an exercise.

XVI, §1. EXERCISES

1. Let f, g be two functions defined on the integers > 0 and $\leqq n + 1$. Assume that $f(n + 1) = 0$. Let $G(k) = g(1) + \cdots + g(k)$. Prove the formula for summation by parts:

$$\sum_{k=1}^{n} f(k)g(k) = \sum_{k=1}^{n} \big(f(k) - f(k + 1)\big)G(k).$$

2. Prove the integral expression for Φ in Proposition 1.4.

3. Let $\{a_n\}$ be a sequence of complex numbers such that $\sum a_n$ converges. Let $\{b_n\}$ be a sequence of real numbers which is increasing, i.e. $b_n \leqq b_{n+1}$ for all n, and $b_n \to \infty$ as $n \to \infty$. Prove that

$$\lim_{N\to\infty} \frac{1}{b_N} \sum_{n=1}^{N} a_n b_n = 0.$$

Does this conclusion still hold if we only assume that the partial sums of $\sum a_n$ are bounded?

4. Let $\{a_n\}$ be a sequence of complex numbers. The series

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

is called a **Dirichlet series**. Let σ_0 be a real number. Prove that if the Dirichlet series converges for some value of s with $\mathrm{Re}(s) = \sigma_0$, then it converges for all s with $\mathrm{Re}(s) > \sigma_0$, uniformly on every compact subset of this region.

5. Let $\{a_n\}$ be a sequence of complex numbers. Assume that there exists a number C and $\sigma_1 > 0$ such that

$$|a_1 + \cdots + a_n| \leqq Cn^{\sigma_1} \qquad \text{for all } n.$$

Prove that $\sum a_n/n^s$ converges for $\mathrm{Re}(s) > \sigma_1$. [Use summation by parts.]

6. Prove the following theorem.

Let $\{a_n\}$ be a sequence of complex numbers, and let A_n denote the partial sum

$$A_n = a_1 + \cdots + a_n.$$

Let $0 \leqq \sigma_1 < 1$, and assume that there is a complex number ρ such that for all n we have

$$|A_n - n\rho| \leqq Cn^{\sigma_1},$$

or in other words, $A_n = n\rho + O(n^{\sigma_1})$. *Then the function* f *defined by the Dirichlet series*

$$f(s) = \sum \frac{a_n}{n^s} \quad \textit{for} \quad \mathrm{Re}(s) > 1$$

has an analytic continuation to the region $\mathrm{Re}(s) > \sigma_1$, *where it is analytic except for a simple pole with residue* ρ *at* $s = 1$.

[*Hint*: Consider $f(s) - \rho\zeta(s)$, use Theorem 1.2, and apply Exercise 5.]

XVI, §2. THE MAIN LEMMA AND ITS APPLICATION

We shall now deal with the more number theoretic applications of the analytic properties. We shall state one more fundamental analytic theorem, and show how it implies the prime number theorem.

Theorem 2.1 (Chebyshev). $\varphi(x) = O(x)$.

Proof. Let n be a positive integer. Then

$$2^{2n} = (1+1)^{2n} = \sum_j \binom{2n}{j} \geqq \binom{2n}{n} \geqq \prod_{n<p\leqq 2n} p = e^{\varphi(2n)-\varphi(n)}.$$

Hence we get the inequality

$$\varphi(2n) - \varphi(n) \leqq 2n \log 2.$$

But if x increases by 1, then $\varphi(x)$ increases by at most $\log(x+1)$, which is $O(\log x)$. Hence there is a constant $C > \log 2$ such that for all $x \geqq x_0(C)$ we have

$$\varphi(x) - \varphi(x/2) \leqq Cx.$$

We apply this inequality in succession to $x, x/2, x/2^2, \ldots, x/2^r$ and sum. This yields

$$\varphi(x) \leqq 2Cx + O(1),$$

which proves the theorem.

We shall now state the main lemma, which constitutes the delicate part of the proof. Let f be a function defined on the real numbers $\geqq 0$, and assume for simplicity that f is bounded, and piecewise continuous. What we prove will hold under much less restrictive conditions: instead of piecewise continuous, it would suffice to assume that the integral

$$\int_a^b |f(t)|\, dt$$

exists for every pair of numbers a, $b \geqq 0$. We shall associate to f the Laplace transform g defined by

$$g(z) = \int_0^\infty f(t)e^{-zt}\,dt \qquad \text{for } \operatorname{Re}(z) > 0.$$

We can then apply the differentiation lemma of Chapter XV, §1, whose proof applies to a function f satisfying our conditions (piecewise continuous and bounded). We conclude that g is analytic for $\operatorname{Re}(z) > 0$. (Do Exercise 1.)

Lemma 2.2 (Main Lemma). *Let f be bounded, piecewise continuous on the reals $\geqq 0$. Let $g(z)$ be defined by the above integral for* $\operatorname{Re}(z) > 0$. *If g extends to an analytic function for* $\operatorname{Re}(z) \geqq 0$, *then*

$$\int_0^\infty f(t)\,dt \textit{ exists and is equal to } g(0).$$

We postpone the proof of the main lemma to the next section, and now give its application. Observe that the function φ is piecewise continuous. In fact, it is locally constant: there is no change in φ between prime numbers.

The application of the main lemma is to prove:

Lemma 2.3. *The integral*

$$\int_1^\infty \frac{\varphi(x) - x}{x^2}\,dx$$

converges.

Proof. Let

$$f(t) = \varphi(e^t)e^{-t} - 1 = \frac{\varphi(e^t) - e^t}{e^t}.$$

Then f is certainly piecewise continuous, and is bounded by Theorem 2.1. Making the substitution $x = e^t$ in the desired integral, $dx = e^t\,dt$, we see that

$$\int_1^\infty \frac{\varphi(x) - x}{x^2}\,dx = \int_0^\infty f(t)\,dt.$$

Therefore it suffices to prove that the integral on the right converges. By the main lemma, it suffices to prove that the Laplace transform of f is analytic for $\operatorname{Re}(z) \geqq 0$, so we have to compute this Laplace transform. We claim that in this case,

$$g(z) = \frac{\Phi(z+1)}{z+1} - \frac{1}{z}.$$

Once we have proved this formula, we can then apply Theorem 1.3 to conclude that g is analytic for $\mathrm{Re}(z) \geqq 0$, thus concluding the proof of Lemma 2.3.

Now to compute $g(z)$, we use the integral formula of Proposition 1.4. By this formula, we obtain

$$\frac{\Phi(s)}{s} - \frac{1}{s-1} = \int_1^\infty \frac{\varphi(x) - x}{x^{s+1}} dx.$$

Therefore

$$\begin{aligned} \frac{\Phi(z+1)}{z+1} - \frac{1}{z} &= \int_1^\infty \frac{\varphi(x) - x}{x^{z+2}} dx \\ &= \int_0^\infty \frac{\varphi(e^t) - e^t}{e^{2t}} e^{-zt} e^t \, dt \\ &= \int_0^\infty f(t) e^{-zt} \, dt. \end{aligned}$$

This gives us the Laplace transform of f, and concludes the proof of Lemma 2.3.

Let f_1 and f_2 be functions defined for all $x \geqq x_0$, for some x_0. We say that f_1 is **asymptotic** to f_2, and write

$$f_1 \sim f_2 \qquad \text{if and only if} \qquad \lim_{x \to \infty} f_1(x)/f_2(x) = 1.$$

Theorem 2.4. *We have $\varphi(x) \sim x$.*

Proof. The assertion of the theorem is logically equivalent to the combination of the following two assertions:

Given $\lambda > 1$, the set of x such that $\varphi(x) \geqq \lambda x$ is bounded;

Given $0 < \lambda < 1$, the set of x such that $\varphi(x) \leqq \lambda x$ is bounded.

Let us prove the first. Suppose the first assertion is false. Then there is some $\lambda > 1$ such that for arbitrarily large x we have $\varphi(x)/x \geqq \lambda$. Since φ is monotone increasing, we get for such x:

$$\int_x^{\lambda x} \frac{\varphi(t) - t}{t^2} dt \geqq \int_x^{\lambda x} \frac{\lambda x - t}{t^2} dt = \int_1^\lambda \frac{\lambda - t}{t^2} dt > 0.$$

The number on the far right is independent of x. Since there are arbitrarily large x satisfying the above inequality, it follows that the integral of Lemma 2.3 does not converge, a contradiction. So the first assertion is proved. The second assertion is proved in the same way and is left to the reader.

Let $\pi(x)$ be the number of primes $\leqq x$.

Theorem 2.5 (Prime Number Theorem). *We have*

$$\pi(x) \sim \frac{x}{\log x}.$$

Proof. We have

$$\varphi(x) = \sum_{p \leqq x} \log p \leqq \sum_{p \leqq x} \log x = \pi(x) \log x;$$

and given $\epsilon > 0$,

$$\varphi(x) \geqq \sum_{x^{1-\epsilon} \leqq p \leqq x} \log p \geqq \sum_{x^{1-\epsilon} \leqq p \leqq x} (1 - \epsilon) \log x$$
$$= (1 - \epsilon) \log x[\pi(x) + O(x^{1-\epsilon})].$$

Using Theorem 2.4 that $\varphi(x) \sim x$ concludes the proof of the prime number theorem.

XVI, §2. EXERCISE

1. Prove the lemma allowing you to differentiate under the integral sign in as great a generality as you can, but including at least the case used in the case of the Laplace transforms used before Lemma 2.2.

XVI, §3. PROOF OF THE MAIN LEMMA

We recall the main lemma.

Let f be bounded, piecewise continuous on the reals $\geqq 0$. Let

$$g(z) = \int_0^\infty f(t)e^{-zt}\,dt \qquad \textit{for } \operatorname{Re}(z) > 0.$$

If g extends to an analytic function for $\operatorname{Re}(z) \geqq 0$, then

$$\int_0^\infty f(t)\,dt \textit{ exists and is equal to } g(0).$$

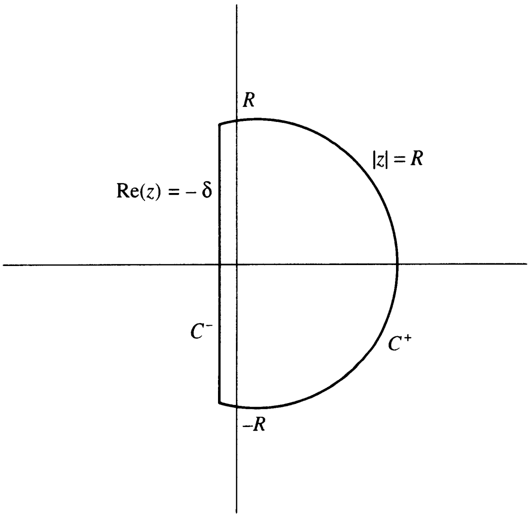
Proof. For $T > 0$ define

$$g_T(z) = \int_0^T f(t)e^{-zt}\,dt.$$

Then g_T is an entire function, as follows at once by differentiating under the integral sign. We have to show that

$$\lim_{T\to\infty} g_T(0) = g(0).$$

Let $\delta > 0$ and let C be the path consisting of the line segment $\mathrm{Re}(z) = -\delta$ and the arc of circle $|z| = R$ and $\mathrm{Re}(z) \geqq -\delta$, as shown on the figure.



By our assumption that g extends to an analytic function for $\mathrm{Re}(z) \geqq 0$, we can take δ small enough so that g is analytic on the region bounded by C, and on its boundary. Then

$$g(0) - g_T(0) = \frac{1}{2\pi i}\int_C (g(z) - g_T(z))e^{Tz}\left(1 + \frac{z^2}{R^2}\right)\frac{dz}{z} = \frac{1}{2\pi i}\int_C H_T(z)\,dz,$$

where $H_T(z)$ abbreviates the expression under the integral sign. Let B be a bound for f, that is $|f(t)| \leqq B$ for all $t \geqq 0$.

Let C^+ be the semicircle $|z| = R$ and $\mathrm{Re}(z) \geqq 0$. Then

$$\left|\frac{1}{2\pi i}\int_{C^+} H_T(z)\,dz\right| \leqq \frac{2B}{R}. \tag{1}$$

Proof. First note that for $\operatorname{Re}(z) > 0$ we have

$$|g(z) - g_T(z)| = \left| \int_T^\infty f(t)e^{-zt}\,dt \right| \leqq B \int_T^\infty |e^{-zt}|\,dt$$
$$= \frac{B}{\operatorname{Re}(z)} e^{-\operatorname{Re}(z)T};$$

and for $|z| = R$,

$$\left| e^{Tz}\left(1 + \frac{z^2}{R^2}\right)\frac{1}{z} \right| = e^{\operatorname{Re}(z)T}\left| \frac{R}{z} + \frac{z}{R} \right| \frac{1}{R} = e^{\operatorname{Re}(z)T} \frac{2|\operatorname{Re}(z)|}{R^2}.$$

Taking the product of the last two estimates and multiplying by the length of the semicircle gives a bound for the integral over the semicircle, and proves the claim.

Let C^- be the part of the path C with $\operatorname{Re}(z) < 0$. We wish to estimate

$$\frac{1}{2\pi i} \int_{C^-} (g(z) - g_T(z))e^{Tz}\left(1 + \frac{z^2}{R^2}\right)\frac{dz}{z}.$$

Now we estimate separately the expression under the integral with g and g_T.

We have

(2) $$\left| \frac{1}{2\pi i} \int_{C^-} g_T(z)e^{Tz}\left(1 + \frac{z^2}{R^2}\right)\frac{dz}{z} \right| \leqq \frac{B}{R}.$$

Proof. Let S^- be the semicircle with $|z| = R$ and $\operatorname{Re}(z) < 0$. Since g_T is entire, we can replace C^- by S^- in the integral without changing the value of the integral, because the integrand has no pole to the left of the y-axis. Now we estimate the expression under the integral sign on S^-. We have

$$|g_T(z)| = \left| \int_0^T f(t)e^{-zt}\,dt \right| \leqq B \int_0^T e^{-\operatorname{Re}(z)t}\,dt$$
$$\leqq \frac{Be^{-\operatorname{Re}(z)T}}{-\operatorname{Re}(z)}.$$

For the other factor we use the same estimate as previously. We take the product of the two estimates, and multiply by the length of the semicircle to give the desired bound in (2).

Third, we claim that

$$\int_{C^-} g(z)e^{Tz}\left(1+\frac{z^2}{R^2}\right)\frac{dz}{z} \to 0 \quad \text{as} \quad T \to \infty. \tag{3}$$

Proof. We can write the expression under the integral sign as

$$g(z)e^{Tz}\left(1+\frac{z^2}{R^2}\right)\frac{1}{z} = h(z)e^{Tz} \qquad \text{where } h(z) \text{ is independent of } T.$$

Given any compact subset K of the region defined by $\mathrm{Re}(z) < 0$, we note that

$$e^{Tz} \to 0 \quad \text{rapidly uniformly for } z \in K, \text{ as } T \to \infty.$$

The word "rapidly" means that the expression divided by any power T^N also tends to 0 uniformly for z in K, as $T \to \infty$. From this our claim (3) follows easily. (Put the details as an exercise.)

We may now prove the main lemma. We have

$$\int_0^\infty f(t)\,dt = \lim_{T\to\infty} g_T(0) \quad \text{if this limit exists.}$$

But given ϵ, pick R so large that $2B/R < \epsilon$. Then by (3), pick T so large that

$$\left|\int_{C^-} g(z)e^{Tz}\left(1+\frac{z^2}{R^2}\right)\frac{dz}{z}\right| < \epsilon.$$

Then by (1), (2), and (3) we get $|g(0) - g_T(0)| < 3\epsilon$. This proves the main lemma, and also concludes the proof of the prime number theorem.

Appendix

The purpose of this appendix is to collect miscellaneous topics which are relevant to the theory of complex analysis, but which were not formally treated in the text for various reasons.

We first start by working out systematically some topics which were assigned as exercises. I think it is better for students to have attempted to solve some non-routine problems. If they have done so unsuccessfully, then they are much more ready to appreciate how the problems are worked out, and they will remember the method of proof better for having tried and possibly failed. The first four sections are basically of this type, and are related to linear algebra.

The last two sections deal with the extension of Cauchy's formula to C^∞ functions, and presents a topic usually omitted from the course entirely, but I think it provides a nice mixture of real and complex analysis which I want to make available for independent reading.

APP., §1. SUMMATION BY PARTS AND NON-ABSOLUTE CONVERGENCE

Although non-absolute convergence is more delicate and peripheral to this course, we shall nevertheless give a brief discussion of one of its most important aspects. We first state a formula.

Lemma 1.1 (Summation by Parts). *Let $\{a_k\}$, $\{b_k\}$ $(k = 0, 1, \dots)$ be two sequences of complex numbers. Let the partial sums for $\{b_k\}$ be*

$$B_n = b_0 + \cdots + b_n = \sum_{k=0}^{n} b_k.$$

Then

$$\sum_{k=0}^{n} a_k b_k = a_n B_n - \sum_{k=0}^{n-1} B_k(a_{k+1} - a_k).$$

This is similar to integration by parts: $\int u\,dv = uv - \int v\,du$. We leave the proof to the reader; it is very easy. We apply summation by parts in the following case.

Proposition 1.2. *Let* $\{a_k\}$ $(k = 0, 1, \ldots)$ *be a monotone decreasing sequence of real numbers whose limit is* 0.

(a) *Let* $\{b_k\}$ *be a sequence of numbers such that the partial sums* B_n *are bounded. Then*

$$\sum_{k=0}^{\infty} a_k b_k$$

converges.

(b) *Let* $\{a_k\}$ *be a monotone decreasing sequence of real functions on a set, converging uniformly to* 0. *Let* $\{f_k\}$ *be a sequence of complex functions on the set, and let*

$$F_n(z) = \sum_{k=0}^{n} f_k(z)$$

be the partial sums of the series $\sum f_k$. *Assume that the partial sums are uniformly bounded, i.e. there exists* $M > 0$ *such that for all* n *we have* $|F_n(z)| \leqq M$ *for all* z. *Then the series* $\sum a_k f_k$ *converges uniformly.*

Proof. Let

$$S_n = \sum_{k=0}^{n} a_k b_k$$

be the partial sum of the series $\sum a_k b_k$. We have to estimate $|S_n - S_m|$ for m, n large. Let $|B_n| \leqq M$ for all n. Then

$$|B_n - B_m| \leqq 2M \qquad \text{for all} \quad m, n.$$

By Lemma 1.1, for $n > m$,

$$\begin{aligned} S_n - S_m = \sum_{k=m+1}^{n} a_k b_k &= a_n(B_n - B_m) - \sum_{k=m+1}^{n} (B_k - B_m)(a_{k+1} - a_k) \\ &= a_n(B_n - B_m) + \sum_{k=m+1}^{n-1} (B_k - B_m)(a_k - a_{k+1}). \end{aligned}$$

By convention, the sum on the right side is 0 if $n = m + 1$. Then we get

$$|S_n - S_m| \leqq a_n 2M + \sum_{k=m+1}^{n-1} 2M(a_k - a_{k+1}) \quad \text{(because } a_k - a_{k+1} \geqq 0\text{)}$$
$$\leqq a_n 2M + 2M(a_{m+1} - a_n) = 2Ma_{m+1}.$$

Given ε, we can select n_0 such that $a_m < \varepsilon/2M$ for $m \geqq n_0$ because of the assumption on the sequence $\{a_k\}$. Then the right side is $< \epsilon$, thus concluding the proof for the series of numbers.

For part (b), concerning the series of functions, we merely replace b_k by $f_k(z)$ for each $z \in S$. Then the above estimate gives the uniform estimate for the partial sums $|S_n(z) - S_m(z)|$, thus concluding the proof of the theorem.

The next proposition is a variation.

Proposition 1.3. *Let $\{a_k\}$ be a sequence of non-negative real numbers, monotone decreasing (not necessarily to 0). Let $\{b_k\}$ be a sequence of complex numbers such that $\sum b_k$ converges. Then $\sum a_k b_k$ converges.*

Proof. This proposition is a corollary of the preceding one, as follows. We let $a = \lim a_k$, and $a'_k = a_k - a$. Then $\{a'_k\}$ is a sequence which decreases monotonically to 0. But

$$\sum a_k b_k = \sum (a_k - a)b_k + \sum ab_k = \sum a'_k b_k + a \sum b_k.$$

Thus if $\sum b_k$ converges, we can apply Proposition 1.2 to conclude the proof.

From the same estimate applied uniformly, we obtain:

Proposition 1.4. *Let $\{a_k\}$ be a monotone decreasing sequence of bounded non-negative real functions on some set. Let $\{f_k\}$ be a sequence of complex functions on the set such that $\sum f_k$ converges uniformly. Then $\sum a_k f_k$ converges uniformly.*

As an example of summation by parts, we work out one of the exercises in the text, namely **Abel's theorem**.

Theorem 1.5. *Let $\{a_n\}$ be a sequence of complex numbers such that $\sum a_n$ converges. Assume that the power series $\sum a_n z^n$ has radius of convergence at least 1. Let $f(x) = \sum a_n x^n$ for $0 \leqq x < 1$. Then*

$$\lim_{x \to 1} f(x) = \sum a_n.$$

Proof. Let $A = \sum_{k=1}^{\infty} a_k$, $A_n = \sum_{k=1}^{n} a_k$. Consider the partial sums

$$s_n(x) = \sum_{k=1}^{n} a_k x^k.$$

We first prove that the sequence of partial sums $\{s_n(x)\}$ converges uniformly for $0 \leqq x \leqq 1$. For $m < n$ we have

$$(1) \qquad s_n(x) - s_m(x) = \sum_{k=m+1}^{n} x^k a_k = x^n(A_n - A_{m+1}) + \sum_{k=m+1}^{n-1} (A_k - A_{m+1})(x^k - x^{k+1}).$$

There exists N such that for k, $m \geqq N$ we have $|A_k - A_{m+1}| \leqq \epsilon$. Hence for $0 \leqq x \leqq 1$,

$$(2) \qquad \begin{aligned} |s_n(x) - s_m(x)| &\leqq \epsilon + \epsilon \sum_{k=m+1}^{n-1} (x^k - x^{k+1}) \\ &= \epsilon + \epsilon(x^{m+1} - x^n) \\ &\leqq 2\epsilon. \end{aligned}$$

This proves the uniformity of the convergence of $\{s_n(x)\}$.

Now given ϵ, pick N as above. Choose δ (depending on N) such that if $|x - 1| < \delta$, then

$$(3) \qquad |s_N(x) - A_N| < \epsilon.$$

By (1), (2), (3), we find that

$$\begin{aligned} |f(x) - A| &\leqq |f(x) - s_n(x)| + |s_n(x) - s_N(x)| + |s_N(x) - A_N| + |A_N - A| \\ &\leqq |f(x) - s_n(x)| + 5\epsilon, \qquad \text{for all} \quad n \geqq N \text{ and } |x - 1| < \delta. \end{aligned}$$

For a given x, pick n so large (depending on x!) so that the first term is also $< \epsilon$, to conclude the proof.

We conclude this section with a theorem on Dirichlet series.

We shall consider series

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s},$$

where $\{a_n\}$ is a sequence of complex numbers, and s is a complex variable. We write $s = \sigma + it$ with σ, t real.

Theorem 1.6. *If the Dirichlet series $\sum a_n/n^s$ converges for some $s = s_0$, then it converges for any s with $\mathrm{Re}(s) > \sigma_0 = \mathrm{Re}(s_0)$, uniformly on any compact subset of this region.*

Proof. Write $n^s = n^{s_0} n^{(s-s_0)}$, and sum the following series by parts:

$$\sum \frac{a_n}{n^{s_0}} \frac{1}{n^{(s-s_0)}}.$$

If $P_n(s_0) = \sum_{m=1}^{n} a_m/m^{s_0}$, then the tail ends of this Dirichlet series are given for $n > m$ by

$$\sum_{k=m+1}^{n} \frac{a_k}{k^{s_0}} \frac{1}{k^{s-s_0}} = \frac{P_n(s_0)}{n^{s-s_0}} + \sum_{k=m+1}^{n-1} P_k(s_0)\left[\frac{1}{k^{s-s_0}} - \frac{1}{(k+1)^{s-s_0}}\right] - \frac{P_m(s_0)}{(m+1)^{s-s_0}}.$$

We have

$$\frac{1}{k^{s-s_0}} - \frac{1}{(k+1)^{s-s_0}} = (s-s_0)\int_k^{k+1} \frac{1}{x^{s-s_0+1}}\,dx,$$

which we estimate easily in absolute value. If $\delta > 0$ and $\mathrm{Re}(s) \geqq \sigma_0 + \delta$, then we conclude that our tail end is small uniformly if $|s - s_0|$ is bounded. This proves the theorem.

Assuming that the Dirichlet series converges for some s, if σ_0 is the smallest real number such that the series converges for $\mathrm{Re}(s) > \sigma_0$, then we call σ_0 the **abscissa of convergence,** and we see that the series converges in the half plane to the right of the line $\sigma = \sigma_0$, but does not converge for any s with $\sigma < \sigma_0$.

If the Dirichlet series converges for $s_1 = \sigma_1 + it_1$, then we must have

$$a_n = O(n^{\sigma_1})$$

because the n-th term of the series a_n/n^{s_1} tends to 0. It follows in particular that the Dirichlet series converges absolutely and uniformly on compacts for

$$\mathrm{Re}(s) \geqq \sigma_1 + 1 + \delta,$$

if $\delta > 0$. This is immediately seen by comparison with the series $\sum 1/n^{1+\delta}$.

APP., §2. DIFFERENCE EQUATIONS

Let $u_1, \ldots, u_d$ be given complex numbers. We want to determine all solutions $(a_0, a_1, \ldots)$ of the equation

$$(*) \qquad a_n = u_1 a_{n-1} + \cdots + u_d a_{n-d} \qquad \text{for all } n \geqq d.$$

Such solutions are therefore infinite vectors

$$(a_0, \ldots, a_{d-1}, a_d, \ldots, a_n, \ldots),$$

and they form a vector space, this being immediately verified directly from the definitions. Observe that a given d-tuple $(a_0, \ldots, a_{d-1})$ determines uniquely the infinite vector which is a solution of $(*)$ and projects on this d-tuple. Let S be the space of solutions and let S_d be the d-dimensional space of the first d coordinates. Then the projection

$$S \to S_d$$

is injective, and it is obviously surjective, so S is a vector space of dimension d. We want to find a basis for S, or even better, given a d-tuple $(a_0, \ldots, a_{d-1})$ we want to express the corresponding solution in terms of this d-tuple and $u_1, \ldots, u_d$.

Let

$$P(X) = X^d - (u_1 X^{d-1} + \cdots + u_d).$$

We call $P(X)$ the **characteristic polynomial** of the equations $(*)$. Let α be a root of $P(X)$. Then it is immediately verified that putting $a_n = \alpha^n$ for $n \geqq d$ gives a solution of $(*)$. Thus we have found a solution for each root of the characteristic polynomial. Note that if $\alpha = 0$, then by convention $0^0 = 1$ and $0^n = 0$ for $n \geqq 1$, so the solution with $n \geqq 0$ is

$$(1, 0, 0, \ldots).$$

Note that $\alpha = 0$ implies $u_d = 0$.

Also we note that the solution $a_n = \alpha^n$ for $n \geqq d$ has the beginning segment

$$(a_0, \ldots, a_{d-1}) = (1, \alpha, \ldots, \alpha^{d-1}),$$

which determines a_n uniquely for $n \geqq d$ by the difference equations $(*)$. In particular, for this solution we have $a_n = \alpha^n$ for all $n \geqq 0$.

One knows from elementary algebra that P has at most d distinct roots.

Theorem 2.1. *Suppose that $P(X)$ has d distinct roots $\alpha_1, \ldots, \alpha_d$. Then: all solutions of $(*)$ are of the form*

$$a_n = b_1 \alpha_1^n + \cdots + b_d \alpha_d^n$$

with arbitrary numbers $b_1, \ldots, b_d$.

Proof. It will suffice to prove that the solutions $(\alpha_1^n), \ldots, (\alpha_d^n)$ are linearly independent, because then they form a basis for the space S of solutions.

Suppose there is a relation of linear dependence, that is

$$b_1\alpha_1^n + \cdots + b_d\alpha_d^n = 0 \qquad \text{for all} \quad n \geqq 0.$$

Looking at the first d equations (with $n = 0, \ldots, d-1$) we see that the determinant of the coefficients is a Vandermonde determinant, which is

$$V = \prod_{i<j} (\alpha_j - \alpha_i) \neq 0$$

by the hypothesis that the roots of the characteristic polynomial are distinct. Hence $b_i = 0$ for all i, thus proving the theorem.

From Theorem 2.1 one can solve explicitly for $b_1, \ldots, b_d$ when $a_0, \ldots, a_{d-1}$ are given. Indeed, we have the system of linear equations

$$\begin{aligned} a_0 &= b_1 + \cdots + b_d, \\ a_1 &= b_1\alpha_1 + \cdots + b_d\alpha_d, \\ &\cdots\cdots\cdot \\ a_{d-1} &= b_1\alpha_1^{d-1} + \cdots + b_d\alpha_d^{d-1}. \end{aligned}$$

Let A be the column vector of the left-hand side, and let $A^1, \ldots, A^d$ be the column vectors of the coefficient matrix on the right-hand side

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_d \\ \vdots & \vdots & \cdots & \vdots \\ \alpha_1^{d-1} & \alpha_2^{d-1} & \cdots & \alpha_d^{d-1} \end{pmatrix}.$$

Then

$$D(A^1, \ldots, A^d) = \prod_{i<j} (\alpha_j - \alpha_i) = V.$$

By Cramer's rule, we obtain

$$b_i = D(A^1, \ldots, A, \ldots, A^d)/V,$$

where the column vector A occurs in the i-th place.

As an application of Theorem 2.1, consider the power series

$$F(T) = \sum_{n=0}^{\infty} a_n T^n.$$

Theorem 2.2. *Assume that $\alpha_1, \ldots, \alpha_d$ are distinct. Then $F(T)$ is a rational function.*

Proof. We have

$$F(T) = b_1 \sum_{n=0}^{\infty} (\alpha_1 T)^n + \cdots + b_d \sum_{n=0}^{\infty} (\alpha_d T)^d$$
$$= \frac{b_1}{1 - \alpha_1 T} + \cdots + \frac{b_d}{1 - \alpha_d T}.$$

This gives the partial fraction decomposition of $F(T)$, showing also that $F(T)$ is a rational function.

We shall now indicate another approach, which we carry out when $d = 2$. In this case, the characteristic polynomial is

$$P(X) = X^2 - u_1 X - u_2 = (X - \alpha_1)(X - \alpha_2).$$

Observe that one then has the factorization

$$1 - u_1 T - u_2 T^2 = (1 - \alpha_1 T)(1 - \alpha_2 T),$$

which comes by putting $X = 1/T$. The rational function for $F(T)$ can then be written down explicitly, as the following proof will show. We have

$$\begin{aligned} F(T) &= a_0 + a_1 T + \sum_{n=2}^{\infty} a_n T^n \\ &= a_0 + a_1 T + \sum_{n=2}^{\infty} (u_1 a_{n-1} + u_2 a_{n-2}) T^n \\ &= a_0 + a_1 T + u_1 T \sum_{n=2}^{\infty} a_{n-1} T^{n-1} + u_2 T^2 \sum_{n=2}^{\infty} a_{n-2} T^{n-2} \\ &= a_0 + a_1 T + u_1 T\big(F(T) - a_0\big) + u_2 T^2 F(T) \\ &= a_0 + (a_1 - a_0 u_1) T + u_1 T F(T) + u_2 T^2 F(T), \end{aligned}$$

from which we get

$$F(T)(1 - u_1 T - u_2 T^2) = a_0 + (a_1 - a_0 u_1) T$$

and so $F(T)$ is the rational function

$$(1) \qquad F(T) = \frac{a_0 + (a_1 - a_0 u_1) T}{(1 - u_1 T - u_2 T^2)} = \frac{a_0 + (a_1 - a_0 u_1) T}{(1 - \alpha_1 T)(1 - \alpha_2 T)}.$$

One can then write down the partial fraction decomposition if $\alpha_1 \neq \alpha_2$, namely

$$\frac{1}{(1-\alpha_1 T)(1-\alpha_2 T)} = \frac{b}{1-\alpha_1 T} + \frac{c}{1-\alpha_2 T}. \tag{2}$$

We solve for b and c from the equations:

$$b + c = 1,$$
$$-\alpha_2 b - \alpha_1 c = 0.$$

Then

$$b = \frac{\alpha_1}{\alpha_1 - \alpha_2} \quad \text{and} \quad c = \frac{\alpha_2}{\alpha_2 - \alpha_1}. \tag{3}$$

We can invert the power series $1 - \alpha_1 T$ and $1 - \alpha_2 T$ by the geometric series:

$$\frac{1}{1-\alpha_1 T} = \sum_{n=0}^{\infty} \alpha_1^n T^n \quad \text{and} \quad \frac{1}{1-\alpha_2 T} = \sum_{n=0}^{\infty} \alpha_2^n T^n.$$

Then it is clear from (1), (2), (3) that there exist numbers b_1, b_2 explicitly given in terms of u_1, u_2, α_1, α_2 such that

$$a_n = b_1\alpha_1^n + b_2\alpha_2^n.$$

Observe that we can also determine b_1, b_2 in a simple way by using these equations for $n = 0$ and $n = 1$, in which case we get:

$$a_0 = b_1 + b_2,$$
$$a_1 = b_1\alpha_1 + b_2\alpha_2.$$

Solving by the high school method yields:

$$b_1 = (\alpha_2 a_0 - a_1)/(\alpha_2 - \alpha_1) \quad \text{and} \quad b_2 = (\alpha_1 a_0 - a_1)/(\alpha_1 - \alpha_2).$$

When $d = 2$, the Vandermonde determinant collapes to one simple factor.

APP., §3. ANALYTIC DIFFERENTIAL EQUATIONS

In this appendix we work out Exercises 6 and 7 of Chapter II, §6.

Theorem 3.1. *Let p be an integer $\geqq 2$. Let $g_1, \ldots, g_{p-1}$ be power series with complex coefficients. Let $a_0, \ldots, a_{p-1}$ be given complex numbers. Then there exists a unique power series $f(T) = \sum a_n T^n$ such that*

$$D^p f = g_{p-1}D^{p-1}f + \cdots + g_1 Df,$$

where D is the derivative. If $g_1, \ldots, g_{p-1}$ converge in a neighborhood of the origin, then so does f.

Proof. The idea is that the coefficient a_n of f for $n \geqq p$ can be determined inductively from the previous coefficients $a_0, \ldots, a_{n-1}$, thus giving the formal power series solution. Then one has to estimate to show the convergence. We carry this out in detail. We let $a_p, a_{p+1}, \ldots,$ be unknown coefficients. Then observe that

$$D^p f(T) = \sum_{n=p}^{\infty} n(n-1)\ldots(n-p+1)a_n T^{n-p},$$

and therefore putting $m = n - p$, we get

$$(1) \qquad D^p f(T) = \sum_{m=0}^{\infty} (m+p)(m+p-1)\ldots(m+1)a_{m+p}T^m.$$

Similarly for every positive integer s with $1 \leqq s \leqq p-1$ we have

$$D^s f(T) = \sum_{k=0}^{\infty} (k+s)(k+s-1)\ldots(k+1)a_{k+s}T^k.$$

It will be useful to use the notation $[k, s] = (k+s)(k+s-1)\ldots(k+1)$.

Next we write down the power series for each g_s, say

$$g_s = \sum_{j=0}^{\infty} b_{s,j}T^j.$$

Then

$$(2) \qquad g_s D^s f(T) = \sum c_{s,m}T^m \qquad \text{where} \qquad c_{s,m} = \sum_{k+j=m} [k, s]a_{k+s}b_{s,j}.$$

Hence once we are given $a_0, \ldots, a_{p-1}$ we can solve inductively for a_m in terms of $a_0, \ldots, a_{m-1}$ and the coefficients of $g_1, \ldots, g_{p-1}$ by the formula

$$(3) \qquad a_{m+p} = \frac{c_{1,m} + \cdots + c_{p-1,m}}{[m, p]} = \sum_{s=1}^{p-1} \sum_{k+j=m} \frac{[k, s]}{[m, p]} a_{k+s}b_{s,j},$$

which determines a_{m+p} uniquely in terms of $a_0, \ldots, a_{m+p-1}, b_{s,1}, \ldots, b_{s,m}$. Hence we have proved that there is a unique power series satisfying the given differential equation.

Assuming that the power series $g_1, \ldots, g_{p-1}$ converge, we must now prove that $f(T)$ converges. We select a positive number K sufficiently large $\geqq 2$ and a positive number B such that

$$|a_0|, \ldots, |a_{p-1}| \leqq K \text{ and } |b_{s,j}| \leqq KB^j \text{ for all } s = 1, \ldots, p-1 \text{ and all } j.$$

We shall prove by induction that for $m \geqq 0$ we have

$$|a_{m+p}| \leqq (p-1)^{m+1}K^2K^mB^m. \tag{4}$$

The standard m-th root test for convergence then shows that $f(T)$ converges.

We note that the expressions (2) for $c_{s,m}$ and hence (3) for a_{m+p} have positive coefficients as polynomials in $a_0, a_1, \ldots$ and the coefficients $b_{s,j}$ of the power series g_s. Hence to make our estimates, we may avoid writing down absolute values by replacing $b_{s,j}$ by KB^j, and we may replace $a_0, \ldots, a_{p-1}$ by K. Then all the values a_{m+p} ($m \geqq 0$) are positive, and we want to show that they satisfy the desired bound (4).

Observe first that for $0 \leqq k \leqq m$ and $1 \leqq s \leqq p-1$ we always have

$$\frac{[k, s]}{[m, p]} \leqq 1.$$

Hence the fraction $[k, s]/[m, p]$ will be replaced by 1 in the following estimates.

Now first we estimate a_{m+p} with $m = 0$. Then $k + j = 0$, so $k = j = 0$, and

$$a_p \leqq \sum_{s=1}^{p-1} \sum_{k+j=0} a_{k+s}b_{s,j} \leqq (p-1)K^2$$

as desired. Suppose by induction that we have proved (4) for all integers $\geqq 0$ and $< n$. Then,

$$\begin{aligned} a_{n+p} \leqq \sum_{s=1}^{p-1} \sum_{k+j=n} a_{k+s}b_j &\leqq (p-1) \sum_{k+j=n} (p-1)^n K^2 K^{k-1} B^k B^j \\ &\leqq (p-1)^{n+1} \sum_{k=0}^{n} K^{k+1}B^n \\ &\leqq (p-1)^{n+1}K^{n+2}B^n, \end{aligned}$$

which is the desired estimate. We have used the elementary inequality

$$\sum_{k=0}^{n} K^k = \frac{K^{n+1}-1}{K-1} \leqq K^{n+1},$$

which is trivial.

Theorem 3.2. *Let $g(T)$ be a power series. Then there is a unique power series $f(T)$ such that $f(T) = a_1 T + \cdots$ satisfying the differential equation*

$$f'(T) = g(f(T)).$$

If g is convergent, so is f.

Proof. Let $g(T) = \sum b_k T^k$ and write $f(T)$ with unknown coefficients

$$f(T) = \sum_{m=1}^{\infty} a_m T^m.$$

Then $f'(T) = \sum m a_m T^{m-1} = \sum (n+1) a_{n+1} T^n$. The given differential equation has the form

$$\sum_{n=0}^{\infty} (n+1) a_{n+1} T^n = b_0 + b_1 f(T) + b_2 f(T)^2 + \cdots.$$

Equating the coefficient of T^n on both sides, we see that $a_1 = b_0$, and

$$(n+1) a_{n+1} = P_n(b_0, \ldots, b_n; a_1, \ldots, a_n)$$

where P_n is a polynomial with positive integer coefficients. In particular, starting with $a_1 = b_0$, we may then solve inductively for a_{n+1} in terms of $a_1, \ldots, a_n$ for $n \geqq 1$. This proves the existence and uniqueness of the power series $f(T)$.

Assume next that $g(T)$ converges. We must prove that $f(T)$ converges. Let B_k be positive numbers such that $|b_k| \leqq B_k$, and such that the power series $G(T) = \sum B_k T^k$ converges. Let $F(T)$ be the solution of the differential equation

$$F'(T) = G(F(T)),$$

and let $F(T) = \sum A_m T^m$, with $A_1 > 0$ and $|a_1| \leqq A_1$. Then

$$(n+1) A_{n+1} = P_n(B_0, \ldots, B_n; A_1, \ldots, A_n),$$

with the same polynomial P_n. Hence $|a_{n+1}| \leqq A_{n+1}$, and if $F(T)$ converges so does $f(T)$.

Since $g(T)$ converges, there exist positive numbers K, B such that

$$|b_k| \leqq KB^k \qquad \text{for all} \quad k = 0, 1, \ldots.$$

We let $B_k = KB^k$. Then

$$G(T) = K \sum_{k=0}^{\infty} B^k T^k = \frac{K}{1 - BT},$$

and so it suffices to prove that the solution $F(T)$ of the differential equation

$$F'(T) = \frac{K}{1 - BF(T)}.$$

converges. This equation is equivalent with

$$F'(T) = K + BF(T)F'(T),$$

which we can integrate to give

$$F(T) = KT + B\frac{F(T)^2}{2}.$$

By the quadratic formula, we find

$$F(T) = \frac{1 - (1 - 2KBT)^{1/2}}{B} = KT + \cdots.$$

We then use the binomial expansion which we know converges, as worked out in the examples of Chapter II, §2 and §3. This proves Theorem 3.2.

APP., §4. FIXED POINTS OF A FRACTIONAL LINEAR TRANSFORMATION

Let M be a 2×2 non-singular complex matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad ad - bc \neq 0.$$

Let $P(t)$ be its characteristic polynomial, so $P(t) = \det(tI - M)$ and

$$P(t) = (t - \lambda)(t - \lambda')$$

where λ, λ' are the eigenvalues. *We assume that* $\lambda \neq \lambda'$ *and* $c \neq 0$. Then M is diagonalizable, but we want an explicit way of doing this. Let W, $W' \in \mathbf{C}^2$ be eigenvectors belonging to the two eigenvalues. We can write W, W' in the form

$$W = \begin{pmatrix} w \\ 1 \end{pmatrix} \quad \text{and} \quad W' = \begin{pmatrix} w' \\ 1 \end{pmatrix}.$$

Indeed, if an eigenvector had the form $W = {}^t(w, 0)$, $w \neq 0$, we could assume $w = 1$. Then directly from the relation $MW = \lambda W$ we find $c = 0$ which we have ruled out.

Let

$$S = (W, W') = \begin{pmatrix} w & w' \\ 1 & 1 \end{pmatrix}.$$

From the relation $M(W, W') = (\lambda W, \lambda' W') = (W, W')\begin{pmatrix} \lambda & 0 \\ 0 & \lambda' \end{pmatrix}$ we find

$$S^{-1}MS = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda' \end{pmatrix}.$$

Thus S is a matrix which conjugates M to a diagonal matrix.

Let $z \in \mathbf{C}$. For $z \neq -d/c$ we define

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

If $z = -d/c$ then the above expression is not defined, but we put $M(z) = \infty$. If $c = 0$ we let $M(\infty) = \infty$. If $c \neq 0$, we let $M(\infty) = a/c$.

One verifies directly that if A, B are non-singular 2×2 matrices, then

$$(AB)(z) = A\big(B(z)\big) \qquad \text{and} \qquad I(z) = z \quad (I = \text{identity matrix}).$$

By a **fixed point** of M we mean a complex number z such that $M(z) = z$.

Theorem 4.1. *Assuming $\lambda \neq \lambda'$ and $c \neq 0$, there are exactly two fixed points, namely w and w'.*

Proof. Directly from the definition, one verifies that $M(w) = w$ and $M(w') = w'$, so w and w' are fixed points. Conversely, the condition that z is a fixed point is expressible as a quadratic equation in z, which has at most two solutions by the quadratic formula. This proves Theorem 4.1.

Theorem 4.2. *Assume $|\lambda| < |\lambda'|$ and $c \neq 0$. Let $z \in \mathbf{C}$ and $z \neq w$. Then*

$$\lim_{k \to \infty} M^k(z) = w'.$$

Proof. Let $\alpha = \lambda/\lambda'$. Then $|\alpha| < 1$. We have

$$S^{-1}M^kS\begin{pmatrix} z \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda^k z \\ \lambda'^k \end{pmatrix} \qquad \text{so} \qquad S^{-1}M^kS(z) = \alpha^k z.$$

Hence $\lim_{k \to \infty} \alpha^k z = 0$ for all complex numbers z.

Next observe that since $z \neq w$, it follows that $S^{-1}(z)$ is a complex number by direct computation (i.e. $S^{-1}(z) \neq \infty$). We then apply the formula we have just derived to $S^{-1}(z)$ instead of z, and we find

$$M^k(z) = S(\alpha^k S^{-1}(z)).$$

Taking the limit yields

$$\lim_{k\to\infty} M^k(z) = S(0) = w'$$

thus proving the theorem.

APP., §5. CAUCHY'S FORMULA FOR C^∞ FUNCTIONS

Let D be an open disc in the complex numbers, and let $\bar{D}$ be the closed disc, so the boundary of $\bar{D}$ is a circle. Cauchy's formula gives us the value as an integral over the circle C:

$$f(z_0) = \frac{1}{2\pi i}\int_C \frac{f(z)\,dz}{z - z_0}$$

if f is holomorphic on $\bar{D}$, that is on some open set containing the closed disc. But what happens if f is not holomorphic but merely smooth, say its real and imaginary parts are infinitely differentiable in the real sense? There is also a formula, which unfortunately is not usually taught in basic courses, although it gives a beautiful application of several notions which arise in both real and complex analysis, and advanced calculus. We shall prove this theorem here. We shall also mention an application, which occurs in the theory of partial differential equations.

Let us write $z = x + iy$. We introduce two new derivatives. Let

$$f(z) = f_1(x, y) + if_2(x, y),$$

where $f_1 = \operatorname{Re} f$ and $f_2 = \operatorname{Im} f$ are the real and imaginary parts of f respectively. We say that f is C^∞ if f_1, f_2 are infinitely differentiable in the naive sense of functions of two real variables x and y. In other words, all partial derivatives of all orders exist and are continuous. We write $f \in C^\infty(\bar{D})$ to mean that f is C^∞ on some open set containing $\bar{D}$.

For such f we define

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

Symbolically, we put

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

The **Cauchy–Riemann equations** can be formulated neatly by saying that f is holomorphic if and only if

$$\frac{\partial f}{\partial \bar{z}} = 0.$$

See Chapter VIII, §1.

We shall need the Stokes–Green formula for a simple type of region. In advanced calculus, one integrates expressions

$$\int_C P\,dx + Q\,dy,$$

where P, Q are C^∞ functions, and C is some curve. The Stokes–Green theorem relates such integrals over a boundary to a double integral

$$\iint_B \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy$$

taken over a region B which is bounded by the curve C. The precise statement is this.

Stokes–Green Formula. *Let B be a region of the plane, bounded by a finite number of curves, oriented so that the region lies to the left of each curve. Let γ be the boundary, so oriented. Let P, Q have continuous first partial derivatives on B and its boundary. Then*

$$\int_\gamma P\,dx + Q\,dy = \iint_B \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy.$$

It is useful to express the Stokes–Green formula in terms of the derivatives $\partial/\partial z$ and $\partial/\partial \bar{z}$. Writing

$$dz = dx + i\,dy \qquad \text{and } d\bar{z} = dx - i\,dy,$$

we can solve for dx and dy in terms of dz and $d\bar{z}$, to give

$$dx = \tfrac{1}{2}(dz + d\bar{z}) \qquad \text{and} \qquad dy = \frac{1}{2i}(dz - d\bar{z}).$$

Then

$$P\,dx + Q\,dy = g\,dz + h\,d\bar{z},$$

where g, h are suitable functions. Let us write symbolically

$$dz \wedge d\bar{z} = -2i\,dx\,dy.$$

Then by substitution, we find the following version of the **Stokes–Green Formula**:

$$\boxed{\int_\gamma g\,dz + h\,d\bar{z} = \iint_B \left(\frac{\partial h}{\partial z} - \frac{\partial g}{\partial \bar{z}}\right) dz \wedge d\bar{z}.}$$

Remark. Directly from the definition of $\partial/\partial z$ and $\partial/\partial \bar{z}$ one verifies that the usual expression for df is given by

$$\frac{\partial f}{\partial z}\,dz + \frac{\partial f}{\partial \bar{z}}\,d\bar{z} = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy.$$

Consider the special case where $B = B(a)$ is the region obtained from the disc $\bar{D}$ by deleting a small disc of radius a centered at a point z_0, as shown on the figure.

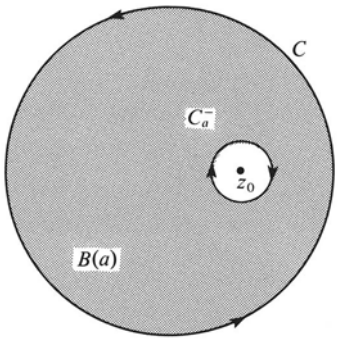


Figure 1

Then the boundary consists of two circles C and C_a^-, oriented as shown so that the region lies to the left of each curve. We have written C_a^- to indicate the circle with clockwise orientation, so that the region $B(a)$ lies to the left of C_a^-. As before, C is the circle around D, oriented counterclockwise. Then the boundary of $B(a)$ can be written

$$\gamma = C + C_a^-.$$

We shall deal with integrals

$$\iint_D \varphi(z)\frac{dz \wedge d\bar{z}}{z - z_0},$$

where $\varphi(z)$ is a smooth function, and where z_0 is some point in the interior of the disc. Such an integral is an improper integral, and is supposed to be interpreted as a limit

$$\lim_{a\to 0}\iint_{B(a)} \varphi(z)\frac{dz \wedge d\bar{z}}{z - z_0},$$

where $B(a)$ is the complement of a disc of radius a centered at z_0. The limit exists, as one sees by using polar coordinates. Letting $z = z_0 + re^{i\theta}$ with polar coordinates around the fixed point z_0, we have

$$dx\, dy = r\, dr\, d\theta$$

and $z - z_0 = re^{i\theta}$, so r cancels and we see that the limit exists, since the integral becomes simply

$$-2i\iint\frac{\varphi(z)\, dr\, d\theta}{e^{i\theta}}.$$

The region $B(a)$ is precisely of the type where we apply the Stokes–Green formula.

Cauchy's Theorem for C^∞ Functions. *Let $f \in C^\infty(\bar{D})$ and let z_0 be a point in the interior D. Let C be the circle around D. Then*

$$f(z_0) = \frac{1}{2\pi i}\int_C \frac{f(z)\, dz}{z - z_0} + \frac{1}{2\pi i}\iint_D \frac{\partial f}{\partial \bar{z}}\frac{dz \wedge d\bar{z}}{z - z_0}.$$

Proof. Let a be a small positive number, and let $B(a)$ be the region obtained by deleting from D a disc of radius a centered at z_0. Then $\partial f/\partial \bar{z}$ is C^∞ on $B(a)$ and we can apply the Stokes–Green formula to

$$\frac{f(z)\, dz}{z - z_0} = g(z)\, dz$$

over this region. Note that this expression has no term with $d\bar{z}$. Furthermore

$$\frac{\partial}{\partial \bar{z}}\left(\frac{1}{z - z_0}\right) = 0 \quad\text{and}\quad \frac{\partial g}{\partial \bar{z}} = \frac{\partial f}{\partial \bar{z}}\frac{1}{z - z_0}$$

because $1/(z - z_0)$ is holomorphic in this region. Then by Stokes–Green we find

$$\int_{C_a^-} \frac{f(z)\,dz}{z - z_0} + \int_C \frac{f(z)\,dz}{z - z_0} = -\iint_{B(a)} \frac{\partial f}{\partial \bar{z}} \frac{dz \wedge d\bar{z}}{z - z_0}.$$

The limit of the integral on the right-hand side as a approaches 0 is the double integral (with a minus sign) which occurs in Cauchy's formula. We now determine the limit of the curve integral over C_a^- on the left-hand side. We parametrize C_a (counterclockwise orientation) by

$$z = z_0 + ae^{i\theta}, \qquad 0 \leq \theta \leq 2\pi.$$

Then $dz = aie^{i\theta}\,d\theta$, so

$$\int_{C_a^-} \frac{f(z)\,dz}{z - z_0} = -\int_{C_a} \frac{f(z)\,dz}{z - z_0} = -\int_0^{2\pi} f(z_0 + ae^{i\theta})i\,d\theta.$$

Since f is continuous at z_0, we can write

$$f(z_0 + ae^{i\theta}) = f(z_0) + h(a, \theta)$$

where $h(a, \theta)$ is a function such that

$$\lim_{a \to 0} h(a, \theta) = 0$$

uniformly in θ. Therefore

$$\begin{aligned}\lim_{a \to 0} \int_{C_a^-} \frac{f(z)\,dz}{z - z_0} &= -2\pi i f(z_0) - \lim_{a \to 0} \int_0^{2\pi} h(a, \theta)i\,d\theta \\ &= -2\pi i f(z_0).\end{aligned}$$

Cauchy's formula now follows at once.

Remark 1. Suppose that f is holomorphic on D. Then

$$\frac{\partial f}{\partial \bar{z}} = 0,$$

and so the double integral disappears from the general formula to give the Cauchy formula as we encountered it previously.

Remark 2. Consider the special case when the function f is 0 on the boundary of the disc. Then the integral over the circle C is equal to 0, and we obtain the formula

$$f(z_0) = \frac{1}{2\pi i} \iint_D \frac{\partial f}{\partial \bar{z}} \frac{dz \wedge d\bar{z}}{z - z_0}.$$

This allows us to recover the values of the function from its derivative $\partial f/\partial \bar{z}$. Conversely, one has the following result.

Theorem. *Let $g \in C^\infty(\bar{D})$ be a C^∞ function on the closed disc. Then the function*

$$f(w) = \frac{1}{2\pi i} \iint_D \frac{g(z)}{z - w} dz \wedge d\bar{z}$$

is defined and C^∞ on D, and satisfies

$$\frac{\partial f}{\partial \bar{w}} = g(w) \qquad \textit{for} \quad w \in D.$$

The proof is essentially a corollary of Cauchy's theorem if one has the appropriate technique for differentiating under the integral sign. However, we have now reached the boundary of this course, and we omit the proof.

APP., §6. CAUCHY'S THEOREM FOR LOCALLY INTEGRABLE VECTOR FIELDS

The main part of Chapter IV is not really a theorem about complex functions, but consists of a purely topological theorem as in Chapter IV, §3, and an application to locally integrable vector fields on an open set of the plane $\mathbf{R}^2$. We now indicate briefly this more general context in a precise way.

Let U be a connected open subset of $\mathbf{R}^2$, and let $F = (f_1, f_2)$ be a continuous real valued vector field on U. If F is of class C^1, i.e. its coordinate functions f_1, f_2 have continuous partial derivatives, and if

$$D_2 f_1 = D_1 f_2,$$

then it is proved in standard courses that F has a potential g on every disc D contained in U, that is a function g such that $\operatorname{grad} g = F$. Cf. for instance my *Undergraduate Analysis*, Springer-Verlag (1983), Chapter XV, Theorem 3.3.

Now in general, for a continuous vector field F, we define F to be **locally integrable** on U if for each point P of U there is a disc D_P

contained in U such that F has a potential on D_P. We let F denote such a vector field for the rest of this section.

Let $\gamma_1 : [a_1, a_2] \to U$ be a continuous curve, whose image is contained in a disc $D \subset U$ such that F has a potential on D. Let $P_i = \gamma_i(a_i)$, so P_1 is the beginning point and P_2 the end point of the curve. We define

$$\int_{\gamma_1} F = g(P_2) - g(P_1).$$

Since a potential is uniquely determined up to an additive constant (on a connected open set, and D in particular), it follows that the value $g(P_2) - g(P_1)$ is independent of the choice of potential g for F on the disc. If the curve γ_1 happens to be C^1, then the above value coincides with the value of the integral computed according to the usual chain rule.

Suppose that γ: $[a, b] \to U$ is an arbitrary continuous curve. We can then define the integral

$$\int_\gamma F$$

just as we defined the integral of a holomorphic function. We can find a partition of $[a, b]$, and a sequence of discs $D_0, \ldots, D_n$ connected by the curve along the partition such that each $D_i \subset U$ and such that F has a potential g_i on D_i. (One uses the uniform continuity of F and the compactness of the image of γ to satisfy these conditions.) The integral is defined to be

$$\int_\gamma F = \sum [g_i(P_{i+1}) - g_i(P_i)].$$

Exactly the same arguments as those of the text in Chapter III, §4 show that this integral is independent of the choices made (partition, discs, potentials) subject to the above conditions.

We can define two continuous curves γ, η from P to Q in U to be **close together** (with respect to F) if there is a sequence of discs as above such that

$$\gamma([a_i, a_{i+1}]) \subset D_i \quad \text{and} \quad \eta([a_i, a_{i+1}]) \subset D_i.$$

The proofs of Theorems 5.1 and 5.2 of Chapter III are then valid in the present context. In particular, if two curves are close together with respect to F, then the integrals of F along these curves are equal.

Furthermore, the homotopy form of Cauchy's theorem is also valid in the present context, replacing the holomorphic function f by the locally integrable vector field F. In particular:

Theorem 6.1. *Let U be simply connected (for instance a disc or a rectangle), and let F be a locally integrable vector field on U. Then F has a potential on U.*

Proof. This comes directly from the homotopy form of Cauchy's theorem. The potential is defined by the integral

$$g(X) = \int_{P_0}^{X} F,$$

taken from a fixed point P_0 in U to a variable point X. The integral may be taken along any continu(curve in U, because the integral is independent of the path between P_0 and X. Thus the analogue of Chapter III, Theorem 6.1 is valid.

In particular, in the definition of an integral of F, one need not specify that F have a potential on each D_i. It suffices that the discs D_i be contained in U.

We may now apply the considerations of Chapter IV, concerning the winding number and homology. Instead of $d\zeta/\zeta$, we use the vector field

$$G(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}\right).$$

By the chain rule, using $r^2 = x^2 + y^2$, $x = r \cos \theta$, $y = r \sin \theta$, one sees that

$$\frac{-y}{x^2 + y^2}\, dx + \frac{x}{x^2 + y^2}\, dy = d\theta.$$

For (x, y) in a disc not containing the origin, G has a potential, which is just θ, plus a constant of integration.

Let U be a connected open set in $\mathbf{R}^2$. For each $P \in \mathbf{R}^2$, $P = (x_0, y_0)$, we may consider the translation of G to P, that is the vector field

$$G_P(x, y) = G_P(X) = G(X - P) = G(x - x_0, y - y_0).$$

Then G_P is locally integrable on every open set not containing P.

Let γ: $[a, b] \to U$ be a closed curve in U. For every P not on γ, the integral

$$\int_\gamma G_P$$

is defined, and from the definition of the integral with discs connected along a partition, we see that the value of this integral is equal to $2\pi k$, for some integer k. We define the **winding number** $W(\gamma, P)$ to be

$$W(\gamma, P) = \frac{1}{2\pi} \int_{\gamma} G_P.$$

We are now in the same position we were for the complex integral, allowing us to define a curve γ **homologous to** 0 in U if $W(\gamma, P) = 0$ for all points P in $\mathbf{R}^2$, $P \notin U$.

The same arguments of Chapter IV, §3, replacing f by F, yield:

Theorem 6.2. *Let U be a connected open set in $\mathbf{R}^2$, and let γ be a closed chain in U. Then γ is homologous to a rectangular chain. If γ is homologous to 0 in U, and F is a locally integrable vector field on U, then*

$$\int_{\gamma} F = 0.$$

Proof. Theorem 3.2 of Chapter IV applies verbatim to the present situation, and we know from Theorem 6.1 that the integral of F around a rectangle is 0, so the theorem is proved.

We also have the analogue of Theorem 2.4 of Chapter IV.

Theorem 6.3. *Let U be an open set and γ a closed chain in U such that γ is homologous to 0 in U. Let $P_1, \ldots, P_n$ be a finite number of distinct points of U. Let γ_k $(k = 1, \ldots, n)$ be the boundary of a closed disc $\bar{D}_k$ contained in U centered at P_k and oriented counterclockwise. We assume that $\bar{D}_k$ does not intersect $\bar{D}_j$ if $k \neq j$. Let*

$$m_k = W(\gamma, P_k).$$

Let U^* be the set obtained by deleting $P_1, \ldots, P_n$ from U. Then γ is homologous to $\sum m_k \gamma_k$ on U^*.

Furthermore, if F is a locally integrable vector field on U^*, then

$$\int_{\gamma} F = \sum m_k \int_{\gamma_k} F.$$

From the above theorem, we also obtain:

Theorem 6.4. *Let U be simply connected, and let $P_1, \ldots, P_n$ be distinct points of U. Let U^* be the open set obtained from U by deleting these points. Let F be a locally integrable vector field on U^*. Let*

$$a_k = \frac{1}{2\pi} \int_{\gamma_k} F,$$

where γ_k is a small circle around P_k, not containing P_j if $j \neq k$. Then the vector field

$$F - \sum a_k G_{P_k}$$

has a potential on U.

Proof. One verifies directly that

$$\int_{\gamma_k} G_{P_j} = \begin{cases} 2\pi & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases}$$

Let γ be a closed curve in U^*. Then immediately from the definition of γ_k and Theorem 6.3, it follows that if we put

$$\eta = \gamma - \sum m_k \gamma_k,$$

then η is homologous to 0 in U^*. Therefore by Theorem 6.3, the integral of F over η is 0. By the definitions, this implies that

$$\int_\gamma \left[F - \sum a_k G_{P_k}\right] = 0.$$

Thus the integral of $F - \sum a_k G_{P_k}$ over any closed curve in U^* is 0, so by a standard result from the calculus in 2 variables, we conclude that

$$F - \sum a_k G_{P_k}$$

has a potential on U^*, thus proving the theorem.

For the "standard result", cf. for instance my *Undergraduate Analysis*, Springer-Verlag, Chapter XV, Theorem 4.2.

A more complete exposition of the material in the present appendix is given in *Undergraduate Analysis*, Springer-Verlag, Second Edition (1997). Chapter XVI.

Theorem 6.4 is the third example that we have encountered of the same type as Theorem 2.4 of Chapter IV, and Theorem 3.9 of Chapter VIII.

Sketch of proof of Theorem 3.9, Chapter VIII. We are given h harmonic on the punctured open set U^*. On each disc W in U^*, h is the real part of an analytic function f_W, uniquely determined up to an additive constant. We want to know the obstruction for h to be the real part of an analytic function on U^*, and more precisely we want to show that there exist constants a_k and an analytic function f on U^* such that

$$h(z) - \sum a_k \log|z - z_k| = \operatorname{Re} f(z).$$

We consider the functions f_W as above. For each W, the derivative f'_W is uniquely determined, and the collection of such functions $\{f'_W\}$ defines an analytic function f' on U^*. Let γ_k be a small circle around z_k, and let

$$a_k = \frac{1}{2\pi i}\int_{\gamma_k} f'(\zeta)\,d\zeta.$$

Let z_0 be a point in U^* and let γ_z be a piecewise C^1 path in U^* from z_0 to a point z in U^*. Let

$$g(z) = f'(z) - \sum_k a_k \frac{1}{z - z_k} \qquad \text{and} \qquad f(z) = \int_{\gamma_z} g(\zeta)\,d\zeta.$$

Show that this last integral is independent of the path γ_z, and gives the desired function. Use Theorem 2.4 of Chapter IV.]

§7. MORE ON CAUCHY-RIEMANN

We give here two more statements about the Cauchy-Riemann equations, which are the heart of some exercises of Chapter VIII.

Theorem 7.1. *Let f be a complex harmonic function on a connected open set U. Let S be the set of points $z \in U$ such that $\partial f/\partial \bar{z} = 0$. Suppose that S has a non-empty interior V. Then $V = U$.*

Proof. From p. 92 Theorem 1.6, we know that an open subset of U which is closed in U is equal to U. Thus it suffices to show that V is closed in U. Let z_0 be a point in $\partial V \cap U$. Let $h = \partial f/\partial \bar{z} = u + iv$, with u, v real. Since partials commute, h is harmonic. Let D_0 be a disc centered at z_0, and $D_0 \subset U$. By Theorem 3.1 or 5.4 of Chapter VIII, there exists an analytic function h_1 on D_0 with $\operatorname{Re}(h_1) = u$. Since $z_0 \in \partial V$, if $z_0 \notin V$ it follows that $D_0 \cap V$ is open, and $u = 0$ on $D_0 \cap V$. Hence h_1 is pure imaginary constant on $D_0 \cap V$, and therefore h_1 is pure imaginary constant on D_0, so $u = 0$ on D_0. Replacing f by if, we conclude that $v = 0$ on D_0 also. Hence f is analytic on D_0, so $D_0 \subset S$ and $z_0 \in V$, qed.

Of course, in the above statement, one could have assumed $\partial f/\partial z = 0$ instead of $\partial f/\partial \bar{z} = 0$. Recall the formula $\overline{\partial f/\partial z} = \partial \bar{f}/\partial \bar{z}$, which shows that in the above situation, f, $\bar{f}$ play a symmetric role.

The next result has to do with the normal derivative, and extends VIII, §2, Exercise 5. Let f, U be open sets in $\mathbf{R}^2$. Let:

$f\colon U \to V$ have coordinate functions satisfying Cauchy-Riemann equations, i.e.

$$\partial_1 f_1 = \partial_2 f_2 \quad \text{and} \quad \partial_2 f_1 = -\partial_1 f_2;$$

$\varphi\colon V \to \mathbf{R}$ a C^1 real valued function on V;

$\gamma\colon [a,b] \to U$ a C^1 curve, so we get a sequence of maps

$$[a,b] \xrightarrow{\gamma} U \xrightarrow{f} V \xrightarrow{\varphi} \mathbf{R}.$$

For each function, we have its derivative at a point as a real linear map. The (non-unitized) normal of γ is

$$N_\gamma = \text{vertical vector} \begin{pmatrix} \gamma_2' \\ -\gamma_1' \end{pmatrix}, \text{ evaluated at each } t \in [a,b].$$

For any point $Q \in V$, calculus tells us that for any vector $Z \in \mathbf{R}^2$ we have

$$(\operatorname{grad} \varphi)(Q) \cdot Z = \varphi'(Q)Z.$$

The normal derivative of φ along the curve $f \circ \gamma$ is by definition

$$D_{N(f\circ\gamma)}\varphi = \varphi'(f \circ \gamma) N_{f\circ\gamma}. \tag{1}$$

Using the chain rule, we also have

$$D_{N(\gamma)}(\varphi \circ f) = (\varphi \circ f)'(\gamma) N_\gamma = \varphi'(f \circ \gamma) f'(\gamma) N_\gamma. \tag{2}$$

Theorem 7.2. *Assume that f_1, f_2 satisfy the Cauchy-Riemann equations. Then*

$$f'(\gamma) N_\gamma = N_{f\circ\gamma}.$$

Proof. On the one hand, writing vectors vertically, we have

$$N_{f\circ\gamma} = \begin{pmatrix} (f_2 \circ \gamma)' \\ -(f_1 \circ \gamma)' \end{pmatrix} = \begin{pmatrix} \partial_1 f_2(\gamma)\gamma_1' + \partial_2 f_2(\gamma)\gamma_2' \\ -\partial_1 f_1(\gamma)\gamma_1' - \partial_2 f_1(\gamma)\gamma_2' \end{pmatrix}.$$

On the other hand, using the matrix representing f' and writing γ' vertically, we get

$$f'(\gamma) N_\gamma = \begin{pmatrix} \partial_1 f_1 & \partial_2 f_1 \\ \partial_1 f_2 & \partial_2 f_2 \end{pmatrix} \begin{pmatrix} \gamma_2' \\ -\gamma_1' \end{pmatrix} = \begin{pmatrix} (\partial_1 f_1)\gamma_2' - (\partial_2 f_1)\gamma_1' \\ (\partial_1 f_2)\gamma_2' - (\partial_2 f_2)\gamma_1' \end{pmatrix}.$$

Using the Cauchy-Riemann equations concludes the proof.

Bibliography

[Ah 66] L. AHLFORS, *Complex Analysis*, McGraw-Hill, 1966

[Ar 65] E. ARTIN, *Collected Papers*, Addison-Wesley 1965; reprinted by Springer-Verlag, 1981

[Ba 75] A. BAKER, *Transcendental Number Theory*, Cambridge University Press, 1975

[BaN 97] J. BAK and D.J. NEWMAN, *Complex Analysis*, Springer-Verlag, 1997

[Co 35] E.T. COPSON, *An Introduction to the Theory of Functions of a Complex Variable*, Oxford University Press, 1935

[Fi 83] S. FISCHER, *Function Theory on Planar Domains*, Wiley–Interscience, 1983, especially Chapter 1, §3 and §4

[Ge 60] A.O. GELFOND, *Transcendental and Algebraic Numbers*, Translation, Dover Publications, 1960 (from the Russian edition, 1951)

[HuC 64] A. HURWITZ and R. COURANT, *Funktionentheorie* (Fourth Edition), Grundlehren Math. Wiss. No. 3, Springer-Verlag, 1964

[La 62] S. LANG, Transcendental points on group varieties, *Topology*, Vol. 1, 1962, pp. 313–318

[La 66] S. LANG, *Introduction to Transcendental Numbers*, Addison-Wesley, 1966

[La 73] S. LANG, *Elliptic Functions*, Springer-Verlag (reprinted from Addison-Wesley), 1973

[La 83] S. LANG, Undergraduate Analysis, Springer-Verlag, 1983, Second edition 1997

[La 93] S. LANG, *Real and Functional Analysis*, Third Edition, Springer-Verlag, 1993

[La 87] S. LANG, *Introduction to Complex Hyperbolic Spaces*, Springer-Verlag, 1987

[Ma 73] J. MARSDEN, *Basic Complex Analysis*, W.H. Freeman, 1973

[Ne 70] R. NEVANLINNA, *Analytic Functions*, Springer-Verlag, 1970 (translated and revised from the first edition in German, 1953)

[Ru 69] W. RUDIN, *Real and Complex Analysis*, McGraw-Hill, 1969

[StW 71] E. STEIN and G. WEISS, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1971, Chapter II, §4

[WhW 63] E.T. WHITTAKER and G. WATSON, *A Course of Modern Analysis*, Cambridge University Press, 1963

Index

Graduate Texts in Mathematics

64 EDWARDS. Fourier Series. Vol. I. 2nd ed.
65 WELLS. Differential Analysis on Complex Manifolds. 2nd ed.
66 WATERHOUSE. Introduction to Affine Group Schemes.
67 SERRE. Local Fields.
68 WEIDMANN. Linear Operators in Hilbert Spaces.
69 LANG. Cyclotomic Fields II.
70 MASSEY. Singular Homology Theory.
71 FARKAS/KRA. Riemann Surfaces. 2nd ed.
72 STILLWELL. Classical Topology and Combinatorial Group Theory. 2nd ed.
73 HUNGERFORD. Algebra.
74 DAVENPORT. Multiplicative Number Theory. 3rd ed.
75 HOCHSCHILD. Basic Theory of Algebraic Groups and Lie Algebras.
76 IITAKA. Algebraic Geometry.
77 HECKE. Lectures on the Theory of Algebraic Numbers.
78 BURRIS/SANKAPPANAVAR. A Course in Universal Algebra.
79 WALTERS. An Introduction to Ergodic Theory.
80 ROBINSON. A Course in the Theory of Groups. 2nd ed.
81 FORSTER. Lectures on Riemann Surfaces.
82 BOTT/TU. Differential Forms in Algebraic Topology.
83 WASHINGTON. Introduction to Cyclotomic Fields. 2nd ed.
84 IRELAND/ROSEN. A Classical Introduction to Modern Number Theory. 2nd ed.
85 EDWARDS. Fourier Series. Vol. II. 2nd ed.
86 VAN LINT. Introduction to Coding Theory. 2nd ed.
87 BROWN. Cohomology of Groups.
88 PIERCE. Associative Algebras.
89 LANG. Introduction to Algebraic and Abelian Functions. 2nd ed.
90 BRØNDSTED. An Introduction to Convex Polytopes.
91 BEARDON. On the Geometry of Discrete Groups.
92 DIESTEL. Sequences and Series in Banach Spaces.
93 DUBROVIN/FOMENKO/NOVIKOV. Modern Geometry—Methods and Applications. Part I. 2nd ed.
94 WARNER. Foundations of Differentiable Manifolds and Lie Groups.
95 SHIRYAEV. Probability. 2nd ed.
96 CONWAY. A Course in Functional Analysis. 2nd ed.
97 KOBLITZ. Introduction to Elliptic Curves and Modular Forms. 2nd ed.
98 BRÖCKER/TOM DIECK. Representations of Compact Lie Groups.
99 GROVE/BENSON Finite Reflection Groups. 2nd ed.
100 BERG/CHRISTENSEN/RESSEL. Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions.
101 EDWARDS. Galois Theory.
102 VARADARAJAN. Lie Groups, Lie Algebras and Their Representations.
103 LANG. Complex Analysis. 3rd ed.
104 DUBROVIN/FOMENKO/NOVIKOV. Modern Geometry—Methods and Applications. Part II.
105 LANG. $SL_2(\mathbf{R})$.
106 SILVERMAN. The Arithmetic of Elliptic Curves.
107 OLVER. Applications of Lie Groups to Differential Equations. 2nd ed.
108 RANGE. Holomorphic Functions and Integral Representations in Several Complex Variables.
109 LEHTO. Univalent Functions and Teichmüller Spaces.
110 LANG. Algebraic Number Theory.
111 HUSEMÖLLER. Elliptic Curves.
112 LANG. Elliptic Functions.
113 KARATZAS/SHREVE. Brownian Motion and Stochastic Calculus. 2nd ed.
114 KOBLITZ. A Course in Number Theory and Cryptography. 2nd ed.
115 BERGER/GOSTIAUX. Differential Geometry: Manifolds, Curves, and Surfaces.
116 KELLEY/SRINIVASAN. Measure and Integral. Vol. I.
117 J.-P. SERRE. Algebraic Groups and Class Fields.
118 PEDERSEN. Analysis Now.
119 ROTMAN. An Introduction to Algebraic Topology.
120 ZIEMER. Weakly Differentiable Functions: Sobolev Spaces and Functions of Bounded Variation.
121 LANG. Cyclotomic Fields I and II. Combined 2nd ed.
122 REMMERT. Theory of Complex Functions. *Readings in Mathematics*
123 EBBINGHAUS/HERMES et al. Numbers. *Readings in Mathematics*
124 DUBROVIN/FOMENKO/NOVIKOV. Modern Geometry—Methods and Applications. Part III
125 BERENSTEIN/GAY. Complex Variables: An Introduction.
126 BOREL. Linear Algebraic Groups. 2nd ed.
127 MASSEY. A Basic Course in Algebraic Topology.

128 RAUCH. Partial Differential Equations.
129 FULTON/HARRIS. Representation Theory: A First Course.
Readings in Mathematics
130 DODSON/POSTON. Tensor Geometry.
131 LAM. A First Course in Noncommutative Rings. 2nd ed.
132 BEARDON. Iteration of Rational Functions.
133 HARRIS. Algebraic Geometry: A First Course.
134 ROMAN. Coding and Information Theory.
135 ROMAN. Advanced Linear Algebra.
136 ADKINS/WEINTRAUB. Algebra: An Approach via Module Theory.
137 AXLER/BOURDON/RAMEY. Harmonic Function Theory. 2nd ed.
138 COHEN. A Course in Computational Algebraic Number Theory.
139 BREDON. Topology and Geometry.
140 AUBIN. Optima and Equilibria. An Introduction to Nonlinear Analysis.
141 BECKER/WEISPFENNING/KREDEL. Gröbner Bases. A Computational Approach to Commutative Algebra.
142 LANG. Real and Functional Analysis. 3rd ed.
143 DOOB. Measure Theory.
144 DENNIS/FARB. Noncommutative Algebra.
145 VICK. Homology Theory. An Introduction to Algebraic Topology. 2nd ed.
146 BRIDGES. Computability: A Mathematical Sketchbook.
147 ROSENBERG. Algebraic K-Theory and Its Applications.
148 ROTMAN. An Introduction to the Theory of Groups. 4th ed.
149 RATCLIFFE. Foundations of Hyperbolic Manifolds.
150 EISENBUD. Commutative Algebra with a View Toward Algebraic Geometry.
151 SILVERMAN. Advanced Topics in the Arithmetic of Elliptic Curves.
152 ZIEGLER. Lectures on Polytopes.
153 FULTON. Algebraic Topology: A First Course.
154 BROWN/PEARCY. An Introduction to Analysis.
155 KASSEL. Quantum Groups.
156 KECHRIS. Classical Descriptive Set Theory.
157 MALLIAVIN. Integration and Probability.
158 ROMAN. Field Theory.
159 CONWAY. Functions of One Complex Variable II.
160 LANG. Differential and Riemannian Manifolds.
161 BORWEIN/ERDÉLYI. Polynomials and Polynomial Inequalities.
162 ALPERIN/BELL. Groups and Representations.
163 DIXON/MORTIMER. Permutation Groups.
164 NATHANSON. Additive Number Theory: The Classical Bases.
165 NATHANSON. Additive Number Theory: Inverse Problems and the Geometry of Sumsets.
166 SHARPE. Differential Geometry: Cartan's Generalization of Klein's Erlangen Program.
167 MORANDI. Field and Galois Theory.
168 EWALD. Combinatorial Convexity and Algebraic Geometry.
169 BHATIA. Matrix Analysis.
170 BREDON. Sheaf Theory. 2nd ed.
171 PETERSEN. Riemannian Geometry.
172 REMMERT. Classical Topics in Complex Function Theory.
173 DIESTEL. Graph Theory. 2nd ed.
174 BRIDGES. Foundations of Real and Abstract Analysis.
175 LICKORISH. An Introduction to Knot Theory.
176 LEE. Riemannian Manifolds.
177 NEWMAN. Analytic Number Theory.
178 CLARKE/LEDYAEV/STERN/WOLENSKI. Nonsmooth Analysis and Control Theory.
179 DOUGLAS. Banach Algebra Techniques in Operator Theory. 2nd ed.
180 SRIVASTAVA. A Course on Borel Sets.
181 KRESS. Numerical Analysis.
182 WALTER. Ordinary Differential Equations.
183 MEGGINSON. An Introduction to Banach Space Theory.
184 BOLLOBAS. Modern Graph Theory.
185 COX/LITTLE/O'SHEA. Using Algebraic Geometry.
186 RAMAKRISHNAN/VALENZA. Fourier Analysis on Number Fields.
187 HARRIS/MORRISON. Moduli of Curves.
188 GOLDBLATT. Lectures on the Hyperreals: An Introduction to Nonstandard Analysis.
189 LAM. Lectures on Modules and Rings.
190 ESMONDE/MURTY. Problems in Algebraic Number Theory.
191 LANG. Fundamentals of Differential Geometry.
192 HIRSCH/LACOMBE. Elements of Functional Analysis.
193 COHEN. Advanced Topics in Computational Number Theory.
194 ENGEL/NAGEL. One-Parameter Semigroups for Linear Evolution Equations.
195 NATHANSON. Elementary Methods in Number Theory.

196 OSBORNE. Basic Homological Algebra.
197 EISENBUD/HARRIS. The Geometry of Schemes.
198 ROBERT. A Course in p-adic Analysis.
199 HEDENMALM/KORENBLUM/ZHU. Theory of Bergman Spaces.
200 BAO/CHERN/SHEN. An Introduction to Riemann–Finsler Geometry.
201 HINDRY/SILVERMAN. Diophantine Geometry: An Introduction.
202 LEE. Introduction to Topological Manifolds.
203 SAGAN. The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions.
204 ESCOFIER. Galois Theory.
205 FÉLIX/HALPERIN/THOMAS. Rational Homotopy Theory. 2nd ed.
206 MURTY. Problems in Analytic Number Theory.
Readings in Mathematics
207 GODSIL/ROYLE. Algebraic Graph Theory.
208 CHENEY. Analysis for Applied Mathematics.
209 ARVESON. A Short Course on Spectral Theory.
210 ROSEN. Number Theory in Function Fields.
211 LANG. Algebra. Revised 3rd ed.
212 MATOUŠEK. Lectures on Discrete Geometry.
213 FRITZSCHE/GRAUERT. From Holomorphic Functions to Complex Manifolds.
214 JOST. Partial Differential Equations.
215 GOLDSCHMIDT. Algebraic Functions and Projective Curves.
216 D. SERRE. Matrices: Theory and Applications.
217 MARKER. Model Theory: An Introduction.
218 LEE. Introduction to Smooth Manifolds.
219 MACLACHLAN/REID. The Arithmetic of Hyperbolic 3-Manifolds.
220 NESTRUEV. Smooth Manifolds and Observables.
221 GRÜNBAUM. Convex Polytopes. 2nd ed.
222 HALL. Lie Groups, Lie Algebras, and Representations: An Elementary Introduction.
223 VRETBLAD. Fourier Analysis and Its Applications.