COMPLEX ANALYSIS

Fall 2006

Christer Bennewitz

Månne ifrån högre zoner
analytiska funktioner
svaret nu dig finna låta
på odödlighetens gåta?

F. Läffler

Preface

These notes are basically a printed version of my lectures in complex analysis at the University of Lund. As such they present a limited view of any of the subject matters brought up, caused by the time constraints one is faced by in a series of lectures. The core of the subject, presented in Chapter 3, is very strongly influenced by the treatment in Ahlfors' *Complex Analysis*, one of the genuine masterpieces of the subject. Any reader who wants to find out more is advised to read this book.

Mathematical prerequisites are in principle the mathematics courses given in the first two semesters in Lund. Most importantly, this includes a reasonably complete discussion of analysis in one and several variables and basic facts about series of functions including absolute and uniform convergence. A course in topology is also useful, but not essential. Primarily, a familiarity with the concept of a connected set is of use.

Egevang, August 2006

Christer Bennewitz

Contents

Preface i

Chapter 1. Complex functions 1
1.1. The complex number system 1
1.2. Polar form of complex numbers 4
1.3. Square roots 5
1.4. Stereographic projection 7
1.5. Möbius transforms 10
1.6. Polynomials, rational functions and power series 16

Chapter 2. Analytic functions 23
2.1. Conformal mappings and analyticity 23
2.2. Analyticity of power series; elementary functions 27
2.3. Conformal mappings by elementary functions 33

Chapter 3. Integration 37
3.1. Complex integration 37
3.2. Goursat's theorem 40
3.3. Local properties of analytic functions 45
3.4. A general form of Cauchy's integral theorem 48
3.5. Analyticity on the Riemann sphere 51

Chapter 4. Singularities 53
4.1. Singular points 53
4.2. Laurent expansions and the residue theorem 56
4.3. Residue calculus 59
4.4. The argument principle 65

Chapter 5. Harmonic functions 73
5.1. Fundamental properties 73
5.2. Dirichlet's problem 79

Chapter 6. Entire functions 85
6.1. Sequences of analytic functions 85
6.2. Infinite products 87
6.3. Canonical products 90
6.4. Partial fractions 94
6.5. Hadamard's theorem 96

Chapter 7. The Riemann mapping theorem 101

Chapter 8. The Gamma function 105

CHAPTER 1

Complex functions

1.1. The complex number system

Recall that a *group* $(G, *)$ is a set G provided with a *binary operation*[1] $*$ satisfying the following properties:

(1) For all elements x, y and $z \in G$ holds $(x * y) * z = x * (y * z)$. (associative law)
(2) There exists a *neutral element* $e \in G$ with the properties $x*e = e * x = x$ for every $x \in G$.
(3) Every element $x \in G$ has an *inverse* x^{-1} with the properties $x * x^{-1} = x^{-1} * x = e$.

Exercise 1.1. Show that a set provided with an associative binary operation can have *at most* one neutral element.
Hint: Show that if the set has a 'left neutral' element and a 'right neutral' element, they must coincide.

Exercise 1.2. Show that if a set has an associative binary operation with neutral element, then any element of the set has *at most* one inverse.
Hint: Show that if an element has a 'left inverse' and a 'right inverse', then these must coincide.

A group may also have the property

(4) For all elements x and $y \in G$ holds $x*y = y*x$. (commutative law)

in which case the group is called commutative or *Abelian* (after Niels Henrik Abel (1802–1829)). Familiar examples of Abelian groups are $(\mathbb{Z}, +)$, the integers under ordinary addition; $(\mathbb{R}, +)$, the real numbers under addition; $(\mathbb{R}^n, +)$, the set of n-tuples of real numbers under (vector) addition; and $(\mathbb{R} \setminus \{0\}, \cdot)$, the non-zero real numbers under multiplication. As an example of a non-Abelian group, consider the set of all rotations around lines through the origin in 3-dimensional space; the binary operation is the ordinary composition of maps. The reader should check these examples carefully; in particular, find the neutral elements and inverses in these groups.

[1]That is, a map $* : G \times G \to G$, so that for every pair of elements x, y of G, there is a unique element of G denoted by $x * y$.

A *field* $(F, +, \cdot)$ is a set F provided with two binary operations $+$ and $\cdot$, such that $(F, +)$ is an Abelian group and, if 0 denotes the neutral element of this group, also $(F \setminus \{0\}, \cdot)$ is an Abelian group. In addition the *distributive laws*

$$\begin{cases} (x+y)\cdot z = x\cdot z + y\cdot z, \\ x\cdot(y+z) = x\cdot y + x\cdot z. \end{cases}$$

hold for all elements x, y and $z \in F$. It is usual to denote the neutral element of $(F \setminus \{0\}, \cdot)$ by 1.

EXERCISE 1.3. Prove that in any field F holds $0 \cdot x = x \cdot 0 = 0$ for all $x \in F$ (as always, 0 denotes the neutral element of the group $(F, +)$).

EXERCISE 1.4. Prove that a field does not have any non-zero *divisors of zero*, *i.e.*, if $xy = 0$, then either $x = 0$ or $y = 0$.

Familiar examples of fields are $(\mathbb{Q}, +, \cdot)$, the rational numbers under ordinary addition and multiplication, and $(\mathbb{R}, +, \cdot)$. We shall show, in this section, that there is precisely one reasonable way of making the Euclidean plane into a field. By introducing Cartesian coordinates this plane may be identified with the Abelian group $(\mathbb{R}^2, +)$, and we will make this into a field by extending the usual multiplication of an element of $\mathbb{R}^2$ by a real number. The resulting field is the field $\mathbb{C}$ of *complex numbers*.

To see how to make the definition, assume we have already managed to construct our field $\mathbb{C}$. Then there is a multiplicative neutral element, which we will for the moment denote by $\mathbf{1}$, to distinguish it from the real number 1. We may identify $\mathbb{R}$ with the set of real multiples of $\mathbf{1}$ (explain!) and may therefore consider $\mathbb{R}$ as a subset of $\mathbb{C}$. Let $\mathbf{e}$ be an element of $\mathbb{R}^2$ which is linearly independent of $\mathbf{1}$, so that $\mathbf{1}, \mathbf{e}$ is a basis in $\mathbb{R}^2$. Any element $z \in \mathbb{C}$ may then be written $z = x\mathbf{1} + y\mathbf{e}$ with real numbers x and y. In particular, there are real numbers a and b such that $\mathbf{e}^2 = a\mathbf{1} + b\mathbf{e}$ so that $z^2 = (x^2 + ay^2)\mathbf{1} + (2xy + by^2)\mathbf{e}$ (note that $\mathbf{1} \cdot \mathbf{1} = \mathbf{1}$, $\mathbf{e} \cdot \mathbf{1} = \mathbf{e}$). Now clearly z^2 is real if $y = 0$ (since actually z itself is, by the identification above). But z^2 will also be real if $x = -\frac{b}{2}y$. We then get $z^2 = (a + \frac{b^2}{4})y^2$. We can not have $a + \frac{b^2}{4} \geq 0$ by Exercise 1.4 since then $(z - y\sqrt{a + \frac{b^2}{4}})(z + y\sqrt{a + \frac{b^2}{4}}) = 0$, but neither of the factors is 0 unless their $\mathbf{e}$-component $y = 0$. Hence $a + \frac{b^2}{4} < 0$. If we set $y = 1/\sqrt{-(a + \frac{b^2}{4})}$ we therefore get $z^2 = -1$.

Roughly, we have seen that if we can define a multiplication in $\mathbb{R}^2$ which makes it into a field with addition being the ordinary vector addition, then there exists an element the square of which is -1 (rather, the additive inverse of the multiplicative neutral element). We pick one such element (we will see later that there are precisely two), denote it by i and call it the *imaginary unit.* If we use $\mathbf{1}, i$ as a basis we may therefore write any element in the plane as $x\mathbf{1} + yi$ with real x, y. For convenience we will actually write it $x + iy$ from now on.

It is important to note that we have not yet shown that it is possible to make a field of the plane; we have just seen that *if* it is possible, then we may identify the x-axis with the real numbers and the y-axis with the multiples of an element, the square of which is -1.

EXERCISE 1.5. Show that if we calculate with symbols $x+iy$, where x and y are real numbers, according to the usual rules for adding and multiplying numbers and in addition use $i^2 = -1$, then all the requirements for a field are satisfied.

From now on the field we have constructed is denoted by $\mathbb{C}$ and called the field of complex numbers. Note that the field of real numbers is an *ordered field.* This means that we have a relation $<$ defined among the real numbers such that

(1) If x and $y \in \mathbb{R}$, then exactly one of $x < y$, $y < x$ and $x = y$ is true.
(2) Sums and products of positive (*i.e.*, > 0) numbers are positive.

We have not introduced anything similar for the complex numbers for the simple reason that it *can not be done.*

EXERCISE 1.6. Show that in an ordered field squares of non-zero elements are always > 0. Use this to show that if it were possible to make $\mathbb{C}$ into an ordered field, then both $1 > 0$ and $-1 > 0$, and hence also $0 > 0$, a contradiction.

As a final note to this first section, the fact that the Euclidean plane can be made into a field is extremely useful in all areas of mathematics and its applications. Since we live in a 3-dimensional (at least) world, it would, from the point of view of applications, be very useful if we could make 3-dimensional space into a field as well. In the early part of the nineteenth century, this is exactly what the famous Irish mathematician W. R. Hamilton tried, unsuccessfully, to do.

EXERCISE 1.7. Try to show that Hamilton was doomed to fail. To simplify things, you may require that the complex plane should be a 2-dimensional restriction of the 3-dimensional field. Show that the existence of divisors of zero can not be avoided.

Hamilton succeeded (1843) to introduce a multiplication in $\mathbb{R}^4$ which makes this into a field, with the minor defect that the multiplicative group is not Abelian (such a structure is called a *skew field*). Hamilton called his structure the *quaternions*; this structure actually strongly hints that it would be profitable, in physics, to consider the world 4-dimensional, with time as the fourth dimension.

EXERCISE 1.8. Consider the set of symbols $x+iy+ju+kv$, where x, y, u and v are real numbers, and the symbols i, j, k satisfy $i^2 = j^2 = k^2 = -1$, $ij = -ji = k$, $jk = -kj = i$ and $ki = -ik = j$. Show that using these relations and calculating with the same formal rules as in dealing with real numbers, we obtain a skew field; this is the set of quaternions.

1.2. Polar form of complex numbers

In the complex number $z = x + iy$ the real number x is called the *real part* of z, $x = \operatorname{Re} z$, and the number y is called the *imaginary part* of z, $y = \operatorname{Im} z$. There is of course nothing imaginary whatever about the imaginary part; the reasons for this curious appellation are historic. If we introduce the notation $\overline{z}$ for the complex number $x - iy$, called the *complex conjugate* of z, we see that $\operatorname{Re} z = \frac{1}{2}(z + \overline{z})$ and $\operatorname{Im} z = \frac{1}{2i}(z - \overline{z})$. In particular, z is real (*i.e.*, has imaginary part 0) precisely if $z = \overline{z}$. If z has real part 0, so that $z = -\overline{z}$, one calls z *purely imaginary*. We define the *absolute value* $|z|$ of $z = x + iy$ to be $|z| = \sqrt{x^2 + y^2}$. This is of course the ordinary length of z, considered as a vector in the plane, provided we draw $1, i$ as orthonormal vectors. A very useful observation is that $z\overline{z} = |z|^2$.

EXERCISE 1.9. Show this and that for any complex numbers z and w we have

(1) $\overline{z+w} = \overline{z} + \overline{w}$,
(2) $\overline{zw} = \overline{z} \cdot \overline{w}$,
(3) $|zw| = |z||w|$.

It is worth remarking how one carries out division by a complex number. Since the complex numbers constitute a field, every non-zero complex number has a multiplicative inverse, *i.e.*, we can divide by it; namely, if $z \neq 0$ and w are complex numbers, then there is a unique complex number u, denoted $\frac{w}{z}$, such that $zu = w$. The question is, how does one write the quotient on the standard form as real part plus i times imaginary part. To see how, multiply through by $\overline{z}$ to obtain $|z|^2 u = \overline{z}w$. Since $|z|^2 \neq 0$ we can divide by this (real) number, and so $u = \overline{z}w/|z|^2$. So, to write w/z on standard form, multiply numerator and denominator by $\overline{z}$.

EXERCISE 1.10. Write $\frac{1+2i}{3+4i}$ on standard form.

The geometric interpretation of addition is already familiar, since this is the ordinary vector addition in the plane. To get a geometric picture of multiplication, we introduce polar coordinates in the plane in the following way. If $z \neq 0$, then $z/|z|$ is located somewhere on the unit circle; hence we can find an angle θ such that $z/|z| = \cos\theta + i \sin\theta$. We may therefore write z on polar form as $z = |z|(\cos\theta + i\sin\theta)$ where θ is called the *argument* of z and is denoted $\theta = \arg z$. It is unfortunate, but *extremely important* that $\arg z$ is **NOT** uniquely determined by z; adding any integer multiple of 2π to θ gives another, equally valid, value for $\arg z$. When one therefore speaks of 'the' argument for a complex number, one means *one of the infinitely many possible values* of the argument. Another, less serious ambiguity, is that we have not assigned an argument to the number 0; it is usual to allow any real number whatsoever as a valid argument for 0.

Now suppose $z = |z|(\cos\theta + i\sin\theta)$ and $w = |w|(\cos\phi + i\sin\phi)$ are complex numbers. Then $zw = |z||w|(\cos\theta\cos\phi - \sin\theta\sin\phi + i(\cos\theta\sin\phi + \sin\theta\cos\phi)) = |zw|(\cos(\theta+\phi) + i\sin(\theta+\phi))$ according to the addition formulas for sin and cos. Thus, when calculating the product of two complex numbers the absolute values are multiplied and the arguments are *added*. In particular, multiplication by a complex number of absolute value 1 is equivalent to a rotation with an angle equal to the argument of the given number.

EXERCISE 1.11. Write the number $z = \sqrt{3} + i$ on polar form and then calculate z^{13} on standard form.

1.3. Square roots

Working with real numbers it is possible to find the square root of any non-negative number; to obtain a unique number the square root is required to be non-negative as well. After introducing complex numbers we can, for *any* given real number, find a real or complex number whose square is the given number. Of course, not much would be gained unless we could actually find the square root of any *complex* number as well. This means that we would like to be able to find a solution to $z^2 = w$ for any complex number w. Suppose $w = u + iv$ and let $z = x + iy$ (in situations like this it is always assumed that u, v, x and y are real numbers). Since $z^2 = x^2 - y^2 + 2ixy$ we need to solve the nonlinear system

$$\begin{cases} x^2 - y^2 = u, \\ 2xy = v. \end{cases} \tag{1.1}$$

in two real unknowns x and y. Squaring and adding the two equations we get, after extracting a (real) square root, that $x^2 + y^2 = \sqrt{u^2+v^2}$ (this simply expresses the fact that $|z|^2 = |w|$, which has to be true in view of Exercise 2.1). Together with the first equation this shows that

$$\begin{cases} x = \pm\sqrt{\frac{1}{2}(\sqrt{u^2+v^2} + u)}, \\ y = \pm\sqrt{\frac{1}{2}(\sqrt{u^2+v^2} - u)}. \end{cases} \tag{1.2}$$

Note that all the expressions within square roots are non-negative no matter what u and v are, so these are ordinary real square roots. (1.2) therefore give all possible solutions of (1.1), and it is easily verified that the first equation is actually satisfied, whereas the second is satisfied if and only if one chooses the right combination of signs, so that there are actually always precisely two distinct complex numbers z satisfying $z^2 = w$, unless $w = 0$ in which case $z = 0$ is the only solution. Since a quadratic equation can be solved by extracting square roots one now easily sees that any quadratic equation with complex coefficients always has a complex root. In fact, if counted by multiplicity there are always exactly two roots (we will return later to the concept of multiplicity for a root).

We have seen that we can always extract square roots of a complex number w, and that there are always (unless $w = 0$) exactly two such numbers. The question arises: Which of the two possibilities are we to denote by the symbol $\sqrt{w}$? Since the complex numbers are not ordered there is no simple answer to this question, as in the real case. To analyze the situation we write $w = |w|(\cos\theta + i\sin\theta)$ on polar form. If $z^2 = w$, then clearly $|z| = \sqrt{|w|}$, and if ϕ is an argument for z, then 2ϕ must be an argument for w. The simplest choice for ϕ is therefore to set $\phi = \theta/2$. Which number z we get this way obviously depends on the choice of θ, which is only determined up to an integer multiple of 2π. If we add 2π to θ we will add π to ϕ, which will replace z by $-z$. Adding or subtracting further multiples of 2π to θ will not yield any more values for z, so we have again seen that there are exactly two square roots of any non-zero number. We can write any complex number w on polar form with an argument θ in the interval $-\pi < \theta \leq \pi$ and choosing the argument of the square root to be $\phi = \theta/2$ we will get $-\frac{\pi}{2} < \phi \leq \frac{\pi}{2}$.

This is one way of assigning a unique value to the square root of any complex number. Considering z as a function of w this is called *the principal branch* of the square root; if w is a non-negative real number it obviously coincides with the usual real square root. The values of the principal branch of the square root are all in the *right half plane, i.e.*, they have non-negative real part. There are, however, other ways of choosing a branch of the square root that are sometimes more convenient. On may for example restrict θ to the interval $0 \leq \theta < 2\pi$, which will give the argument of the square root in the interval $0 \leq \phi < \pi$, *i.e.*, this branch of the root has all its values in the upper half plane.

Why can one not, once and for all like in the case of the real square root, choose a particular branch and stick to it? The reason is problems with continuity. Suppose we have a nice curve in the w-plane which intersects the negative real axis. If we take the square root of this, using the principal branch, the image of the curve in the z-plane will jump from a point on the negative imaginary axis to a point on the positive imaginary axis; we have lost the continuity of the curve. Another choice of branch might solve the problem for a particular curve, but it is clear that no choice of branch will be suitable for all curves. Since there is no choice of branch which will work best in all situations *one must not use* the notation $\sqrt{}$ without specifying which branch of the square root one is talking about.

The need to deal with several different branches occurs for all kinds of other complex functions and is a major complicating factor in the theory. There is a sophisticated and completely satisfactory solution to the problem, namely the introduction of the concept of a Riemann surface. Unfortunately we can not go into that here.

1.4. Stereographic projection

Since we have a notion of distance (*i.e.*, $d(z,w) = |z-w|$) in $\mathbb{C}$ we may view $\mathbb{C}$ as a metric space. It is clear that this space is *complete* in the sense that any Cauchy sequence converges; to see this note that since $|\operatorname{Re} z| \le |z|$ and $|\operatorname{Im} z| \le |z| \le |\operatorname{Re} z| + |\operatorname{Im} z|$ for any $z \in \mathbb{C}$ it follows that if $z_j = x_j + iy_j$, $j = 1, 2, \dots$ is a Cauchy sequence in $\mathbb{C}$, then x_j, $j = 1, 2, \dots$ and y_j, $j = 1, 2, \dots$ are Cauchy sequences in $\mathbb{R}$. Furthermore, if $x_j \to x \in \mathbb{R}$ and $y_j \to y \in \mathbb{R}$ as $j \to \infty$, then $x_j + iy_j \to x + iy \in \mathbb{C}$ as $j \to \infty$. Thus the completeness of $\mathbb{C}$ follows from that of $\mathbb{R}$.

From the point of view of topology, it would be even better if $\mathbb{C}$ were *compact, i.e.*, any open cover of $\mathbb{C}$ should have a finite subcover. This is not true, however, as can be seen by considering the open cover of $\mathbb{C}$ consisting of all open balls $|z| < R$ centered at 0, which obviously has no finite subcover. One can make $\mathbb{C}$ compact without changing its topology by adding (at least) one 'ideal' point and modifying the metric. This *one-point compactification* of the complex plane is very important in the theory of functions of a complex variable and we will give a very enlightening geometric interpretation of it in this section.

Imagine $\mathbb{C}$ as the x_1x_2-plane in $\mathbb{R}^3$ and let S_2 be the unit sphere; it will intersect $\mathbb{C}$ along the unit circle. Call the point $(0,0,1)$ on the sphere the *North pole* N (so that $(0,0,-1)$ is the South pole). We can map $\mathbb{C}$ in a one-to-one fashion onto $S_2 \setminus \{N\}$ by mapping $z \in \mathbb{C}$ onto the point $(x_1, x_2, x_3) \in S_2$ such that the straight line connecting z with N goes through (x_1, x_2, x_3). This map is called *stereographic projection* and has many interesting properties, as we shall see. In this connection S_2 is called the *Riemann sphere*.

It is nearly obvious that this stereographic projection is a bi-continuous map, using the topology induced by the metric of $\mathbb{R}^3$. To make absolutely sure, let us find the mapping explicitly. The line through N and $z = x + iy \in \mathbb{C}$ is $(x_1, x_2, x_3) = (0,0,1) + t(x, y, -1)$. The intersection with S_2 is given by t satisfying $t^2(x^2+y^2) + (1-t)^2 = 1$ which gives $t = 0$, *i.e.*, N, and the more interesting $t = 2/(x^2+y^2+1)$.

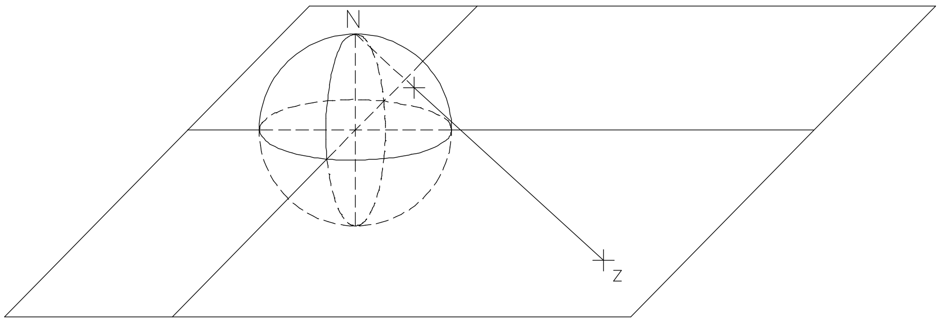

FIGURE 1. Stereographic projection

We therefore get

$$\begin{cases} x_1 = \dfrac{2\,\mathrm{Re}\,z}{|z|^2+1} \\ x_2 = \dfrac{2\,\mathrm{Im}\,z}{|z|^2+1} \\ x_3 = \dfrac{|z|^2-1}{|z|^2+1} \end{cases} .$$

Since $|z|^2 + 1 = 2/(1 - x_3)$ by the third equation the inverse is easily seen to be given by

$$z = \frac{x_1 + ix_2}{1 - x_3} .$$

It is clear that these maps are both continuous (note that $(x_1, x_2, x_3) \in S_2 \setminus \{N\}$ so $x_3 \neq 1$). We may now introduce a new metric in $\mathbb{C}$ by setting the distance between points in $\mathbb{C}$ equal to the Euclidean distance between their image points on S_2.

EXERCISE 1.12. Show that this metric is given by

$$d(z, w) = 2\frac{|z - w|}{(|z|^2 + 1)^{1/2}(|w|^2 + 1)^{1/2}}.$$

Also show that the distance between the image of z and N is $\frac{2}{(|z|^2+1)^{1/2}}$.

In view of Exercise 1.12 we may now add to $\mathbb{C}$ an 'ideal' point ∞, the image of which under stereographic projection is N. This new set is called the *extended complex plane* and we denote it by $\mathbb{C}^*$. Using the metric of Exercise 1.12 in $\mathbb{C}^*$ the extended plane becomes homeomorphic to the Riemann sphere with the topology of Euclidean distance. Since S_2 is compact, so is the extended plane; we have compactified the plane. For the statement of the next theorem, note that a circle in S^2 is the intersection of S^2 by a non-tangential plane, and any such (non-empty) intersection is a circle.

THEOREM 1.13. *The image of a straight line in $\mathbb{C}$ under stereographic projection is a circle through N, with N excluded. The image of a circle in $\mathbb{C}$ under stereographic projection is a circle not containing N. The inverse image of any circle on S_2 is a straight line together with ∞ if the circle passes through N, otherwise a circle.*

PROOF. Since a straight line in the x_1x_2-plane together with N determines a unique plane, the intersection of which with S_2 is the image of the straight line we only need to consider the case of a circle in $\mathbb{C}$. If it has center a and radius r its equation is $|z-a|^2 = r^2$ or $|z|^2 - 2\operatorname{Re}(\overline{a}z) + |a|^2 = r^2$. Substituting $z = \frac{x_1+ix_2}{1-x_3}$ into this, using that $x_1^2 + x_2^2 + x_3^2 = 1$ and $x_3 \neq 1$, we get $1 + x_3 - 2x_1 \operatorname{Re} a - 2x_2 \operatorname{Im} a + (1-x_3)(|a|^2 - r^2) = 0$ which is the equation of a plane. Conversely, a circle on the Riemann sphere is determined by three distinct points. The inverse images of these three points determine a circle in $\mathbb{C}$. The image of this circle is clearly the original circle. $\square$

In view of this theorem we will by a *circle in the extended plane* mean either a line together with ∞, or an actual circle.

A map is called *conformal* if it preserves angles and their orientation (we will give a more exact definition in Chapter 2.1). A surface is given an orientation by assigning to each point a normal direction which varies continuously with the point. For example, the usual orientation of $\mathbb{C}$ is given by letting at each point the normal point upwards, *i.e.*, in our present picture in the direction of the x_3-axis. Similarly, we may give the Riemann sphere an orientation by letting the normal point towards the origin.

The angle between two smooth curves in an oriented surface at a point of intersection of the curves is the angle between the tangents at the point. There are two such angles, the sum of which is π. If the curves are given in a certain order, the *positively oriented* angle between them is that angle through which one has to turn the first tangent vector so as to coincide with the second tangent vector, turning counterclockwise as seen from the normal to the surface. A strict definition would of course have to be freed from such obviously intuitive geometric concepts, but we will not attempt this here.

THEOREM 1.14. *Stereographic projection is conformal.*

PROOF. Consider two curves intersecting at z and their tangents at z in $\mathbb{C}$. Together with N the tangents determine two planes that intersect the Riemann sphere in two circles through N. The tangents to the circles at N are in these planes and also in the plane through N parallel to $\mathbb{C}$. It follows that they are parallel to the original tangent vectors so that viewed from *inside* the sphere they give rise to an angle equal to but of opposite orientation to the original angle.

The circles intersect also at the image of z on the sphere, and are tangent to the images of the curves there. The angles at the two points where the circles intersect are equal but of opposite orientation by symmetry (the two angles are images of each other under reflection in the plane through the origin and parallel to the normals of the planes of the circles). The theorem now follows. □

Although the proof above is very geometric in nature it is actually not difficult to make it analytic, using the fact that stereographic projection is a differentiable map, but we will not do that here.

1.5. Möbius transforms

A *Möbius transform* (also called a *linear fractional transformation*) is a non-constant mapping of the form $z \mapsto f(z) = \dfrac{az+b}{cz+d}$ for complex numbers a, b, c and d. To begin with we consider this defined in $\mathbb{C}$ except, if $c \neq 0$, for $z = -d/c$. The fact that the mapping is non-constant means that (a, b) is not proportional to (c, d). This can be expressed by requiring $ad - bc \neq 0$ which is always assumed from now on. Clearly we get the same mapping if we multiply all the coefficients a, b, c, d by the same non-zero number so that although the mapping is determined by the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ any non-zero multiple of this matrix gives the same mapping. The requirement $ad - bc \neq 0$ means that the determinant is $\neq 0$ so multiplying by an appropriate number we may always assume that the determinant is 1. This determines the coefficients up to a change in sign of all of them.

It is clear that if $c = 0$, then $f(z) \to \infty$ as $z \to \infty$. On the other hand, if $c \neq 0$, then $f(z) \to \infty$ as $z \to -d/c$ and $f(z) \to a/c$ as $z \to \infty$. We may therefore extend the definition of f to all of the extended plane $\mathbb{C}^*$ in such a way that the extended function is a continuous function of $\mathbb{C}^*$ into $\mathbb{C}^*$. We will always consider Möbius transforms as defined in the extended plane, or equivalently on the Riemann sphere, in this way. We have the following interesting proposition.

PROPOSITION 1.15. *If f and g are Möbius transforms corresponding to the matrices A and B, then the composed map $f \circ g$ is a Möbius transform corresponding to the matrix AB.*

EXERCISE 1.16. Prove Proposition 1.15.

Since the set of all non-singular 2×2 matrices is a group under matrix multiplication, it follows that so are the Möbius transforms. This means that any Möbius transform has an inverse which is also a Möbius transform.

EXERCISE 1.17. Find all Möbius transforms T for which $T^2 = T$.

Among other things this means that a Möbius transform is a *homeomorphism* of the extended plane onto itself, *i.e.*, a continuous one-to-one and onto map whose inverse is also continuous. But Möbius transforms have more surprising properties. Recall that we by a circle *in the extended plane* mean either an actual circle in the plane *or* a straight line together with ∞.

THEOREM 1.18. *Möbius transforms are conformal and* ***circle-preserving,*** i.e., *any circle in the extended plane is mapped onto a circle in the extended plane.*

PROOF. The theorem is obvious for certain simple special cases, namely a *translation* $z \mapsto z + b$, a *rotation* $z \mapsto az$ where $|a| = 1$ and a *dilation* $z \mapsto az$ where $a > 0$. It is therefore also true for a *multiplication* $z \mapsto az$ where $0 \neq a \in \mathbb{C}$, since this is the composite of the rotation $z \mapsto \frac{a}{|a|}z$ and the dilation $z \mapsto |a|z$. Composing a multiplication with a translation the theorem follows for a *linear* map $z \mapsto az+b$ where $a \neq 0$, and hence for any Möbius transform for which $c = 0$. If $c \neq 0$ we have $\frac{az+b}{cz+d} = \frac{bc-ad}{c}\frac{1}{cz+d} + \frac{a}{c}$ so that this map is the composite of three maps, the first and last being linear and the middle map is the *inversion* $z \mapsto 1/z$. The theorem therefore follows if we can prove it for an inversion.

If the image of z under stereographic projection is (x_1, x_2, x_3), then we have $z = \frac{x_1+ix_2}{1-x_3}$ so that $1/z = \frac{1-x_3}{x_1+ix_2} = (1-x_3)\frac{x_1-ix_2}{x_1^2+x_2^2}$. Since (x_1, x_2, x_3) is on the unit sphere we have $x_1^2 + x_2^2 = 1 - x_3^2$ so that $1/z = \frac{x_1-ix_2}{1+x_3}$. Therefore $1/z$ is the inverse stereographic projection of $(x_1, -x_2, -x_3)$. The map that takes (x_1, x_2, x_3) into this is a rotation around the x_1-axis by an angle π. This is obviously a circle-preserving and conformal map, and since we know that also stereographic projection is circle-preserving and conformal it follows that the inversion has the desired properties. The proof is now complete. □

EXERCISE 1.19. Prove the theorem by calculation, not using stereographic projection.

Note that since removing a circle from the extended plane leaves a set with exactly two components, and since Möbius transforms are continuous in the extended plane, the interior of any circle in the plane is mapped *either* onto the interior *or* onto the exterior, including ∞, of another circle. This follows from the fact that continuous maps preserve connectedness.

EXERCISE 1.20. Prove the statement above in detail.

Sets that are left invariant under a mapping are obviously important characteristics of the map. For a Möbius transform one may for example ask which circles it leaves invariant, or conversely, which Möbius transforms leave a given circle invariant. We will consider some such problems later. Right now we will instead ask for *fixpoints* of a given transform, *i.e.*, points left invariant by the map. By our definition of the image of ∞, this is a fixpoint if and only if the map is linear. A linear map $z \mapsto az + b$ also has the finite fixpoint $z = b/(1-a)$, except if $a = 1$. Thus, a translation which is not the identity has only the fixpoint ∞, but any other linear map which is not the identity has exactly one finite fixpoint as well. For a Möbius transform $z \mapsto \frac{az+b}{cz+d}$ with $c \neq 0$ the equation for a fixpoint becomes $z(cz+d) = az+b$ which is a quadratic equation. It therefore has either two distinct roots or a double root. We have therefore proved the following proposition.

PROPOSITION 1.21. *A Möbius transform different from the identity has either one or two fixpoints, as a map defined on the extended plane.*

EXERCISE 1.22. Find the fixed points of the linear transformations

$$w = \frac{z}{2z-1}\ , \quad w = \frac{2z}{3z-1}\ , \quad w = \frac{3z-4}{z-1}\ , \quad w = \frac{z}{2-z}\ .$$

In particular, a Möbius transform that leaves three distinct points invariant is the identity. It also follows that there can be at most one Möbius transform that takes three given, distinct points into three specified, distinct points. Because, if there were two, say f and g, then $f^{-1} \circ g$ would be a transform different from the identity and leaving the given points invariant. Conversely, we will prove that there actually always exists a Möbius transform that takes the given points into the specified ones. To see this, define the *cross ratio* of four distinct points z_0, z_1, z_2, z_3 in $\mathbb{C}^*$ by

$$(z_0, z_1, z_2, z_3) = \frac{z_0 - z_2}{z_0 - z_3} \Big/ \frac{z_1 - z_2}{z_1 - z_3}$$

when all the points are finite. If one of them is ∞, the cross ratio is defined as the appropriate limit of the expression above. The following proposition follows by inspection.

PROPOSITION 1.23. *Suppose z_1, z_2, z_3 are distinct points in $\mathbb{C}^*$. The unique Möbius transform taking these points to $1, 0, \infty$ in order is $z \mapsto (z, z_1, z_2, z_3)$.*

It is now clear that to find the unique Möbius transform taking the distinct points z_1, z_2, z_3 into the distinct points w_1, w_2, w_3 in order, one simply has to solve for w in $(w, w_1, w_2, w_3) = (z, z_1, z_2, z_3)$.

EXERCISE 1.24. Find the Möbius transformation that carries 0, i, $-i$ in order into 1, -1, 0.

EXERCISE 1.25. Show that any Möbius transformation which leaves $\mathbb{R} \cup \{\infty\}$ invariant may be written with real coefficients.

EXERCISE 1.26. Show that the map $z \mapsto \frac{z-1}{z+1}$ maps the right half-plane (*i.e.*, the set $\operatorname{Re} z > 0$) onto the interior of the unit circle.

Two points z and z^* are said to be *symmetric with respect to* $\mathbb{R}$ if $z^* = \overline{z}$. If T is a Möbius transform that maps $\mathbb{R} \cup \{\infty\}$ onto itself, then according to Exercise 1.25 one may write T with real coefficients. It follows that Tz and $T(z^*)$ are symmetric with respect to the real axis if and only if z and z^* are. To generalize the concept of symmetry with respect to the real axis to symmetry with respect to any circle in the extended plane we make the following definition.

DEFINITION 1.27. Let Γ be a circle in $\mathbb{C}^*$. Two points z and z^* are said to be *symmetric with respect to* Γ if there is a Möbius transform T which maps Γ onto the real axis for which $T(z^*) = \overline{Tz}$.

By the reasoning just before the definition it is clear that this is a genuine extension of the notion of conjugate points and that z and z^* are symmetric with respect to Γ precisely if $T(z^*) = \overline{Tz}$ for *any* Möbius transform T that takes Γ to the real axis. For, if T and S both take Γ onto the real axis and $T(z^*) = \overline{Tz}$, then $U = ST^{-1}$ maps the real axis onto itself so that $S(z^*) = UT(z^*) = U(\overline{Tz}) = \overline{UTz} = \overline{Sz}$. There is therefore for every z precisely one point z^* so that z, z^* are symmetric with respect to Γ. A similar calculation proves the next theorem.

THEOREM 1.28. *Suppose S is a Möbius transform that takes the circle $\Gamma \in \mathbb{C}^*$ onto the circle $\Gamma' \in \mathbb{C}^*$. Then the points z and z^* are symmetric with respect to Γ if and only Sz and $S(z^*)$ are symmetric with respect to Γ'.*

PROOF. If T maps Γ onto the real axis, then $U = TS^{-1}$ maps Γ' onto the real axis. But $US(z^*) = T(z^*)$ and $USz = Tz$ so that $US(z^*) = \overline{USz}$ if and only if $T(z^*) = \overline{Tz}$. The theorem follows. □

In short, Theorem 1.28 says that symmetry is preserved by Möbius transforms. The next theorem allows us to calculate the symmetric point to any given z and circle.

THEOREM 1.29. *If Γ is a straight line, then z and z^* are symmetric with respect to Γ precisely if they are each others mirror image in Γ. If Γ is a genuine circle with center a and radius R, then a and ∞ are symmetric with respect to Γ. If z is finite and $\neq a$, then z and z^* are symmetric precisely if $(z^* - a)\overline{(z - a)} = R^2$.*

PROOF. If Γ is a straight line it is mapped onto the real axis by a translation or a rotation and these transformations obviously preserve mirror images.

If Γ is a circle with center a and radius R the map $z \mapsto i\frac{z-a-R}{z-a+R}$ takes Γ onto the real axis (since $a + R \mapsto 0$, $a - R \mapsto \infty$ and $a - iR \mapsto 1$). Now a and ∞ are mapped onto $-i$ and i respectively, so they are a symmetric pair. If z has neither of these values a simple calculation shows that z and z^* are mapped onto conjugate points precisely if $(z^* - a)\overline{(z - a)} = R^2$. □

In particular the fact that the center of a circle and ∞ are symmetric with respect to the circle are often very helpful in trying to find maps that take a given circle into another.

EXERCISE 1.30. Find the Möbius transform which carries the circle $|z| = 2$ into $|z + 1| = 1$, the point -2 into the origin, and the origin into i.

EXERCISE 1.31. Find all Möbius transforms that leave the circle $|z| = R$ invariant. Which of these leave the *interior* of the circle invariant?

EXERCISE 1.32. Suppose a Möbius transform maps a pair of concentric circles onto a pair of concentric circles. Is the ratio of the radii invariant under the map?

EXERCISE 1.33. Find all circles that are orthogonal to $|z| = 1$ and $|z - 1| = 4$.

We will end this section by discussing *conjugacy classes* of Möbius transforms.

DEFINITION 1.34. Two Möbius transforms S and T are called *conjugate* if there is a Möbius transform U such that $S = U^{-1}TU$.

Conjugacy is obviously an equivalence relation, *i.e.*, if we write $S \sim T$ when S is conjugate to T, then we have:

(1) $S \sim S$ for any Möbius transform S. (reflexive)
(2) If $S \sim T$, then $T \sim S$ (symmetric)
(3) If $S \sim T$ and $T \sim W$, then $S \sim W$. (transitive)

It follows that the set of all Möbius transforms is split into *equivalence classes* such that every transform belongs to exactly one equivalence class and is equivalent to all the transforms in the same class, but to no others.

EXERCISE 1.35. Prove the three properties above and the statement about equivalence classes. What are the elements of the equivalence class that contains the identity transform?

The concept of conjugacy has importance in the theory of (discrete) *dynamical systems.* This is the study of sequences generated by the iterates of some map, *i.e.*, if S is a map of some set M into itself, one studies sequences of the form $z, Sz, S^2z, \dots$ where $z \in M$. This sequence is called the (forward) *orbit* of z under the map S. One is particularly interested in what happens 'in the long run', *e.g.*, for which z's the sequence has a limit (and what the limit then is), for which z's the sequence is periodic and for which z's there seems to be no discernible pattern at all ('chaos'). Note that if $S = U^{-1}TU$, then $S^n = U^{-1}T^nU$ so that all maps in the same conjugacy class behave qualitatively in the same way, at least with respect to the properties listed above. It therefore seems natural to try to find, in each conjugacy class, some particularly simple map for which the questions above are particularly simple to answer. In other words, one looks for a 'canonical representative' in each equivalence class. We will carry out this for the case of Möbius transforms.

If $S = U^{-1}TU$ and z is a fixpoint of S, then Uz is a fixpoint of T since $TUz = USz = Uz$. If S has only one fixpoint z_0 we may choose V so that $Vz_0 = \infty$. Then VSV^{-1} has only the fixpoint ∞ and is therefore a translation $z \mapsto z + b$ for some $b \neq 0$. If we set $U = \frac{1}{b}V$ it follows that $USU^{-1}z = z + 1$. If S has two fixpoints z_1 and z_2 we may choose U so that $Uz_1 = 0$ and $Uz_2 = \infty$. Then $T = USU^{-1}$ has the fixpoint ∞, so it is linear, $Tz = az + b$, and it also has the fixpoint 0, so $b = 0$. Now set, for $\lambda \neq 0$,

$$T_\lambda z = \begin{cases} z + 1 \text{ for } \lambda = 1 \ , \\ \lambda z \quad \text{ for } 0 \neq \lambda \neq 1 \ . \end{cases}$$

We have then proved most of the following theorem.

THEOREM 1.36. *For every Möbius transform S different from the identity there exists $\lambda \neq 0$ such that $S \sim T_\lambda$. If $T_\lambda \sim T_\mu$, then either $\lambda = \mu$ or $\lambda = 1/\mu$.*

PROOF. It only remains to prove the last statement. But this is clear if $\lambda = 1$, since this is the only value for which T_λ has just one fixpoint. We may therefore assume that λ and μ are both $\neq 1$ (and of course non-zero). But if $UT_\lambda = T_\mu U$ and $Uz = \frac{az+b}{cz+d}$ we obtain

$$\frac{a\lambda z + b}{c\lambda z + d} = \mu \frac{az + b}{cz + d} \tag{1.3}$$

for all z. Since $ad - bc \neq 0$ we can not have $d = c = 0$. If $d \neq 0$, setting $z = 0$ gives $b/d = \mu b/d$ so that $b = 0$ and therefore $a \neq 0$. If now $c \neq 0$, setting $z = -d/c$ we get ∞ on the right but not the left. It follows that $c = 0$ and (5.2) becomes $\lambda = \mu$. On the other hand, if $d = 0$ we must have $c \neq 0$ and so $z = \infty$ gives $a/c = \mu a/c$. It follows that $a = 0$. In this case (5.2) becomes $\lambda = 1/\mu$ and the proof is complete. □

What we have proved is that each conjugacy class different from the class of the identity contains one of the operators T_λ and also $T_{1/\lambda}$, but no other operators of this form. We may therefore with any Möbius transform S associate the corresponding unique (non-ordered) pair $(\lambda, 1/\lambda)$ of reciprocal complex numbers, called the *multiplier* of S. The multiplier is thus a conjugacy invariant. Note that some T_λ leave the interior of certain circles in the extended plane invariant. Namely, T_1 leaves all halfplanes above or below a horizontal line invariant. If $\lambda > 0$ (but $\neq 1$), then T_λ leaves all halfplanes bounded by a line through the origin invariant. Finally, if $|\lambda| = 1$ but $\lambda \neq 1$, then T_λ leaves the interiors and exteriors of any circle concentric with the origin invariant.

On the other hand, if λ is neither positive nor of absolute value 1 there is no disk which is invariant under T_λ. Show this as an exercise! The transforms in the conjugacy class of T_1 are called *parabolic*, those in the conjugacy class of T_λ for some $\lambda > 0$ but $\neq 1$ are called *hyperbolic* and those in the conjugacy class of T_λ for some $\lambda \neq 1$ with $|\lambda| = 1$ are called *elliptic*. The reason for these names will be clear from the result of Exercise 1.37. The remaining Möbius transforms are called *loxodromic*. This is because they are conjugate to a T_λ for which the sequence of iterates $z, T_\lambda z, T_\lambda^2 z, \dots$ lie on a logarithmic spiral, which under stereographic projection becomes a curve known as a loxodrome.

EXERCISE 1.37. Suppose that the coefficients of the transformation

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

are normalized by $ad-bc = 1$. Show that S is elliptic if $0 \leq (a+d)^2 < 4$, parabolic if $(a+d)^2 = 4$, hyperbolic if $(a+d)^2 > 4$ and loxodromic in all other cases. *Hint:* The determinant and the trace $a+d$ of a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is invariant under conjugation by an invertible matrix.

EXERCISE 1.38. Show that a linear transformation which satisfies $S^n = S$ for some integer n is necessarily elliptic.

EXERCISE 1.39. If S is hyperbolic or loxodromic, show that $S^n z$ converges to a fixpoint as $n \to \infty$, the same for all z which are not equal to the other fixpoint. The exceptional fixpoint is called *repelling*, the other one *attractive*. What happens when $n \to -\infty$? What happens in the parabolic and elliptic cases?

EXERCISE 1.40. Find all linear transformations that are rotations of the Riemann sphere.
Hint: The *antipodal point* to a point on the unit sphere is obtained by multiplication by -1. Use the fact that an antipodal pair is mapped onto an antipodal pair by a rotation.

1.6. Polynomials, rational functions and power series

We define a *polynomial* to be a complex-valued function p of a complex variable given by a formula $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ where the coefficients $a_0, \dots a_n$ are complex numbers, $a_n \neq 0$, and n is a non-negative integer, called the *degree* of the polynomial, $\deg p$. The function identically equal 0 is also a polynomial, of degree $-\infty$. The sum of two polynomials of degrees n and m is a polynomial of degree $\leq \max(n, m)$. The product of two polynomials of degrees n and m is a polynomial of degree $n+m$. The *division algorithm* says that if p and q are polynomials, then there are unique polynomials k and r with $\deg r < \deg q$ such that $p = kq + r$. From this follows the *factor theorem* which states that if $p(a) = 0$, then $z - a$ divides p. The proof is simply the observation that since $p(z) = k(z)(z-a)+r$ where r is constant (of degree < 1), then $r = 0$ if and only if $p(a) = 0$. It is of course possible that the quotient k is also divisible by $z - a$. If j is the largest integer such that $(z-a)^j$ divides p, then j is called the *multiplicity* of a as a zero of p.

It also follows from the factor theorem that two polynomials p, q for which $p(z) = q(z)$ for all $z \in \mathbb{C}$ have to be identical, *i.e.*, have the same coefficients.

A very important fact about polynomials (which is only true if we consider polynomials in the complex domain) is the *fundamental theorem of algebra* which says that any non-constant polynomial has a zero. We will prove this later, but assume it for the present. Combining the fundamental theorem of algebra with the factor theorem it easily follows that if we add up the multiplicities of all the zeros of a polynomial p ('count the zeros with their multiplicities'), the sum will be the degree of p.

Also for complex functions the concepts of limit and continuity are of central importance. However, since complex numbers are just vectors in $\mathbb{R}^2$, where we in addition has defined a multiplication, we can take these concepts over from the calculus of several real variables. For reference we nevertheless state the definitions

Definition 1.41. Suppose f is a complex-valued function of either a real or complex variable, with domain $\Omega \subset \mathbb{R}$ or $\Omega \subset \mathbb{C}$.

- If a is a point in the closure of Ω, we say that $\lim_{z\to a} f(z) = A$ if A is a complex number such that for every $\varepsilon > 0$ there is a $\delta > 0$ with the property that $|f(z) - A| < \varepsilon$ whenever $z \in \Omega$ and $0 < |z-a| < \delta$.
- If $a \in \Omega$ we say that f is continuous at a if $\lim_{z\to a} f(z) = f(a)$.

All the standard calculation rules for limits and continuity familiar from calculus continue to hold in this context, with exactly the same proofs, so we will not dwell on this. We also remind the reader of the concept of *uniform convergence* for a sequence of functions.

Definition 1.42. Suppose f and $f_1, f_2, \dots$ are complex-valued function of either a real or complex variable, with domain $\Omega \subset \mathbb{R}$ or $\Omega \subset \mathbb{C}$. If $K \subset \Omega$ we say that $f_j \to f$ uniformly on K if for every $\varepsilon > 0$ there is a real number N such that $|f_j(z) - f(z)| < \varepsilon$ for all $z \in K$ if $j \geq N$.

As a function in $\mathbb{C}$ a polynomial is continuous; this follows easily since constant polynomials and the polynomial z obviously are continuous, and any other polynomial can be built up from these by multiplications and additions so the continuity follows from the standard calculation rules for limits.

A *rational* function is a quotient $r(z) = p(z)/q(z)$ where p and q are polynomials and q not identically 0 (if q is constant r is a polynomial).

It follows that r is continuous as a function in $\mathbb{C}$ in all points which are not zeros of q. We may assume that p and q have no common non-constant polynomial factors (the common divisor to two polynomials of largest degree can always be found by a purely algebraic device, the Euclidean algorithm). Hence p and q have no common zeros. It follows that $r(z) \to \infty$ as z tends to any zero of q. As $z \to \infty$ we have $r(z) \to 0$ if $\deg p < \deg q$ and $r(z) \to \infty$ if $\deg p > \deg q$. If $\deg p = \deg q$, then $r(z) \to a/b$ where a and b are the highest order coefficients of p and q respectively.

A *power series* is a series

$$\sum_{n=0}^{\infty} a_n(z-a)^n \tag{1.4}$$

where $a, a_0, a_1, a_2, \dots$ are given complex numbers and z a complex variable. In many respects such series behave like 'polynomials of infinite order' and that is actually how they were viewed until the end of the 19:th century. The very first question to ask is of course: For which values of z does the series converge? In order to answer this question we make the following definition.

Definition 1.43. Let the *radius of convergence* for (1.4) be

$$R = \sup\{r \geq 0 \mid a_0, a_1 r, a_2 r^2, \dots \text{ is a bounded sequence }\}\ .$$

Then R is either a number ≥ 0 or $R = \infty$.

The explanation for the definition is in the following theorem.

Theorem 1.44. *For $|z-a| > R$ the series* (1.4) *diverges and for $|z-a| < R$ it converges absolutely. The convergence is uniform on every compact subset of $|z-a| < R$.*

In order to prove the theorem we need a few results which should be well known in the context of functions of a real variable.

Theorem 1.45. *An absolutely convergent complex series is convergent.*

Proof. For any complex number z we have $|\operatorname{Re} z| \leq |z|$ and $|\operatorname{Im} z| \leq |z| \leq |\operatorname{Re} z| + |\operatorname{Im} z|$. Hence, if $\sum |a_n|$ is convergent, then by comparison the real series $\sum \operatorname{Re} a_n$ and $\sum \operatorname{Im} a_n$ are absolutely convergent, to x and y say. The theorem now follows from

$$|\sum_{n=0}^{N} a_n - x - iy| \leq |\sum_{n=0}^{N} \operatorname{Re} a_n - x| + |\sum_{n=0}^{N} \operatorname{Im} a_n - y| \to 0 \text{ as } N \to \infty\ .$$

□

The next theorem is the complex version of what is usually known under the silly name of *Weierstrass' M-test.*

THEOREM 1.46. *Let A be a subset of $\mathbb{C}$ and $f_1, f_2, \ldots$ a sequence of complex functions defined on A and such that $|f_n(z)| \le a_n$ for all $z \in A$ and $n = 1, 2, \ldots$. If $\sum_{n=0}^{\infty} a_n$ converges, then $\sum_{n=0}^{\infty} f_n(z)$ converges uniformly in A.*

PROOF. By Theorem 1.45 the series $\sum f_n(z)$ converges absolutely for every $z \in A$; call the sum $s(z)$. Then

$$|s(z) - \sum_{n=0}^{N} f_n(z)| = |\sum_{n=N+1}^{\infty} f_n(z)| \le \sum_{n=N+1}^{\infty} |f_n(z)| \le \sum_{n=N+1}^{\infty} a_n \,.$$

The last member does not depend on z and tends to 0 as $N \to \infty$. The theorem follows. □

PROOF OF THEOREM 1.44. If $|z - a| > R$ then $a_n(z-a)^n$, $n = 0, 1, 2, \ldots$ is an unbounded sequence and hence can not converge to 0. Hence the power series diverges.

If $r < R$, then there exists $\rho > r$ such that $a_n \rho^n$, $n = 0, 1, 2, \ldots$ is a bounded sequence; let C be a bound. Then if $|z - a| \le r$ we have $|a_n(z-a)^n| \le |a_n| r^n = |a_n \rho^n| (r/\rho)^n \le C(r/\rho)^n$. Since a geometric series with quotient $0 \le r/\rho < 1$ is convergent, the theorem follows from Theorem 1.46 (any compact subset of $|z - a| < R$ is a subset of $|z - a| \le r$ for some $r < R$). □

Here is the complex version of another well known theorem.

THEOREM 1.47. *Suppose $f_1, f_2, \ldots$ is a sequence of continuous, complex functions converging uniformly to f on the set M. Then f is continuous on M.*

The proof is word for word the same as in the case of real functions so we will not repeat it here. We have the following corollary of Theorems 1.44 and 1.47.

COROLLARY 1.48. *If R is the radius of convergence of* (1.4)*, then* (1.4) *is a continuous function of z for $|z - a| < R$.*

PROOF. The partial sums of a power series are polynomials and therefore continuous. Since any z in the disk $|z - a| < R$ is an interior point of a compact subset of the disk the claim follows from Theorems 1.44 and 1.47. □

So far we have said nothing about convergence on the boundary of the circle of convergence. There is a good reason for this; nothing much can be said in general. One can have divergence at every point of the circle, convergence at some points and divergence at others or one can have absolute convergence at every point of the circle. A general result by Carleson (1966) says that if $\sum_{n=0}^{\infty} |a_n R^n|^2$ converges, then (1.4) will converge 'almost everywhere' on the circle, in the sense of Lebesgue integration. On the other hand, there are examples (the first one given by Kolmogorov in 1926) for which $a_n R^n \to 0$ such that (1.4) diverges for *every* point on the circle.

EXERCISE 1.49. Show that $\sum_{n=0}^{\infty} z^n$ diverges at every point of its circle of convergence, that $\sum_{n=0}^{\infty} z^n/n$ converges for some but not all points on its circle of convergence and that $\sum_{n=0}^{\infty} z^n/n^2$ converges absolutely for all points on its circle of convergence.

It is often possible to find the radius of convergence for a given power series by inspection and use of the definition. As an aid in cases where this might be difficult we have the following two theorems.

THEOREM 1.50. $\overline{\lim_{n\to\infty}} |a_n|^{1/n} = 1/R$. *This is to be interpreted by using the conventions* $1/0 = \infty$ *and* $1/\infty = 0$.

Here we have defined $\overline{\lim_{n\to\infty}} c_n = \limsup_{n\to\infty} c_n = \lim_{n\to\infty} \sup_{k\geq n} c_k$ for a real sequence $c_0, c_1, \dots$.

PROOF. Let $L = \overline{\lim_{n\to\infty}} |a_n|^{1/n}$. If $r < 1/L$, then $|a_n|^{1/n} < 1/r$ for all sufficiently large n. Hence $|a_n r^n| < 1$ for such n, so the sequence $a_n r^n$, $n = 0, 1, 2, \dots$ is bounded. Hence $1/L \leq R$.

If $r > 1/L$, then there exists ρ, $r > \rho > 1/L$, so that $|a_n|^{1/n} > 1/\rho$ for infinitely many n. Hence $|a_n r^n| = |a_n \rho^n|(r/\rho)^n > (r/\rho)^n$ for infinitely many n. Since $(r/\rho)^n \to \infty$ the sequence $a_n r^n$, $n = 0, 1, 2, \dots$ can not be bounded and so $1/L \geq R$ and the proof is complete (check the cases $L = 0$ and $L = \infty$ separately). □

THEOREM 1.51. *If* $\lim_{n\to\infty} |a_n/a_{n+1}|$ *exists it is equal to* R.

PROOF. Let $L = \lim_{n\to\infty} |a_n/a_{n+1}|$. If $r < L$, then $r < |a_n/a_{n+1}|$ for all sufficiently large n. It follows that $|a_{n+1} r^{n+1}| < |a_n r^n|$ for such n so that $a_n r^n$, $n = 0, 1, 2, \dots$, being of eventually decreasing absolute value, is bounded. Hence $L \leq R$.

On the other hand, if $r > L$, choose ρ so that $r > \rho > L$. Then there exists N such that $\rho > |a_n/a_{n+1}|$ for $n \geq N$. It follows for such n that $|a_{n+1}\rho^{n+1}| > |a_n \rho^n| > \dots > |a_N \rho^N|$. We therefore have $|a_n r^n| = |a_n \rho^n|(r/\rho)^n > |a_N \rho^N|(r/\rho)^n \to \infty$ as $n \to \infty$ (we may assume $a_N \neq 0$ since this must be true eventually for $\lim_{n\to\infty} |a_n/a_{n+1}|$ to exist). It follows that $L \geq R$ so the theorem is proved. □

EXERCISE 1.52. Find the radius of convergence for the following power series:

(a) $\sum_{n=0}^{\infty} n^2 z^2$

(b) $\sum_{n=0}^{\infty} n^n z^n$

(c) $\sum_{n=0}^{\infty} \frac{z^n}{(n!)^2}$

(d) $\sum_{n=0}^{\infty} 2^{-n} z^n$

(e) $\sum_{n=2}^{\infty} \frac{z^n}{\ln n}$

(f) $\sum_{n=1}^{\infty} \frac{2^n}{n^2} z^n$

(g) $\sum_{n=1}^{\infty} \frac{z^n}{\arctan n}$

(h) $\sum_{n=0}^{\infty} (n^2 + 2^n) z^n$

(i) $\sum_{n=0}^{\infty} \cos(n\pi/4) z^n$

(j) $\sum_{n=1}^{\infty} \frac{2^n + 2^{-n}}{n} z^n$

(k) $\sum_{n=1}^{\infty} n^{1/n} z^n$

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

(m) $\sum_{n=0}^{\infty} (\sqrt{n^2+1} - \sqrt{n^2-1}) z^n$

(n) $\sum_{n=0}^{\infty} \frac{(n+2)^3}{3^{2n+1}} z^n$

(o) $\sum_{n=0}^{\infty} \frac{2^{1/n}}{2^n} z^n$

(p) $\sum_{n=1}^{\infty} \frac{2^n}{\sqrt{n}} z^{2n}$

(q) $\sum_{n=1}^{\infty} \frac{z^n}{1 + 2 + \cdots + n}$

(r) $\sum_{n=1}^{\infty} (1 + \frac{1}{2} + \cdots + \frac{1}{n}) z^n$

(s) $\sum_{n=0}^{\infty} z^{n^2}$

(t) $\sum_{n=1}^{\infty} \frac{z^{n^2}}{n^2}$

(u) $\sum_{n=1}^{\infty} \frac{z^{n^2}}{n}$

(v) $\sum_{n=0}^{\infty} \frac{z^{4n}}{2^n + n^2}$

(w) $\sum_{n=0}^{\infty} (\sin n) z^n$

(x) $\sum_{n=1}^{\infty} \frac{n^n}{n!} z^n$

(y) $\sum_{n=1}^{\infty} (\arctan \frac{1}{n}) z^n$

(z) $\sum_{n=0}^{\infty} \binom{a}{n} z^n$

(å) $\sum_{n=1}^{\infty} \binom{2n}{n} z^n$

(ä) $\sum_{n=0}^{\infty} \frac{(3n+1)!}{(n!)^4} z^n$

(ö) $\sum_{n=0}^{\infty} 2^{\sqrt{n}} z^n$

(ööh) $\sum_{n=1}^{\infty} (1 + \frac{1}{n})^{n^2} z^n$

CHAPTER 2

Analytic functions

2.1. Conformal mappings and analyticity

DEFINITION 2.1. A map $f : \Omega \to \mathbb{C}$, where Ω is an open subset of $\mathbb{C}$, is called *conformal* if it satisfies the following conditions:

(1) As a map from a subset of $\mathbb{R}^2$ into $\mathbb{R}^2$, f is differentiable.
(2) f preserves angles of intersection between smooth curves.
(3) f preserves orientation in the sense that the determinant of the total derivative of the map is > 0.

To explain the definition in more detail, note that if $z = x + iy$, where x and y are real, then $f(z) = u(x, y)+iv(x, y)$ where u and v are real-valued functions of two real variables, so the action of f can also be described by the mapping $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \mapsto \begin{pmatrix} u(x,y) \\ v(x,y) \end{pmatrix}$. The first condition of the definition then says that this map should be differentiable. Recall that this implies that the partial derivatives u_x, u_y, v_x and v_y exist and that the chain rule can be applied when composing with other differentiable maps. Also recall that the existence of the partials is not enough to guarantee differentiability, but if the partials are *continuous*, then the map is differentiable.

We measure the angle between two non-zero vectors α and $\beta \in \mathbb{R}^n$ by the expression $\frac{\langle \alpha,\beta \rangle}{\|\alpha\|\|\beta\|}$, where $\langle \cdot, \cdot \rangle$ is the usual scalar product and $\| \cdot \|$ the Euclidean norm (the actual angle is arccos of this). If $t \mapsto \gamma(t) = \gamma_1(t) + i\gamma_2(t)$ is a differentiable curve in Ω, then its tangent vector is γ' or, expressed as a column vector, $\begin{pmatrix} \gamma_1' \\ \gamma_2' \end{pmatrix}$. The image $f \circ \gamma$ of γ under f is another differentiable curve. According to the chain rule its tangent vector is $J \begin{pmatrix} \gamma_1' \\ \gamma_2' \end{pmatrix}$ where $J = \left(\begin{smallmatrix} u_x & u_y \\ v_x & v_y \end{smallmatrix}\right)$ is the *Jacobi matrix* or *total derivative* of the map. The second point of the definition then means that the linear map given by the Jacobi matrix maps any two vectors onto two vectors which make the same angle as the original vectors. The third point simply means that the Jacobian $\left|\begin{smallmatrix} u_x & u_y \\ v_x & v_y \end{smallmatrix}\right| = u_x v_y - u_y v_x \geq 0$ in Ω.

EXERCISE 2.2. Show that the map $z \mapsto \overline{z}$ satisfies the two first points of Definition 2.1, but reverses the orientation (*i.e.*, the Jacobian is < 0). Such a map is called *anti-conformal*. Show that any anti-conformal map is of the form $z \mapsto \overline{f(z)}$ where f is conformal.

This shows that there is really no need to study anti-conformal maps separately from conformal maps.

We have the following basic theorem.

THEOREM 2.3. *Suppose $f = u + iv$ is conformal in Ω. Then the partials of u and v satisfy the* Cauchy-Riemann equations

$$\begin{cases} u_x = v_y, \\ u_y = -v_x. \end{cases} \tag{2.1}$$

Conversely, if $\binom{u}{v}$ satisfy the Cauchy-Riemann equations, the corresponding map is differentiable and its Jacobi matrix does not vanish at any point in Ω, then the map is conformal.

The slight asymmetry in the statement of Theorem 2.3 will be removed later; in fact, we will later show (see the discussion after Corollary 4.23) that the Jacobi matrix can not vanish in any domain where the map is conformal.

PROOF. Suppose f is conformal and let α and β be the column vectors in the Jacobi matrix. Since multiplication by the Jacobi matrix preserves angles the vectors $\binom{1}{0}$ and $\binom{0}{1}$ are mapped onto orthogonal vectors, *i.e.*, α and β are orthogonal. Similarly, the vectors $\binom{1}{1}$ and $\binom{1}{-1}$ are mapped onto orthogonal vectors. Since the scalar product of $\alpha+\beta$ and $\alpha-\beta$ is $\|\alpha\|^2 - \|\beta\|^2$ it follows that α and β also have the same length, so that $\binom{u_x}{v_x} = \pm\binom{v_y}{-u_y}$. The Jacobian is therefore $\pm(u_x^2 + u_y^2)$. To preserve orientation we must choose the plus sign. It follows that any conformal map satisfies the Cauchy-Riemann equations.

Conversely, if the map satisfies the Cauchy-Riemann equations, is differentiable, and has non-vanishing Jacobi matrix, then this matrix is $\sqrt{u_x^2 + u_y^2}\,\mathcal{O}$ where $\mathcal{O}$ is an orthogonal matrix with determinant one, *i.e.*, a rotation. The map is therefore conformal. □

EXERCISE 2.4. Show that the map $z \mapsto z^2$ is conformal in any open set not containing the origin.

We will now connect the geometric notion of a conformal map with the analytic notion of complex derivative. We first need a definition.

DEFINITION 2.5. A complex-valued function f defined in an open subset of $\mathbb{C}$ is said to be *differentiable at a* if

$$\lim_{z \to a} \frac{f(z) - f(a)}{z - a}$$

exists. The limit is called the derivative of f at a and is denoted by $f'(a)$.

All the elementary properties of derivatives that we know from the theory of a real function of one variable continue to hold, with essentially the same proofs. We collect some such properties in the next theorem.

THEOREM 2.6. *Suppose that f is differentiable at a. Then*

1. *f is continuous at a.*
2. *Cf is differentiable at a with derivative $Cf'(a)$ for any constant C.*
3. *If g is differentiable at a, then so is $f+g$, fg and, if $g(a) \neq 0$, f/g and*
$$(f+g)'(a) = f'(a) + g'(a)$$
$$(fg)'(a) = f'(a)g(a) + f(a)g'(a)$$
$$(f/g)'(a) = (f'(a)g(a) - f(a)g'(a))/g(a)^2$$
4. *If g is differentiable at $f(a)$, then $g \circ f$ is differentiable at a and the* ***chain rule*** *$(g \circ f)'(a) = g'(f(a))f'(a)$ is valid.*
5. *If $f'(a) \neq 0$ and the inverse f^{-1} is defined in a neighborhood of $f(a)$ and is continuous at $b = f(a)$, then the inverse is differentiable at b and $(f^{-1})'(b)) = 1/f'(a)$.*
6. *Polynomials and rational functions are differentiable where they are defined (as functions in $\mathbb{C}$) and their derivatives are calculated in the same way as in the case of real polynomials and rational functions.*

EXERCISE 2.7. Prove Theorem 2.6.

EXERCISE 2.8. Show that any branch of $\sqrt{z}$ is differentiable for $z \neq 0$ and calculate the derivative.

We will later prove (see Corollary 4.23) that if f is differentiable in a *neighborhood* of a, then the assumption $f'(a) \neq 0$ implies all the other assumptions of Theorem 2.6 (5). We are now ready to state the second main result of this section.

THEOREM 2.9. *$f = u + iv$ has a complex derivative at $z = a + ib$ if and only the map $\binom{x}{y} \mapsto \binom{u(x,y)}{v(x,y)}$ is differentiable and the Cauchy-Riemann equations (2.1) are satisfied at (a, b). We also have $f'(z) = u_x(x, y) + iv_x(x, y) = v'_y(x, y) - iu'_y(x, y)$.*

PROOF. For f to be differentiable with derivative $a+ib$ at $z = x+iy$ means that

$$|f(z + w) - f(z) - (a + ib)w| = |w|r(w) \text{ where } r(w) \to 0 \text{ as } w \to 0 .$$

Similarly, for $\binom{u}{v}$ to be differentiable at (x, y) with a Jacobi matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ satisfying the Cauchy-Riemann equations means that

$$\left\| \begin{pmatrix} u(x+h, y+k) - u(x, y) \\ v(x+h, y+k) - v(x, y) \end{pmatrix} - \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} \right\| = \|(h, k)\| \|\rho(h, k)\|$$

where $\rho(h, k) \to 0$ as $(h, k) \to 0$. But if $f = u + iv$ and $w = h + ik$ the left hand sides of these two relations are equal so the theorem follows. $\square$

If g is a complex-valued function of a real variable with real and imaginary parts u and v respectively, we say that g is differentiable if u and v are, and define $g' = u' + iv'$. Using the equivalence in Theorem 2.9 it then follows from the chain rule for vector-valued functions of several variables that if g is a complex-valued, differentiable function of one variable with range in the domain of a complex differentiable function f, then the chain rule $\frac{d}{dt} f(g(t)) = f'(g(t))g'(t)$ is valid.

There are some alternative ways of expressing the Cauchy-Riemann equations which are sometimes used. If we view f as a function of $x = \operatorname{Re} z$ and $y = \operatorname{Im} z$ it is clear that the Cauchy-Riemann equations are equivalent to $f'_x + if'_y = 0$. Note also that this means that if the complex derivative f' exists, then $f' = f'_x = -if'_y$.

The differential of f as a function of (x, y) is $df = f'_x dx + f'_y dy$, in particular $dz = dx + idy$ and $d\bar{z} = dx - idy$. We can therefore write $df = \frac{1}{2}(f'_x - if'_y)dz + \frac{1}{2}(f'_x + if'_y)d\bar{z}$ and for this reason one introduces the notation $\frac{\partial f}{\partial z} = \frac{1}{2}(f'_x - if'_y)$ and $\frac{\partial f}{\partial \bar{z}} = \frac{1}{2}(f'_x + if'_y)$. The Cauchy-Riemann equations may then be expressed as $\frac{\partial f}{\partial \bar{z}} = 0$, and then $\frac{\partial f}{\partial z} = f'$.

We also have $df = \frac{\partial f}{\partial z}\, dz + \frac{\partial f}{\partial \bar{z}}\, d\bar{z}$, so if we introduce the *holomorphic differential* ∂ by $\partial f = \frac{\partial f}{\partial z}\, dz$ and the *anti-holomorphic differential* $\overline{\partial}$ by $\overline{\partial} f = \frac{\partial f}{\partial \bar{z}}\, d\bar{z}$ we have $d = \partial + \overline{\partial}$, and the Cauchy-Riemann equations may also be written as $\overline{\partial} f = 0$. An analytic function is therefore a solution of the homogeneous $\overline{\partial}$ equation (pronounced d-bar equation). This also means that f is analytic if $df = \partial f = \frac{\partial f}{\partial z}\, dz = f'(z)\, dz$.

DEFINITION 2.10. A function $f : M \to \mathbb{C}$ where $M \subset \mathbb{C}$ is called *analytic in* M if it is defined and differentiable in some open set containing M.

Note that to say that f is analytic in a means more than just having a derivative in a; f has to be differentiable in a whole *neighborhood* of a. A function which is analytic in an open set Ω is often said to be *holomorphic* in Ω, and the set of functions which are holomorphic in Ω is often denoted $H(\Omega)$.

Practically always the only domains of analyticity that are of interest are *connected*. There are two notions of connectivity in common use, arcwise connectivity which is used in calculus, and the more general notion of connectivity from topology. Since we are always considering open domains, it makes no difference whether you use one or the other, since they are equivalent for open sets. For convenience, we will use the word *region* to denote an open, connected subset of the complex plane (or, occasionally, of the Riemann sphere).

We end by a simple result that we will use in the next section.

THEOREM 2.11. *Suppose f is analytic in a region Ω and that $f'(z) = 0$ for all $z \in \Omega$. Then f is constant in Ω.*

We will actually prove much stronger results later; in fact it will be enough to assume that the zeros of f' has a point of accumulation in Ω for the conclusion to be valid.

PROOF. If $z \in \Omega$ and $w \in \mathbb{C}$ is sufficiently close to z, then the line segment between z and w is entirely in Ω. For $0 \le t \le 1$ we then obtain $\frac{d}{dt} f(z + t(w - z)) = f'(z + t(w - z))(w - z) = 0$ using the remark after the proof of Theorem 2.9. Thus $\operatorname{Re} f$ and $\operatorname{Im} f$ are constant on the line segment. In particular, $f(z) = f(w)$ so that f is *locally* constant. Now pick $a \in \Omega$ and let $A = \{z \in \Omega \mid f(z) = f(a)\}$. Then A is open by what we just saw. But also $\Omega \setminus A$ is open for the same reason. Since $a \in A$ we have $A \neq \varnothing$. Since Ω is connected we therefore must have $\Omega \setminus A = \varnothing$, *i.e.*, $A = \Omega$. In other words, f is constant. □

2.2. Analyticity of power series; elementary functions

We will first continue the study of power series begun in Chapter 1.6. First of all, if a power series really behaves 'like a polynomial of infinite order', then we should be able to differentiate the series like a finite sum, *i.e.*, term by term, and actually obtain the derivative of the sum of the series.

In order to prove this, we first note that the usual derivative and integral of a function of one variable extends to the case of a complex-valued function of a real variable in an obvious manner. If f is such a function, with real and imaginary parts u and v, we simply define $f'(t) = u'(t) + iv'(t)$ and $\int_a^b f(t)\,dt = \int_a^b u(t)\,dt + i\int_a^b v(t)\,dt$. This means that we define f as differentiable respectively integrable if its real and imaginary parts have these properties.

It immediately follows that the fundamental theorem of calculus $\frac{d}{dt}\int_a^t f = f(t)$ holds also for complex-valued, continuous functions f. It is also more or less obvious that the usual calculation rules for derivatives and integrals continue to hold. In particular, the integral is interval additive and linear, so that

$$\int_a^b f = \int_a^c f + \int_c^b f$$

$$\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g$$

for $a < c < b$ and arbitrary constants α and β if f and g are both integrable on $[a, b]$. We also have the *triangle inequality*

$$\Big|\int_a^b f(t)\,dt\Big| \le \int_a^b |f(t)|\,dt. \tag{2.2}$$

This is less obvious, but follows from

$$\operatorname{Re}\left(e^{i\theta}\int_a^b f\right) = \int_a^b \operatorname{Re}(e^{i\theta} f) \leq \int_a^b |f|$$

by choosing $\theta = -\arg(\int_a^b f)$.

As already mentioned we also have the chain rule $\frac{d}{dt}f(g(t)) = f'(g(t))g'(t)$ if f is analytic and g is a differentiable complex-valued function of a real variable. Thus, if f is analytic in a region containing the line segment connecting z and $z+h$, then $\frac{d}{dt}f(z+th) = hf'(z+th)$ so that $\frac{1}{h}(f(z+h)-f(z)) = \int_0^1 f'(z+th)\,dt$ if the derivative is continuous. An immediately consequence is the following lemma.

Lemma 2.12. *Suppose f is analytic with continuous derivative in a compact set K containing the line segment connecting z and $z+h$ where $h \neq 0$. Then we have $|\frac{1}{h}(f(z+h)-f(z)| \leq \sup_K |f'|$.*

Proof. By the triangle inequality we obtain

$$\Big|\frac{f(z+h)-f(z)}{h}\Big| = \Big|\int_0^1 f'(z+th)\,dt\Big| \leq \int_0^1 |f'(z+th)|\,dt \leq \sup_K |f'|.$$

□

Exercise 2.13. Prove the theorem without assuming f' to be continuous.
Hint: Use the mean value theorem on $\operatorname{Re}(e^{i\theta}f(z+th))$.

We can now state our theorem about differentiating power series.

Theorem 2.14. *If the series $f(z) = \sum_{k=0}^{\infty} a_k(z-a)^k$ has convergence radius R, then f has derivatives of all orders for $|z-a| < R$. The derivatives are calculated by term by term differentiation, and the resulting series all have radius of convergence R. In particular, $f'(z) = \sum_{k=1}^{\infty} ka_k(z-a)^{k-1}$.*

Proof. We will prove the statement for the first derivative. The statement for the higher derivatives then follows immediately. Clearly

$$g(z) = \sum_{k=1}^{\infty} ka_k(z-a)^{k-1}$$

has the same radius of convergence as $\sum_{k=1}^{\infty} ka_k(z-a)^k$ and since $\sqrt[k]{k} \to 1$ as $k \to \infty$, it follows from Theorem 1.50 that g has the same radius of convergence as f.

If $r < R$ the series $\sum_{k=1}^{\infty} k|a_k|r^{k-1}$ converges, and

$$\frac{1}{h}(f(z+h)-f(z)) = \sum_{k=1}^{\infty} \frac{a_k}{h}((z+h-a)^k - (z-a)^k).$$

Now fix z, $|z-a|<r$. Then the terms of this series are continuous functions of h, with value $ka_k(z-a)^{k-1}$ for $h=0$. By Lemma 2.12 the terms have absolute value less than $k|a_k|r^{k-1}$ if $|z-a|<r$ and $|z+h-a|\le r$, so according to Theorems 1.46 and 1.47 the sum is a continuous function of h in $|z+h-a|\le r$. For $h\neq 0$ its value is $\frac{1}{h}(f(z+h)-f(z))$ and for $h=0$ the value is $g(z)$. Thus f is differentiable and $f'(z)=g(z)$ for any z satisfying $|z-a|<R$. □

We will use Theorem 2.14 to introduce some more elementary functions. It is clear that the series $\sum_{k=0}^{\infty}\frac{z^k}{k!}$ converges for all z so that the following definition is meaningful.

Definition 2.15. For any $z\in\mathbb{C}$, let

(1) $e^z=\sum_{k=0}^{\infty}\frac{z^k}{k!}$,
(2) $\sin z=\frac{e^{iz}-e^{-iz}}{2i}$,
(3) $\cos z=\frac{e^{iz}+e^{-iz}}{2}$.

These are all analytic functions in the whole plane. Such a function is called *entire*. From the definition follows immediately that $e^0=1$, $\sin 0=0$ and $\cos 0=1$. Furthermore, $\frac{d}{dz}e^z=e^z$, $\frac{d}{dz}\cos z=-\sin z$ and $\frac{d}{dz}\sin z=\cos z$. It also follows that sin is odd ($\sin(-z)=-\sin z$) and cos even ($\cos(-z)=\cos z$) and that we have the power series expansions $\sin z=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}z^{2k+1}$ and $\cos z=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}z^{2k}$.

Theorem 2.16. *The functions of Definition 2.15 satisfy the following functional equations:*

(1) $e^{z+w}=e^ze^w$, *for any complex numbers z and w.*
(2) $\sin(z+w)=\sin z\cos w+\cos z\sin w$,
(3) $\cos(z+w)=\cos z\cos w-\sin z\sin w$

for any complex numbers z and w.

Note that the particular case $w=-z$ of (1) shows that $e^{-z}e^z=1$ so that $e^z\neq 0$ for all $z\in\mathbb{C}$.

Proof. Given $w\in\mathbb{C}$, let $f(z)=e^{-z}e^{z+w}$. This is an entire function with derivative $f'(z)=-e^{-z}e^{z+w}+e^{-z}e^{z+w}=0$ so it is constant by Theorem 2.11. Setting $z=0$ we obtain $e^{-z}e^{z+w}=e^w$ for all z and w. The special case $w=0$ shows that $e^{-z}e^z=1$ so that $e^{-z}=1/e^z$. The first formula now follows; the other formulas follow immediately from this by inserting the definitions of sin and cos in the formulas to the right. □

Theorem 2.17. *We have $|e^z|=e^{\operatorname{Re} z}$. There exists a smallest real number $\pi>0$ such that $\sin\pi=0$, and this number satisfies $2.8<\pi<3.2$. The exponential function has period $2\pi i$, the functions* sin *and* cos *period 2π. All other periods are integer multiples of these.*

PROOF. First note that $e^{x+iy} = e^x(\cos y + i\sin y)$ by Theorem 2.16. Since the coefficients in the power series for e^z are all real it follows that as a function of a real variable, e^x is real-valued. It is also > 0, since it is continuous, never $= 0$ and $e^0 = 1 > 0$. Since it is also its own derivative it follows that it is strictly increasing (and strictly convex). For similar reasons cos and sin are real-valued for real arguments and since

$$\cos^2 z + \sin^2 z = 1 \tag{2.3}$$

(take $w = -z$ in Theorem 2.16 (3)) it follows that $\cos y + i\sin y$ is a point on the unit circle for $y \in \mathbb{R}$. Hence $|e^z| = e^{\operatorname{Re} z}$.

We next note that if a real, continuous and non-constant function is periodic, then all its periods are integer multiples of its *smallest* positive period. First of all, since y is a period if and only if $-y$ is, there are positive periods if there are any. Next, if there are arbitrarily small periods > 0, then given x and $\varepsilon > 0$ we can find a period a, $0 < a < \varepsilon$ and an integer p such that $|ap - x| < \varepsilon$. Now $f(0) = f(ap)$ and by continuity $f(ap) \to f(x)$ as $\varepsilon \to 0$, so f is constant. Also note that the set of periods of f is closed, since if $y_j \to y$ and all y_j are periods, then $f(y + x) = \lim f(y_j + x) = f(x)$, so that also y is a period. If f is non-constant it therefore has a smallest positive period a, and if b is another period, then for any integer q, $b - aq$ is a period. But if q is the integer quotient and r the remainder when dividing b by a, then $0 \le r = b - aq < a$. So, a can not be the smallest positive period unless $r = 0$.

If now w is a period for the exponential function so that $e^z = e^{z+w}$ for all z, we see that this is equivalent to $e^w = 1$. Taking absolute values it follows that $\operatorname{Re} w = 0$. Setting $w = iy$ we see that y is a real period for both sin and cos. Note that neither of these functions can have non-real periods since we immediately obtain from Theorem 2.16 (2) respectively 2.16 (3) that w is a period of either of these functions if and only if $\sin w = 0$ and $\cos w = 1$. By (2.3) the first of these relations follows from the second, which may be rewritten as $(e^{iw} - 1)^2 = 0$. This is true if and only if iw is a period for e^z. Therefore, sin and cos have the same periods, they are real and y is the smallest positive period of the trigonometric functions if and only if it is the smallest positive number for which $\cos y = 1$.

Now $\cos y = 1 - 2\sin^2 \frac{y}{2}$ according to Theorem 2.16 (3) and (2.3). It follows that y is the smallest positive number for which $\cos y = 1$ if and only if $\frac{y}{2}$ is the smallest positive zero of sin. According to (2.3) we must then have $\cos \frac{y}{2} = -1$ and there can be no smaller positive numbers for which cos takes the value -1. Now $\cos \frac{y}{2} = 2\cos^2 \frac{y}{4} - 1$ so that $\frac{y}{2}$ has this property if and only if $\frac{y}{4}$ is the smallest positive zero for cos. It now only remains to show that cos actually has a smallest positive zero, and to estimate its value.

Since cos is continuous the set of its non-negative zeros is a closed set and therefore has a smallest element if it is non-empty. By (2.3) we have $\cos x \leq 1$ for real x and integrating this from 0 to $x > 0$ four times we get in turn $\sin x \leq x$, $1 - \cos x \leq \frac{x^2}{2}$, $x - \sin x \leq \frac{x^3}{6}$ and $\frac{x^2}{2} - 1 + \cos x \leq \frac{x^4}{24}$. It follows that for $x > 0$ we have $1 - \frac{x^2}{2} \leq \cos x \leq 1 - \frac{x^2}{2} + \frac{x^4}{24}$ (this may also be deduced from the fact that the power series for cos is an alternating series). The first positive zeros of the two polynomials are $\sqrt{2} > 1.4$ and $\sqrt{6 - 2\sqrt{3}} < 1.6$ respectively. It follows that cos has a smallest positive zero which is in the interval $(1.4, 1.6)$. The proof is now complete. □

One may easily continue to define strictly all the usual (real) functions of elementary calculus and prove all the usual properties of them. We will assume this done; in particular x is the arclength of the arc of the unit circle beginning at 1 and ending at $e^{ix} = \cos x + i \sin x$, so x is the angle between the rays through these points starting at 0. We will also use the common properties of the inverse tangent function.

If we want to extend the definition of the logarithm to the complex domain, we should find the inverse of the exponential function. However, since the exponential function is periodic it has no inverse unless we restrict its domain appropriately (*cf.* the definition of the inverse trigonometric functions). To see how to do this, let us attempt to calculate the inverse of the exponential function, *i.e.*, to solve the equation $z = e^w$ for a given z.

We first note that we must assume $z \neq 0$, since the exponential function never vanishes. Taking absolute values we find that $|z| = e^{\operatorname{Re} w}$ so that $\operatorname{Re} w = \ln |z|$, where ln is the usual natural logarithm of a positive real number. We also obtain $\frac{z}{|z|} = e^{i \operatorname{Im} w}$. Now $\cos 0 = 1$ and $\cos \pi = -1$ and since cos is continuous, it takes all values in $[-1, 1]$ in the interval $[0, \pi]$. Since $\operatorname{Re} \frac{z}{|z|} \in [-1, 1]$ we can find $x \in [0, \pi]$ such that $\cos x = \operatorname{Re} \frac{z}{|z|}$. It follows that $\sin x = \pm \operatorname{Im} \frac{z}{|z|}$. Changing the sign of x changes the sign on $\sin x$ but leaves $\cos x$ unchanged. Therefore either e^{ix} or e^{-ix} equals $\frac{z}{|z|}$.

We may therefore solve the equation for w given any $z \neq 0$. If w_1 and w_2 are two solutions, it follows that $e^{w_1 - w_2} = 1$, so that w_1 and w_2 differ by an integer multiple of $2\pi i$. We call any permissible value of $\operatorname{Im} w$ an *argument* for z, and denote any such number by $\arg z$. We should therefore define $\log z = \ln |z| + i \arg z$. To get an actual (single-valued) function, we must make particular choices of $\arg z$ for each z. We shall see later that in order to be able to this and obtain a continuous function, we can not allow all of $\mathbb{C} \setminus \{0\}$ in the domain. Intuitively it is clear that we must choose the domain such that there are no closed curves in it that 'go around' the origin, since following such a curve we would have changed the argument continuously by an integer multiple of 2π when we arrive back at the starting point. This leads to the following concept.

DEFINITION 2.18. A connected subset of the Riemann sphere is called *simply connected* if its complement is connected.

If Ω is a region where we want to define a single-valued, continuous argument function, it must not contain 0 or ∞, and to exclude the possibility of having a closed curve in Ω that 'winds around' 0, we should exclude from Ω a connected set containing both 0 and ∞. Now suppose we have selected a region Ω which is simply connected in $\mathbb{C}$ and does not contain 0, and one of the possible arguments for some point in Ω. It seems plausible that this should determine a single-valued, continuous logarithm in Ω. We shall show later (see Theorem 3.19) that this is the case; we call such a function a *branch* of the logarithm.

The most important example is obtained when one chooses Ω to be $\mathbb{C}$ with the non-positive part of the real axis removed, and fixes the argument at 1 to be 0. This is called the *principal branch* of the logarithm. The argument of any number in Ω is determined by the requirement that it is in $(-\pi, \pi)$. The notation Log with a capital L is sometimes used for this branch.

Another important case is when one instead removes the non-negative real axis and fixes the argument at -1 to be π. The argument is then in the interval $(0, 2\pi)$. Other choices are obtained when one removes from $\mathbb{C}$ any smooth, non self-intersecting curve starting at 0 and ending at ∞. In any case, it is not possible to talk about the complex logarithm without specifying which branch one is dealing with.

THEOREM 2.19. *Any branch of the logarithm is analytic with derivative* $1/z$.

PROOF. For any $z = x + iy$ we have $\log(x + iy) = u(x, y) + iv(x, y)$ where $u(x, y) = \frac{1}{2}\ln(x^2 + y^2)$ and $v(x, y) = \arctan\frac{y}{x} + k\pi$, where k is some integer except if $x = 0$ in which case $v(x, y) = \frac{\pi}{2} - \arctan\frac{x}{y} + k\pi$. By continuity the same value of k has to be used in any sufficiently small neighborhood of z. Differentiating we therefore get $u_x(x, y) = \frac{x}{x^2+y^2}$, $u_y(x, y) = \frac{y}{x^2+y^2}$, $v_x(x, y) = -\frac{y}{x^2+y^2}$ and $v_y(x, y) = \frac{x}{x^2+y^2}$ so that the Cauchy-Riemann equations are satisfied. Since the partials are all continuous for $(x, y) \neq (0, 0)$, the function is analytic by Theorem 2.3. The derivative is $u_x + iv_x = \frac{x-iy}{x^2+y^2} = \frac{1}{z}$ so the proof is complete. □

EXERCISE 2.20. Prove Theorem 2.19 by use of Theorem 2.6 (5).

We are now able to define arbitrary powers of any complex number $w \neq 0$. We set $w^z = e^{z\log w}$, where log is some branch of the logarithm, giving rise to a branch of the power. By varying the branch, there is therefore in general infinitely many values of the power; *e.g.*, $i^i = e^{i\log i}$ and since $\log i = \ln 1 + i\arg i$ the possible values of $\log i$ are $i(\frac{\pi}{2} + 2k\pi)$, where k is an integer. Hence the possible values of i^i are $e^{-\frac{\pi}{2}-2k\pi}$. There are therefore infinitely many possible values (note that they are all real!). In some cases the situation is simpler, however; if w is real > 0 one always uses the principal branch of the logarithm so that w^z for real z is the elementary exponential function with base w.

One can of course also view the exponent as fixed and the base as the variable; these are the *power functions* $z \mapsto z^\alpha$. If α is an integer it is clear that the choice of branch for the logarithm is irrelevant; the function coincides with the elementary concept of a power function. If α is rational $= \frac{n}{m}$ where $m > 0$ and n are integers with no common factors, there are exactly m possible values of z^α for each $z \neq 0$; one usually says that there are m branches. This agrees with our discussion of the square root in Chapter 1.3. If α is irrational or non-real, however, there are always infinitely many branches of the power. Different powers are said to be of the same branch if they are defined through the same branch of the logarithm.

THEOREM 2.21. *Any branch of z^α is analytic (in its domain) with derivative $\alpha z^{\alpha-1}$, using the same branch.*

PROOF. $\frac{d}{dz}e^{\alpha \log z} = \frac{\alpha}{z}e^{\alpha \log z} = \alpha e^{(\alpha-1)\log z}$. □

If α is real and > 0 the power function is also defined for $z = 0$, with value 0 and if $\alpha = 0$ the power function is the constant 1.

2.3. Conformal mappings by elementary functions

We will here only give some examples of mappings induced by power functions and by the exponential function and their combinations with Möbius transforms.

Suppose $\alpha \in \mathbb{R}$. That a branch of $w = z^\alpha$ is defined in an open set Ω means that $z^\alpha = e^{\alpha \log z}$ for an appropriately chosen branch of the logarithm. Note that those $z \in \Omega$ for which $|z| = r$ are mapped onto $|w| = r^\alpha$ so that circular arcs centered at the origin are mapped onto (other) circular arcs centered at the origin. Similarly, if z is on a ray $\arg z = \theta$ we have $\arg w = \alpha\theta$ so rays from the origin are mapped onto other rays from the origin. Also, angles at the origin are multiplied by a factor α so that the map is certainly not conformal there unless $\alpha = 1$. This is true even if α is an integer so that z^α is well defined in the whole plane. Note that the derivative vanishes at 0 then.

These observations show that a wedge domain, bounded by two rays from the origin making the angle ϕ may be mapped onto a half plane by applying a branch of z^α where $\alpha = \pi/\phi$. More generally, any region with a corner at the origin may have this corner 'straightened out' by applying an appropriate power function. Since any region bounded by two intersecting circular arcs may be mapped onto a wedge by a Möbius transform taking the points of intersection to 0 and ∞ respectively, any such region may be mapped onto a half plane by composing a Möbius transform and a power function.

EXERCISE 2.22. Construct a conformal mapping that takes the region

$$\begin{cases} |z+3| < \sqrt{10}, \\ |z-2| < \sqrt{5}, \end{cases}$$

onto the interior of the first quadrant.

EXERCISE 2.23. Map the region

$$0 < \arg z < \pi/\alpha \ , \qquad 0 < |z| < 1 \ ,$$

onto the interior of the unit circle ($\alpha \geq 1/2$).

Since $e^z = e^x(\cos y + i \sin y)$ if $z = x + iy$ it is clear that the exponential function takes any line parallel to the real axis into a ray from the origin. Similarly, any vertical line segment is taken into a circular arc centered at the origin and with angular opening equal to the length of the line segment. This means that an infinite strip parallel to the real axis, *i.e.*, a region of the type $a < \operatorname{Im} z < b$, is mapped onto a wedge domain by the exponential function. A half infinite strip defined by $a < \operatorname{Im} z < b$, $\operatorname{Re} z < c$ is similarly mapped onto a circular sector centered at the origin.

EXERCISE 2.24. What is the image of the region $0 < \operatorname{Im} z < 2\pi$ under the map $w = e^z$?

EXERCISE 2.25. Construct a conformal mapping of the region

$$\begin{cases} |z-1| > 1, \\ \quad |z| < 2, \end{cases}$$

onto the upper half plane.

EXERCISE 2.26. Construct a conformal map of the region

$$\begin{cases} -\pi/4 < \arg z < \pi/4 \ , \\ \quad 0 < |z| < R \ , \qquad (R > 1) \end{cases}$$

onto the interior of the unit circle, so that $z = 1$ is mapped onto the origin. Calculate the length of the arc of the unit circle which is the image of the arc

$$\begin{cases} -\pi/4 \leq \arg z \leq \pi/4 \ , \\ \quad |z| = R \ . \end{cases}$$

EXERCISE 2.27. What is the image of the unit disk under the map $w = F(z) = (z + 1/z)/2$, $F(0) = \infty$.
Hint: Introduce $W = \frac{w-1}{w+1}$ and $Z = \frac{z-1}{z+1}$.

EXERCISE 2.28. Map the region $a < \arg z < b$, where $0 < a < b < 2\pi$, conformally onto the complex plane with the positive real axis removed.

EXERCISE 2.29. Find a conformal map that takes

$$\Omega = \begin{cases} |(1-i)z + (1+i)\overline{z}| < 2 \ , \\ (1+i)z + (1-i)\overline{z} > 0 \ , \end{cases}$$

onto

$$\Omega' = \begin{cases} |z - 1 - i| < 2 \ , \\ -i(z - \overline{z}) > 2 \ . \end{cases}$$

EXERCISE 2.30. Consider the conformal map given by $\cos z$. What are the images of lines parallel to the real and imaginary axes? What is the image of the strip $-\pi < \operatorname{Re} z < \pi$?

CHAPTER 3

Integration

3.1. Complex integration

Complex integration is at the core of the deeper facts about analytic functions. Here we will discuss the basic definitions.

Let γ be a *piecewise differentiable curve* in $\mathbb{C}$. This means a complex-valued, continuous function defined on a compact real interval which is continuously differentiable except at a finite number of points, where at least the left and right hand limits of the derivative exist. Thus it is described by an equation $z = z(t)$ where $a \leq t \leq b$ for some real numbers a and b and z' is continuous except for a finite number of jump discontinuities. For convenience we will in the sequel call such a curve an *arc*.

If f is a continuous, complex-valued function of a complex variable defined on an arc γ, then the composite function $f(z(t))$ is continuous and we make the following definition.

Definition 3.1. $\displaystyle\int_\gamma f(z)\,dz = \int_a^b f(z(t))z'(t)\,dt.$

If you know about line integrals and $f = u + iv$, $z = x + iy$ you will realize that $\int_\gamma f(z)\,dz$ is the line integral

$$\int_\gamma u\,dx - v\,dy + i\int_\gamma v\,dx + u\,dy,$$

but we will not use this. It is, however, very important that the complex integral is *independent of the parametrization* of the arc γ. This means the following. A change of parameter is given by a piecewise differentiable, increasing function $t(s)$ mapping an interval $[c, d]$ onto $[a, b]$. The usual change of variables formula then shows that $\int_\gamma f(z)\,dz = \int_c^d f(z(t(s)))z'(t(s))t'(s)\,ds$. Here $z'(t(s))t'(s)$ is, by the chain rule, the derivative of $z(t(s))$, so that the definition of the complex integral gives the same value whether we parametrize γ by $z(t)$ or $z(t(s))$.

Note that the arc γ has an *orientation*, in that it begins at $z(a)$ and ends at $z(b)$. If $t(s)$ is a *decreasing* piecewise differentiable function, mapping $[c, d]$ onto $[a, b]$, then the equation $z = z(t(s))$ will give a parametrization of the *opposite arc* to γ, which we denote by $-\gamma$, in that the initial point is now $z(t(c)) = z(b)$ and the final point $z(t(d)) = z(a)$. Thus we have $\int_{-\gamma} f(z)\,dz = -\int_\gamma f(z)\,dz$.

It is clear by the definition that the integral is linear in f, and also that if we divide an arc γ into two sub-arcs γ_1 and γ_2 by splitting the parameter interval into two subintervals with no common interior points (keeping the correct orientation), then $\int_\gamma f(z)\,dz = \int_{\gamma_1} f(z)\,dz + \int_{\gamma_2} f(z)\,dz$. It is now an obvious step to consider the sum of two (or more) arcs γ_1 and γ_2 even if they are not sub-arcs of another arc, and define the integral over such a sum as the sum of the integrals over the individual terms. Such a formal sum of arcs is called a *chain*. Given arcs $\gamma_1, \dots, \gamma_n$ we may integrate over chains of the form $\gamma = a_1\gamma_1 + \cdots + a_n\gamma_n$, where the coefficients $a_1, \dots, a_n$ are arbitrary integers, indicating that the integral $\int_{\gamma_j} f$ enters in $\int_\gamma f$ with the coefficient a_j. If $a_j = 0$ the arc γ_j can of course be left out of γ.

Note that our notation for the opposite of an arc makes sense, in that integrating over $-\gamma$ amounts to integrating over $(-1)\gamma$. Very often we will integrate over *closed* arcs. This means an arc where the initial and final points coincide. A *simple* arc is one without self-intersections; for a closed arc this means no self-intersections apart from the common initial and final point.

There is also a triangle inequality for complex integrals. From the definition of integral and the triangle inequality (2.2) it immediately follows that

$$\Big|\int_\gamma f(z)\,dz\Big| \le \int_\gamma |f(z)|\,|dz|,$$

where the last integral is defined by $\int_\gamma |f(z)|\,|dz| := \int_a^b |f(z(t)||z'(t)|\,dt$ and is called an *integral with respect to arc-length*. The reason for this is, of course, that $\int_\gamma |dz| = \int_a^b |z'(t)|\,dt$ gives the length of the arc γ. If you don't know this already, you may take it as a definition of length. Note that a very similar calculation to the one we did earlier shows that an integral with respect to arc length is independent of the parametrization, and in this case also of the orientation of the arc.

Example 3.2. Suppose γ is the circle $|z - a| = r$, oriented by running through it counter-clockwise. A parametrization is $z(t) = a + re^{it}$, $0 \le t \le 2\pi$. We obtain $z'(t) = ire^{it}$ so that $|z'(t)| = r$. The length of the circle is therefore $\int_0^{2\pi} r dt = 2\pi r$, as expected.

It is possible to integrate along more general curves than those that are piece-wise differentiable, so called *rectifiable* curves. There is seldom any reason to do this in complex analysis, however. In fact, when integrating analytic functions the integral is, as we shall see later, independent of small changes in the path we integrate over, so it is practically always enough to consider piece-wise differentiable arcs.

The elementary way of calculating integrals of a function of a real variable is by finding a primitive of the integrand. This method can be used also for complex integrals. Suppose that a continuous function f has a primitive, *i.e.*, a function F analytic on a continuously differentiable curve γ such that $F' = f$. Suppose $z = z(t)$, $a \le t \le b$, is a parametrization of γ. Then

$$\int_\gamma f(z)\,dz = \int_a^b f(z(t))z'(t)\,dt = \int_a^b (F(z(t)))'\,dt = \big[F(z(t))\big]_a^b = F(z_1) - F(z_0), \tag{3.1}$$

where $z_0 = z(a)$ is the initial and $z_1 = z(b)$ the final point of γ. If γ is just piece-wise continuously differentiable the same formula holds; one only has to use (3.1) on each differentiable piece and add the resulting formulas. The evaluations of the primitive at the intermediate points will then cancel, and we obtain (3.1) again.

So far the theory of analytic functions closely parallels the theory of functions of a real variable. This is quite misleading, as we shall see in the next section. The first indication that the theory of analytic functions is very different from one-variable real analysis comes when one asks the question of which functions f of a complex variable have a primitive. This turns out to require that f is analytic, but not even this is enough. There are also requirements on the nature of the domain of f, and these questions are a central theme for the theory of analytic functions. The starting point is the following theorem.

THEOREM 3.3. *Suppose f is continuous in a region Ω. Then f has a primitive F in Ω if and only if* $\int_\gamma f(z)\,dz = 0$ *for every closed arc $\gamma \subset \Omega$. It is enough if this is true for arcs made up solely of vertical and horizontal line segments.*

PROOF. If F is a primitive of f in Ω and γ a closed arc with initial and final points $z_1 = z_0$, then $\int_\gamma f(z)\,dz = F(z_1) - F(z_0) = 0$ since $z_1 = z_0$.

Conversely, if the integral along closed arcs vanishes, pick a point $z_0 \in \Omega$ and define $F(z) = \int_\gamma f$, where γ is an arc in Ω starting at z_0 and ending at z. This gives an unambiguous definition of F, since if $\tilde\gamma$ is another such arc, then the arc $\gamma - \tilde\gamma$ is a closed arc in Ω. Thus the integral along γ has the same value as the integral along $\tilde\gamma$. We may restrict ourselves to arcs of the special type of the statement of the theorem, since in an open, connected set Ω every pair of points may be connected by an arc of this kind in Ω (show this as an exercise!).

It now remains to show that F is a primitive of f in Ω.

Writing $z = x + iy$ with real x, y we shall calculate the partial derivatives of F with respect to x and y. To do this, let $h \in \mathbb{R}$ be so small that the line segment between z and $z+h$ is contained in Ω. Then $F(z+h) - F(z) = \int_0^h f(z+t)\,dt$. This is seen by choosing an arc γ starting at z_0 and ending at z to calculate $F(z)$, and then calculating $F(z+h)$ by adding to γ the line segment connecting z to $z+h$, which we parametrize by $z(t) = z + t$, $0 \le t \le h$.

By the fundamental theorem of calculus, differentiating with respect to h gives $F'_x(z) = f(z)$. Similarly, considering $F(z+ih) - F(z) = i\int_0^h f(z+it)\,dt$ we obtain $F'_y(z) = if(z)$. Thus the Cauchy-Riemann equation $F'_x + iF'_y = 0$ is satisfied, and $F' = F'_x = f$, so that F is a primitive of f. □

3.2. Goursat's theorem

In this section we shall begin to explore properties of analytic functions which show them to be very different in nature to differentiable functions of a real variable.

We first prove a fundamental theorem by Goursat (1905). We then consider integrals along the boundary of a rectangle. A rectangle is of course a set defined by inequalities $a \le \operatorname{Re} z \le b$, $c \le \operatorname{Im} z \le d$, and the boundary consists of four line segments with endpoints at the points $a + ic$, $b + ic$, $b + id$ and $a + id$. The boundary is therefore a closed arc, and we orient it by running through the vertices in the order described, ending up finally with $a + ic$ again. This means we run through the boundary in the direction which has the interior of the rectangle to the left of the boundary. This orientation of the boundary is called *positive.*

THEOREM 3.4. *Suppose f is analytic in a closed rectangle (*i.e.*, in an open set containing the rectangle) and let γ be the positively oriented boundary of the rectangle. Then* $\displaystyle\int_\gamma f(z)\,dz = 0.$

PROOF. Let R be the rectangle and I be the value of the integral. Now divide R into four congruent rectangles by one horizontal and one vertical cut, and let the integrals over the positively oriented boundaries of the sub-rectangles be I^j, $j = 1, \dots, 4$. A common side to two of the rectangles will then be given opposite orientation in the corresponding integrals. It follows that $I = I^1 + I^2 + I^3 + I^4$, since the contributions from integrating over the cuts will cancel. Thus the absolute value of at least one of the I^j will be $\ge |I|/4$. Let R_1 be a corresponding sub-rectangle and I_1 the associated integral, so that $|I| \le 4|I_1|$. We can now repeat the process with R_1, and then repeat this process indefinitely. We obtain a nested[1] sequence $R_1, R_2, \dots$ of rectangles and

[1] *i.e.*, each rectangle is contained in the previous one

a corresponding sequence $I_1, I_2, \ldots$ of integrals such that $|I| \leq 4^n |I_n|$ for $n = 1, 2, \ldots$.

The sequences of lower left corner real and imaginary parts in R_n are both increasing, because the rectangles are nested, and bounded from above, because all rectangles are contained in R. It follows that the sequence of lower left corners converge to a point $w \in R$. Let d be the diameter of R, *i.e.*, the length of the diagonal. Then it is clear that the diameter of R_n is $d_n = 2^{-n} d$, so that given any neighborhood of w, R_n will be contained in this neighborhood for all sufficiently large n.

Now f is differentiable at w, so that $|\frac{f(z)-f(w)}{z-w} - f'(w)| < \varepsilon$ if z is sufficiently close to w. Denoting the expression inside the absolute value signs by $\rho(z)$ we obtain $f(z) = f(w) + f'(w)(z-w) + \rho(z)(z-w)$, where $|\rho(z)| < \varepsilon$ if z is in a sufficiently small neighborhood of w. Choose n so large that R_n is contained in such a neighborhood, and let γ_n be the positively oriented boundary of R_n, so that

$$\begin{aligned} I_n &= \int_{\gamma_n} f(z)\, dz \\ &= \int_{\gamma_n} f(w)\, dz + \int_{\gamma_n} f'(w)(z-w)\, dz + \int_{\gamma_n} \rho(z)(z-w)\, dz. \end{aligned}$$

Now, the constant $f(w)$ has primitive $f(w)z$ and the first order polynomial $f'(w)(z-w)$ has primitive $\frac{1}{2} f'(w)(z-w)^2$, so that the first two integrals in the second line vanish. The third integral is estimated as follows:

$$\Big| \int_{\gamma_n} \rho(z)(z-w)\, dz \Big| \leq \varepsilon \int_{\gamma_n} |z-w||dz| \leq \varepsilon d_n L_n,$$

where L_n is the length of γ_n and d_n as before is the diameter of R_n. The estimate follows by the triangle inequality and since $|z-w| \leq d_n$, both of z and w being in R_n. However, we have $d_n = 2^{-n} d$, and it is equally clear that $L_n = 2^{-n} L$, where L is the length of the boundary of R. Thus we have $|I| \leq 4^n |I_n| \leq dL\varepsilon$. Since $\varepsilon > 0$ is arbitrary, it follows that $I = 0$. □

We will also need a slight extension of Goursat's theorem.

COROLLARY 3.5. *Suppose f is analytic in a closed rectangle R except for at an interior point p, where $(z-p)f(z) \to 0$ as $z \to p$. If γ is the positively oriented boundary of R, then* $\int_\gamma f(z)\, dz = 0$.

PROOF. Let $\varepsilon > 0$ and $R_0 \subset R$ be a square centered at p and so small that $|(z-p)f(z)| < \varepsilon$ for $z \in R_0$. If γ_0 is the positively oriented boundary of R_0 we obtain

$$\Big|\int_{\gamma_0} f(z)\,dz\Big| \le \varepsilon \int_{\gamma_0} \frac{|dz|}{|z-p|} \le 8\varepsilon.$$

The last inequality is due to the facts that $|z-p| \ge \ell/2$ if ℓ is the side length of R_0, and that the length of γ_0 is 4ℓ.

Now extend the sides of R_0 until they cut R into 9 rectangles, one of which is R_0. The other 8 satisfy the assumptions of Theorem 3.4. It follows that $|\int_\gamma f| = |\int_{\gamma_0} f| \le 8\varepsilon$, and since $\varepsilon > 0$ is arbitrary the integral over γ must be 0. □

We can now prove a first version of a fundamental theorem known as the *Cauchy integral theorem.*

COROLLARY 3.6 (Cauchy's integral theorem for a disk). *Suppose f is analytic in an open disk D, except possibly at a point p where $(z-p)f(z) \to 0$ as $z \to p$. Then f has a primitive in $D \setminus \{p\}$, and for every closed curve γ in $D \setminus \{p\}$ we then have* $\displaystyle\int_\gamma f(z)\,dz = 0$.

PROOF. In view of Theorem 3.3 it is enough to show that f has a primitive in $D \setminus \{p\}$.

Let z_0 be a fixed point in D with both $\operatorname{Re} z_0 \ne \operatorname{Re} p$ and $\operatorname{Im} z_0 \ne \operatorname{Im} p$. We may also assume that the center of D has both real and imaginary parts closer to those of p than to those of z_0. Let $z \ne p$ be another point in D.

Suppose first that the boundary of the rectangle with opposite corners at z_0 and z is in D and does not contain p. We then define $F(z)$ as the integral of f along first the horizontal side of the rectangle starting at z_0, and then the vertical side ending at z. It is clear, reasoning as in the proof of Theorem 3.3, that $F'_y(z) = if(z)$. However, by Corollary 3.5 $F(z)$ will have the same value if we first integrate along the vertical side starting at z_0 and then along the horizontal side ending at z, and with this definition we see that $F'_x(z) = f(z)$, so that F is a primitive of f wherever it is defined.

It remains to define F at points for which p is on the boundary of the rectangle, or one of the corners of the rectangle is outside D. Then first note that we could have started our path of integration by first moving vertically, then horizontally and finally vertically again until we reach z, and the horizontal path may be chosen anywhere between $\operatorname{Im} z$ and $\operatorname{Im} z_0$, as long as the path stays in D and doesn't contain p. This change will not affect the value of F in view of Corollary 3.5.

Suppose now that either p is on the horizontal side ending at z, or else that the other endpoint of this side is outside D. We then define $F(z)$ just as before and obtain $F'_y(z) = if(z)$. However, when calculating F'_x we modify our path by first following the horizontal side starting at z_0 some distance, then following a vertical path until we reach the horizontal side ending at z, and then following this side until we reach z. This can be done so that the path is inside D and does not contain p. The value of the integral will again equal $F(z)$ because of Corollary 3.5, and we get just as before that $F'_x(z) = f(z)$.

It is clear that a similar construction will work if p is on the vertical side ending at z, or if this side is not in D. Thus F is a well defined analytic function in $D \setminus \{p\}$ with derivative f. You should draw a picture of the various cases and convince yourself that the construction will give an unambiguous definition of F! $\square$

The conclusion of Corollary 3.6 can not be drawn with weaker assumptions on f at the point p. To illustrate this, let $f(z) = 1/(z-p)$ which is analytic in any disk centered at p, except at $z = p$, and let γ be the positively oriented boundary of such a circle. We may parametrize γ by $z(t) = p + re^{it}$, $0 \le t \le 2\pi$. Then $z'(t) = ire^{it}$ so that

$$\int_\gamma \frac{dz}{z-p} = \int_0^{2\pi} \frac{ire^{it}\,dt}{re^{it}} = i\int_0^{2\pi} dt = 2\pi i \neq 0. \tag{3.2}$$

This example is actually more crucial than is immediately obvious, and we use it as the basis for the notion of *index* or *winding number* of a point with respect to a closed arc.

Definition 3.7.

(1) A *cycle* is a chain (a formal sum of arcs) which may be written as a sum of finitely many *closed* arcs.
(2) The *index* of a point $p \notin \gamma$ with respect to a cycle γ is
$$n(\gamma, p) = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z-p}.$$

Note that the range of γ is a compact set, being finite union of continuous images of the compact parameter intervals, so its complement is open. An open set may be split into *open, connected components*[2]. Clearly there is precisely one unbounded component in the complement of γ.

Lemma 3.8. *The index has the following properties.*

(1) *$n(\gamma, p)$ is always an integer.*
(2) *$n(-\gamma, p) = -n(\gamma, p)$.*
(3) *$n(\gamma_1 + \gamma_2, p) = n(\gamma_1, p) + n(\gamma_2, p)$ if γ_1 and γ_2 are both cycles not containing p.*
(4) *$n(\gamma, p)$ is constant as a function of p in any connected component of the complement of the range of γ.*

[2]This is true whether you think of connected as meaning arcwise connected, or use the more general concept of connectedness introduced in topology.

(5) $n(\gamma, p) = 0$ *for all* p *in the unbounded component of the complement of the range of* γ.

Proof. Let $z = z(t)$, $a \leq t \leq b$, be a parametrization of a closed arc γ and set $g(t) = \int_a^t \frac{z'(s)\,ds}{z(s)-p}$ for $t \in [a, b]$. Then $g(b) = 2\pi i\, n(\gamma, p)$ and $g'(t) = z'(t)/(z(t) - p)$ so that the derivative of $h(t) = e^{-g(t)}(z(t) - p)$ is identically 0. We have $h(a) = z(a) - p$, so h is constant equal to $z(a) - p$. For $t = b$ we obtain $e^{-g(b)}(z(b) - p) = z(a) - p$. Since γ is a closed arc we have $z(b) = z(a) \neq p$ so that $e^{-g(b)} = 1$. Thus $g(b)$ is an integer multiple of $2\pi i$. Since a finite sum of integers is an integer this proves (1).

(2) and (3) are obvious from the definition, and it is also obvious that $n(\gamma, p)$ depends continuously on $p \notin \gamma$ (give detailed reasons yourself!). But a continuous, real-valued function in a region assumes intermediate values, so since the index is integer-valued, (4) follows.

Finally, it is clear that $n(\gamma, p) \to 0$ as $p \to \infty$, and since $n(\gamma, p)$ is independent of p for p in the unbounded component of the complement of γ, (5) follows. □

A circle has a complement consisting of exactly two components, and since we saw in (3.2) that the index of the center of a positively oriented circle is 1, all other points in the open disk will also have index 1 with respect to the boundary circle.

Theorem 3.9 (Cauchy's integral formula). *Suppose* f *is analytic in an open set* D *for which the conclusion of Corollary 3.6 is correct, and that* γ *is a cycle in* D. *Then, if* $p \notin \gamma$,

$$n(\gamma, p) f(p) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)\,dz}{z - p}.$$

In particular, this is true if D *is a disk.*

Proof. Put $g(z) = \frac{f(z)-f(p)}{z-p}$. Then g is analytic in $D \setminus \{p\}$ and $(z - p)g(z) = f(z) - f(p) \to 0$ as $z \to p$ since f is continuous at p. Thus $\int_\gamma g(z)\,dz = 0$ by Corollary 3.6. But by the definition of g this means that $\int_\gamma \frac{f(z)}{z-p}\,dz = f(p) \int_\gamma \frac{dz}{z-p}$. The theorem follows. □

For the special case when γ is a positively oriented circle we obtain $f(p) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z-p}\,dz$ when p is inside the circle. When p is outside the circle, the integral equals 0. The situation for analytic functions is therefore radically different than for differentiable functions of a real variable, since Cauchy's integral formula shows that if you know the values of an analytic function on a circle, then all the values inside the circle are determined. We shall see many more instances of how the behavior of an analytic function in one place determines the behavior in other locations.

Note that so far we only know the conclusion of the theorem in the case when D is a disk. More general regions will be discussed in Section 3.4.

3.3. Local properties of analytic functions

We start with a useful result about analytic dependence on a parameter in certain integrals.

LEMMA 3.10. *Suppose f is continuous on a circle γ with equation $|z-p|=r$. Then the function*

$$g(z)=\frac{1}{2\pi i}\int_\gamma \frac{f(\zeta)}{\zeta-z}\,d\zeta \tag{3.3}$$

is analytic in the corresponding open disk $|z-p|<r$. In fact, we may expand the function in a power series $g(z)=\sum_{k=0}^\infty a_k(z-p)^k$ with radius of convergence at least equal to r. The coefficients in the series are given by $a_k=\frac{1}{2\pi i}\int_\gamma \frac{f(z)}{(z-p)^{k+1}}\,dz$.

PROOF. The denominator in the integral is $\zeta-z=(\zeta-p)(1-\frac{z-p}{\zeta-p})$. The reciprocal of this is the sum of a convergent geometric series since $|\frac{z-p}{\zeta-p}|<1$, z being closer to p than ζ. A partial sum of this series has the sum

$$\sum_{k=0}^{n-1}(\zeta-p)^{-1}\Big(\frac{z-p}{\zeta-p}\Big)^k=(\zeta-p)^{-1}\frac{1-(\frac{z-p}{\zeta-p})^n}{1-\frac{z-p}{\zeta-p}}$$
$$=\frac{1}{\zeta-z}-\frac{(z-p)^n}{(\zeta-p)^n(\zeta-z)}.$$

Solving for $1/(\zeta-z)$ and inserting in (3.3) we obtain

$$g(z)=\sum_{k=0}^{n-1}a_k(z-p)^k+\frac{(z-p)^n}{2\pi i}\int_\gamma \frac{f(\zeta)\,d\zeta}{(\zeta-p)^n(\zeta-z)},$$

where a_k are given in the statement of the theorem. We can estimate the absolute value of the last term by

$$\Big|\frac{z-p}{r}\Big|^n\frac{1}{2\pi}\int_\gamma \frac{|f(\zeta)|}{|\zeta-z|}\,|d\zeta|$$

which obviously tends to 0 as $n\to\infty$ since $|z-p|<r$. The lemma follows. □

Essentially all results about the local behavior of analytic functions, *i.e.*, properties valid in a neighborhood of a point of analyticity, can be deduced from the following theorem, which is an easy consequence of Lemma 3.10.

THEOREM 3.11. *Suppose f is analytic in a disk $|z-p|<R$. Then f has derivatives of all orders and one may expand f in a power series $f(z)=\sum_{k=0}^{\infty} a_k(z-p)^k$, with radius of convergence at least equal to R. We have $a_k = f^{(k)}(p)/k! = \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{(z-p)^{k+1}}\,dz$, where γ is any positively oriented circle centered at p such that f is analytic in the corresponding closed disk.*

PROOF. Let γ be the circle $|z-p|=r$, $0<r<R$. If z is inside the circle Cauchy's integral formula gives

$$f(z) = \frac{1}{2\pi i}\int_\gamma \frac{f(\zeta)}{\zeta - z}\,d\zeta.$$

We may now apply Lemma 3.10. We may choose r as close to R as we wish, so the radius of convergence is at least R. Since any power series may be differentiated term by term as many times as we wish, the differentiability follows. This also shows that $f^{(k)}(p) = k!a_k$. □

If the largest open disk centered at p in which there is an analytic function that agrees with f near p has radius R ($\leq \infty$), then the radius of convergence of the power series is $\geq R$. But we can not have strict inequality here, since f then has an analytic extension to a larger disk. We conclude that *the circle of convergence has at least one singularity of f on its boundary.* In particular, if f is entire, it may be expanded in a power series around any $p \in \mathbb{C}$, and the radius of convergence will always be infinite.

Another observation is that if all derivatives of a function analytic in a disk vanishes at the center of the disk, then the function is identically zero in the disk, since all coefficients in the power series vanish. We can generalize this.

THEOREM 3.12. *Suppose f is analytic in a region Ω and that all derivatives of f vanish at a point $p \in \Omega$. Then f vanishes identically.*

PROOF. The set of points where all derivatives vanish is, as we just saw, open. But so is the set of points where at least one derivative does not vanish, since all derivatives are continuous. Thus Ω is the union of two disjoint open sets, one of which therefore has to be empty[3]. The theorem follows. □

The power series is called the *Taylor series for f at p* and the formula

$$f(z) = \sum_{k=0}^{n-1} a_k(z-p)^k + \frac{(z-p)^n}{2\pi i}\int_\gamma \frac{f(\zeta)\,d\zeta}{(\zeta-p)^n(\zeta-z)},$$

[3]We use the topological notion of connectedness. Modify the proof yourself to apply to the notion of arcwise connectedness

obtained from Lemma 3.10 is known as *Taylor's formula with n terms and remainder.*

Theorem 3.11 gives integral formulas for the derivatives of an analytic function at the center of a disk. This may be generalized.

COROLLARY 3.13. *Suppose f is analytic in a region Ω for which Cauchy's theorem is valid, and that $p \in \Omega$. Then the derivatives of f at p are given by*

$$f^{(n)}(p) = \frac{n!}{2\pi i} \int_\gamma \frac{f(\zeta)}{(\zeta - p)^{n+1}}\, d\zeta, \tag{3.4}$$

where γ is any cycle in $\Omega \setminus \{p\}$ and such that $n(\gamma, p) = 1$.

In particular, (3.4) is true for any positively oriented circle γ containing p and such that f is analytic in the corresponding closed disk.

PROOF. Suppose $f(z) = \sum_{k=0}^\infty a_k(z-p)^k$ near p. The function $g(z) = (f(z) - \sum_{k=0}^{n-1} a_k(z-p)^k)/(z-p)^n$ is analytic in Ω, since this is obvious away from p, and near p it follows from the power series expansion, which also shows that $g(p) = a_n = f^{(n)}(p)/n!$.

Now $(z-p)^{k-n-1}$ has a primitive $(z-p)^{k-n}/(k-n)$ in $\Omega \setminus \{p\}$ for $k < n$ so applying Cauchy's integral formula to g, all the terms from the sum contribute 0 to the integral. The corollary follows. □

We are also able to give a kind of converse to the Cauchy integral theorem which is sometimes useful.

THEOREM 3.14 (Morera). *Suppose f is continuous in a region Ω and $\int_\gamma f(z)\, dz = 0$ for all cycles in Ω. Then f is analytic in Ω. It is actually enough if each point of Ω has a neighborhood such that the condition is satisfied when γ is the boundary of any rectangle contained in the neighborhood.*

PROOF. The assumption shows that f has a primitive in a neighborhood of every point of Ω, according to Theorem 3.3, or with the less restrictive assumptions, according to the proof of the Cauchy integral theorem. Thus f is near each point the derivative of an analytic function, so it is itself differentiable in Ω, *i.e.*, analytic. □

COROLLARY 3.15. *Zeros of an analytic function not identically 0 are isolated points in the domain of analyticity.*

PROOF. Suppose f is analytic at p and $f(p) = 0$. According to Theorem 3.11 we may expand f in power series $\sum_{k=0}^\infty a_k(z-p)^k$. Since $f(p) = 0$ the first term in the series vanishes, and if n is the first index for which $a_n \neq 0$ we obtain $f(z) = (z-p)^n g(z)$, where g is the analytic function $\sum_{k=0}^\infty a_{n+k}(z-p)^k$, so that $g(p) = a_n \neq 0$. The positive integer n is called to *order* or *multiplicity* of the zero p.

Since g is continuous and $g(p) \neq 0$ there is a neighborhood of p in which g doesn't vanish. Since $(z-p)^n$ only vanishes for $z = p$ it follows that there is a neighborhood of p in which p is the only zero of f. □

Note that the fact that the zeros do not accumulate anywhere in the domain of analyticity does not prevent them from accumulating at some point of the boundary of the domain. An example is $\sin(1/z)$, which is analytic in $z \neq 0$ and has zeros $1/(k\pi)$, $k = \pm1, \pm2, \dots$, which accumulate at 0.

A fundamental theorem for entire functions is named after Liouville.

Theorem 3.16 (Liouville). *Suppose f is an entire function such that $|f(z)| \leq C|z|^n$ for all sufficiently large $|z|$. Then f is a polynomial of degree $\leq n$. In particular, if f is bounded in all of $\mathbb{C}$, then f is constant.*

Proof. Suppose γ is a circle centered at 0 of radius r, and consider (3.4) for $p = 0$. If $M(r) = \sup_{|z|=r} |f(z)|$ we obtain $|f^{(k)}(0)| \leq k!r^{-k}M(r)$, $k = 0, 1, 2, \dots$. These estimates are called *Cauchy's estimates.* Our assumption is that $M(r) \leq Cr^n$ for large r, so that $|f^{(k)}(0)| \leq k!Cr^{n-k}$ if r is large enough. As $r \to \infty$ we obtain $f^{(k)}(0) = 0$ for $k > n$, so that the Taylor expansion of f is a polynomial of degree $\leq n$. □

The fact that the only bounded, entire functions are constants is often very useful. We can for example now give a very simple proof of the fundamental theorem of algebra.

Theorem 3.17 (Fundamental theorem of algebra). *Any non-constant polynomial has at least one zero.*

Proof. Suppose P is a polynomial without zeros. Then $1/P(z)$ is an entire function, and we shall see that it is bounded, so that Liouville's theorem will show it to be constant.

If $P(z) = a_n z^n + a_{n-1}z^{n-1} + \dots + a_0$ with $a_n \neq 0$ we may write $P(z) = z^n(a_n + a_{n-1}/z + \dots + a_0/z^n)$. Here the expression in brackets tends to a_n as $z \to \infty$. Since $z^n \to \infty$ if $n > 0$ we have $1/P$ bounded for large $|z|$. Thus P is constant. The theorem follows. □

3.4. A general form of Cauchy's integral theorem

The aim of this section is to prove a version of Corollary 3.6 valid in more general regions than disks. Note that as soon as we do this we also have a more general version of Cauchy's integral formula, Theorem 3.9. We remind the reader that a *chain* is a formal sum of arcs, which can be integrated over by integrating over each term separately, and adding the results. Similarly, a chain which may be written as a sum where all terms are *closed* arcs is called a *cycle.*

We also remind the reader about the notion of a simply connected set, according to Definition 2.18. This property is closely connected with the notion of index.

THEOREM 3.18. *A region Ω is simply connected if and only if $n(\gamma, p) = 0$ for all cycles $\gamma \subset \Omega$ and all points $p \notin \Omega$.*

PROOF. Suppose the complement of Ω (with respect to the extended plane) is connected. Then the complement is contained in the unbounded region determined by γ, so Lemma 3.8 (5) shows that $n(\gamma, p) = 0$ if $p \notin \Omega$

Conversely, if $\complement\Omega$ is not connected, we can write it as the disjoint union of two closed sets[4], one of which contains ∞, and the other therefore being bounded. If the bounded set is K, it is compact and will therefore have a smallest distance $d > 0$ to the other part of the complement. It is left to the reader to prove this.

Let $p \in K$ and cover the whole plane by a net of closed squares with side $d/2$, such that p is the center of one of the squares. Clearly only finitely many of the squares have at least one point in common with K, since K is bounded. Let these squares have positively oriented boundaries $\gamma_1, \dots, \gamma_n$ and consider the cycle $\gamma = \sum \gamma_j$. Clearly $n(\gamma, p) = 1$ since p is in exactly one of the squares. It is also clear that $\gamma \subset \Omega \cup K$, since the diameters of the squares are $< d$ and they contain points from K. Now, certain sides of the squares occur twice in γ, being common to two adjacent squares. Any side that has a point in common with K is of this type, and the contributions from these sides in an integral cancel, since they are run through in opposite directions. Removing these sides will therefore not change γ, and then $\gamma \subset \Omega$. It follows that if Ω is not simply connected, then indices for points outside Ω with respect to cycles in Ω are not always 0. □

We shall use this characterization of simply connected regions to prove the following general version of Cauchy's theorem.

THEOREM 3.19 (Cauchy's integral theorem). *If f is analytic in a simply connected region Ω, then $\int_\gamma f(z)\,dz = 0$ for any cycle $\gamma \subset \Omega$.*

PROOF. We will show that the assumptions imply that f has a primitive in Ω. This follows if we can show that the integral of f along a cycle γ in Ω consisting only of vertical and horizontal line segments always vanishes, since then the integral from a fixed point z_0 to z along a path of this type is independent of the particular path, so that we obtain a well defined primitive in the usual way.

If γ is such a cycle, extend all line segments in γ indefinitely. We obtain a rectangular net consisting of some rectangles with positively oriented boundaries $\gamma_1, \dots, \gamma_n$, and some unbounded regions. We may

[4]This is a topological fact

assume $n > 0$, and pick a point p_j in the interior of each rectangle. We shall first show that γ is the cycle $\gamma' = \sum n(\gamma, p_j)\gamma_j$. It is clear by construction that $n(\gamma - \gamma', p_j) = 0$ and also that $n(\gamma - \gamma', p) = 0$ for every point p in one of the unbounded regions determined by the net, since these points are obviously all in the unbounded regions determined by γ respectively γ'.

We shall show that no side of any rectangle is in $\gamma - \gamma'$. Suppose to the contrary that a side σ of the rectangle bounded by γ_j is contained in $\gamma - \gamma'$ with coefficient $a \neq 0$. There are at least one region determined by the net in addition to the rectangle γ_j which have some part of σ on its boundary. Let p be a point in in such a region. Now σ is not contained in $\gamma - \gamma' - a\gamma_j$, so that the indices of p and p_j are the same with respect to this cycle. But by construction the indices are actually 0 and $-a$, respectively, so that $a = 0$, and $\gamma - \gamma'$ is the empty cycle.

Next we prove that all γ_j for which $n(\gamma_j, p_j) \neq 0$ bound rectangles contained in Ω. For suppose p is in the closed rectangle, but not in Ω. Then $n(\gamma, p) = 0$, since Ω is simply connected. On the other hand, the line segment connecting p and p_j does not intersect γ, so p and p_j are in the same component of the complement of γ, and therefore have the same index with respect to γ. It follows that $n(\gamma, p_j) = 0$ unless the rectangle bounded by γ_j is contained in Ω. It follows by Goursat's theorem that $\int_{\gamma_j} f = 0$ for all j for which γ_j is part of γ. Thus also $\int_\gamma f = 0$. □

We end this section with a very important consequence of the previous theorem.

Corollary 3.20. *Suppose f is analytic and has no zeros in a simply connected region Ω. Then one may define a branch of* $\log(f(z))$ *in Ω.*

Proof. Since f has no zeros in Ω the function $f'(z)/f(z)$ is analytic in Ω so that Cauchy's integral theorem applies to it. According to Theorem 3.3 there is therefore a primitive g of this function defined in Ω, and $\frac{d}{dz}(f(z)e^{-g(z)}) = f'(z)e^{-g(z)} - f(z)\frac{f'(z)}{f(z)}e^{-g(z)} = 0$, so that $f(z)e^{-g(z)} = C$, where $C \neq 0$ since neither f nor the exponential function vanishes. Thus we may find $A \in \mathbb{C}$ so that $e^A = C$. It follows that $f(z) = e^{g(z)+A}$, so that $g(z) + A$ is a branch of $\log(f(z))$. □

Since one may define a branch of the logarithm one may also define branches of any power function in Ω. We shall use this in proving the Riemann mapping theorem in Chapter 7.

Remark 3.21. To obtain a version of Cauchy's theorem valid in arbitrary regions we would have to discuss homology of cycles, and we will abstain from this. We sometimes have to deal with regions which are not simply connected, but the cycles we integrate over are then always very simple and explicitly given and therefore never cause any problem.

For example, suppose f is analytic in a circular ring Ω defined by $0 \leq r_0 < |z-a| < R_0 \leq \infty$ and suppose $r_0 < r < R < R_0$ and let γ be the cycle consisting of the two circles $|z-a| = r$ and $|z-a| = R$, the first negatively and the second positively oriented. Then $\int_\gamma f(z)\,dz = 0$, and we also have Cauchy's formula $f(w) = \frac{1}{2\pi i}\int_\gamma \frac{f(z)\,dz}{z-w}$ for any w satisfying $r < |w-a| < R$.

To see this, let σ_1 be the vertical ray going upwards from a and σ_2 the opposite ray. Then $\Omega \setminus \sigma_1$ is simply connected. Let γ_1 be a positively oriented cycle obtained by taking the part of γ in the set $\operatorname{Im} z \leq \operatorname{Im} a$ and connecting the pieces by radial line segments. Then Theorem 3.19 shows that $\int_{\gamma_1} f = 0$.

Similarly, the region $\Omega \setminus \sigma_2$ is simply connected, and if γ_2 is the part of γ in $\operatorname{Im} z \geq \operatorname{Im} a$, made into a positively oriented cycle by adding radial line segments, we also have $\int_{\gamma_2} f = 0$. But $\gamma = \gamma_1 + \gamma_2$, since the radial line segments will be run through twice and in opposite directions.

It is also easy to see that if $\operatorname{Im} w > \operatorname{Im} a$, then $n(\gamma_1, w) = 0$ and $n(\gamma_2, w) = 1$ so that $n(\gamma, w) = 1$. The reader should modify the construction for other locations of w to see that $n(\gamma, w) = 1$ as soon as $r < |w-a| < R$.

3.5. Analyticity on the Riemann sphere

Viewing analytic functions as defined on the Riemann sphere, where all points including ∞ look the same, one should be able to define analyticity at infinity. This leads to the following definition.

Definition 3.22. Suppose f is analytic in a neighborhood $|z| > r > 0$ of ∞ Then we say that f is analytic at ∞ if $z \mapsto f(1/z)$, which is analytic in $0 < |z| < 1/r$, extends to a function analytic also at 0.

Similarly, if f is analytic in a neighborhood of a and $f(z) \to \infty$ as $z \to a$ it would be tempting to say that f is analytic at a if $1/f(z)$ extends to a function analytic at a. We will not use this terminology since it may lead to confusion, but it is a perfectly reasonable point of view. In fact, in the next section we will show that if $f(z) \to \infty$ as $z \to a$, then $1/f(z)$ *always* has an analytic extension to a.

Any Möbius transform is in this sense analytic everywhere on the Riemann sphere, and the reader should should carry out the simple verification, and also show that the same is true for any rational function.

CHAPTER 4

Singularities

4.1. Singular points

An *isolated singularity* of a complex function f is a point a such that it has a neighborhood $\mathcal{O}$ with f analytic in $\mathcal{O} \setminus \{a\}$ (a so called *punctured neighborhood* of a). In some cases a is an isolated singularity simply because we do not know that f is analytic there, or that f is not analytic at a but will become so provided we assign the correct value to $f(a)$. In that case a is said to be a *removable singularity* for f. A typical example would be $z \mapsto \frac{\sin z}{z}$ which is not defined at 0, but where it is clear from the power series expansion of $\sin z$ that the function becomes entire once we assign it the value 1 at the origin. The main fact about removable singularities is contained in the following theorem.

THEOREM 4.1. *Suppose that f is analytic in a punctured neighborhood of a. Then a is a removable singularity for f if and only if $(z-a)f(z) \to 0$ as $z \to a$.*

Thus the singularities we allowed in Corollaries 3.4, 3.6 are actually removable, and may be ignored.

PROOF. The 'only if' part of the theorem is trivial, since in that case f must have a finite limit at a. To prove the other direction, let γ and ω be the positively oriented boundaries of disks centered at a and such that f is analytic in the punctured disks. If ω is the smaller disk f is analytic in the ring-shaped region between ω and γ, so by Remark 3.21

$$f(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta - \frac{1}{2\pi i} \int_\omega \frac{f(\zeta)}{\zeta - z}\, d\zeta$$

if z is in the ring-shaped region. Note that the first integral is analytic in the disk bounded by γ according to Lemma 3.10. If we can show that the integral over ω is zero we have therefore proved the theorem, since we may remove the singularity at a by defining $f(a)$ to be the value of the first integral at $z = a$.

Actually, the integral over ω does not depend on the radius of the disk it bounds, as long as that radius is smaller than $|z-a|$. To show that the integral is 0 it is therefore sufficient to show that its limit as the radius tends to 0 is 0. To see this, let $\varepsilon > 0$ and choose $\delta > 0$ so small that $|(\zeta-a)f(\zeta)|<\varepsilon$ if $|\zeta-a|<\delta$. Then, if the radius of ω is $r<\delta$ and $r<|z-a|/2$, we obtain $|\zeta-z|\geq|z-a|-|\zeta-a|=|z-a|-r\geq|z-a|/2$ so that

$$\left|\int_\omega \frac{f(\zeta)}{\zeta-z}\,d\zeta\right| \leq \varepsilon\int_\omega \frac{|d\zeta|}{|\zeta-a||\zeta-z|} \leq \frac{4\pi}{|z-a|}\,\varepsilon\ .$$

The proof is now complete. □

Let us now consider an arbitrary isolated singularity a for f. Then one of the following three cases must obtain:

(1) There is a real number α such that $|z-a|^\alpha f(z)\to 0$ as $z\to a$.
(2) There is a real number α such that $|z-a|^\alpha f(z)\to\infty$ as $z\to a$.
(3) Neither of the first two cases hold.

Consider first case (1). If $\alpha\leq 1$, then f has a removable singularity at a by Theorem 4.1. Otherwise, if n is the largest integer $<\alpha$ we have $(z-a)^{n+1}f(z)\to 0$ as $z\to a$. By Theorem 4.1 it follows that the function $(z-a)^n f(z)$ has a removable singularity at a. This function may have a zero at a, but ignoring the trivial case when f is identically zero, we may lower the value of n until $(z-a)^n f(z)$ has a non-zero value at a. If $n\leq 0$ it follows that f is analytic at a. If $n>0$ and the power series expansion around a of $(z-a)^n f(z)$ is $\sum_{k=0}^\infty a_k(z-a)^k$ it follows that

$$f(z)=\sum_{k=-n}^{-1} b_k(z-a)^k+\sum_{k=0}^{\infty} b_k(z-a)^k,$$

where $b_k=a_{n+k}$. The first sum above is called the *singular part* of f at a. Note that the singular part is analytic everywhere (even at ∞) except at a. Therefore, if we subtract the singular part from f we get a function which is analytic wherever f is, and also at a. Subtracting the singular part at a therefore removes the singularity at a. The fact that the singular part, in this case, consists of a finite sum of very simple functions makes this type of singularity rather harmless. It is called a *pole of order n*.

A pole of order $n>0$ is characterized by the fact that $(z-a)^n f(z)$ has a non-zero limit as $z\to a$, just as a zero of order n is characterized by the fact that $(z-a)^{-n}f(z)$ has a non-zero limit as $z\to a$. Note that $f(z)\to\infty$ as z approaches a pole so that $1/f$ has a removable singularity there. We may therefore view a pole as a point where f is 'analytic with the value ∞'; this agrees completely with our point of view when we discussed functions analytic on the Riemann sphere. Also note that *poles, like zeros, are isolated points.* We finally note that if f has a pole or zero of order $|n|$ at a, then case (1) holds exactly if $\alpha>n$ and case (2) holds exactly if $\alpha<n$.

Now let us consider case (2). If n is the smallest integer $\geq\alpha$, then $(z-a)^n f(z)\to\infty$ as $z\to a$ so that $(z-a)^{-n}/f(z)$ has a removable singularity at a. It is clear that this function has a zero at a, say of order $k > 0$. It is then clear that $(z-a)^{n+k}f(z)$ has a removable singularity with a non-zero value at a. Therefore, if $n + k \leq 0$ f has a removable singularity at a, and otherwise f has a pole of order $n + k$ at a. So, also in case (2) we have at worst a pole at a.

Unless we have case (3) we therefore have at worst a pole at a and a singular part consisting of a finite linear combination of negative integer powers of $z - a$. Conversely, this can *not* be the case in case (3) since a pole or a regular point immediately puts us in the cases (1) and (2). We call the singularity at a *essential* when we have case (3). It clearly is a less simple situation, since we can not have a finite singular part in this case. We shall see in the next section that there actually is a singular part, but it has infinitely many terms. Another indication of how complicated the behavior of an analytic function is near an essential singularity is given by the following theorem.

THEOREM 4.2 (Casorati-Weierstrass). *The range of the restriction of an analytic function to an arbitrary punctured neighborhood of an essential singularity is dense in* $\mathbb{C}$.

PROOF. Suppose f is analytic in the punctured neighborhood Ω of a, and that there is a complex number b such that all values of f in Ω has distance at least $d > 0$ from b. Consider the function $g(z) = (f(z) - b)^{-1}$. It is analytic in Ω and bounded by $1/d$ there. By Theorem 4.1 it therefore has a removable singularity at a so that $1/g(z)$ has at most a pole at a (if g has a zero of order n at a, then the pole has order n). So, $f(z) = b + 1/g(z)$ has at worst a pole at a. $\square$

EXAMPLE 4.3. The function $e^{1/z}$, $z \neq 0$, has an essential singularity at 0. To see this, note that if $z \to 0$ along the positive real axis the function tends to ∞, so the function can not have a removable singularity at 0. On the other hand, $e^{1/z} \to 0$ as $z \to 0$ along the negative real axis, so the origin can not be a pole either. The only remaining possibility is an essential singularity. Note that by the usual power series expansion for the exponential function we have $e^{1/z} = 1 + \sum_{k=1}^{\infty} \frac{1}{k! z^k}$. Hence this function actually has a singular part, but it consists of infinitely many terms.

Let us end this section by a short discussion of poles at infinity. Naturally f is said to have a pole of order n at ∞ if $z \mapsto f(1/z)$ has a pole of order n at 0. It therefore has a singular part which is a polynomial $p(1/z)$ of order n in $1/z$. In particular, it follows that $f(1/z) - p(1/z)$ has a removable singularity at 0, so is bounded there. It follows that $f(z) - p(z)$ is bounded at infinity. The singular part of a function which has a pole of order n at infinity is therefore a polynomial of order n.

Definition 4.4. A function is said to be *meromorphic* in a region Ω if it is analytic in Ω except for poles at certain points.

Suppose f is meromorphic in the extended plane $\mathbb{C}^*$. Since the extended plane is a compact set, f can only have a finite number of poles; by Bolzano Weierstrass' theorem there would otherwise be a point of accumulation of poles in the extended plane. This would have to be a non-isolated singularity. We may therefore subtract the singular parts for all the poles from f and will then be left with a function analytic in the extended plane. In particular, a bounded function. By Liouville's theorem it will have to be constant. We have proved the following theorem.

Theorem 4.5. *A function is meromorphic in the extended plane if and only if it is rational.*

As a special case it follows that an entire function which is not a polynomial has an essential singularity at ∞. The elementary functions e^z, $\cos z$ and $\sin z$ therefore have essential singularities at ∞.

4.2. Laurent expansions and the residue theorem

In this section we will give an expansion generalizing the power series expansion of an analytic function. In particular we will see that a function has a singular part at any isolated singularity, analogous to what we discussed in the previous section, but now possibly consisting of infinitely many terms. We will then use this expansion to prove the residue theorem, which gives a particularly simple way to calculate many complex integrals. We will finally apply this to several types of real integrals that are difficult or impossible to calculate by elementary means.

Consider a function f which is analytic in a region containing the ring $0 \leq R_0 < |z-a| < R_1 \leq \infty$. The case $R_0 = 0$ corresponds to the case when we have an isolated singularity at a. If $R_0 < r < R < R_1$, then it follows from Remark 3.21 that the Cauchy integral formula holds in the form

$$f(z) = \frac{1}{2\pi i} \int\limits_{|\zeta-a|=R} \frac{f(\zeta)}{\zeta - z}\, d\zeta - \frac{1}{2\pi i} \int\limits_{|\zeta-a|=r} \frac{f(\zeta)}{\zeta - z}\, d\zeta$$

for any z satisfying $r < |z-a| < R$. If we set

$$f_1(z) = \frac{1}{2\pi i} \int\limits_{|\zeta-a|=R} \frac{f(\zeta)}{\zeta - z}\, d\zeta, \qquad f_2(z) = -\frac{1}{2\pi i} \int\limits_{|\zeta-a|=r} \frac{f(\zeta)}{\zeta - z}\, d\zeta,$$

then $f(z) = f_1(z) + f_2(z)$ for such values of z. However, by Lemma 3.10 f_1 is analytic in $|z-a| < R$. It follows that f_2 is analytic in $r < |z-a| < R$. Actually, f_2 is analytic in $|z-a| > r$, even at ∞, which is seen similarly to the proof of Lemma 3.10. In fact, setting

$z = a+1/w$ we may write the denominator in f_1 as $\zeta-z = \zeta-a-1/w = -(1-(\zeta-a)w)/w$, and since $|(z-a)w| < 1$ the reciprocal of this is the sum of a convergent geometric series, and reasoning just as in the proof of Lemma 3.10 we obtain

$$f_2(a+1/w) = -\sum_{k=0}^{\infty} \frac{w^{k+1}}{2\pi i} \int\limits_{|\zeta-a|=r} f(\zeta)(\zeta-a)^k \, d\zeta,$$

a power series in w which converges for $|w| < 1/r$, corresponding to $|z-a| > r$. It follows that f_2 is analytic in $|z-a| > r$, including $z = \infty$. Setting

$$a_k = \frac{1}{2\pi i} \int\limits_{|\zeta-a|=r} \frac{f(\zeta)}{(\zeta-a)^{k+1}} \tag{4.1}$$

also for $k = -1, -2, \dots$ we can write this as $f_2(z) = \sum_{k=-\infty}^{-1} a_k(z-a)^k$. Adding up we obtain the following theorem.

THEOREM 4.6 (Laurent expansion). *Suppose f is analytic in $0 \le R_0 < |z-a| < R_1 \le \infty$. Then f has a **Laurent expansion** around a of the form*

$$f(z) = \sum_{k=-\infty}^{\infty} a_k(z-a)^k \ ,$$

converging at least for $R_0 < |z-a| < R_1$, where the coefficients a_k are given by (4.1).

*The **singular part of** f **at** a is $\sum_{k=-\infty}^{-1} a_k(z-a)^k$ and is analytic for $|z-a| > R_0$, including at ∞. In particular, if a is an isolated singularity for f, the singular part expansion converges everywhere except at $z = a$. The difference of f and its singular part is analytic wherever f is and also for $|z-a| < R_1$.*

Note that wherever the series converges the function $f(z) - \frac{a_{-1}}{z-a} = \sum_{k\neq -1} a_k(z-a)^k$ is the derivative of the function $\sum_{k\neq-1} \frac{a_k}{k+1}(z-a)^{k+1}$, so that its integral along any closed curve in the domain of convergence is 0. It follows that the integral of f around a positively oriented circle γ in $R_0 < |z-a| < R_1$ is equal to $2\pi i a_{-1}$. The coefficient a_{-1} in the Laurent expansion of f around a is called the *residue* of f at a, since it determines what remains after integration around a closed curve. We will denote the residue of f at an isolated singularity a by $\operatorname{Res} f(a)$. This is of course 0 unless a actually is a singularity of f. A slight generalization of the above gives the following important theorem.

THEOREM 4.7 (Residue theorem). *Suppose Ω is a simply connected region, and that f is analytic in Ω except for a finite number of isolated singularities. Then, if γ is a cycle in Ω not passing through any of the* *singularities,*

$$\int_\gamma f(z)\,dz = 2\pi i \sum_{z\in\Omega} n(\gamma, z)\,\mathrm{Res}\,f(z).$$

PROOF. Subtract from f the singular parts for all singularities. This leaves a function analytic in Ω, so that its integral is zero. As we saw above, the singular parts are analytic outside the corresponding singularity, and removing the term with index -1 the rest of the singular part has a primitive defined outside the singularity, so that their integrals vanish. It remains only to integrate the terms of index -1 for each singularity, which gives the result by the definition of the index. □

In all our applications of the residue theorem we will choose the cycle γ so that the indices of all the singularities with respect to γ are either 1 or 0.

A formula for the residue at an isolated singularity is of course given by (4.1) for $k = -1$. Actually, this formula is not of much practical value; on the contrary, one tries to find the residues without integration and then uses this to evaluate integrals. It is clear, however, that for this to be possible we need methods not involving integration to find residues. No such generally applicable method is known in the case of an essential singularity, even though there are of course many cases when we will know the Laurent expansion, as we saw in the case of $e^{1/z}$. The situation is different in the case of a pole, and we have the following theorem.

THEOREM 4.8. *Suppose that f has a pole of order n at a. Then* $\mathrm{Res}\,f(a) = \frac{1}{(n-1)!}\lim_{z\to a}\frac{d^{n-1}}{dz^{n-1}}((z-a)^n f(z))$. *In particular, for a simple pole the residue is $\lim_{z\to a}(z-a)f(z)$. If $f = p/q$ where p and q are analytic at a, $p(a) \neq 0$ and q has a simple zero at a, then the residue at a is $p(a)/q'(a)$. Similarly, if q has a double zero at a the residue is*

$$\frac{6p'(a)q''(a) - 2p(a)q'''(a)}{3(q''(a))^2}\ . \tag{4.2}$$

Similar, even more complicated, formulas hold for higher order poles.

PROOF. According to assumption $f(z) = \sum_{k=-n}^{\infty} a_k(z-a)^k$ for z close to a, so $g(z) = (z-a)^n f(z)$ has a removable singularity at a and a_{-1} is the coefficient of $(z-a)^{n-1}$ in the corresponding power series expansion. But this coefficient is $g^{(n-1)}(a)/(n-1)!$ and since $g^{(n-1)}$ is continuous at a the first claim follows. If now q has a simple zero at a, then $(z-a)p(z)/q(z) = p(z)\frac{z-a}{q(z)-q(a)} \to p(a)/q'(a)$ since $q(a) = 0$, $q'(a) \neq 0$. Finally, if q has a double zero at a, then $q(z) = (z-a)^2 q_2(z)$ where $q_2(a) = q''(a)/2$ and $q_2'(a) = q'''(a)/6$, as is easily verified. Hence

$((z-a)^2 f(z))' = (p(z)/q_2(z))' = (p'(z)q_2(z) - p(z)q_2'(z))(q_2(z))^{-2}$. Letting $z \to a$ (4.2) follows, and the final claim is left to the reader to verify. □

The conclusion of all this is that simple poles cause little problem in determining the residue, whereas higher order poles are considerably more messy to deal with. In the next section we shall see how one may use the residue theorem to calculate certain real integrals.

4.3. Residue calculus

In this section we shall see how one may use the residue theorem to calculate certain real integrals. We will only discuss a few types of integrals that can be handled; many others exist.

1. Let us first consider an integral of the form $\int_0^{2\pi} \rho(\cos\theta, \sin\theta)\, d\theta$. Here $\rho(x, y)$ is a rational function of two variables with no poles for (x, y) on the unit circle. If we think of this integral as the result of calculating an integral around the unit circle by the parametrization $z = e^{i\theta}$, $0 \le \theta \le 2\pi$, we note that by Euler's formulas we have $\cos\theta = \frac{1}{2}(z + 1/z)$ and $\sin\theta = \frac{1}{2i}(z - 1/z)$. Furthermore, $dz = ie^{i\theta}\, d\theta$ so that $d\theta = -i\, dz/z$. The integral therefore equals $-i \int_{|z|=1} \rho(\frac{z+1/z}{2}, \frac{z-1/z}{2i})\, dz/z$. The integrand is rational and it only remains to find those poles that are inside the unit circle and evaluate their residues.

EXAMPLE 4.9. Consider the integral $\int_0^{2\pi} \frac{d\theta}{a+\cos\theta}$ where $a > 1$. Setting $z = e^{i\theta}$ as above, the integral equals $-i \int_{|z|=1} (a + \frac{1}{2}(z + 1/z))^{-1}\, dz/z$ which after simplification becomes

$$-2i \int_{|z|=1} \frac{dz}{z^2 + 2az + 1} .$$

The zeros of the denominator are $z = -a \pm \sqrt{a^2 - 1}$. Since their product is 1, precisely one root is inside the unit circle; a being > 1 this root is $\sqrt{a^2 - 1} - a$. Since the pole comes from a simple zero in the denominator, we can use the method of Theorem 4.8 to calculate the residue. The residue is therefore the value of $(2z + 2a)^{-1}$ at the root. By the residue theorem the original integral therefore equals $\frac{2\pi}{\sqrt{a^2-1}}$.

2. We next consider an integral $\int_{-\infty}^{\infty} \rho(x)\, dx$ where ρ is a rational function with no real poles and the degree of the denominator at least 2 higher than the degree of the numerator, so that the integral certainly converges. To calculate this using residue calculus, let γ be a half circle in the upper half plane, centered at the origin and with radius R, together with the real line segment $[-R, R]$. We give γ positive orientation. For R sufficiently large, all the poles of ρ which are in the upper half plane will be inside γ so that $\int_\gamma \rho(z)\, dz = 2\pi i \sum_{\operatorname{Im} z > 0} \operatorname{Res} \rho(z)$.

On the other hand, along the part of γ which is a half circle we can estimate the integral by

$$\Big| \int\limits_{\substack{|z|=R\\ \operatorname{Im} z>0}} \rho(z)\,dz \Big| \le \sup_{|z|=R} |z^2\rho(z)| \int\limits_{\substack{|z|=R\\ \operatorname{Im} z>0}} \frac{|dz|}{|z|^2} = 2\pi \sup_{|z|=R} |z^2\rho(z)|/R \to 0$$

as $R \to \infty$, since $z^2\rho(z)$ is bounded for large values of $|z|$, by the assumption on the degree of ρ. It follows that

$$\int\limits_{-\infty}^{\infty} \rho(x)\,dx = 2\pi i \sum_{\operatorname{Im} z>0} \operatorname{Res} \rho(z)\ .$$

EXAMPLE 4.10. Consider the integral $\int_{-\infty}^{\infty} \frac{x^2}{x^4+1}\,dx$ which satisfies all the requirements above. The poles of the integrand are given by the zeros of its denominator, so they are the roots of $z^4+1=0$. Setting $z = Re^{i\theta}$ we easily see that the roots are $z_k = e^{i(\pi/4+k\pi/2)}$, $k=0,1,2,3$. The roots in the upper half plane are the two first ones. Since the zeros are simple ones, the residues are obtained by evaluating $\frac{z^2}{4z^3} = \frac{1}{4z}$ at these points. It follows that

$$\int\limits_{-\infty}^{\infty} \frac{x^2}{x^4+1}\,dx = 2\pi i(e^{-i\pi/4} + e^{-3i\pi/4})/4 = \pi\sqrt{2}/2\ .$$

3. We next consider an integral $\int_{-\infty}^{\infty} \rho(x)e^{iax}\,dx$ where a is real and ρ a rational function without real poles. This is the Fourier transform of the function ρ at $-a$. We assume that the degree of the denominator of ρ is higher than the degree of the numerator. This does not guarantee absolute convergence of the integral, but as we shall see it does imply conditional convergence if $a \neq 0$. In the calculations below we shall assume that $a > 0$, but the case $a < 0$ can be treated very similarly. This is done either by replacing the upper half plane by the lower half plane in the considerations below, or else by first making the change of variable $t = -x$ in the integral, which has the effect of replacing a by $-a$.

Let A, B and C be positive real numbers. We consider a contour γ starting with a segment $[-A, B]$ of the real axis, continuing with a vertical line segment from B to $B+iC$, then a horizontal line segment from $B+iC$ to $-A+iC$ and finally a vertical line segment from $-A+iC$ to $-A$. If A, B and C are sufficiently large, this rectangle will contain all poles of ρ which are in the upper half plane so that

$$\int\limits_{\gamma} \rho(z)e^{iaz}\,dz = 2\pi i \sum_{\operatorname{Im} z>0} \operatorname{Res}(\rho(z)e^{iaz})\ .$$

Our assumptions guarantee that $z\rho(z)$ is bounded for $|z|$ sufficiently large, say $|z| > R$. Let a corresponding bound be M. If we parametrize the vertical line segment from A to $A+iC$ by $z=A+it$, $0\le t\le C$ the absolute value of the corresponding integral may be estimated by

$$\int_0^C \frac{M}{|A+it|}e^{-at}\,dt \le \frac{M}{A}\int_0^\infty e^{-at}\,dt = \frac{M}{aA}\ ,$$

provided $A>R$. Similarly, the integral over the other vertical side may be estimated by $\frac{M}{aB}$ provided $B>R$. Note that these estimates are independent of C. Assuming that $C>R$ and parametrizing the upper side of the rectangle by $z=-t+iC$, $-B\le t\le A$, we can similarly estimate the corresponding part of the integral by

$$\int_{-B}^{A} \frac{M}{|-t+iC|}e^{-aC}\,dt \le \frac{M(A+B)}{Ce^{aC}}\ .$$

This clearly tends to 0 as $C\to\infty$. It follows that

$$\Big|\int_{-A}^{B} \rho(x)e^{iax}\,dx - 2\pi i\sum_{\operatorname{Im} z>0}\operatorname{Res}(\rho(z)e^{iaz}\Big| \le \frac{M}{a}(\frac{1}{A}+\frac{1}{B})\ .$$

This shows that the original integral indeed converges, at least conditionally, and that its value is given by the residues in the upper half plane.

Example 4.11. Consider the integral $\int_{-\infty}^{\infty}\frac{\cos(x\xi)}{x^2+1}\,dx$, where ξ is real. First note that the cos function is even. We may therefore replace ξ by $|\xi|$ without affecting the value of the integral. Next, note that the integral is the real part of $\int_{-\infty}^{\infty}\frac{e^{ix|\xi|}}{x^2+1}\,dx$ which we may evaluate using the method above, and then take the real part of. Actually, since it is easily seen that the integrand of the imaginary part is an odd function, the imaginary part is zero anyway. Note, however, that we can *not* evaluate the present integral, or integrals similar to it, by considering the residues of $\frac{\cos(z\xi)}{z^2+1}$, since the function $\cos(z\xi)$ is large for large $|\operatorname{Im} z|$, in both upper and lower half planes.

According to our deliberations above, we have

$$\int_{-\infty}^{\infty}\frac{e^{ix|\xi|}}{x^2+1}\,dx = 2\pi i\sum_{\operatorname{Im} z>0}\operatorname{Res}\frac{e^{iz|\xi|}}{z^2+1}\ .$$

In the upper half plane there is only one singularity, a simple pole at $z=i$. Thus, the residue is obtained by evaluating $\frac{1}{2z}e^{iz|\xi|}$ at $z=i$. This gives $\frac{e^{-|\xi|}}{2i}$ so that we obtain

$$\int_{-\infty}^{\infty}\frac{\cos(x\xi)}{x^2+1}\,dx = \pi e^{-|\xi|}\ .$$

EXAMPLE 4.12. We will consider one more example of this type of integral, with an added difficulty. The integral we want to evaluate is $\int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx$. According to our present strategy we ought to relate this to the residues of e^{iz}/z. Unfortunately this function has a pole at the origin; note that there is no such problem with $\frac{\sin x}{x}$, which is an entire function. This is immediately seen from the power series expansion. To circumvent the difficulty we modify the path so that the line segment $[-A, B]$ on the real axis is replaced by the two line segments $[-A, -r]$ and $[r, B]$, connected by a half circle in the upper half plane, centered at the origin and with radius r. If γ denotes this half circle, but run through counterclockwise, estimates of the same kind as before show that

$$\int_{|x|>r} \frac{\sin x}{x}\, dx = \operatorname{Im}\Big\{ \int_{\gamma} \frac{e^{iz}}{z}\, dz + 2\pi i \sum_{\operatorname{Im} z>0} \operatorname{Res} \frac{e^{iz}}{z} \Big\} .$$

Since there are no poles in the upper half plane we only need to consider the integral over γ. If we parametrize γ by $z = re^{i\theta}$, $0 \le \theta \le \pi$, the integral equals $i \int_0^{\pi} \exp(ire^{i\theta})\, d\theta$ which tends to $i\pi$ as $r \to 0$. It follows that $\int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx = \pi$.

4. Next consider the integral $\int_0^{\infty} x^{\alpha}\rho(x)\, dx$ where $0 < \alpha < 1$ and ρ is rational with no positive real poles. For convergence we must assume that the degree of the denominator in ρ is at least 2 more than the degree of the numerator. Similarly, we may allow at most a simple pole at the origin. If we want to relate the value of this integral to the residues of the function $z^{\alpha}\rho(z)$, note that we now have a branch point at the origin. However, instead of causing difficulties this is actually what will allow us to evaluate the integral.

Suppose we choose the branch of z^{α} where the plane is cut along the positive real axis and for which $0 < \arg z < 2\pi$. This means that $z^{\alpha} = e^{\alpha \log z}$ where log is the corresponding branch of the logarithm. As we approach a point $x > 0$ on the real axis from above we then obtain the usual real power x^{α}. However, as we approach the same point from below we instead obtain $e^{\alpha(\ln x + i2\pi)} = e^{i2\pi\alpha}\, x^{\alpha}$.

Intuitively, we would therefore like to choose a contour γ which starts at $r > 0$ on the 'upper edge' of the real axis, continues to $R > r$, then follows the circle with radius R and centered at the origin, counterclockwise, until we reach R on the 'lower edge' of the real axis, then back to r, still on the 'lower edge', and finally along the circle with radius r and centered at the origin, clockwise, until we reach the initial point again (you will need to draw a picture of this). The two contributions from integrating along the real axis will not cancel since the power has different values along the 'upper' and 'lower' edges. The catch is, there is no such thing as an upper or lower edge of the positive real axis; in fact, since we have cut the plane along the positive real axis, we can't integrate along it at all.

The problem can be avoided in the following way. First cut the plane along some ray from the origin other than the positive real axis. In the remaining part of the plane define a branch of z^α by requiring its values to be real on the positive real axis. Now pick a contour, starting at $r > 0$, continuing to $R > r$, then along the circle of radius R as before, but stop before you reach the branch cut and go back along a ray ω, which does not contain any pole of ρ, until you reach the circle with radius r again. Finally, continue clockwise along this circle until you reach the point r again.

Next, pick another branch of z^α by cutting the plane along a ray which comes before ω, counted counterclockwise from the positive real axis. Fix the branch by requiring its values on ω to coincide with the values of the earlier branch. This will give the branch the value $e^{i2\pi\alpha}x^\alpha$ for a real, positive x. We integrate the new branch along a contour which starts at R, continuous along the positive real axis to r, then follows the circle with radius r clockwise until it reaches the ray ω, then follows this ray outwards until it reaches the circle with radius R. Finally, it follows this circle counterclockwise until it reaches the point R on the positive real axis again.

If we add the two integrals constructed above, the contributions to them along the ray ω will cancel, and the total effect will be exactly as if we could integrate as we originally did, letting z^α have different values along the 'upper' and 'lower' edges of the positive real axis. In view of this, we will not commit any errors of we think of this as being possible.

Now let us estimate the integrals along the circles. Note that $|z^\alpha| = |z|^\alpha$ whatever branch we use (as long as α is real). Our assumption on the degree of ρ means that $z^2\rho(z)$ is bounded, say by M, for large $|z|$. If R is sufficiently large we therefore get

$$|\int_{|z|=R} z^\alpha \rho(z)\,dz| \le MR^{\alpha-2} \int_{|z|=R} |dz| = 2\pi MR^{\alpha-1} \to 0$$

as $R \to \infty$. Similarly, $z\rho(z)$ is bounded, say by m, for $|z|$ sufficiently small. If $r > 0$ is sufficiently small we therefore get

$$|\int_{|z|=r} z^\alpha \rho(z)\,dz| \le mr^{\alpha-1} \int_{|z|=R} |dz| = 2\pi mr^\alpha \to 0$$

as $r \to 0$. The integrals along the positive real axis together contribute $(1 - e^{i2\pi\alpha}) \int_r^R x^\alpha \rho(x)\,dx$. It follows that

$$\int_0^\infty x^\alpha \rho(x)\,dx = (1 - e^{i2\pi\alpha})^{-1} 2\pi i \sum_{z\neq 0} \operatorname{Res}(z^\alpha \rho(z)) \ .$$

Note that it is *extremely important* that one uses the branch of z^α where the plane is cut along the positive real axis and $0 < \arg z < 2\pi$.

EXAMPLE 4.13. Consider the integral $\int_0^\infty \frac{x^\alpha\,dx}{x^2+1}$, $0 < \alpha < 1$. Using the appropriate branch the function $\frac{z^\alpha}{z^2+1}$ has simple poles at $\pm i$, so the residues at these points are the values of $\frac{z^\alpha}{2z}$ at $\pm i$.

The integrand satisfies all the conditions above, so the integral equals $(1-e^{i2\pi\alpha})^{-1}\pi(i^\alpha-(-i)^\alpha)$. We have to choose the argument of i to be $\pi/2$ and that of $-i$ to be $3\pi/2$ so $i^\alpha-(-i)^\alpha = e^{i\alpha\pi/2}-e^{i3\alpha\pi/2} = e^{i\alpha\pi}(e^{-i\alpha\pi/2}-e^{i\alpha\pi/2})$. Similarly, $1-e^{i2\alpha\pi} = e^{i\alpha\pi}(e^{-i\alpha\pi}-e^{i\alpha\pi})$.

The second factor equals $(e^{-i\alpha\pi/2}-e^{i\alpha\pi/2}) \times (e^{-i\alpha\pi/2}+e^{i\alpha\pi/2}) = (e^{-i\alpha\pi/2}-e^{i\alpha\pi/2})2\cos(\alpha\pi/2)$. It follows that

$$\int_0^\infty \frac{x^\alpha\,dx}{x^2+1} = \frac{\pi}{2\cos(\alpha\pi/2)}\ .$$

This is actually true for $-1 < \alpha < 0$ too, since in that case we may write the integrand as $\frac{x^{\alpha+1}}{x(x^2+1)}$ where now $0 < \alpha+1 < 1$ so all the assumptions are satisfied. But the residues remain the same since $z^{\alpha+1}/z = z^\alpha$, using the same branches of the powers. The formula is of course also true for $\alpha = 0$, for elementary reasons.

5. We finally consider an integral $\int_0^\infty \rho(x)\ln x\,dx$, where again ρ is rational without positive real poles. We still need to assume that the degree of the denominator in ρ is at least 2 more than that of the numerator. In contrast to integrals of type **4**, however, we can no longer allow a pole at the origin. On the other hand, we use the same contour, justifying the use of different values of the logarithm along the 'upper' and 'lower' edges of the positive real axis as before. If we consider $\rho(z)\log z$, using the branch of the logarithm where $0 < \arg z < 2\pi$, its values at $x > 0$ on the 'upper' edge of the positive real axis is $\rho(x)\ln x$, where ln is the usual real logarithm. For $x > 0$ on the 'lower' edge we instead get $\rho(x)(\ln x + 2\pi i)$. The difference is therefore $-2\pi i\rho(x)$, so we will not get the integral we are looking for. So, instead we consider the function $(\log z)^2\rho(z)$ which is $(\ln x)^2\rho(x)$ on the upper and $(\ln^2 x + 4i\pi\ln x - 4\pi^2)\rho(x)$ on the lower edge of the positive real axis. The difference is therefore $-4\pi i\rho(x)\ln x + 4\pi^2\rho(x)$. It follows that

$$\int_0^\infty \rho(x)\ln x\,dx + i\pi\int_0^\infty \rho(x)\,dx = -\frac{1}{2}\sum_{z\neq 0}\operatorname{Res}((\log z)^2\rho(z))\ .$$

If, as we normally assume, ρ has real coefficients we can therefore calculate the desired integral by taking the real part of the right hand side. Otherwise, we would first have to calculate the second integral by integrating $\rho(z)\log z$ as in our first attempt.

EXAMPLE 4.14. Consider the integral $\int_0^\infty \frac{\ln x}{x^2+1}\,dx$ which satisfies the assumptions above. Using the appropriate branch in the plane cut along the positive real axis, the function $\frac{(\log z)^2}{z^2+1}$ has simple poles at $\pm i$ so the residues are the values of $\frac{(\log z)^2}{2z}$ at these points. The sum of the residues is therefore $\frac{1}{2i}((i\pi/2)^2 - (i3\pi/2)^2) = -i\pi^2$ which is purely imaginary. It follows that

$$\int_0^\infty \frac{\ln x}{x^2+1}\,dx = 0\ .$$

Incidentally, by taking the imaginary part, we also get $\int_0^\infty \frac{dx}{1+x^2} = \pi/2$, but there is of course an easier way of getting this.

EXAMPLE 4.15. As a final example we consider an integral of type 5, but with an added difficulty. Since $\ln x$ has a simple zero at $x = 1$ it ought to be possible to allow a simple pole of ρ at 1. This causes problems, however, since the branch of the logarithm we use on the 'lower edge' of the positive real axis does not have a zero at 1. To circumvent the difficulty, we replace the part of the integral along the lower edge between $1+r$ and $1-r$ by a half circle of radius $r > 0$ in the lower half plane, centered at 1. As an example, consider $\int_0^\infty \frac{\ln x}{x^2-1}\,dx$. The integral along the half circle is then

$$-i\int_\pi^{2\pi} \frac{(\log(1+re^{i\theta}))^2}{(1+re^{i\theta})^2-1} re^{i\theta}\,d\theta \to 2\pi^3 i \qquad \text{as } r \to 0\ .$$

Note that the logarithm approaches $2\pi i$ as $r \downarrow 0$! The integrand has a simple pole at $z = -1$ and nowhere else, so the sum of the residues is $\frac{(\log(-1))^2}{-2} = \pi^2/2$. It follows that

$$\lim_{r\to 0} \int_{\substack{x\ge 0\\ |x-1|\ge r}} \left(-4\pi i \frac{\ln x}{x^2-1} + 4\pi^2 \frac{1}{x^2-1}\right) dx = -2\pi^3 i + 2\pi i \pi^2/2 = -\pi^3 i\ .$$

Taking imaginary parts and dividing by -4π we finally obtain

$$\int_0^\infty \frac{\ln x}{x^2-1}\,dx = \frac{\pi^2}{4}\ .$$

4.4. The argument principle

The following theorem is a simple consequence of the residue theorem.

THEOREM 4.16. *Suppose that f is meromorphic in a simply connected region Ω and that γ is a cycle in Ω.*

Assume further that f has zeros $a_1, a_2, \dots, a_n$ and poles $b_1, b_2, \dots, b_k$ in Ω, each repeated according to multiplicity and none of them on γ. Then

$$\frac{1}{2\pi i}\int_\gamma \frac{f'(z)}{f(z)}\,dz = \sum_{j=1}^{n} n(\gamma, a_j) - \sum_{j=1}^{k} n(\gamma, b_j).$$

We normally choose γ so that the index of each zero and pole with respect to γ equals one, and then the right hand side becomes the *difference between the number of zeros and the number of poles in* Ω, each counted by multiplicity.

PROOF. If $f(z) = (z-a)^n g(z)$ where n is a non-zero integer, g is analytic near a and $g(a) \neq 0$, then $f'(z) = n(z-a)^{n-1}g(z) + (z-a)^n g'(z)$ so that $\frac{f'(z)}{f(z)} = \frac{n}{z-a} + \frac{g'(z)}{g(z)}$. The last term is analytic near a, so the residue of the left hand side at a is n. Since $|n|$ is the multiplicity of a as a zero respectively pole, the theorem follows from the residue theorem. □

The theorem is usually known as the *argument principle*, for the following reason. If γ is a closed arc, the integral $\int_\gamma \frac{f'(z)}{f(z)}\,dz$ equals $\int_{f\circ\gamma} \frac{dz}{z}$, as is easily seen using the definition of the integral through a parametrization. Thus the integral is $2\pi i\, n(f\circ\gamma, 0)$. But this is the *variation of the argument* of z as z runs through $f\circ\gamma$. To see this, note that the principal logarithm is a primitive of $1/z$ away from the negative real axis. Now, $f\circ\gamma$ may intersect the negative real axis at certain points; assume for simplicity that they are finitely many. At every such intersection we have to add or subtract 2π from the argument of z, depending on whether we intersect from below or above. Between two intersections we may calculate the integral by using the principal branch of the logarithm. Adding everything up, the real parts will cancel, and what remains is an integer multiple of $2\pi i$, in other words i times the variation of the argument along the curve. Clearly the variation of the argument of z along $f\circ\gamma$ is the same as the variation of the argument of $f(z)$ along γ.

The integral is therefore (i times) the variation of the argument of $f(z)$ as z runs through γ. Since one can often find the variation of argument without calculating the integral, this gives information on the number of zeros or poles in a region. Used this way, the argument principle is of great importance to many applications in control theory and related subjects. We give a few examples of how this is done.

EXAMPLE 4.17. We wish to find the number of zeros in the right half plane of the polynomial $p(z) = z^5 + z + 1$.

If $z = iy$ is purely imaginary $p(z) = iy(y^4+1) + 1$ has real part 1, so is never zero. If ϕ is the argument of $p(iy)$ we have $\tan\phi = y(y^4+1)$ which tends to $+\infty$ as $y \to +\infty$ and $-\infty$ as $y \to -\infty$. Running through the imaginary axis from iR to $-iR$ for a large $R > 0$ the argument thus decreases by nearly π. On a large circle $|z| = R$ we have $p(z) = z^5(1+z^{-4}+z^{-5})$, where the second factor is nearly one, so that its argument varies very little, whereas the argument of the first factor increases by 5π as we follow the circle in a positive direction from $-iR$ to iR. The variation of argument of p along the boundary of the large halfdisk is therefore nearly 4π, and since it must be an integer multiple of 2π it is exactly 4π. There are therefore exactly two zeros inside the halfcircle if it is sufficiently large. In other words, there are precisely two zeros in the right half plane!

EXAMPLE 4.18. We wish to find the number of zeros in the first quadrant of the polynomial $f(z) = z^4 - z^3 + 13z^2 - z + 36$.

First note that there are no zeros on either the real or imaginary axes since for $z = x \in \mathbb{R}$ we have

$$x^4 - x^3 + 13x^2 - x + 36 = (x^2+1)(x-\frac{1}{2})^2 + \frac{47}{4}x^2 + \frac{143}{4} > 0$$

and for $z = iy$, $y \in \mathbb{R}$ we have

$$z^4 - z^3 + 13z^2 - z + 36 = y^4 - 13y^2 + 36 + i(y^3 - y).$$

The imaginary part vanishes only for $y = 0$ and $y = \pm 1$, neither of which is a zero for the real part. Now let γ be the line segment from 0 to $R > 0$, followed by a quarter circle of radius R centered at 0 and ending at iR, and finally the vertical line segment from iR to 0. For R sufficiently large, all the zeros in the first quadrant will be inside γ, so we only need to calculate the variation of argument for the polynomial along γ. Since $f > 0$ on the real axis, the argument stays equal to 0 along the horizontal part of γ. For $|z| = R$ we write $f(z) = z^4(1 - \frac{1}{z} + \frac{13}{z^2} - \frac{1}{z^3} + \frac{36}{z^4})$. Note that the bracketed expression tends to 1 as $R \to \infty$ so its argument varies only a little around 0. The argument of the first factor varies 4 times the variation of the argument of z, *i.e.*, by $4\frac{\pi}{2} = 2\pi$. So, along the circular arc the argument varies close to 2π.

It remains to find the variation of the argument along the imaginary axis. If ϕ denotes the argument of $f(z)$, then $\tan\phi = \frac{y^3-y}{y^4-13y^2+36}$. For $y = 0$ this is 0, and for $y \to \infty$ we get $\tan\phi \to 0$. The argument variation along the vertical part of γ is therefore close to some integer multiple of π. To go from one multiple to the next, $\tan\phi$ will have to become ∞ in between. This happens at the zeros of $y^4 - 13y^2 + 36 = (y^2-9)(y^2-4) = (y+3)(y+2)(y-3)(y-2)$. The first two factors stay positive for $y \geq 0$ so the denominator in $\tan\phi$ passes from positive to negative as y decreases through 3, and from negative to positive as y decreases past 2. In both these points the numerator is positive, so $\tan\phi$ passes from $+\infty$ to $-\infty$ as y decreases through 3 and then from $-\infty$ back to $+\infty$ as y decreases through 2.

Hence, if we start at $y = R$ for a large value of R, the variation in argument along the vertical line segment is close to 0. Therefore, for large $R > 0$ the variation in argument of f along γ is close to 2π, and since it has to be an integer multiple of 2π, it is exactly 2π. There is therefore exactly one zero of f in the first quadrant.

A useful consequence of the argument principle is the following theorem.

THEOREM 4.19 (Rouché's theorem). *Suppose f and g are analytic in a simply connected region Ω and that γ is a cycle in Ω such that $n(\gamma, z)$ is 0 or 1 for every $z \in \Omega$.*

*Also assume that $|f(z) - g(z)| < |f(z)|$ for $z \in \gamma$. Then f and g have the same number of zeros, counted with multiplicity, enclosed by γ (*i.e.*, for which the index with respect to γ is 1).*

PROOF. The inequality shows that neither f nor g can have a zero on γ. If we set $F(z) = \frac{g(z)}{f(z)}$, then the zeros for F are the zeros for g and the poles for F are the zeros for f. We therefore need to show that F has the same number of zeros and poles, *i.e.*, that the variation of argument of F along γ is 0. Note that this is true even if f and g have common zeros so that there is some cancellation in F.

However, by assumption $|F(z) - 1| < 1$ for $z \in \gamma$. Hence F has all its values on γ in the disk with radius 1 centered at 1, which does not contain the origin. Hence the variation of argument is 0 (give a detailed motivation, for example using the principal logarithm). ☐

EXAMPLE 4.20. We shall determine the number of zeros in the right half plane of the function $g(z) = a - z - e^{-z}$, where $a > 1$.

It is clear that the function $f(z) = a - z$ has only the zero $z = a$, which is in the right half plane. If γ is a positively oriented half circle in the right half plane, with radius R and centered at the origin, this zero is inside γ as soon as $R > a$. For $z = iy$ on the imaginary axis we have $|f(z)| = \sqrt{a^2 + y^2} \geq a > 1$, and on the circular arc we have $|f(Re^{i\theta})| = |Re^{i\theta} - a| \geq R - a > 1$ if $R > 1 + a$. But $|f(z) - g(z)| = |e^{-z}| = e^{-\operatorname{Re} z} \leq 1$ for z in the right half plane. Hence $|f(z) - g(z)| < |f(z)|$ for $z \in \gamma$ as soon as $R > 1 + a$. Therefore, by Rouché's theorem, g has exactly one zero in the right half plane, and this zero has absolute value $\leq 1 + a$.

The next theorem demonstrates a very important topological property of an analytical map.

THEOREM 4.21. *Suppose f is analytic at z_0 and that $f(z_0) = w_0$ with multiplicity n,* i.e., *$f(z) - w_0$ has a zero of multiplicity n for $z = z_0$. Then, for every sufficiently small $\varepsilon > 0$ there exists a $\delta > 0$ such that if $|a - w_0| < \delta$, then $f(z) = a$ has exactly n roots (counted with multiplicity) in $|z - z_0| < \varepsilon$.*

PROOF. Since zeros of analytic functions are isolated, we may require $\varepsilon > 0$ to be so small that z_0 is the only point in $|z - z_0| \leq \varepsilon$ where

$f(z) - w_0 = 0$. If $\delta = \min_{|z-z_0|=\varepsilon} |f(z) - w_0|$ it follows that the integral

$$\frac{1}{2\pi i} \int_{|z-z_0|=\varepsilon} \frac{f'(z)}{f(z) - a}\, dz$$

is continuous as a function of a for $|a - w_0| < \delta$. But it is also an integer, so it must be constant in this disk. Since it equals n for $a = w_0$ the theorem follows from the argument principle. $\square$

We restate the most important part of the conclusion of Theorem 4.21 as the *open mapping theorem.*

COROLLARY 4.22. *Suppose f is analytic in some region and not constant. Then f is an open mapping,* i.e., *the images of open sets are open.*

PROOF. If z_0 is in the domain of f, then by Theorem 4.21 the image of any sufficiently small neighborhood of z_0 contains a neighborhood of $f(z_0)$. Hence f is an open mapping. $\square$

Note that $n = 1$ in Theorem 4.21 exactly if $f'(z_0) \neq 0$, and that $n = 1$ means that the inverse function f^{-1} is defined in $|z - w_0| < \delta$. By Corollary 4.22 the inverse function has the property that the inverse image of an open set under f^{-1} is open; in other words, the inverse is continuous. But by Theorem 2.6 (5) this implies that f^{-1} is analytic, with $(f^{-1})'(z) = 1/f'(f^{-1}(z))$ (note that the denominator is $\neq 0$ here). We therefore also have the following corollary.

COROLLARY 4.23. *If $f'(z_0) \neq 0$, then f maps a neighborhood of z_0 conformally and topologically (*i.e., *continuously and with continuous inverse) onto a neighborhood of $f(z_0)$.*

It remains to see what type of mapping we have in a neighborhood of a point z_0 where $f'(z_0) = 0$. We have one very well known example; the function $z \mapsto z^n$ where n is an integer > 1. This function has an n-fold zero at $z = 0$, and the image of a neighborhood of 0 covers a neighborhood exactly n times. This is, in fact, what happens in general. To see this, consider a function f such that $f(z) - w_0$ has a zero of order n at z_0. We may then write $f(z) = w_0 + (z - z_0)^n g(z)$, where g is analytic where f is, and $g(z_0) \neq 0$. According to Corollary 3.20 we may therefore define a single-valued branch $h(z)$ of $\sqrt[n]{g(z)}$ which is analytic in a neighborhood of $z = z_0$.

Note that $\frac{d}{dz}(z-z_0)h(z) = h(z) + (z-z_0)h'(z)$ which equals $h(z_0) \neq 0$ for $z = z_0$. The function $z \mapsto (z - z_0)h(z)$ therefore maps a neighborhood of z_0 conformally onto a neighborhood of 0. We may therefore view $f(z) = w_0 + ((z - z_0)h(z))^n$ as a composite of this function, of the function $z \mapsto z^n$, and a translation. It follows that the image of a small neighborhood of z_0 under f covers a neighborhood of w_0 exactly n times.

We turn now from these general considerations to a very useful and very specific result.

THEOREM 4.24 (Maximum principle). *Suppose f is analytic in a region Ω. If $|f|$ has a (local) maximum in Ω, then f is constant.*

A variant of this states that if f is analytic in a compact set, then the maximum of $|f|$ on the set is taken on the boundary unless f is constant. This follows from Theorem 4.24 and the fact that a function continuous on a compact set, in this case $|f|$, takes a maximum value.

PROOF. Suppose f is not constant. According to the open mapping theorem, given any neighborhood $\mathcal{O}$ of z_0, all values in a sufficiently small neighborhood of $f(z_0)$ are taken in $\mathcal{O}$. Some of these values will be further from the origin than $f(z_0)$, so $|f(z_0)|$ can not be a local maximum value of $|f|$. □

A rather special, but as it turns out, very useful, consequence of the maximum principle is the following.

THEOREM 4.25 (Schwarz' lemma). *Suppose f is analytic in $|z| < 1$, that $|f(z)| < 1$ and $f(0) = 0$. Then $|f(z)| \leq |z|$ for $|z| < 1$, $|f'(0)| \leq 1$, and if equality occurs in either of these inequalities, then $f(z) = cz$ for some c with $|c| = 1$.*

PROOF. The function $g(z) = f(z)/z$ has a removable singularity at 0; we must set $g(0) = f'(0)$. For $|z| = R < 1$ we have $|g(z)| < |z|^{-1} = 1/R$ so by the maximum principle we have $|g(z)| < 1/R$ for $|z| < R$. Given any z with $|z| < 1$ we therefore have $|g(z)| < 1/R$ for all R, $|z| < R < 1$. Letting $R \to 1$ we get $|g(z)| \leq 1$ in the unit disk. The maximum principle finally tells us that if we have equality anywhere, *i.e.*, a local maximum of $|g|$, then g is constant. The theorem follows. □

Schwarz' lemma has a very important application in determining to what extent conformal maps are unique. Later we shall show that any simply connected region can be mapped conformally and bijectively onto the unit disk. This immediately shows that any two simply connected regions may be mapped conformally and bijectively onto each other, since one may first map both conformally and bijectively onto the unit disk, and then compose the inverse of one map with the other map. The resulting function then maps one region onto the other conformally and bijectively.

It is clear that uniqueness questions can also be answered if they can be resolved for the special case of a map onto the unit disk. It is immediately clear that if there is a conformal map of Ω onto the unit disk, then we can pick any point $z_0 \in \Omega$ and require it to be mapped to

0. For, by assumption there is a conformal map of Ω onto the unit disk; suppose the image of z_0 is w_0. We can then find a Möbius transform that maps the unit disk onto itself and takes w_0 to 0. Composing the original map with this Möbius transform we obtain a map of Ω which takes z_0 to 0. Is this map unique?

Suppose f and g both map Ω conformally onto the unit disk, and both map $z_0 \in \Omega$ onto 0. Then $f \circ g^{-1}$ maps the unit disk onto itself and keeps 0 fixed. By Schwarz' lemma $|f \circ g^{-1}(w)| \leq |w|$. But setting $z = f \circ g^{-1}(w)$ this means $|g \circ f^{-1}(z)| \geq |z|$. On the other hand, Schwarz' lemma again tells us that $|g \circ f^{-1}(z)| \leq |z|$ so that in fact equality holds throughout the unit disk. A final use of Schwarz' lemma tells us that $f \circ g^{-1}(z) = cz$ where $|c| = 1$.

Note that $c = (f \circ g^{-1})'(0) = \frac{f'(z_0)}{g'(z_0)}$ so that if we specify the argument of the derivative at z_0 as well, the map is unique. A particular case is of course when Ω is the unit disk itself; it follows that the only automorphisms of the unit disk (bijective conformal maps of the unit disk onto itself) are the Möbius transforms with this property. More generally, given any two regions that are circles or half planes, the only bijective conformal maps of one onto the other are Möbius transforms. Similar statements can be made with respect to the other special regions for which we found explicit conformal maps (wedges, infinite strips, *etc.*).

CHAPTER 5

Harmonic functions

5.1. Fundamental properties

Suppose f is analytic in some region Ω and u, v are its real and imaginary parts, so that $f(x+iy) = u(x,y) + iv(x,y)$. Then u and v are *harmonic* in Ω, according to the following definition.

DEFINITION 5.1. A function u defined in an region $\Omega \subset \mathbb{C}$ is called harmonic if it is twice continuously differentiable in Ω and satisfies $\Delta u = 0$, where $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$.

This follows since u, v satisfy the Cauchy-Riemann equations

$$\begin{cases} u_x = v_y \ , \\ u_y = -v_x \ . \end{cases}$$

Since f is infinitely differentiable, we can differentiate the first equation with respect to x, the second with respect to y, and add the results to obtain $\Delta u = 0$, using that $v_{xy} = v_{yx}$. Similarly one shows that v is harmonic.

If a function u, harmonic in Ω, is given, then another harmonic function v is called a *conjugate* function to u in Ω if $u + iv$ is analytic in Ω. Note that if u has a conjugate function in some region, then it is determined up to an additive real constant. For suppose $u + iv$ and $u + i\tilde{v}$ are both analytic. Then so is the difference $i(v - \tilde{v})$ which has real part 0. It follows that the imaginary part $v - \tilde{v}$ is constant (this follows from the Cauchy-Riemann equations, but also directly from the open mapping theorem).

Note that if v is the harmonic conjugate of u, then $-u$ is the harmonic conjugate of v, since $v - iu = -i(u + iv)$ is analytic if $u + iv$ is. A harmonic function does not necessarily have a conjugate function defined in all of its domain; consider for example $\ln \sqrt{x^2 + y^2}$ which is the real part of any branch of the logarithm and therefore harmonic in $\mathbb{R}^2 \setminus \{(0,0)\}$. It can not have a conjugate function in this set, because that would imply that we could define a single-valued branch of the logarithm in the plane with just the origin removed. But we can't. On the other hand, *locally* there is always a conjugate function. In fact, the following theorem holds.

THEOREM 5.2. *If u is harmonic in a disk, then it has a conjugate function there.*

PROOF. Suppose (x_0, y_0) is the center of the disk and set $v(x,y) = \int_{y_0}^{y} u_x(x,t)\,dt - \int_{x_0}^{x} u_y(t,y_0)\,dt$ for any (x,y) in the disk. Note that v is well defined since we are only evaluating u at points in the disk. By the fundamental theorem of calculus we have $v_y = u_x$ and differentiating under the integral sign we obtain

$$v_x(x,y) = \int_{y_0}^{y} u_{xx}(x,t)\,dt - u_y(x,y_0)$$

$$= -\int_{y_0}^{y} u_{yy}(x,t)\,dt - u_y(x,y_0) = -u_y(x,y),$$

using the fact that u is harmonic. So, v is a harmonic conjugate of u. □

Since analytic functions are infinitely differentiable we immediately obtain the following corollary.

COROLLARY 5.3. *Harmonic functions are infinitely differentiable.*

There is a much more general version of the theorem, which states that any function harmonic in a simply connected region has a harmonic conjugate there. However, since this follows from Exercise 5.5 below and the Riemann mapping theorem, which we will prove later, we will not attempt a proof here.

COROLLARY 5.4. *Suppose f is analytic in Ω and u is harmonic in the range of f. Then $u \circ f$ is harmonic in Ω.*

PROOF. In a neighborhood of any point in its domain u has a conjugate function, so it is the real part of some analytic function g defined near the point. Since the composite $g \circ f$ is analytic, its real part $u \circ f$ is harmonic. □

EXERCISE 5.5. Suppose u is harmonic in the region Ω and that one can find a bijective conformal mapping of Ω onto the unit disk. Show that u has a harmonic conjugate in Ω.

The next theorem is also a simple corollary of Theorem 5.2, but it is so important it is a theorem anyway.

THEOREM 5.6 (Mean value property). *Suppose u is harmonic in the open disk centered at z with radius R, and continuous in the closed disk. Then*

$$u(z) = \frac{1}{2\pi} \int_{0}^{2\pi} u(z + Re^{i\theta})\,d\theta\ .$$

PROOF. In the open disk u is the real part of an analytic function f, by Theorem 5.2. If $0 < r < R$ Cauchy's integral formula implies that $f(z) = \frac{1}{2\pi i} \int_{|\zeta - z| = r} \frac{f(\zeta)}{\zeta - z} \, d\zeta$. Parametrizing the circle by $\zeta = z + re^{i\theta}$, $0 \leq \theta \leq 2\pi$, gives

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} f(z + re^{i\theta}) \, d\theta \; .$$

Taking the real part of this gives the desired formula with R replaced by r. By the continuity of the integrand, however, we may now let $r \to R$ and so obtain the desired result. □

Clearly one can calculate mean values in the above sense for any continuous function. Interestingly enough, any continuous function having the mean value property has to be harmonic (and is therefore also infinitely differentiable). We will show this in Theorem 5.11.

THEOREM 5.7 (Maximum principle). *Suppose u is continuous on the closure of a bounded region Ω and satisfies the mean value property in Ω. Then u takes its largest and smallest value in $\overline{\Omega}$ on $\partial\Omega$, and if either is assumed in an interior point, then u is constant.*

PROOF. Suppose $a \in \Omega$ and $\sup_{\overline{\Omega}} = u(a)$. There is a disk $|z - a| < R$ contained in Ω, and $u(a + re^{i\theta}) \leq u(a)$ for all θ and $0 < r < R$. If there is strict inequality for some choice of r, θ, then there is strict inequality in a neighborhood by continuity, and $\frac{1}{2\pi} \int_0^{2\pi} u(a + re^{i\theta} \, d\theta < u(a)$, violating the mean value property.

Thus the set $M = \{z \in \Omega \mid u(z) = u(a)\}$ is open, as is the complement $\{z \in \Omega \mid u(z) \neq u(a)\}$ by continuity. Since Ω is connected and $M \neq \varnothing$ it follows that $M = \Omega$, so that u is constant.

Since $-u$ satisfies the mean value property if u does, the statement about smallest value follows as well. □

Harmonic functions satisfy the mean value property, so the theorem applies to harmonic functions. We obtain a corollary, which is also referred to as the maximum principle.

COROLLARY 5.8. *Suppose u is harmonic and not constant in a region Ω. Then u has no local extrema in Ω.*

PROOF. By Theorem 5.7 u is constant in a neighborhood of a local extremum point a. Consider the set

$$M = \{z \in \Omega \mid u(\zeta) = u(a) \text{ for } \zeta \text{ in a neighborhood of } z\}.$$

Clearly M is open. But if $z_j \in M$, $z_j \to z \in \Omega$, then any neighborhood of z contains a disk where u is identically $u(a)$. Therefore, near z the function u is the real part of an analytic function which is constant on an open set and therefore is constant. It follows that $z \in M$ so M is also relatively closed in Ω. Since Ω is connected and $M \neq \varnothing$, it follows that $M = \Omega$, *i.e.*, u is constant in Ω. □

A problem of great importance both for the theory of harmonic functions and their applications is *Dirichlet's problem.* It concerns the possibility of finding a function harmonic in a given region, continuous on its closure, and taking prescribed values on the boundary. There are also other, more general formulations which will not concern us here. Note that if we can solve Dirichlet's problem for some region Ω, and if we can find a conformal map of Ω onto some other region ω which extends continuously as an invertible map of the closure of Ω onto the closure of ω, then by Corollary 5.4, we can also solve Dirichlet's problem for the region ω. We first note:

THEOREM 5.9. *If Dirichlet's problem has a solution for a bounded region Ω, then it is unique.*

PROOF. Suppose u and v are harmonic in Ω, continuous in the closure and agree on $\partial\Omega$. Then $u-v$ is harmonic in Ω and vanishes on the boundary. But according to Theorem 5.7 it takes both its largest and smallest value on the boundary; we therefore have $u = v$ throughout Ω. □

To prove the existence of a solution is much harder, and requires additional assumptions. We will here give a solution for the simple case when Ω is a disk centered at the origin. In the next section we will show the existence of a solution in much more general circumstances.

We start by assuming that we have a function u, harmonic in $|z| < R$ and continuous in $|z| \leq R$. We should like to express the values of u in the interior of the disk in terms of its values on the boundary. The mean value property gives us such a formula for the center of the circle. An obvious way of trying to get a formula for other interior points would be to use a Möbius transform to map the unit disk onto the given disk in such a way as to map the origin to a given point a in the disk. The map $T(\zeta) = R\frac{R\zeta+a}{R+\bar{a}\zeta}$, which has inverse $\zeta = R\frac{z-a}{R^2-\bar{a}z}$, does exactly that. Hence

$$u(a) = u(T(0)) = \frac{1}{2\pi}\int_0^{2\pi} u(T(e^{i\phi}))\,d\phi = \frac{1}{2\pi i}\int_{|\zeta|=1} u(T(\zeta))\,\frac{d\zeta}{\zeta}\ .$$

Setting $z = T(\zeta)$ in this integral gives (note that $z\bar{z} = R^2$)

$$u(a) = \frac{1}{2\pi i}\int_{|z|=R} u(z)\frac{R^2-|a|^2}{|z-a|^2}\,\frac{dz}{z}\ .$$

Setting $z = Re^{i\theta}$ finally gives *Poisson's integral formula*

$$u(a) = \frac{1}{2\pi}\int_0^{2\pi} u(Re^{i\theta})\frac{R^2-|a|^2}{|Re^{i\theta}-a|^2}\,d\theta\ . \tag{5.1}$$

What we have just seen is this: If Dirichlet's problem for the disk $|z| < R$ has a solution, it must be given by Poisson's integral formula. Note that $\frac{R^2-|a|^2}{|z-a|^2} = \operatorname{Re}\frac{z+a}{z-a}$ and that for $|a| < R$ the integral

$$\frac{1}{2\pi}\int_0^{2\pi} u(Re^{i\theta})\frac{Re^{i\theta}+a}{Re^{i\theta}-a}\,d\theta \tag{5.2}$$

is an analytic function of a, as is seen by differentiating under the integral sign. The real part of this integral is Poisson's integral, so that the imaginary part is a conjugate harmonic function to u in $|z| < R$. But (5.2) is an analytic function whether u is harmonic or not, as long as it behaves well enough on the boundary $|z| = R$ for us to be allowed to differentiate under the integral sign. Continuity is certainly enough. It follows that Poisson's integral represents a harmonic function for any function u defined and continuous on $|z| = R$. We denote this function by P_u, so that we know that $P_u = u$ in the disk if u is known to be harmonic in the interior and continuous on the closed disk.

If u is only defined and continuous on the boundary we still know that P_u is harmonic in the interior. To show that P_u solves Dirichlet's problem, it only remains to show that it assumes the correct boundary values. First note that, since a constant is harmonic, the integral of the *Poisson kernel* $\frac{1}{2\pi}\frac{R^2-|a|^2}{|Re^{i\theta}-a|^2}$ is 1 for all a, $|a| < R$. Since the Poisson kernel is also positive, it follows that

$$|P_u(a) - u(Re^{i\phi})| \le \frac{1}{2\pi}\int_0^{2\pi} |u(Re^{i\theta}) - u(Re^{i\phi})|\frac{R^2-|a|^2}{|Re^{i\theta}-a|^2}\,d\theta\ .$$

Given $\varepsilon > 0$ we may find $\delta > 0$ so that $|u(Re^{i\theta}) - u(Re^{i\phi})| < \varepsilon$ for $\phi - \delta < \theta < \phi + \delta$. The integral over $[0, \phi - \delta] \cup [\phi + \delta, 2\pi]$ (if $\phi = 0$, over $[\delta, 2\pi - \delta]$) clearly tends to 0 as $a \to Re^{i\phi}$, and the integral over $[\phi - \delta, \phi + \delta]$ (respectively $[0, \delta] \cup [2\pi - \delta, 2\pi]$) is $< \varepsilon$. It follows that $|P_u(a) - u(Re^{i\phi})| \to 0$ as $a \to Re^{i\phi}$ so that actually P_u tends to the correct boundary values. We have proved the following theorem.

THEOREM 5.10. *Suppose u is a continuous function defined on $|z| = R$. Then the function which equals $P_u(z)$ for $|z| < R$ and $u(z)$ for $|z| = R$ is harmonic in $|z| < R$ and continuous in $|z| \le R$.*

In the process of solving Dirichlet's problem we also obtained (5.2) which expresses the values of a function analytic in the disk $|z| < R$ in terms of the boundary values of its real part, in the case when these are assumed continuous. This is a well known theorem by H. A. Schwarz.

Theorem 5.11. *Suppose u is continuous in a region $\Omega \subset \mathbb{C}$ and has the mean value property there. Then u is harmonic.*

Proof. Let $|z - z_0| < R$ be an open disk with closure contained in Ω and P_u the Poisson integral applied to $u(\cdot + z_0)$. Then $P_u(\cdot + z_0)$ is harmonic in the disk so that $P_u - u$ satisfies the mean value property in the disk and is continuous in its closure. Therefore $P_u - u$ satisfies the maximum principle Theorem 5.7 in the closed disk. But $P_u - u$ vanishes on the boundary of the disk and is therefore identically 0. Thus $P_u = u$ in the disk, so that u is harmonic. □

We finally consider the *reflection principle*. In order to formulate the theorem, let us call a region Ω *symmetric* with respect to the real axis if for each z it contains z if and only if it contains $\overline{z}$. We denote the intersection of Ω with the real axis by σ and the part of Ω which is in the (open) upper half plane by Ω_+.

Theorem 5.12 (Reflection principle). *Suppose v is continuous in $\Omega_+ \cup \sigma$, vanishes on σ and is harmonic in Ω_+. Then v has a harmonic extension to Ω satisfying the symmetry $v(\overline{z}) = -v(z)$. If v is the imaginary part of a function f analytic in Ω_+, then f has an analytic extension to Ω satisfying $f(z) = \overline{f(\overline{z})}$.*

Proof. If we define the extension of v by setting $v(z) = -v(\overline{z})$ for $z \in \Omega \cap \{z \mid \operatorname{Im} z < 0\}$ it is clear that v is continuous in Ω and harmonic except possibly on σ. Let p be an arbitrary point on σ. We need to show that v is harmonic in a neighborhood of p. Let $R > 0$ be so small that the disk $|z - p| \leq R$ is contained in Ω, and let P_v be the Poisson integral corresponding to this disk, extended by continuity to the boundary of the disk. Then P_v is harmonic in $|z - p| < R$ and we will be done if we can prove that P_v coincides with v there.

Now P_v vanishes on the real diameter of $|z - p| < R$ because of the symmetry in v, and the boundary values of P_v on $|z - p| = R$ coincide with those of v, by Theorem 5.10. Hence the function $P_v - v$, which is harmonic in the half disk $|z - p| < R$, $\operatorname{Im} z > 0$, has vanishing boundary values in this half disk. By the maximum principle $P_v = v$ in the half disk, so P_v is a harmonic extension of v to the whole disk, and obviously has the same symmetry as v. It follows that P_v coincides with v in the disk, so that v is harmonic there.

Now suppose f is analytic in Ω_+ with imaginary part v there. Consider a disk as before with center on σ. In this disk v has a harmonic conjugate $-u$ so that $g = u + iv$ is analytic in the disk. Now $\overline{g(\overline{z})}$ is also analytic in the disk and the function $g(z) - \overline{g(\overline{z})}$ is analytic in the disk, has zero imaginary part and vanishes on the real axis. It follows that this function is identically zero so that g has the appropriate symmetry. Since $f - g$ has zero imaginary part in the upper half circle it is a real constant there. It follows that f can be extended analytically as claimed. □

5.2. Dirichlet's problem

In this section we will solve the Dirichlet problem by *Perron's method.* Recall first that one version of Dirichlet's problem is the following:

> **Find a function u harmonic in a given region Ω such that $u(z) \to f(\zeta)$ as $\Omega \ni z \to \zeta \in \partial\Omega$ where f is a given function on $\partial\Omega$.**

It is not hard to see that this problem can not be solved in general without assumptions both on the boundary values f and the nature of the boundary $\partial\Omega$. We will impose such conditions later. Perron's method, like many other methods for solving Dirichlet's problem, consists in converting the problem of finding a solution to a maximization problem. To explain how, we need to make the following definition.

Definition 5.13. A real-valued function v, defined and continuous in a region Ω, is called *subharmonic* if for every u harmonic in a subregion $\tilde{\Omega}$ of Ω the function $v - u$ satisfies the maximum principle in $\tilde{\Omega}$.

That $v - u$ satisfies the maximum principle in $\tilde{\Omega}$ means that $v - u$ has no maximum in $\tilde{\Omega}$ unless it is constant. The following theorem gives a more concrete characterization of subharmonicity.

Theorem 5.14. *A continuous function v is subharmonic in Ω if and only if*

$$v(z_0) \le \frac{1}{2\pi} \int_0^{2\pi} v(z_0 + re^{i\theta})\, d\theta$$

whenever the disk $|z - z_0| \le r$ is contained in Ω.

Proof. If the inequality holds, then it holds also for $v - u$ since u has the mean value property. But the inequality is all that is needed to prove the maximum principle (*cf.* the proof of Theorem 5.7) so that one direction of the theorem follows.

Conversely, if $v - u$ satisfies the maximum principle for every harmonic u we may for u pick the Poisson integral P_v belonging to v on the circle $|z - z_0| = r$. Then $v(z) - P_v(z)$ approaches 0 as z approaches the circle from its interior (Theorem 5.10). By the maximum principle $v - P_v \le 0$ in the disk; in particular, for $z = z_0$ we obtain the desired inequality. □

We list some elementary properties of subharmonic functions.

(1) If v is subharmonic in Ω, then so is kv for any non-negative constant k.
(2) If v_1 and v_2 are subharmonic in Ω, then so is $v_1 + v_2$.
(3) If v_1 and v_2 are subharmonic in Ω, then so is $\max(v_1, v_2)$.
(4) If v is subharmonic in Ω, D is a disk whose closure is in Ω and P_v is the Poisson integral corresponding to this disk with boundary values given by v, put $\tilde{v} = P_v$ in D and $\tilde{v} = v$ in $\Omega \setminus D$. Then $\tilde{v}$ is subharmonic in Ω.

The first two properties are immediate consequences either of the definition or of Theorem 5.14. The other properties are only a little less obvious.

Proof of (3). Let $v = \max(v_1, v_2)$ and suppose $v - u$ has a maximum at $z_0 \in \tilde{\Omega} \subset \Omega$, where u is harmonic in $\tilde{\Omega}$. We may assume $v(z_0) = v_1(z_0)$. We then have

$$v_1(z) - u(z) \leq v(z) - u(z) \leq v(z_0) - u(z_0) = v_1(z_0) - u(z_0)$$

for $z \in \tilde{\Omega}$. It follows, first that $v_1 - u$ is constant, and then from the same inequality that $v - u$ is constant. Hence v is subharmonic. ☐

Proof of (4). By Theorem 5.10 $\tilde{v}$ is continuous. We have $v \leq P_v$ in D so $v \leq \tilde{v}$ throughout Ω. Since P_v is harmonic and v subharmonic it follows that $\tilde{v}$ is subharmonic except possibly on ∂D. But if $\tilde{v} - u$ has a maximum at a point on ∂D, then so has $v - u$ so that $v - u$ is constant. But then it follows that also $P_v - u$ and hence $\tilde{v} - u$ is constant. ☐

Note that any harmonic function is also subharmonic. It follows by the maximum principle that it is greater than any subharmonic function with smaller boundary values. If we therefore let $\mathcal{F}$ denote the set of all functions v subharmonic in Ω which have the additional property that $\overline{\lim}_{\Omega \ni z \to \zeta} v(z) \leq f(\zeta)$ for every $\zeta \in \partial\Omega$, then the solution of Dirichlet's problem, if it exists, ought to be the largest element of $\mathcal{F}$. To make sure that $\mathcal{F}$ is not empty we now assume that f is *bounded*, $|f(\zeta)| \leq M$ for all $\zeta \in \partial\Omega$. It follows that any constant $\leq -M$ is in $\mathcal{F}$, so $\mathcal{F}$ is definitely not empty. A less important, but convenient, assumption we will make is that also Ω is bounded. We now set

$$u(z) = \sup_{v \in \mathcal{F}} v(z)\ , \quad z \in \Omega\ ,$$

expecting this to be the solution of Dirichlet's problem, if it exists. In fact, with no further assumptions, u is harmonic in Ω.

Lemma 5.15. *The function u defined above is harmonic in Ω.*

To be able to prove Lemma 5.15 we need the following important lemma.

THEOREM 5.16 (Harnack's principle). *Suppose $u_1, u_2, \ldots$ is an increasing sequence of functions harmonic in a region Ω. Then either $u_n \to +\infty$ locally uniformly in Ω, or else u_n converges locally uniformly to a function u which is harmonic in Ω.*

PROOF. Suppose u is harmonic in a closed disk $|z - z_0| \le \rho$. The Poisson integral formula then states that for z in the open disk

$$u(z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\rho^2 - r^2}{|\rho e^{i\theta} - (z - z_0)|^2} \, u(z_0 + \rho e^{i\theta}) \, d\theta \ ,$$

where $r = |z - z_0|$. Since $\rho - r \le |\rho e^{i\theta} - (z - z_0)| \le \rho + r$ by the triangle inequality the first factor in the integral can be estimated by

$$\frac{\rho - r}{\rho + r} \le \frac{\rho^2 - r^2}{|\rho e^{i\theta} - (z - z_0)|^2} \le \frac{\rho + r}{\rho - r} \ .$$

If now u is *non-negative* in the disk we obtain *Harnack's inequality*

$$\frac{\rho - r}{\rho + r} \, u(z_0) \le u(z) \le \frac{\rho + r}{\rho - r} \, u(z_0) \ ,$$

by the Poisson integral formula and the mean value property. Now suppose $r \le \rho/2$. Then Harnack's inequality shows that

$$\frac{1}{3} \, u(z_0) \le u(z) \le 3u(z_0) \ . \tag{5.3}$$

Now consider the sequence $u_1, u_2, \ldots$. Since the sequence is increasing it has a *pointwise* limit everywhere in Ω, which is either finite or $+\infty$. If $n > m$ the function $u_n - u_m$ is positive and harmonic in Ω so we can apply Harnack's inequality to it. It follows from (5.3) that if $u_n(z_0) \to +\infty$, then $u_n \to +\infty$ uniformly in a neighborhood of z_0. It also follows that the set where u_n tends to $+\infty$ is an open subset of Ω. Similarly, if $u_n(z_0)$ has a finite limit, then the limit is finite in a neighborhood of z_0 so the set where the limit is finite is also open. Since Ω is connected it follows that either u_n tends locally uniformly to $+\infty$ in Ω, or else the limit function u is finite everywhere. Applying (5.3) to $u_n - u_m$ and letting $n \to \infty$ we get

$$0 \le u(z) - u_m(z) \le 3(u(z_0) - u_m(z_0))$$

so that the convergence is locally uniform. Finally, to see that u is harmonic we may apply the Poisson integral formula to u_n over any circle contained in Ω and take the limit under the integral sign, by uniform convergence. It follows that locally u is given by its Poisson integral so that u is harmonic. The proof is complete ☐

LEMMA 5.17. *Suppose v is subharmonic in Ω and for some constant K we have $\overline{\lim}_{\Omega \ni z \to \zeta} v(z) \le K$ for every $\zeta \in \partial\Omega$. Then $v \le K$ in Ω.*

PROOF. If $\varepsilon > 0$ there is a neighborhood of $\partial\Omega$ where $v < K + \varepsilon$. It follows that the set $E = \{z \in \Omega \mid v(z) \geq K + \varepsilon\}$ is closed and since it is bounded (as a subset of the bounded set Ω), it is in fact compact. If $E \neq \varnothing$ it follows that v has a maximum in E, which will also be an interior maximum in Ω. It would follow that v is constant $\geq K+\varepsilon$ which contradicts the assumption about the boundary behavior. Hence E is empty, and since $\varepsilon > 0$ is arbitrary the desired conclusion follows. □

PROOF OF LEMMA 5.15. First note that by Lemma 5.17 $v \leq M$ for all $v \in \mathcal{F}$. It follows that u is finite everywhere in Ω. Now let $z_0 \in \Omega$. We may then choose a sequence $v_1, v_2, \dots$ from $\mathcal{F}$ such that $v_n(z_0) \to u(z_0)$. We also have $v_n(z_0) \leq u(z_0)$, $n = 1, 2, \dots$. Now let $V_n = \max(v_1, \dots, v_n)$. By property (3) of subharmonic functions $V_n \in \mathcal{F}$ and $v_n(z_0) \leq V_n(z_0) \leq u(z_0)$ so we have $V_n(z_0) \to u(z_0)$. In addition the sequence $V_1, V_2, \dots$ is increasing. Now choose a disk D containing z_0 and whose closure is in Ω and let $\tilde{V}_n$ equal V_n outside D and the Poisson integral of V_n over ∂D in D. By property (4) of subharmonic functions also $\tilde{V}_n \in \mathcal{F}$ so $\tilde{V}_n \leq u$ and it is $\geq V_n$ by the maximum principle. Hence $\tilde{V}_n(z_0) \to u(z_0)$ and $\tilde{V}_1, \tilde{V}_2, \dots$ is increasing. Since $\tilde{V}_n$ is harmonic in D we may apply Harnack's principle, and since $\tilde{V}_n(z_0) \to u(z_0) < \infty$ it follows that $\tilde{V}_n \to U$ locally uniformly in D, where U is a harmonic function for which $U(z_0) = u(z_0)$.

Now let z_1 be an arbitrary point of D. As before we can then find a sequence $w_1, w_2, \dots$ in $\mathcal{F}$ such that $w_n(z_1) \to u(z_1)$. If we set $\tilde{w}_n = \max(w_n, v_n)$ we still have elements of $\mathcal{F}$, the limit at z_1 is unchanged and we also have $\tilde{w}_n \geq v_n$. We continue similar to what we did above, setting $W_n = \max(\tilde{w}_1, \dots, \tilde{w}_n)$ and then $\tilde{W}_n$ equal to W_n outside D and equal to the corresponding Poisson integral inside D. The sequence $\tilde{W}_1, \tilde{W}_2, \dots$ is then in $\mathcal{F}$, harmonic in D, increasing and $\tilde{W}_n(z_1) \to u(z_1)$. We also have $\tilde{W}_n \geq \tilde{V}_n$ so that $\tilde{W}_n(z_0) \to u(z_0)$. As before it follows that in D we have $\tilde{W}_n \to U_1$ locally uniformly, where U_1 is harmonic, $U_1(z_1) = u(z_1)$, $U \leq U_1$ and also $U_1(z_0) = u(z_0) = U(z_0)$. The harmonic function $U - U_1$ is therefore non-positive but 0 in z_0. By the maximum principle it is constant and therefore identically 0. It follows that $U(z_1) = u(z_1)$. Since z_1 is an arbitrary point of D it follows that $U = u$ in D so that u is harmonic in a neighborhood of every point $z_0 \in \Omega$. The proof is complete. □

To deal with the question whether u assumes the desired boundary values we need to introduce the concept of a *barrier* function.

DEFINITION 5.18. A *barrier for* Ω *at a point* $\zeta \in \partial\Omega$ is a function w harmonic in Ω and continuous in $\overline{\Omega}$, and such that $w(\zeta) = 0$ but w is strictly positive in all other points of $\overline{\Omega}$.

The following lemma reduces the question of whether u takes the desired boundary values to the question of finding barriers.

LEMMA 5.19. *Suppose f is continuous at a point $\zeta_0 \in \partial\Omega$ and there is a barrier for Ω at ζ_0. Then $u(z) \to f(\zeta_0)$ as $\Omega \ni z \to \zeta_0$.*

PROOF. We will show that we have $\overline{\lim}_{\Omega \ni z \to \zeta_0} u(z) \leq f(\zeta_0) + \varepsilon$ and that $\underline{\lim}_{\Omega \ni z \to \zeta_0} u(z) \geq f(\zeta_0) - \varepsilon$ for every $\varepsilon > 0$ from which the theorem follows.

Let $\varepsilon > 0$ and choose a neighborhood $\mathcal{O}$ of ζ_0 such that $|f(\zeta) - f(\zeta_0)| < \varepsilon$ for $\zeta \in \mathcal{O} \cap \partial\Omega$. Furthermore, let w_0 be the minimum of w over the (compact) set $\overline{\Omega} \setminus \mathcal{O}$. By the properties of w we have $w_0 > 0$. Now put $V(z) = f(\zeta_0) + \varepsilon + \frac{w(z)}{w_0}(M - f(\zeta_0))$. Then V is harmonic in Ω and continuous in the closure. For $\zeta \in \mathcal{O} \cap \partial\Omega$ we have $V(\zeta) \geq f(\zeta_0) + \varepsilon > f(\zeta)$. For $\zeta \in \partial\Omega \setminus \mathcal{O}$ we have $w(\zeta) \geq w_0$ so we get $V(\zeta) \geq M + \varepsilon > f(\zeta)$. If $v \in \mathcal{F}$ and $\zeta \in \partial\Omega$ we therefore have $\overline{\lim}_{\Omega \ni z \to \zeta}(v(z) - V(z)) < 0$ so by Lemma 5.17 $v \leq V$ in Ω. It follows that also $u \leq V$ in Ω so that $\overline{\lim}_{\Omega \ni z \to \zeta_0} u(z) \leq V(\zeta_0) = f(\zeta_0) + \varepsilon$.

To prove the other inequality, set $W(z) = f(\zeta_0) - \varepsilon - \frac{w(z)}{w_0}(M + f(\zeta_0))$. Again W is harmonic in Ω and continuous in the closure. For $\zeta \in \mathcal{O} \cap \partial\Omega$ we have $W(\zeta) \leq f(\zeta_0) - \varepsilon < f(\zeta)$ and for $\zeta \in \partial\Omega \setminus \mathcal{O}$ we have $w(\zeta) \geq w_0$ so that we get $W(\zeta) \leq -M - \varepsilon < f(\zeta)$. It follows that $W \in \mathcal{F}$ so that $W \leq u$. Hence $\underline{\lim}_{\Omega \ni z \to \zeta_0} u(z) \geq W(\zeta_0) = f(\zeta_0) - \varepsilon$. The proof is complete. □

It is sometimes easy to find a barrier. For example, suppose a point $\zeta \in \partial\Omega$ has a *supporting line*, *i.e.*, a line which intersects the closure of Ω only in ζ, and let α be the direction of the line, chosen so that Ω is to the left of it. Then $\operatorname{Im}(e^{-i\alpha}(z - \zeta)$ is a barrier for Ω at ζ. Show this! If Ω is *strictly convex*, then every boundary point has a supporting line so there is a barrier for Ω at every boundary point. To state a more general result, we make the following definition.

DEFINITION 5.20. A region Ω is said to have the *segment property* at a boundary point ζ if there exists a line segment exterior to Ω except that one endpoint is ζ.

A continuous curve $\gamma \subset \partial\Omega$ without self-intersections is called a *free boundary arc* of Ω if every point on γ is the center of a disk which is split in *exactly two* components by $\partial\Omega$. It is called *one-sided* if one of the components is always in Ω and the other not.

It is clear that if γ is a free onesided boundary arc of Ω and γ has a normal at a point $\zeta \in \gamma$, then Ω has the segment property at ζ; one only has to choose a sufficiently short piece of the exterior normal.

LEMMA 5.21. *A region Ω has a barrier at any boundary point where it has the segment property.*

PROOF. Suppose Ω has the segment property at $\zeta \in \partial\Omega$ and that the other endpoint of the corresponding segment is p. We can then choose a complex number a such that the segment is mapped onto the negative real axis by $z \mapsto a\frac{z-\zeta}{z-p}$, the image of ζ being 0. Using the principal branch of the root it is then obvious that $\operatorname{Re}\sqrt{a\frac{z-\zeta}{z-p}}$ is a barrier for Ω at ζ. $\square$

We collect our results about Dirichlet's problem in the following theorem.

THEOREM 5.22. *Suppose Ω is a bounded region having the segment property at each of its boundary points. Then Dirichlet's problem has a unique solution in Ω for arbitrary boundary values f continuous on $\partial\Omega$.*

PROOF. We have proved everything claimed in this section except the uniqueness; but this is Theorem 5.9. $\square$

CHAPTER 6

Entire functions

6.1. Sequences of analytic functions

In this section we shall consider sequences of analytic functions which are uniformly convergent. We will use the notation $H(\Omega)$ for the functions holomorphic (analytic) in the region $\Omega \subset \mathbb{C}$. By a region we will always mean an open, connected set. Recall that we say that a sequence $f_1, f_2, \dots$ of real or complex-valued functions defined on a set E is *uniformly convergent on* E to another function f defined on E provided that for each $\varepsilon > 0$ we can find a number N such that if $n \geq N$ then $|f_n(z) - f(z)| < \varepsilon$ for every $z \in E$. If one introduces the *maximum-norm* $\| \cdot \|_E$ by setting $\|f\|_E = \sup_{z \in E} |f(z)|$ the uniform convergence of f_n to f on E is equivalent to $\|f_n - f\|_E \to 0$ as $n \to \infty$. When dealing with functions defined in an open set $\Omega \subset \mathbb{C}$ (or $\Omega \subset \mathbb{R}^n$) one often talks about *locally uniform* convergence. A sequence of functions $f_1, f_2, \dots$ defined in Ω is said to converge locally uniformly to f in Ω if every $x \in \Omega$ has a neighborhood in which the sequence converges uniformly to f. Equivalently, this means that $f_n \to f$ uniformly *on every compact subset* of Ω. This is an immediate consequence of the Heine-Borel lemma.

Exercise 6.1. Show this equivalence!

Recall that in Flervariabelanalys it is proved that the uniform limit of continuous functions is continuous. It immediately follows that the same is true of *locally* uniform limits of continuous functions (explain why this is obvious!). When dealing with analytic functions one can say a lot more.

The main result of the section is the following.

Theorem 6.2. *Suppose $f_n \in H(\Omega)$, $n = 1, 2, 3, \dots$ and that $f_n \to f$ locally uniformly in Ω. Then $f \in H(\Omega)$. Furthermore, $f_n^{(j)} \to f^{(j)}$ locally uniformly in Ω for $j = 1, 2, 3, \dots$.*

Proof. Let γ be a positively oriented circle such that the corresponding closed disk is contained in Ω. For z in the open disk we then have (Corollary 3.13)

$$f_n^{(j)}(z) = \frac{j!}{2\pi i} \int_\gamma \frac{f_n(w)\, dw}{(w - z)^{j+1}} \,. \qquad (6.1)$$

Since $f_n \to f$ uniformly on the closed disk the integral on the right converges to $\int_\gamma \frac{f(w)\,dw}{(w-z)^{j+1}}$ as $n \to \infty$. For $j = 0$ the left hand side converges to $f(z)$, so that f satisfies the Cauchy integral formula; thus by Lemma 3.10 f is analytic in a neighborhood of every point of Ω so that $f \in H(\Omega)$. By uniform convergence the right hand side of (6.1) converges (pointwise) to $f^{(j)}(z)$. Suppose γ has radius r. I claim that this convergence is uniform for z in the disk of radius $r/2$ concentric to γ, which would prove locally uniform convergence and thus finish the proof.

To verify the claim, note that for z in the sub-disk and $w \in \gamma$ we have $|z - w| \geq r/2$ so that

$$\Big| \int_\gamma \frac{f_n(w)\,dw}{(w-z)^{j+1}} - \int_\gamma \frac{f(w)\,dw}{(w-z)^{j+1}} \Big|$$
$$\leq (r/2)^{-j-1} \int_\gamma |f_n(w) - f(w)|\,|dw| \leq 2\pi r (r/2)^{-j-1} \|f_n - f\|_\gamma \,.$$

Since $f_n \to f$ uniformly on γ this shows the uniform convergence. □

Theorem 6.2 was first proved by Weierstrass in a slightly different formulation which we state as a corollary.

Corollary 6.3. *Suppose $f_1, f_2, \dots$ are all in $H(\Omega)$ and the series $\sum_{k=1}^\infty f_k$ converges locally uniformly on Ω. Then the series converges to a function in $H(\Omega)$, it may be differentiated termwise any number of times, and the differentiated series all converge locally uniformly in Ω.*

This is obviously equivalent to Theorem 6.2. We prove one more result (by A. Hurwitz) on uniform convergence.

Theorem 6.4. *Suppose $f_n \in H(\Omega)$ for $n = 1, 2, \dots$ and that $f_n \to f$ locally uniformly on Ω as $n \to \infty$. Suppose furthermore that none of the functions f_n assume the value w in Ω. Then neither does f, unless f is constant ($= w$).*

Proof. Replacing f_n by $f_n - w$ and f by $f - w$ we may as well assume that $w = 0$. Assume that f is not identically zero. We must then prove that f has no zeros in Ω. We know, since $f \in H(\Omega)$, that the zeros of f are *isolated*, so any point of Ω is the center of a closed disk contained in Ω and such that f has no zeros on the boundary circle. If γ is the positively oriented boundary of such a disk, then the number of zeros of f in the open disk is given by $\frac{1}{2\pi i}\int_\gamma f'/f$, so we need to show that any such integral is 0.

Since $z \mapsto |f(z)|$ is continuous and γ compact, $|f|$ assumes a minimum m on γ which is > 0 since f has no zeros on γ. Since $f_n \to f$

uniformly on γ we have $|f_n| \geq m/2$ on γ for all sufficiently large n. So, for $z \in \gamma$ and sufficiently large n we have

$$|\frac{1}{f_n(z)} - \frac{1}{f(z)}| = \frac{|f(z) - f_n(z)|}{|f_n(z)f(z)|} \leq \frac{2}{m^2}\|f_n - f\|_\gamma \ .$$

Thus $1/f_n \to 1/f$ uniformly on γ, so by Theorem 1.2 it follows that $f_n'/f_n \to f'/f$ uniformly on γ. Thus $\int_\gamma f_n'/f_n \to \int_\gamma f'/f$. But all the integrals on the left equal 0, because f_n are all zero-free in Ω. It follows that the limit is also 0, and the proof is complete. □

As an almost immediate consequence we have the following interesting corollary about so called *univalent* functions. A univalent function is an injective (one-to-one) analytic function.

COROLLARY 6.5. *Suppose $f_n \in H(\Omega)$, $n = 1, 2, \dots$, and $f_n \to f$ locally uniformly in Ω. If all f_n are univalent, then so is f, unless it is constant.*

PROOF. Assume f is not constant. Then, if $f(z_0) = w$, we must show that $f(z) - w \neq 0$ for $z \in \Omega \setminus \{z_0\}$. Setting $g_n(z) = f_n(z) - f_n(z_0)$ we have $g_n \to f - w$ locally uniformly in Ω. Since by assumption g_n does not vanish in $\Omega \setminus \{z_0\}$, neither does $f - w$ by Theorem 6.4. □

EXERCISE 6.6. Show that for any $\varepsilon > 0$ there exists N such that all Taylor polynomials of $\sin x$ (partial sums of $\sum \frac{(-1)^k}{(2k+1)!}\, x^{2k+1}$) of degree at least N has exactly one zero in $(\pi - \varepsilon, \pi + \varepsilon)$.

EXERCISE 6.7. A famous theorem by Weierstrass states that any function continuous on a real interval $[a, b]$ is the uniform limit of a sequence of polynomials. Why does this not contradict Theorem 6.2?

6.2. Infinite products

Any analytic function may be expanded in a power series centered at any point of the domain of analyticity; the radius of convergence is such that on the boundary of the disk of convergence there is at least one singularity of the function. If the function is analytic everywhere in $\mathbb{C}$, the radius of convergence is therefore infinite. Such a function is called *entire* (in Britain often also *integral*). A power series used to be viewed as a 'polynomial of infinite order', especially if the radius of convergence is infinite. The reason is of course that many properties of polynomials have their counterpart for entire functions.

One of the more fundamental properties of a polynomial is that, according to the fundamental theorem of algebra and the factor theorem, it may be factored into a product of first degree polynomials, each of which vanishes at one of the zeros of the polynomial. If p is a polynomial of degree n one usually writes $p(z) = A\prod_{k=1}^n (z - z_k)$, where $z_1, z_2, \dots$ are the zeros of p, repeated according to multiplicity, and A

is the highest order coefficient of p. Clearly this does not generalize to entire functions; a polynomial of infinite degree can hardly have a highest order coefficient. But one may also write

$$p(z) = Bz^j \prod_{z_k \neq 0} (1 - \frac{z}{z_k}) \ , \tag{6.2}$$

where B is the coefficient of the non-zero term in p with lowest degree, and j is the multiplicity of $z = 0$ as a zero of p (so that $j = 0$ if $p(0) \neq 0$). As we shall see in the next section this expansion has a generalization to arbitrary entire functions. In this section we shall prepare the ground for this by considering infinite products.

What meaning should one assign to $\prod_{k=1}^{\infty} A_k$? The obvious answer is to consider the partial products $P_n = \prod_{k=1}^{n} A_k$ and then assign to the infinite product the value $\lim_{n\to\infty} P_n$ if the limit exists. This is almost right, but note that the limit is 0 if just one factor is zero, completely independent of the values of all the other factors. This does not seem reasonable, so one makes the following modified definition.

Definition 6.8. The infinite product $\prod_{k=1}^{\infty} A_k$ is said to converge to P if

(1) The sequence of partial products converge to P.
(2) There are only a finite number of zero factors in the product, and the sequence of partial products obtained by excluding these factors converge to a **non-zero** number.

If $P_n \to P \neq 0$ as $n \to \infty$ it follows that $A_n = P_n/P_{n-1} \to P/P = 1$ as $n \to \infty$, so the factors in a convergent product always tend to 1. It is therefore convenient to write infinite products on the form

$$\prod_{k=1}^{\infty} (1 + a_k) \ , \tag{6.3}$$

so that the necessary condition for convergence just derived takes the following form.

Proposition 6.9. *A necessary (but not sufficient) condition for convergence of the infinite product* (6.3) *is that $a_k \to 0$ as $k \to \infty$.*

Since a sequence has a non-zero limit precisely if the sequence of logarithms has a finite limit it is natural to compare the infinite product (6.3) with the series with terms $\log(1 + a_k)$. Recall that the *principal branch* of the logarithm is $\operatorname{Log} z = \ln|z| + i \arg z$, where $-\pi < \arg z \leq \pi$.

Theorem 6.10. *If $a_k \neq -1$, $k = 1, 2, \dots$, then the infinite product* (6.3) *converges if and only if the series*

$$\sum_{k=1}^{\infty} \operatorname{Log}(1 + a_k) \tag{6.4}$$

converges. Here Log *denotes the principal branch of the logarithm.*

PROOF. Since the terms of a convergent series must tend to 0, and by Proposition 2.2, we must have $a_k \to 0$ if either the product (6.3) or the series (6.4) converges. If S_n denotes the partial sum of the series we have $P_n = e^{S_n}$ so that the convergence of the product follows from that of the series.

Conversely, assume that $P_n \to P \neq 0$ and choose a branch of the logarithm which is continuous in a neighborhood of P. Then $\log P_n \to \log P$. We have $S_n = \log P_n + 2k_n\pi i$, where k_n is an integer. Thus $S_n - S_{n-1} = \log P_n - \log P_{n-1} + 2(k_n - k_{n-1})\pi i$. But since $a_k \to 0$ the imaginary part of $S_n - S_{n-1}$ tends to 0. Since also $\log P_n - \log P_{n-1} \to 0$ it follows that $k_n - k_{n-1} \to 0$ so that, since k_n is an integer, all k_n equal a fixed integer k from a certain value of n on. This means that $S_n \to \log P + 2k\pi i$ so the proof is complete. □

We will, by definition, say that the product (6.3) is *absolutely convergent* if the infinite product $\prod_{k=1}^{\infty}(1+|a_k|)$ converges. By Theorem 2.3 this is equivalent to the convergence of the positive series

$$\sum_{k=1}^{\infty} \log(1 + |a_k|) \ . \tag{6.5}$$

Noting that $\frac{\mathrm{Log}(1+z)}{z} \to 1$ as $z \to 0$ (using the principal branch of the logarithm) it follows by a standard comparison theorem that the series (6.4) (omitting terms for which $a_k = -1$) is absolutely convergent if and only if $\sum_{k=1}^{\infty} a_k$ is absolutely convergent (note that if either of the two series are convergent, then we must have $a_k \to 0$ as $k \to \infty$). In particular it follows that (6.5) converges if and only $\sum_{k=1}^{\infty} |a_k|$ converges. So we have proved the following proposition.

PROPOSITION 6.11. *The product* (6.3) *converges absolutely if and only if* $\sum_{k=1}^{\infty} a_k$ *converges absolutely. This is also equivalent to the series* (6.4) *converging absolutely, after omitting (the finite number of) terms for which* $a_k = -1$.

We now turn to the case when the factors of (6.3) are functions of $z \in \mathbb{C}$. By inspection of the proofs it is clear that all the results obtained so far remain true if we replace 'convergence' by 'locally uniform convergence'. So by Theorem 1.2, if $a_k \in H(\Omega)$ for every k, then (6.3) converges locally uniformly to a function in $H(\Omega)$ if $\sum |a_k|$ converges locally uniformly in Ω. In particular, by Weierstrass' majorization theorem (Weierstrass' M-test in most English language books) it follows that this is the case if $\sum \|a_k\|_K$ converges for every compact $K \subset \Omega$.

We can now return to the problem of generalizing the polynomial factorization (6.2) to an arbitrary entire function. Suppose that we have an entire function for which 0 is a zero of multiplicity j which also has other zeros $a_1, a_2, \ldots,$ repeated according to multiplicity. By analogy with (6.2) our candidate for this function would then be a constant multiple of $z^j \prod_{k=1}^{\infty}(1-\frac{z}{a_k})$. This may not be so, however. First of all, there are entire functions with no zeros at all. One example is e^z; more generally, $e^{g(z)}$ is such a function for any entire function g. We would certainly have to allow such a factor in front of the product to obtain a generally valid factorization. Furthermore, for the product to converge absolutely for some $z \neq 0$ we must require that $\sum \frac{1}{|a_k|}$ converges; this may not always hold, although it is true that we always have $a_k \to \infty$ as $k \to \infty$ (Exercise 6.12). For example, the function $\sin(\pi z)$ has zeros $0, \pm 1, \pm 2, \dots$ and $\sum 1/k$ is divergent. A little more effort is therefore required to obtain a general factorization formula for entire functions. We will carry this out in the next section.

EXERCISE 6.12. Prove that if $a_1, a_2, \dots$ are the zeros of an entire function, repeated according to multiplicity, then $a_k \to \infty$ as $k \to \infty$.

6.3. Canonical products

Consider first the case of an entire function f with only finitely many non-vanishing zeros $a_1, \dots, a_n$, as always counted with multiplicities. If the multiplicity of 0 as a zero is $j \geq 0$ it is clear that

$$h(z) = f(z)z^{-j} / \prod_{k=1}^{n}(1-\frac{z}{a_k})$$

is an entire function without zeros. Thus also $h'(z)/h(z)$ is entire, so it has an entire primitive g. Differentiating $h(z)e^{-g(z)}$ we obtain $h'(z)e^{-g(z)} - h(z)\frac{h'(z)}{h(z)}e^{-g(z)} = 0$ so that h is a constant multiple of e^g. By adding an appropriate constant to g, if necessary, we may assume that $h = e^g$. Thus we obtain $f(z) = z^j e^{g(z)} \prod_{k=1}^{n}(1-\frac{z}{a_k})$ for some entire function g. If f has infinitely many zeros the same reasoning gives the representation

$$f(z) = z^j e^{g(z)} \prod_{k=1}^{\infty}(1-\frac{z}{a_k}) \ , \tag{6.6}$$

with an entire function g, *provided that the infinite product converges locally uniformly.* This is ensured, by the previous chapter, if $\sum |a_k|^{-1}$ converges.

EXAMPLE 6.13. The function $f(z) = \frac{\sin\sqrt{z}}{\sqrt{z}}$, where $\sqrt{z}$ is any branch of the root, has zeros $(k\pi)^2$, $k = 1, 2, \dots$ and no others. It is an entire function since it has a power series expansion $\sum \frac{(-1)^k}{(2k+1)!} z^k$. This follows immediately from the expansion of $\sin z$. From this Euler proved his famous formula

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \tag{6.7}$$

using the following argument. According to (6.2) the coefficient of z in any polynomial with constant coefficient 1 is the negative of the sum of the reciprocals of the zeros of the polynomial. The power series for f has constant coefficient 1 and the coefficient of z is $-1/6$. The sum of the reciprocals of the zeros of f is $\sum_{k=1}^{\infty}(k\pi)^{-2}$. Hence (6.7) follows. Well...

The argument can be quite easily justified if we assume that we know that

$$\frac{\sin\sqrt{z}}{\sqrt{z}} = \prod_{k=1}^{\infty}\left(1 - \frac{z}{k^2\pi^2}\right) . \tag{6.8}$$

It is at least clear that the product converges absolutely and locally uniformly since $\sum 1/k^2$ converges. The partial product P_n is a polynomial and its z-coefficient is $-\sum_{k=1}^{n}(k\pi)^{-2}$. This is the value of $P_n'(0)$, and according to Theorem 6.2 this converges to the derivative at 0 of the infinite product. Assuming (6.8) the power series for f shows that this is $-1/6$ and so Euler would be vindicated.

Instead of proving (6.8) we will show the equivalent statement

$$\frac{\sin\pi z}{\pi z} = \prod_{k=1}^{\infty}\left(1 - \frac{z^2}{k^2}\right) , \tag{6.9}$$

obtained by replacing z in (6.8) by $\pi^2 z^2$. Since the infinite product converges absolutely locally uniformly we at least know from (6.6) that

$$\frac{\sin\pi z}{\pi z} = e^{g(z)}\prod_{k=1}^{\infty}\left(1 - \frac{z^2}{k^2}\right) ,$$

where g is entire. Taking the logarithmic derivative of both sides we obtain

$$\pi\cot\pi z - 1/z = g'(z) + \sum_{k=1}^{\infty}\left\{\frac{1}{z-k} + \frac{1}{z+k}\right\} , \tag{6.10}$$

and differentiating once more we obtain

$$\frac{\pi^2}{\sin^2\pi z} = -g''(z) + \sum_{k=-\infty}^{\infty}\frac{1}{(z+k)^2} .$$

However, both the infinite sum and the left hand side are of period 1 here, so g'' also has period 1. Writing $z = x + iy$ it is easy to see that both the infinite sum and the left hand side tend to 0 locally uniformly in x as $y \to \pm\infty$. Thus the same is true of g'' which is therefore bounded and entire, hence by Liouville's theorem constant; since $g''(iy) \to 0$ as $y \to +\infty$ the constant is 0. Therefore g' is constant. But from (6.10) follows that g' is odd, so that also g' vanishes. Thus g is constant, so that (6.9) holds apart from a constant factor. Since both sides tend to 1 as $z \to 0$ this factor is 1, and we have finally established (6.9), and thereby Euler's formula.

EXERCISE 6.14. Justify all unproved claims at the end of Example 6.13.

What is one to do to obtain a factorization for an entire function where the sum of the reciprocals of the zeros is not absolutely convergent? The idea is to replace the factor $1 - \frac{z}{a_k}$ in the product by $(1 - \frac{z}{a_k})e^{p_k(z)}$, where p_k is an entire function which promotes convergence without introducing new zeros. As we shall see, one can always choose p_k to be a polynomial. Convergence is obtained by choosing p_k so that $(1 - \frac{z}{a_k})e^{p_k(z)}$ is sufficiently close to 1, so the ultimate choice would be $-\log(1 - \frac{z}{a_k})$. Unfortunately this is not an entire function. It is therefore natural to attempt to choose p_k as a Taylor polynomial of this function of sufficiently high degree. Now for the principal branch of the logarithm $-\operatorname{Log}(1-z) = \sum_{k=1}^{\infty} \frac{z^k}{k}$ and the series converges for $|z| < 1$. In fact, if we set $L_n(z) = \sum_{k=1}^{n} \frac{z^k}{k}$ we have $-\operatorname{Log}(1-z) = L_n(z) + R_n(z)$, where an easy estimate gives

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

for $|z| < 1$. According to Proposition 6.11 the product

$$\prod_{k=1}^{\infty}(1 - \frac{z}{a_k})e^{p_k(z)} \tag{6.11}$$

converges absolutely and locally uniformly in z precisely if the series

$$\sum_{k=1}^{\infty}\{\operatorname{Log}(1 - \frac{z}{a_k}) + p_k(z)\} \tag{6.12}$$

does (check this carefully!). We assume now $|z| < R$. There are only finitely many factors in (6.11) for which $|a_k| < 2R$ (Exercise 2.5) so excluding these factors from the product will not affect convergence. We may thus also assume that $|a_k| \geq 2R$. If we choose $p_k(z) = L_{n_k}(z/a_k)$ the absolute value of the term in (6.12) is $|R_{n_k}(z/a_k)|$ and may therefore be estimated by 2^{1-n_k}, using the facts that $R/|a_k| \leq 1/2$ and $1 - R/|a_k| \geq 1/2$. We conclude that (6.12) converges absolutely and uniformly in $|z| < R$ if we can choose n_k for every k such that the series $\sum_{k=1}^{\infty} 2^{-n_k}$ converges. A obvious choice that works is $n_k = k$. Since R is arbitrary we conclude that the choice $p_k(z) = L_k(z/a_k)$ makes (6.11) absolutely and locally uniformly convergent. We have proved the following theorem by Weierstrass.

THEOREM 6.15. *There exists an entire function with arbitrarily prescribed non-vanishing zeros $a_1, a_2, \dots$ (repeated according to multiplicity), provided they are either finitely many or else $a_k \to \infty$ as $k \to \infty$.*

Every entire function with these and no other zeros may be written

$$f(z) = z^j e^{g(z)} \prod_{a_k \neq 0} (1 - \frac{z}{a_k}) e^{L_{n_k}(\frac{z}{a_k})} \ , \tag{6.13}$$

where g is an entire function, $L_n(z) = \sum_{k=1}^{n} \frac{z^k}{k}$, and n_k, $k = 1, 2, \dots$ are certain (sufficiently large) positive integers. A possible choice is $n_k = k$.

The theorem has a very important corollary concerning meromorphic functions. Recall that a function f is called meromorphic in Ω if it is analytic in Ω except for isolated singularities which are poles.

COROLLARY 6.16. *Every function which is meromorphic in the whole plane is the quotient of two entire functions.*

PROOF. If f is meromorphic in the whole plane we may, according to Theorem 3.3, find an entire function g so that all the poles of f are zeros of g, and with the same multiplicities. Thus $h = fg$ is an entire function and $f = h/g$. □

The expansion (6.13) becomes particularly interesting if one may choose $n_k = h$ independent of k. This is the case if $\sum |R_h(z/a_k)|$ converges absolutely uniformly for $|z| \leq R$ for any R. Since $a_k \to \infty$ this happens if $\sum (R/|a_k|)^{h+1} = R^{h+1} \sum 1/|a_k|^{h+1}$ converges. In other words, if the zeros do not tend too slowly to infinity. Suppose now that h is the smallest integer for which $\sum 1/|a_k|^{h+1}$ converges (so that $h \geq 0$). Then

$$\prod_{k=1}^{\infty} (1 - \frac{z}{a_k}) e^{L_h(z/a_k)}$$

is called the *canonical product* associated with the sequence $a_1, a_2, \dots$, and the integer h is called the *genus* of the canonical product. If possible we use the canonical product in the expansion (6.13). In that case the expansion becomes uniquely determined by f. If it then happens that g is a polynomial, one says that the function f has *finite genus*, and the genus of f is the degree of g or the genus of the canonical product, whichever is the largest. This means for example that the function $\sin\sqrt{z}/\sqrt{z}$ considered in Example 6.13 and with the product expansion (6.8) is of genus 0.

EXAMPLE 6.17. The function $\sin \pi z$ has all the integers as its zeros, and since $\sum 1/n$ diverges but $\sum 1/n^2$ converges we obtain an expansion of the form

$$\sin \pi z = z e^{g(z)} \prod_{k \neq 0} (1 - \frac{z}{k}) e^{z/k} \ .$$

If we group the factors for $\pm k$ together and compare the result to (6.9) it follows that g is the constant $\log \pi$. Consequently, $\sin \pi z$ is of genus

1 and has the canonical expansion

$$\sin \pi z = z\pi \prod_{k\neq 0}(1-\frac{z}{k})e^{z/k} \tag{6.14}$$

EXERCISE 6.18. If f has genus h, what is the possible range for the genus of $z \mapsto f(z^2)$?

EXERCISE 6.19. Let $a_1, a_2, \dots$ be a sequence satisfying $0 < |a_k| < 1$ for all k for which $\sum_{k=1}^{\infty}(1-|a_k|)$ converges. Show that the product

$$\prod_{k=1}^{\infty} \frac{\overline{a_k}}{|a_k|} \frac{a_k - z}{1-\overline{a_k}z}$$

(a so called *Blaschke product*) converges to a function holomorphic in the unit disk with the given sequence as zeros.

6.4. Partial fractions

As we have seen a meromorphic function is the quotient of two entire functions, and thus the analogue of a rational function. A fundamental fact about rational functions is that they allow a *partial fractions expansion.* In fact, if $r(z) = p(z)/q(z)$ where p and q are polynomials without common factors, then one may write

$$r(z) = g(z) + \sum_{k=1}^{n} P_k(\frac{1}{z-a_k})$$

where g and all P_k are polynomials, $a_1, \dots, a_n$ the different zeros of q, and $\deg P_k = n_k$ where n_k is the multiplicity of a_k as a zero of q. Note that $P_k(\frac{1}{z-a_k})$ is the singular part of r at a_k as a meromorphic function. For a function meromorphic in the whole plane one would therefore expect a similar expansion, where now g is entire and n may be infinite. This leads to *Mittag-Leffler's theorem*, although the sum has to be slightly modified to ensure convergence.

THEOREM 6.20 (Mittag-Leffler). *Let $a_1, a_2, \dots$ be a sequence converging to ∞ and let P_k be polynomials without constant terms. Then there are functions meromorphic in the whole plane with poles precisely at a_k and corresponding singular part $P_k(\frac{1}{z-a_k})$. The most general such meromorphic function may be written*

$$f(z) = g(z) + \sum_{k=1}^{\infty}(P_k(\frac{1}{z-a_k}) - q_k(z)) \ , \tag{6.15}$$

where g is entire and q_k suitably chosen polynomials.

PROOF. If $a_k = 0$ we choose $q_k = 0$. If $a_k \neq 0$ the function $h(z) = P_k(\frac{1}{z-a_k})$ is analytic at 0 and we will choose for q_k the corresponding

Taylor polynomial of degree n_k. If γ is the circle $|\zeta| = |a_k|/2$ and z a point inside the circle we then have

$$h(z) - \sum_{k=0}^{n_k} \frac{h^{(k)}(0)}{k!} z^k = \frac{1}{2\pi i} \int_\gamma h(\zeta)\{\frac{1}{\zeta - z} - \sum_{k=0}^{n_k} \frac{z^k}{\zeta^{k+1}}\} \, d\zeta \; .$$

Summing the geometric series we obtain

$$h(z) - q_k(z) = \frac{z^{n_k+1}}{2\pi i} \int_\gamma \frac{h(\zeta)\, d\zeta}{(\zeta - z)\zeta^{n_k+1}} \; .$$

Supposing $|P_k(\frac{1}{z-a_k})| \leq M_k$ for $|z| = |a_k|/2$ we obtain

$$|h(z) - q_k(z)| \leq 3M_k(\frac{2|z|}{|a_k|})^{n_k+1}$$

for $|z| \leq |a_k|/3$. Consider a disk $|z| \leq R$. There are only finitely many of the a_k with $|a_k| < 3R$, and it is clear that after removing the corresponding terms (6.15) converges uniformly in $|z| \leq R$ if the series $\sum M_k(\frac{2R}{|a_k|})^{n_k}$ converges. We may consider this a power series in R, and it will then have infinite radius of convergence if the terms tend to 0 for every $R > 0$. Choosing $n_k \geq \log M_k$ we have $M_k(\frac{2R}{|a_k|})^{n_k} \leq (\frac{2eR}{|a_k|})^{n_k} \to 0$ as $k \to \infty$ since $a_k \to \infty$. Thus the sum in (6.15) represents a meromorphic function with the same singular parts as f in all poles, so the theorem follows. □

EXAMPLE 6.21. The function $\pi \cot \pi z$ has simple poles with residue 1 at every integer. Since $\sum \frac{1}{z-k}$ does not converge, we must include a convergence term in the partial fractions expansion. The constant term in the Taylor expansion of $\frac{1}{z-k}$ is $-\frac{1}{k}$, and since $\sum(\frac{1}{z-k} + \frac{1}{k}) = \sum \frac{z}{k(z-k)}$ converges like $\sum \frac{1}{k^2}$, we must have

$$\pi \cot \pi z = g(z) + \frac{1}{z} + \sum_{k \neq 0} (\frac{1}{z-k} + \frac{1}{k}) \tag{6.16}$$

with an entire function g. Comparing this with (6.10) it is clear that $g \equiv 0$. Differentiating the expansion we obtain the partial fractions expansion

$$\frac{\pi^2}{\sin^2 \pi z} = \sum_{k=-\infty}^{\infty} \frac{1}{(z-k)^2} \; .$$

As a final example, consider the function $\frac{\pi}{\sin \pi z}$ which has simple poles at the integers, with residue 1 at even and -1 at odd integers. The partial fractions expansion is thus of the form

$$\frac{\pi}{\sin \pi z} = g(z) + \sum_{k=-\infty}^{\infty} \frac{(-1)^k}{z-k}$$

with an entire function g. The series is not absolutely convergent here, which would be ensured by choosing $q_k = (-1)^{k+1}\frac{1}{k}$ for $k \neq 0$. However, the alternating series $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ converges, so the series is locally uniformly convergent (away from the integers) as it stands. To determine g, note that the sum may be rewritten

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

$$= \frac{1}{z} + \sum_{k=1}^{\infty}\Big(\frac{1}{z-2k} + \frac{1}{z+2k}\Big) - \frac{1}{z-1} - \sum_{k=1}^{\infty}\Big(\frac{1}{z-1-2k} + \frac{1}{z-1+2k}\Big)$$

Thus comparing to (6.16) we obtain

$$g(z) = \frac{\pi}{2}\cot\frac{\pi z}{2} - \frac{\pi}{2}\cot\frac{\pi(z-1)}{2} - \frac{\pi}{\sin \pi z}\ .$$

But the right hand side is easily seen to vanish identically, so we have the partial fractions expansion

$$\frac{\pi}{\sin \pi z} = \sum_{k=-\infty}^{\infty} \frac{(-1)^k}{z-k}$$

EXERCISE 6.22. Evaluate

$$\sum_{k=-\infty}^{\infty} \frac{1}{(z+k)^2 + a^2}\ .$$

6.5. Hadamard's theorem

In this section we will prove a fundamental theorem by Hadamard connecting the growth rate at infinity of an entire function with the distribution of its zeros. As we know, the genus of an entire function gives information about the distribution of its zeros $a_1, a_2, \ldots$, since if the genus is h the function either has only finitely many zeros, or else the series $\sum 1/|a_k|^{h+1}$ converges. We now introduce a measure for the growth at infinity of an entire function f. First denote by $M(r)$ the maximum of $|f(z)|$ on the circle $|z| = r$.

DEFINITION 6.23. The *order* λ of an entire function is defined by

$$\lambda = \overline{\lim_{r\to\infty}} \frac{\log\log M(r)}{\log r}\ .$$

This means that λ is the smallest number such that $|f(z)| < e^{|z|^{\lambda+\varepsilon}}$ for any given $\varepsilon > 0$ as soon as $|z|$ is sufficiently large. Consequently, polynomials have order 0, e^z and $\sin z$ have order 1, $e^{p(z)}$ has order n if p is a polynomial of degree n, and e^{e^z} has infinite order. Note that whereas the genus is always a natural number (or infinity), the order may be any non-negative number (or infinity); for example, the entire function $\frac{\sin\sqrt{z}}{\sqrt{z}}$ we discussed earlier has order 1/2 (show this!).

THEOREM 6.24 (Hadamard). *The genus h and order λ of an entire function satisfy $h \le \lambda \le h+1$.*

The proof needs a bit of preparation. Recall that the real and imaginary parts of an analytic function of $z = x+iy$ are harmonic functions. This means in particular that $\log|f(z)|$ is a harmonic function wherever f is analytic and $\neq 0$, since in a neighborhood of such a point one may define a branch of $\log f(z)$, which has real part $\log|f(z)|$. Furthermore, if u is harmonic in a neighborhood of $|z| \le \rho$, then it satisfies the *Poisson integral formula* (see (5.1))

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{\rho^2 - |z|^2}{|z-\rho e^{it}|^2}\, u(\rho e^{it})\, dt$$

for $|z| < \rho$. In particular, we have the mean value property $u(0) = \frac{1}{2\pi}\int_0^{2\pi} u(\rho e^{it})\, dt$.

If f is analytic in the disk $|z| \le \rho$ and never 0, we can apply Poisson's integral formula to $\log|f(z)|$. If f has zeros inside the circle we instead obtain *Poisson-Jensen's formula.*

THEOREM 6.25. *Suppose $f \in H(\Omega)$ where Ω contains the disk $|z| \le \rho$ and that f has only the zeros $a_1, \dots, a_n$ in $|z| < \rho$, and no zeros on $|z| = \rho$. Then the Poisson-Jensen formula*

(6.17)
$$\log|f(z)| = -\sum_{k=1}^{n} \log\Big|\frac{\rho^2 - \overline{a}_k z}{\rho(z-a_k)}\Big| + \frac{1}{2\pi}\int_0^{2\pi} \operatorname{Re}\frac{\rho e^{it}+z}{\rho e^{it}-z}\, \log|f(\rho e^{it})|\, dt$$

is valid if $|z| < \rho$ is not one of the zeros. In particular, if $f(0) \neq 0$ we have Jensen's formula

$$\log|f(0)| = -\sum_{k=1}^{n} \log\frac{\rho}{|a_k|} + \frac{1}{2\pi}\int_0^{2\pi} \log|f(\rho e^{it})|\, dt$$

PROOF. Note that if $|z| = \rho$, then $\frac{\rho^2 - \overline{a}_k z}{\rho(z-a_k)} = \frac{z}{\rho}\frac{\overline{z}-\overline{a}_k}{z-a_k}$ has absolute value 1. Hence, if we set

$$F(z) = f(z)\prod_{k=1}^{n} \frac{\rho^2 - \overline{a}_k z}{\rho(z-a_k)}\ ,$$

then F has no zeros in $|z| < \rho$ and $|F(z)| = |f(z)|$ for $|z| = \rho$. Thus (6.17) follows on applying Poisson's integral formula to $\log|F(z)|$. □

We can now turn to Hadamard's theorem.

PROOF OF THEOREM 6.24. Assume first that the entire function f has finite genus h. This means that $\sum 1/|a_k|^{h+1}$ converges, where $a_1, a_2, \dots$ are the zeros of f. The exponential factor in (6.13) is clearly of order $\le h$, and since the order of a product clearly does not exceed the order of the factors, we need only consider the canonical product. Using the notation from page 92 it is $P(z) = \prod e^{-R_h(z/a_k)}$. To estimate the size of this we shall prove that

$$|\operatorname{Re} R_h(z)| \le (2h+1)|z|^{h+1} \tag{6.18}$$

for all z. This is true for $h = 0$, since $\log|1-z| \le \log(1+|z|) \le |z|$. By the definition of R_h it is obvious that we have

$$|\operatorname{Re} R_h(z)| \le |\operatorname{Re} R_{h-1}(z)| + |z|^h \tag{6.19}$$

for all z. If $|\operatorname{Re} R_{h-1}(z)| \le (2h-1)|z|^h$ then clearly (6.18) follows if $|z| \ge 1$. But if $|z| < 1$ we have the estimate $(1-|z|)|\operatorname{Re} R_h(z)| \le |z|^{h+1}$ from page 92. Multiplying (6.19) by $|z|$ and adding we get $|\operatorname{Re} R_h(z)| \le |z||\operatorname{Re} R_{h-1}(z)| + 2|z|^{h+1}$ from which again (6.18) follows by the induction assumption. We can now estimate

$$\log|P(z)| = \sum(-\operatorname{Re} R_h(z/a_k)) \le (2h+1)|z|^{h+1} \sum \frac{1}{|a_k|^{h+1}}$$

which shows that the order of $P(z)$ is at most $h+1$.

Conversely we have to prove that if the function f has finite order λ then $\sum 1/|a_k|^{h+1}$ converges, where h is the integer part of λ. If the number of zeros of f in $|z| < \rho$ is denoted $n(\rho)$, then applying Jensen's formula for the disk $|z| \le 2\rho$ we obtain

$$n(\rho)\log 2 \le \frac{1}{2\pi}\int_0^{2\pi} \log|f(2\rho e^{it})|\,dt - \log|f(0)|\ ,$$

where we have ignored the terms coming from zeros satisfying $\rho \le |a_k| < 2\rho$. Given $\varepsilon > 0$ the integrand is here bounded by $\rho^{\lambda+\varepsilon}$ for sufficiently large ρ, so if we order the zeros according to size $|a_1| \le |a_2| \le \dots$ we have $k \le n(|a_k|) \le |a_k|^{\lambda+\varepsilon}$ for large k. Thus we have a bound $1/|a_k|^{h+1} \le 1/k^{(h+1)/(\lambda+\varepsilon)}$. If we choose ε so small that $\alpha = (h+1)/(\lambda+\varepsilon) > 1$ the series $\sum 1/k^\alpha$ converges, so the genus of the canonical product is at most h.

We finally need to show that the function g in the exponential factor in (6.13) is a polynomial of degree $\le h$. To this end, note that if $f = u + iv$ is analytic, then $f' = u'_x + iv'_x = u'_x - iu'_y$ according to the Cauchy-Riemann equations. Applying $\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}$ to (6.17) we therefore obtain

$$\frac{f'(z)}{f(z)} = \sum_{k=1}^{n(\rho)} \frac{1}{z-a_k} + \sum_{k=1}^{n(\rho)} \frac{\overline{a_k}}{\rho^2 - \overline{a_k}z} + \frac{1}{2\pi}\int_0^{2\pi} \frac{2\rho e^{it}}{(\rho e^{it}-z)^2} \log|f(\rho e^{it})|\,dt\ .$$

Differentiating this h times gives

$$\frac{d^h}{dz^h}\frac{f'(z)}{f(z)} = -\sum_{k=1}^{n(\rho)} \frac{h!}{(a_k - z)^{h+1}} + \sum_{k=1}^{n(\rho)} \frac{h!\,\overline{a_k}^{h+1}}{(\rho^2 - \overline{a_k}z)^{h+1}} + \frac{1}{2\pi}\int_0^{2\pi} \frac{2(h+1)!\rho e^{it}}{(\rho e^{it} - z)^{h+2}} \log|f(\rho e^{it})|\,dt\ . \tag{6.20}$$

We will show that the two last terms tend to 0 as $\rho \to \infty$. Note first that the integral vanishes if f is constant, so that the integral is unchanged if we divide f by $M(\rho)$. If $|z| \leq \rho/2$ the absolute value of the integral is therefore at most a constant multiple of

$$\rho^{-h-1}\int_0^{2\pi} \log\frac{M(\rho)}{|f(\rho e^{it})|}\,dt\ .$$

By Jensen's formula we have $\frac{1}{2\pi}\int_0^{2\pi}\log|f| \geq \log|f(0)|$ and since by assumption $\frac{\log(M(\rho))}{\rho^{h+1}} \to 0$ as $\rho \to \infty$, it follows that the integral in (6.20) vanishes as $\rho \to \infty$. Similarly, the penultimate term in (6.20) may, for $|z| \leq \rho/2$ be estimated by $n(\rho)/\rho^{h+1}$ which, as we have already seen, tends to 0 as $\rho \to \infty$. It follows that

$$\frac{d^h}{dz^h}\frac{f'(z)}{f(z)} = -\sum_{k=1}^{\infty} \frac{h!}{(a_k - z)^{h+1}}\ .$$

If we write $f(z) = e^{g(z)}P(z)$, where P is the canonical product, then clearly the sum to the right is $\frac{d^h}{dz^h}\frac{P'(z)}{P(z)}$, so that it follows that $g^{(h+1)}(z) = 0$. Thus g is a polynomial of degree at most h, and the proof is finally complete. □

As an indication of the power of Hadamard's theorem, we have the following corollary.

Corollary 6.26. *An entire function of non-integer order assumes every finite value infinitely many times.*

Proof. Since $f(z)$ and $f(z) - w$ obviously have the same order, as functions of z, for every complex number w, it is enough to show that the function f has infinitely many zeros if it is of non-integer order. If f only has finitely many zeros, then the canonical product is a polynomial and thus of order 0. Thus f is a polynomial times e^p where p also is a polynomial (the genus being finite by Theorem 6.24). If p has degree n, then clearly f has order n, which is an integer. The corollary follows. □

Note that the most useful way to interpret Theorem 6.24 is as a factorization theorem for functions of finite order. If the order is not an integer, the genus, and thus the form of the factorization, is uniquely determined, whereas there is an ambiguity if the order is an integer.

EXERCISE 6.27. Let f be entire and $M(r)$ as before. Suppose $\overline{\lim\limits_{r\to\infty}} \frac{\log M(r)}{r^\lambda} = A$ is finite and not 0. Show that f is of order λ, but that the existence of the limit does not follow from assuming f to have order λ. An entire function for which A is finite and > 0 is said to be of order λ and *normal type*. Extend Corollary 6.26 to show that *an entire function of finite order has infinitely many zeros unless it is of integer order and normal type.*

CHAPTER 7

The Riemann mapping theorem

In this chapter we will prove the Riemann mapping theorem by a limiting procedure. We will then need to know that the sequence of mappings constructed, or at least a subsequence of it, has a limit. To see this, the sequence needs to have a compactness property, analogous to the Bolzano-Weierstrass' theorem for sequences of numbers. The appropriate concept is given by the following definition.

DEFINITION 7.1. A family (*i.e.*, a set) $\mathcal{F}$ of analytic functions defined on a region Ω is called *normal* if every sequence of functions in $\mathcal{F}$ has a subsequence locally uniformly convergent in Ω.

EXERCISE 7.2. Prove this equivalence (use the Heine-Borel theorem)!

The main result about normal families is the following characterization.

THEOREM 7.3. *A family $\mathcal{F}$ of functions analytic on a region Ω is normal if and only if it is* ***locally equibounded****.*

Here *locally equibounded* means that for each compact subset E of Ω there is a constant K_E such that $|f(z)| \leq K_E$ for every $f \in \mathcal{F}$ and $z \in E$. Equivalently, every point in Ω has a neighborhood E such that this holds. The proof of Theorem 7.3 is a fairly simple consequence of a more general compactness theorem by Arzela and Ascoli. Before we can state this theorem we need to make a definition.

DEFINITION 7.4. A family $\mathcal{F}$ of complex valued functions defined in a complex region Ω is called *locally equicontinuous* if for every $\varepsilon > 0$ and compact subset E of Ω there is a $\delta > 0$ such that $|f(z) - f(w)| < \varepsilon$ for every $f \in \mathcal{F}$ and all z, $w \in E$ satisfying $|z - w| < \delta$.

Note that δ as given in the definition above depends *only* on $\mathcal{F}$, E and ε. In other words, it does not depend on the particular function f we are dealing with.

THEOREM 7.5 (Arzela-Ascoli). *Suppose $f_1, f_2, \ldots$ is a sequence of complex-valued functions defined on a region $\Omega \subset \mathbb{C}$, and assume the sequence is locally equibounded and equicontinuous in Ω. Then there is a locally uniformly convergent subsequence.*

PROOF. The set of points in Ω with rational real and imaginary parts is *countable* and *dense* in Ω. That the set is countable means that there is a sequence $z_1, z_2, \dots$ consisting precisely of these points, and that it is dense means that any neighborhood of any point in Ω contains a point from the sequence $z_1, z_2, \dots$. Consider now the sequence $f_1(z_1), f_2(z_1), f_3(z_1), \dots$ of complex numbers. This is a bounded sequence since the set $\{z_1\}$ is compact, so by the Bolzano-Weierstrass' theorem it has a convergent subsequence, given by evaluating a subsequence $f_{11}, f_{12}, f_{13}, \dots$ of $f_1, f_2, \dots$ at z_1; call the limit $f(z_1)$. The sequence $f_{11}(z_2), f_{12}(z_2), f_{13}(z_2), \dots$ is again bounded, so we can find a subsequence $f_{21}, f_{22}, f_{23}, \dots$ of $f_{11}, f_{12}, f_{13}, \dots$ which converges when evaluated at z_2; call the limit $f(z_2)$. Since a subsequence of a convergent sequence converges to the same thing as the sequence itself, we still have $\lim_{n\to\infty} f_{2n}(z_1) = f(z_1)$. Continuing in this fashion we get a sequence of sequences $f_{k1}, f_{k2}, f_{k3}, \dots,\ k = 1, 2, \dots$ such that each sequence is a subsequence of the ones coming before it, and such that $\lim_{n\to\infty} f_{kn}(z_j) = f(z_j)$ for $j \leq k$. Now consider the 'diagonal sequence' $f_{11}, f_{22}, f_{33}, \dots$. This is a subsequence of the sequence $f_{j1}, f_{j2}, f_{j3}, \dots$ from its j:th element onwards, so $\lim_{k\to\infty} f_{kk}(z_j) = f(z_j)$ for *any* j. We shall finish the proof by showing that in fact $f_{11}, f_{22}, f_{33}, \dots$ converges locally uniformly on Ω.

Let a compact subset E of Ω and a number $\varepsilon > 0$ be given. By local equicontinuity we can then find $\delta > 0$ so that $|f_{nn}(z) - f_{nn}(w)| < \varepsilon/3$ for z, $w \in E$ and $|z - w| < \delta$. Now consider the open cover of E given by the balls of radius δ and centered at z_j, $j = 1, 2, \dots$. This is a cover since $z_1, z_2, \dots$ is dense in Ω. By the Heine-Borel theorem there is a finite number of balls, say centered at $z_1, z_2, \dots, z_k$ which already cover E. Given $z \in E$ we can therefore find z_j with $j \leq k$ such that $|z - z_j| < \delta$ and therefore get

$$\begin{aligned} &|f_{nn}(z) - f_{mm}(z)| \\ &\quad \leq |f_{nn}(z) - f_{nn}(z_j)| + |f_{nn}(z_j) - f_{mm}(z_j)| + |f_{mm}(z_j) - f_{mm}(z)| \\ &\qquad\qquad\qquad < \varepsilon/3 + |f_{nn}(z_j) - f_{mm}(z_j)| + \varepsilon/3. \end{aligned}$$

By Cauchy's convergence principle (for complex numbers) and our construction it follows that for every j there is a number N_j such that $|f_{nn}(z_j) - f_{mm}(z_j)| < \varepsilon/3$ if n, $m > N_j$. If we choose N as the largest of $N_1, \dots, N_k$ it follows that

$$|f_{nn}(z) - f_{mm}(z)| < \varepsilon \quad \text{if } n \text{ and } m > N\ .$$

Using the other direction of Cauchy's convergence principle it follows that $f(z) = \lim_{n\to\infty} f_{nn}(z)$ exists for every $z \in \Omega$, and letting $m \to \infty$ in the expression above we get $|f_{nn}(z) - f(z)| \leq \varepsilon$ for every $z \in E$ if $n > N$. This shows that $f_{nn} \to f$ locally uniformly in Ω. □

Proof of Theorem 7.3. It is clear by Theorem 7.5 that all we have to do is show local equicontinuity of $\mathcal{F}$. So let $z_0 \in \Omega$ and choose $r > 0$ such that the closed disk with radius $2r$ and centered at z_0 is in Ω. The boundary of the disk is a compact subset of Ω so we can find a uniform bound M on this set for all $f \in \mathcal{F}$, by assumption. If z and w are in the disk $B(r, z_0)$ with radius r and center z_0 we obtain

$$\begin{aligned}
|f(z) - f(w)| &= \Big|\frac{1}{2\pi i} \int_{|\zeta - z_0| = 2r} f(\zeta)\Big(\frac{1}{\zeta - z} - \frac{1}{\zeta - w}\Big)\, d\zeta\Big| \\
&= \frac{|z-w|}{2\pi}\Big| \int_{|\zeta - z_0| = 2r} \frac{f(\zeta)}{(\zeta - z)(\zeta - w)}\, d\zeta\Big| \\
&\le \frac{M|z-w|}{2\pi r^2} \int_{|\zeta - z_0| = 2r} |d\zeta| = \frac{2M}{r}|z-w|
\end{aligned}$$

since $|\zeta - z| > r$, $|\zeta - w| > r$. It follows that choosing $\delta = \frac{r}{2M}\varepsilon$ makes $|f(z) - f(w)| < \varepsilon$ if z and $w \in B(r, z_0)$ and $|z - w| < \delta$. The local equicontinuity of the family $\mathcal{F}$ follows and the theorem is therefore a corollary to Theorem 7.5. □

Theorem 7.6 (Riemann mapping theorem). *Given a simply connected region Ω which is not the entire complex plane $\mathbb{C}$ and a point $z_0 \in \Omega$ there is precisely one univalent conformal map f of Ω onto the unit disk such that $f(z_0) = 0$ and $f'(z_0) > 0$.*

Note that Liouville's theorem shows that it is not possible to map the *entire* plane $\mathbb{C}$ conformally onto the unit disk; the only bounded entire functions are the constants.

Proof. We have already proved the uniqueness in Chapter 4.4 after Schwarz' lemma (p.70–71). To see how to get existence, note that if g solves the problem and f is a map of Ω into the unit disk mapping z_0 onto 0 and with positive derivative at z_0, then $f \circ g^{-1}$ satisfies the conditions of Schwarz' lemma so $|(f \circ g^{-1})'(0)| \le 1$. Calculating the derivative we see that this means that $f'(z_0) \le g'(z_0)$. If we have equality it follows from Schwarz' lemma that $f \equiv g$.

Now let $\mathcal{F}$ be the family of univalent functions f analytic in Ω such that $f(z_0) = 0$, $|f(z)| \le 1$ for $z \in \Omega$ and $f'(z_0) > 0$. We just saw that if our problem has a solution it is the element of $\mathcal{F}$ which maximizes the derivative at z_0. To complete the proof along these lines we need to: **(1)** Show that $\mathcal{F}$ is not empty, **(2)** See that $\mathcal{F}$ has an element f maximizing the derivative at z_0 and, finally, **(3)** Show that this f actually solves the mapping problem.

(1) Since Ω is not all of $\mathbb{C}$ there is a (finite) point $a \notin \Omega$. Since Ω is simply connected we can define a single-valued branch h of $\sqrt{z-a}$ in Ω. Clearly h can not take the value $-w$ if it somewhere takes the value

w. But by the open mapping theorem there is a disk $|w - h(z_0)| < \rho$ contained in the image $h(\Omega)$. It follows that $|h(z) + h(z_0)| \geq \rho$ for all $z \in \Omega$; in particular $2|h(z_0)| \geq \rho$. The function

$$\frac{h(z) - h(z_0)}{h(z) + h(z_0)} = h(z_0)\Big(\frac{1}{h(z_0)} - \frac{2}{h(z) + h(z_0)}\Big)$$

maps z_0 to 0 and is bounded by $4|h(z_0)|/\rho$. Its derivative at z_0 is $\frac{h'(z_0)}{2h(z_0)}$. If we now put

$$g(z) = \frac{\rho}{4} \frac{|h'(z_0)|}{|h(z_0)|^2} \frac{h(z_0)}{h'(z_0)} \frac{h(z) - h(z_0)}{h(z) + h(z_0)}$$

it follows that g is univalent, $g(z_0) = 0$, $|g(z)| \leq 1$, and $g'(z_0) > 0$ so that $g \in \mathcal{F}$. Hence $\mathcal{F} \neq \varnothing$.

(2) Since all elements of $\mathcal{F}$ have their values in the unit disk it follows that $\mathcal{F}$ is an equibounded family, and therefore by Theorem 7.3 a normal family. Now let $B = \sup_{f \in \mathcal{F}} f'(z_0)$ so that $0 < B \leq \infty$. We can then find a sequence $f_1, f_2, \dots$ in $\mathcal{F}$ so that $f_j'(z_0) \to B$ as $j \to \infty$. Since $\mathcal{F}$ is normal we can find a locally uniformly convergent subsequence; call the limit function f. It is then clear that $f'(z_0) = B$ so that actually $B < \infty$ and f is not constant. By Corollary 6.5 f is univalent. It is clear that $f(z_0) = 0$ and f has its values in the closed unit disk; but by the open mapping theorem the values are then in the open unit disk.

(3) We need to prove that $f(\Omega)$ is the unit disk. Suppose to the contrary that w_0 is in the unit disk but $w_0 \notin f(\Omega)$. Since Ω is simply connected we may define a single-valued branch of

$$G(z) = \sqrt{\frac{f(z) - w_0}{1 - \overline{w_0} f(z)}} \,.$$

Since the Möbius transform $w \mapsto \frac{w - w_0}{1 - \overline{w_0} w}$ preserves the unit disk, the function G maps Ω univalently into the unit disk. To obtain a member of $\mathcal{F}$ we now set

$$F(z) = \frac{|G'(z_0)|}{G'(z_0)} \frac{G(z) - G(z_0)}{1 - \overline{G(z_0)} G(z)} \,.$$

It is again clear that F has its values in the unit disk and maps z_0 to 0. The derivative at z_0 is easily calculated to be $F'(z_0) = B \frac{1 + |G(z_0)|^2}{2|G(z_0)|} > B$ so that $F \in \mathcal{F}$. But this contradicts the definition of B.

Note that it is no accident that we get $F'(z_0) > f'(z_0)$; this just expresses the fact that the inverse of the map $f \mapsto F$ takes the unit disk into itself with 0 fixed so that Schwarz' lemma shows that the derivative at 0 is < 1 (clearly the map is no rotation). $\square$

CHAPTER 8

The Gamma function

In earlier courses you may have encountered the function

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\,dt \ ,$$

the *Gamma function*. The integral converges locally uniformly in z for $\operatorname{Re} z > 0$, since the absolute value of the integrand is $t^{x-1}e^{-t}$ if $z = x + iy$. If $0 < r \le x \le R$ this shows that on the interval $(0, 1]$ the integrand may be estimated by t^{r-1}, the integral of which converges on $(0, 1]$. Similarly, on the interval $[1, \infty)$ the integrand may be estimated by $t^{R-1}e^{-t} = t^{R-1}e^{-t/2} \cdot e^{-t/2}$. Here the first factor tends to 0 as $t \to \infty$ and is therefore bounded on $[0, \infty)$, say by M, so the integrand may be estimated by $Me^{-t/2}$ which has convergent integral. It follows that Γ is analytic in $\operatorname{Re} z > 0$, since the integrand is analytic.

Integration by parts shows that the functional equation

$$\Gamma(z + 1) = z\,\Gamma(z) \tag{8.1}$$

is valid for $\operatorname{Re} z > 0$ (check this). Since clearly $\Gamma(1) = 1$ it follows by induction that $\Gamma(n+1) = n!$ for natural numbers n, so one may view the gamma-function as an extension of the factorial to non-natural numbers. Another very important consequence of (8.1) is that it allows one to extend Γ analytically to the left of $\operatorname{Re} z = 0$. If Γ is already defined in $z + 1$ we may define $\Gamma(z) = \frac{1}{z}\,\Gamma(z + 1)$. Clearly this works as long as $z \ne 0$. By induction we may therefore define Γ everywhere except at the non-positive integers. In these points the extended gamma-function has simple poles. In this way the gamma-function is extended to a meromorphic function in the whole complex plane, with poles at $0, -1, -2, \dots$ and nowhere else.

Exercise 8.1. Calculate the residues of Γ at the non-positive integers!

To obtain a product expansion of Γ, let us first construct an entire function with simple zeros where Γ has poles. Since $\sum 1/n$ diverges but $\sum 1/n^2$ does not, we set

$$F(z) = z \prod_{k=1}^{\infty} (1 + \frac{z}{k}) e^{-z/k} \ .$$

Clearly $F(z)F(-z) = -z^2 \prod_{k\neq 0}(1-\frac{z}{k})e^{z/k}$ so comparison with (6.14) shows that

$$F(z)F(-z) = -\frac{z\sin \pi z}{\pi} \ . \tag{8.2}$$

It is also clear that $F(z+1)$ has the same zeros, apart from $z=0$, as $F(z)$, so we have

$$F(z)/z = e^{\gamma(z)}F(z+1) \tag{8.3}$$

with an entire function γ. To determine γ we take the logarithmic derivative of both sides to obtain

$$\sum_{k=1}^{\infty}(\frac{1}{z+k}-\frac{1}{k}) = \gamma'(z) + \frac{1}{z+1} + \sum_{k=1}^{\infty}(\frac{1}{z+1+k}-\frac{1}{k}) \ .$$

If we replace k by $k+1$ in the first sum we obtain, after simplification

$$\gamma'(z) = \sum_{k=1}^{\infty}(\frac{1}{k}-\frac{1}{k+1}) - 1 \ .$$

Since the series telescopes to the sum 1 we have $\gamma'(z)=0$ so that γ is constant. To determine the value of γ, we note that $F(z)/z \to 1$ as $z \to 0$ so we obtain from (8.3) that $1 = e^{\gamma}F(1)$. But the n:th partial product of $F(1)$ is

$$\frac{2}{1}\frac{3}{2}\cdots\frac{n+1}{n}e^{-1-\frac{1}{2}-\cdots-\frac{1}{n}} = (n+1)\exp(-\sum_{k=1}^{n}\frac{1}{k}) \ ,$$

so that $\gamma = \lim_{n\to\infty}(\sum_{k=1}^{n}\frac{1}{k} - \log n)$. The constant is called *Euler's constant* and equals approximately 0.5772. As far as I know it is not known whether γ is rational (though it seems unlikely). If we set $G(z) = e^{-\gamma z}/F(z)$ we have the expansion

$$G(z) = \frac{e^{-\gamma z}}{z}\prod_{k=1}^{\infty}(1+\frac{z}{k})^{-1}e^{z/k} \ , \tag{8.4}$$

and from (8.3) follows

$$G(z+1) = zG(z) \ ,$$

the same functional equation that Γ satisfies. One might now guess that $G = \Gamma$. We will show this, which is surprisingly difficult. Note that we obtain from (8.2) and the functional equation the so called *reflection formula*

$$G(z)G(1-z) = \frac{\pi}{\sin \pi z} \ .$$

Since F has no poles the function G has no zeros, and since it has the same poles as Γ, the function $\Gamma(z)/G(z)$ is entire. If we can show that it is bounded, then by Liouville's theorem it is constant and since $\Gamma(1) = G(1)$ we would be done. Note that by the functional equations that G and Γ satisfy we have $\Gamma(z+1)/G(z+1) = \Gamma(z)/G(z)$ so that

Γ/G is periodic with period 1. We therefore only need to bound Γ/G in a period strip, say $1 \le \operatorname{Re} z \le 2$. But it is immediately clear that, in this strip, $|\Gamma(x+iy)| \le \Gamma(x)$ so Γ is bounded in the strip by the maximum of Γ in the real interval $[1,2]$. We now need a lower bound for $G(x+iy)$ for $1 \le x \le 2$ and $|y|$ large. Such a bound can be obtained from *Stirling's formula*

$$G(z) = \sqrt{2\pi} z^{z-1/2} e^{-z} e^{J(z)} , \tag{8.5}$$

where $J(z) \to 0$ as $z \to \infty$ in a half-plane $\operatorname{Re} z \ge c > 0$. We will prove this formula later; for the moment let us show that it implies the desired lower bound for G and hence the identity of G and Γ. If (8.5) is true we obtain, for $z = x+iy$,

$$\log|G(z)| = \tfrac{1}{2}\log 2\pi - x + (x - \tfrac{1}{2})\log|z| - y\arg z + \operatorname{Re} J(z) .$$

All terms are here bounded from below except $-y\arg z$, which is at least bounded from below by $-\pi|y|/2$. It follows that Γ/G is bounded in the period strip by a constant multiple of $e^{\pi|y|/2}$. For a function of period 1 this is enough to show boundedness, since such a function may be viewed as a function of $\zeta = e^{2\pi iz}$, the possible values of $z = \frac{1}{2\pi i}\log\zeta$ differing by integers. As a function of ζ the function Γ/G has isolated singularities at 0 and ∞, but our bound on Γ/G is $e^{|\log|\zeta||/4}$, *i.e.*, for small $|\zeta|$ a multiple of $|\zeta|^{-1/4}$ and for large $|\zeta|$ a multiple $|\zeta|^{1/4}$. Thus both singularities are removable (see the following exercise), Γ/G is bounded, and we are done.

Exercise 8.2. Recall that if f is analytic with an isolated singularity at $z = w$, then the singularity is removable if (and only if) $(z-w)f(z) \to 0$ as $z \to w$. State a similar condition for singularities at infinity.
Hint: Look at the discussion just before Theorem 4.2.

Let us now turn to Stirling's formula, so assume $\operatorname{Re} z > 0$. According to (8.4) the logarithmic derivative of G is $-\gamma - \frac{1}{z} - \sum_1^\infty(\frac{1}{z+k} - \frac{1}{k})$ and differentiating once more we get

$$\frac{d}{dz}\frac{G'(z)}{G(z)} = \sum_{k=0}^{\infty}\frac{1}{(z+k)^2} .$$

For fixed z in the right half-plane the terms of the sum are the residues in the right half-plane of the function $H(\zeta) = \dfrac{\pi\cot\pi\zeta}{(z+\zeta)^2}$. Note that $\zeta = -z$ is not in the right half-plane, and $\frac{\pi\zeta\cos\pi\zeta}{\sin\pi\zeta}$ is analytic and equals 1 at 0. Thus the residue at 0 is $\frac{1}{z^2}$. By periodicity the residue at k is $\frac{1}{(z+k)^2}$; thus the residues of H are as stated.

Now let γ be the contour consisting of a rectangle with corners $\pm iY$ and $n+\frac{1}{2}\pm iY$, except for avoiding $\zeta = 0$ by a small semicircle of radius r centered at 0, such that 0 is inside the contour. Consider $\frac{1}{2\pi i}\int_\gamma H$.

This is independent of r for small r and equals $\sum_{k=0}^{n} \frac{1}{(z+k)^2}$. On the horizontal sides the factor $\cot \pi\zeta$ tends uniformly to $\pm i$ as $Y \to \infty$, and the other factor $(z+\zeta)^{-2}$ tends uniformly to 0, so the corresponding integrals also tend to 0. Our contour now consists of two infinite vertical lines, apart from the little semicircle. On the line $\operatorname{Re}\zeta = n+\frac{1}{2}$ the factor $\cot \pi\zeta$ is bounded, independently of the integer n, so the corresponding integral is less than a multiple of $\int_{\operatorname{Re}\zeta=n+\frac{1}{2}} |z+\zeta|^{-2}\, d\operatorname{Im}\zeta$ which tends to 0 as $n \to \infty$. The integrals over the straight line parts of the remaining part of the contour may be written

$$-\frac{1}{2}\int_{-\infty}^{-r} \frac{\cot(i\pi\eta)}{(i\eta+z)^2}\, d\eta - \frac{1}{2}\int_{r}^{\infty} \frac{\cot(i\pi\eta)}{(i\eta+z)^2}\, d\eta$$
$$= \frac{1}{2}\int_{r}^{\infty} \cot(i\pi\eta)\Big(\frac{1}{(i\eta-z)^2} - \frac{1}{(i\eta+z)^2}\Big)\, d\eta\ ,$$

and the integral over the semi-circle tends to $\frac{1}{2z^2}$ as $r \to 0$ so we finally obtain

$$\frac{d}{dz}\frac{G'(z)}{G(z)} = \frac{1}{2z^2} + \frac{1}{2}\int_0^\infty \cot(i\pi\eta)\frac{4i\eta z}{(\eta^2+z^2)^2}\, d\eta\ . \tag{8.6}$$

Exercise 8.3. Verify all calculations and claims above!

Using Euler's formulas we may write $i\cot(i\pi\eta) = 1 + \frac{2}{\exp(2\pi\eta)-1}$, and the part of the integral coming from the term 1 has the value $1/z$. In this way we obtain

$$\frac{d}{dz}\frac{G'(z)}{G(z)} = \frac{1}{z} + \frac{1}{2z^2} + \int_0^\infty \frac{4\eta z}{(\eta^2+z^2)^2}\,\frac{d\eta}{e^{2\pi\eta}-1}\ .$$

We need to integrate this twice to obtain Stirling's formula. A first integration gives, for $\operatorname{Re} z > 0$,

$$\frac{G'(z)}{G(z)} = C + \log z - \frac{1}{2z} - \int_0^\infty \frac{2\eta}{\eta^2+z^2}\,\frac{d\eta}{e^{2\pi\eta}-1}\ .$$

Give a justification for changing the order of integration in the integral! To integrate once more we first make an integration by parts in the integral. Noting that a primitive of the second factor is $\frac{1}{2\pi}\log(1-e^{-2\pi\eta})$ we obtain

$$\int_0^\infty \frac{2\eta}{\eta^2+z^2}\,\frac{d\eta}{e^{2\pi\eta}-1} = \frac{1}{\pi}\int_0^\infty \frac{\eta^2-z^2}{(\eta^2+z^2)^2}\log(1-e^{-2\pi\eta})\, d\eta\ ,$$

so that another integration gives

$$\log G(z) = D + Cz + (z - \tfrac{1}{2}) \log z - \frac{1}{\pi} \int_0^\infty \frac{z}{\eta^2 + z^2} \log(1 - e^{-2\pi\eta})\, d\eta \ .$$

The last term (including the minus-sign) we define to be

$$J(z) = \frac{1}{\pi} \int_0^\infty \frac{z}{\eta^2 + z^2} \log \frac{1}{1 - e^{-2\pi\eta}}\, d\eta \ ,$$

so it only remains to show that $J(z)$ has the claimed behavior and to determine the constants of integration C, D. But we have $|\eta^2 + z^2| = |z - i\eta||z + i\eta| \geq c|z|$ if $\operatorname{Re} z \geq c > 0$ so the integral over $[N, \infty)$ may be estimated by the integral $\frac{1}{c\pi} \int_N^\infty \log \frac{1}{1-e^{-2\pi\eta}}\, d\eta$ which is convergent and therefore $< \varepsilon$ for sufficiently large N. But if $|z| > N$ we can estimate the integral over $(0, N]$ by $\frac{|z|}{\pi(|z|^2 - N^2)} \int_0^\infty \log \frac{1}{1-e^{-2\pi\eta}}\, d\eta$, which tends to 0 as $z \to \infty$. Thus $J(z) \to 0$ if $z \to \infty$ in $\operatorname{Re} z \geq c > 0$. The functional equation for G may be expressed $\log G(z+1) = \log z + \log G(z)$, at least if $z > 0$. Substituting (8.6) in this gives, after simplification,

$$C = -(z + \frac{1}{2}) \log(1 + \frac{1}{z}) + J(z) - J(z+1) \ .$$

Letting $z \to +\infty$ it follows that $C = -1$. To determine D we substitute (8.6) in the reflection formula $G(z)G(1-z) = \pi / \sin \pi z$ for $z = \frac{1}{2} + iy$ to obtain, after simplification,

$$\frac{\pi}{\cosh \pi y} = (e^D)^2$$
$$\times \exp(-1 + iy(\log(\tfrac{1}{2} + iy) - \log(\tfrac{1}{2} - iy)) + J(\tfrac{1}{2} + iy) + J(\tfrac{1}{2} - iy)),$$

where the logarithms have their principal value. Further simplification gives

$$(e^D)^2 = 2\pi \exp(1 + 2y \arctan(2y) - J(\tfrac{1}{2} + iy) - J(\tfrac{1}{2} - iy))/(e^{y\pi} + e^{-y\pi})$$
$$= 2\pi \exp(y\pi - 2y \arctan \tfrac{1}{2y} + 1 - J(\tfrac{1}{2} + iy) - J(\tfrac{1}{2} - iy))/(e^{y\pi} + e^{-y\pi})$$
$$\to 2\pi \quad \text{as} \quad y \to +\infty \ .$$

Since $G(x) > 0$ for $x > 0$ it follows that $e^D = \sqrt{2\pi}$ so we have finally proved Stirling's formula for G. Since this implies the identity of G and Γ we have also proved the reflection formula

$$\Gamma(z)\, \Gamma(1-z) = \frac{\pi}{\sin \pi z} \ .$$

and Stirling's formula

$$\Gamma(z) = \sqrt{2\pi} z^{z-1/2} e^{-z} e^{J(z)}.$$

EXERCISE 8.4. Verify the calculations above. Then show that the integrand in $J(z)$ may be developed as a finite sum of odd powers of $1/z$ plus a remainder and that the result may be integrated to yield an expansion

$$J(z) = \sum_{k=1}^{n} \frac{A_k}{z^{2k-1}} + J_n(z) \ ,$$

where the remainder $J_n(z)$ may be estimated by a constant multiple of $1/z^{2n+1}$ for large z satisfying $\operatorname{Re} z \geq c > 0$. Also show that for fixed z the remainder $J_n(z)$ has no limit as $n \to \infty$. An expansion of this kind is called an *asymptotic expansion* (as $z \to \infty$ in $\operatorname{Re} z \geq c > 0$). One may express the constants A_k explicitly in terms of the so called Bernoulli numbers.