Michael Papadimitrakis

Notes on Complex Analysis

Department of Mathematics

University of Crete

Contents

Chapter 1

The complex plane and the sphere of Riemann.

1.1 The complex plane.

In $\mathbb{R}^2$, besides the usual vector space *addition*, which is defined by

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

there is the operation of *multiplication*, defined by

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

We can easily prove that $\mathbb{R}^2$ equipped with these two binary operations is an algebraic *field*. The neutral element of multiplication is $(1, 0)$ and the inverse of $(x, y) \neq (0, 0)$ is $\big(\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}\big)$.

We denote $\mathbb{C}$ the set $\mathbb{R}^2$ equipped with the above addition and multiplication.

It is easy to prove that the function

$$\mathbb{R} \ni x \mapsto (x, 0) \in \mathbb{C}$$

is a one-to-one *field homomorphism* from $\mathbb{R}$ into $\mathbb{C}$. This permits the *identification* of $\mathbb{R}$ with the subset $\{(x, 0) \,|\, x \in \mathbb{R}\}$ of $\mathbb{C}$. In other words, we may identify every $x \in \mathbb{R}$ with the corresponding $(x, 0) \in \mathbb{C}$ and consider $\mathbb{R}$ as a subset of $\mathbb{C}$. This is exactly the same as the identification we make when we want to view $\mathbb{R}$ as the real line, the x-axis, in the two-dimensional plane identified with $\mathbb{R}^2$. From now on we do not distinguish between x and $(x, 0)$, i.e.

$$x = (x, 0).$$

We define i, the *imaginary unit*, to be the element $(0, 1)$:

$$i = (0, 1).$$

Now we have

$$(x, 0) + i(y, 0) = (x, 0) + (0, 1)(y, 0) = (x, 0) + (0, y) = (x, y).$$

If we replace $(x, 0)$ and $(y, 0)$ with the corresponding x and y, we get

$$x + iy = (x, y).$$

From now on we shall write the elements of $\mathbb{C} = \mathbb{R}^2$ in both forms: $x + iy$ and (x, y). We shall prefer the first, $x + iy$, the *complex form* of the elements of $\mathbb{C}$. We say that $x + iy$ is a **complex number** and that $\mathbb{C}$ is the set of complex numbers.

Now the definitions of addition and multiplication take the forms:

$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2),$$

$$(x_1 + iy_1)(x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2).$$

In particular we have

$$(\pm i)^2 = -1.$$

We shall prove later that, besides the polynomial equation z^2+1 which has as solutions the complex numbers $\pm i$, every polynomial equation with coefficients in $\mathbb{C}$ is solvable in $\mathbb{C}$. In other words, we shall prove that $\mathbb{C}$ is an *algebraically closed field*.

The usual *order relation* $<$, which makes $\mathbb{R}$ an ordered field, cannot be extended in $\mathbb{C}$. In fact, $\mathbb{C}$ cannot be equipped with any order relation so that it becomes an ordered field (with the addition and muptiplication already defined in $\mathbb{C}$). Indeed, no matter what the order relation is, we must have that an element of the form $z^2 = z\,z$ is "positive" if $z \neq 0$, and then we end up with the contradiction: $1 = 1^2$ is "positive" and $-1 = i^2$ is also "positive". Therefore, when we write inequalities like $z \leq w$ or $z < w$ we always accept that z, w are real numbers.

It is customary to use symbols like x, y, u, v, t, ξ, η for real numbers, and symbols like z, w, ζ for complex numbers. For instance, we write: $z = x + iy$, $w = u + iv$, $\zeta = \xi + i\eta$.

For every $z = x + iy = (x, y)$ we introduce the symbols

$$\operatorname{Re} z = x, \quad \operatorname{Im} z = y, \quad \overline{z} = x - iy = (x, -y), \quad |z| = \sqrt{x^2 + y^2}$$

These are called **real part**, **imaginary part**, **conjugate**, and **absolute value** (or **modulus**) of z, respectively.

The useful identities

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

are trivial to prove. We also have the trivial inequalities

$$|\operatorname{Re} z| \leq |z|, \quad |\operatorname{Im} z| \leq |z|, \quad |z| \leq |\operatorname{Re} z| + |\operatorname{Im} z|,$$

and the **triangle inequality**

$$||z| - |w|| \leq |z \pm w| \leq |z| + |w|.$$

The geometrical model for $\mathbb{C}$ is the same as for $\mathbb{R}^2$, i.e. the cartesian plane with two perpendicular axes: every $z = x + iy = (x, y)$ corresponds to the point of the plane with abscissa x and ordinate y. The horizontal axis of all points $x = (x, 0)$ is the **real axis**. The vertical axis of all points $iy = (0, y)$ is the **imaginary axis**. In this framework, the cartesian plane is also called **complex plane**.

We recall that the cartesian equation of the general *line* in the plane is

$$ax + by = c,$$

where $a, b, c \in \mathbb{R}$, $a^2 + b^2 \neq 0$. If we set $z = x + iy$ and $w = a + ib \neq 0$, then the above equation takes the form

$$\operatorname{Re}(\overline{w}z) = c.$$

Similarly, the defining inequalities $ax + by < c$ and $ax + by > c$ of the two *halfplanes* on the two sides of the line with equation $ax+by = c$ become $\operatorname{Re}(\overline{w}z) < c$ and $\operatorname{Re}(\overline{w}z) > c$, respectively.

We shall denote

$$[z_1, z_2] = \{(1 - t)z_1 + tz_2 \,|\, 0 \leq t \leq 1\}$$

the *linear segment* joining the points z_1, z_2. When we say *interval* we mean a linear segment on the real line: $[a,b]\subseteq\mathbb{R}$.

The *euclidean distance* between the points $z_1=(x_1,y_1)$ and $z_2=(x_2,y_2)$ is

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

Therefore, the *circle*, the *open disc*, and the *closed disc* with center $z=(x,y)$ and radius $r>0$ take the form

$$C_z(r)=\{w\,|\,|w-z|=r\},\quad D_z(r)=\{w\,|\,|w-z|<r\},\quad \overline{D}_z(r)=\{w\,|\,|w-z|\leq r\}.$$

We recall the special symbols

$$\mathbb{T},\quad \mathbb{D},\quad \overline{\mathbb{D}}$$

for the unit circle $C_0(1)$, the unit disc $D_0(1)$ and the closed unit disc $\overline{D}_0(1)$, respectively.

The *real part* and the *imaginary part* of a complex function $f:A\to\mathbb{C}$, where A is any nonempty set, are the functions $u=\operatorname{Re} f:A\to\mathbb{R}$ and $v=\operatorname{Im} f:A\to\mathbb{R}$, respectively, defined by

$$u(a)=\operatorname{Re} f(a)=\tfrac{1}{2}(f(a)+\overline{f(a)}),\qquad v(a)=\operatorname{Im} f(a)=\tfrac{1}{2i}(f(a)-\overline{f(a)}).$$

Of course, we have $f(a)=u(a)+iv(a)=(u(a),v(a))$ for $a\in A$.

Now, $\mathbb{C}=\mathbb{R}^2$ has the familiar *euclidean metric space* structure. We have the notions of: interior point, boundary point, limit point, and accumulation point of a set; interior A°, boundary ∂A, and closure $\overline{A}$ of a set A; open set, closed set, compact set, and connected set. We also have the notions of convergence of sequences of complex numbers, and limits and continuity of functions defined in $\mathbb{C}$ or taking values in $\mathbb{C}$.

We only recall the following very simple properties of limits. The variable points z,w may represent the terms of a sequence or the values of an independent variable or the values of a function, and then we get the familiar algebraic properties of limits of sequences and of functions.

Of course, the convergence $z\to z_0$ is equivalent to $|z-z_0|\to 0$. Also, $z\to z_0$ is equivalent to $\operatorname{Re} z\to\operatorname{Re} z_0$, $\operatorname{Im} z\to\operatorname{Im} z_0$. This can be proved by using the inequalities

$$|\operatorname{Re} z-\operatorname{Re} z_0|=|\operatorname{Re}(z-z_0)|\leq|z-z_0|,\quad |\operatorname{Im} z-\operatorname{Im} z_0|=|\operatorname{Im}(z-z_0)|\leq|z-z_0|,$$

$$|z-z_0|\leq|\operatorname{Re}(z-z_0)|+|\operatorname{Im}(z-z_0)|=|\operatorname{Re} z-\operatorname{Re} z_0|+|\operatorname{Im} z-\operatorname{Im} z_0|.$$

Moreover, if $z\to z_0$ and $w\to w_0$, then $z+w\to z_0+w_0$ and $zw\to z_0w_0$. Both can be proved either by reducing them to convergence of real and imaginary parts or -preferably- by using the triangle inequality:

$$|(z+w)-(z_0+w_0)|=|(z-z_0)+(w-w_0)|\leq|z-z_0|+|w-w_0|$$

and

$$\begin{aligned}|zw-z_0w_0|&=|(z-z_0)(w-w_0)+(z-z_0)w_0+(w-w_0)z_0|\\&\leq|z-z_0||w-w_0|+|z-z_0||w_0|+|w-w_0||z_0|.\end{aligned}$$

If $z\to z_0\neq 0$, we can prove that $\frac{1}{z}\to\frac{1}{z_0}$ using the equality $|\frac{1}{z}-\frac{1}{z_0}|=\frac{|z-z_0|}{|z||z_0|}$. We use the equality $|\overline{z}-\overline{z_0}|=|z-z_0|$ to prove that $z\to z_0$ implies $\overline{z}\to\overline{z_0}$. Similarly, we use the triangle inequality $||z|-|z_0||\leq|z-z_0|$ to prove that $z\to z_0$ implies $|z|\to|z_0|$.

We shall consider the limit $z\to\infty$ in section 1.3 where the point ∞ will be introduced.

We also mention the standard examples of polynomial functions

$$p(z)=a_nz^n+\cdots+a_1z+a_0$$

and rational functions

$$r(z)=\tfrac{p(z)}{q(z)}=\tfrac{a_nz^n+\cdots+a_1z+a_0}{b_mz^m+\cdots+b_1z+b_0}.$$

A polynomial function is continuous in $\mathbb{C}$, and a rational function is also continuous in $\mathbb{C}$ except at the roots of the polynomial in its denominator. Again, we shall consider the limits of p and r at infinity, and the limits of r at the roots of its denominator in section 1.3.

1.2 Argument and polar representation.

The trigonometric functions

$$\sin : \mathbb{R} \to \mathbb{R}, \quad \cos : \mathbb{R} \to \mathbb{R}$$

are defined and their properties are studied in the theory of functions of a real variable. In particular, we know that sin and cos are periodic in $\mathbb{R}$, with smallest positive period 2π, i.e. $\sin(\theta+2\pi) = \sin\theta$ and $\cos(\theta + 2\pi) = \cos\theta$.

Let I be any interval of length 2π, which contains only one of its endpoints, e.g. $[0, 2\pi)$ or $(-\pi, \pi]$. Then we know that for every $a, b \in \mathbb{R}$ with $a^2 + b^2 = 1$ there exists a unique $\theta \in I$ so that $\cos\theta = a$ and $\sin\theta = b$. Equivalently, for every $\zeta \in \mathbb{C}$ with $|\zeta| = 1$ there exists a unique $\theta \in I$ so that $\zeta = \cos\theta + i\sin\theta$. Therefore, the function

$$\cos + i\sin : \mathbb{R} \to \mathbb{T}$$

is periodic, with 2π as its smallest positive period, and its restriction

$$\cos + i\sin : I \to \mathbb{T}$$

is one-to-one and onto $\mathbb{T}$. Thus, for every $\zeta \in \mathbb{T}$ the equation $\cos\theta + i\sin\theta = \zeta$ has infinitely many solutions in $\mathbb{R}$, and exactly one solution in I.

Now, for every $z \in \mathbb{C}$, $z \neq 0$, we have $\frac{z}{|z|} \in \mathbb{T}$, and so the equation $\cos\theta + i\sin\theta = \frac{z}{|z|}$ has infinitely many solutions in $\mathbb{R}$, and exactly one solution in the interval I. The set of all solutions in $\mathbb{R}$ is called **argument** or **angle** of z and it is denoted $\arg z$, i.e.

$$\arg z = \big\{\theta \in \mathbb{R} \,\big|\, \cos\theta + i\sin\theta = \tfrac{z}{|z|}\big\}.$$

So we have the equivalence:

$$\theta \in \arg z \quad \Leftrightarrow \quad \theta \in \mathbb{R} \text{ and } \cos\theta + i\sin\theta = \tfrac{z}{|z|}.$$

Thus, $\arg z$ has infinitely many elements and it is clear, by the 2π-periodicity of sin and cos, that these elements form a (two-sided) arithmetical progression of step 2π. In other words, if θ is an arbitrary element of $\arg z$, then all elements of $\arg z$ are described by $\theta + k2\pi$, $k \in \mathbb{Z}$.

On the other hand, the unique solution of the equation $\cos\theta + i\sin\theta = \frac{z}{|z|}$ in the specific interval $I = (-\pi, \pi]$ is called **principal argument** or **principal angle** of z and it is denoted $\mathrm{Arg}\, z$:

$$\theta = \mathrm{Arg}\, z \quad \Leftrightarrow \quad -\pi < \theta \leq \pi \text{ and } \cos\theta + i\sin\theta = \tfrac{z}{|z|}.$$

Thus, $\mathrm{Arg}\, z$ is one of the elements of $\arg z$, the one which is contained in $(-\pi, \pi]$.

Examples. (i) $\mathrm{Arg}\, 3 = 0$ and $\arg 3 = \{k2\pi \,|\, k \in \mathbb{Z}\}$.
(ii) $\mathrm{Arg}(4i) = \frac{\pi}{2}$ and $\arg(4i) = \{\frac{\pi}{2} + k2\pi \,|\, k \in \mathbb{Z}\}$.
(iii) $\mathrm{Arg}(-2) = \pi$ and $\arg(-2) = \{\pi + k2\pi \,|\, k \in \mathbb{Z}\}$.
(iv) $\mathrm{Arg}(1 + i) = \frac{\pi}{4}$ and $\arg(1 + i) = \{\frac{\pi}{4} + k2\pi \,|\, k \in \mathbb{Z}\}$.
(v) $\mathrm{Arg}(-1 - i\sqrt{3}) = -\frac{2\pi}{3}$ and $\arg(-1 - i\sqrt{3}) = \{-\frac{2\pi}{3} + k2\pi \,|\, k \in \mathbb{Z}\}$.

We remark that *we do not define argument or angle for the number* 0.

Since the elements of $\arg z$ form an arithmetical progression of step 2π, is is obvious that, if $z_1 \neq 0, z_2 \neq 0$, then either $\arg z_1 = \arg z_2$ or $\arg z_1 \cap \arg z_2 = \emptyset$. More precisely, it is easy to see that $\arg z_1 = \arg z_2$ if and only if $\frac{z_2}{z_1} > 0$ or, equivalently, if and only if z_1, z_2 belong to the same halfline with vertex 0.

Comparing real and imaginary parts of the two sides of the following identity, we see that it is equivalent to the well known addition formulas of sin and cos:

$$\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2) = (\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2).$$

A direct consequence by induction is the familiar **formula of de Moivre**:

$$\cos(n\theta) + i\sin(n\theta) = (\cos\theta + i\sin\theta)^n$$

for every $n \in \mathbb{Z}$.

Proposition 1.1. *For every nonzero z_1, z_2 we have*

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

By this we mean that the sum of any element of $\arg z_1$ *and any element of* $\arg z_2$ *is an element of* $\arg(z_1 z_2)$ *and, conversely, any element of* $\arg(z_1 z_2)$ *is the sum of an element of* $\arg z_1$ *and an element of* $\arg z_2$*.*

Proof. We take any $\theta_1 \in \arg z_1$ and any $\theta_2 \in \arg z_2$ and $\theta = \theta_1 + \theta_2$. Then by the addition formulas,

$$\cos\theta + i\sin\theta = (\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2) = \tfrac{z_1}{|z_1|}\tfrac{z_2}{|z_2|} = \tfrac{z_1 z_2}{|z_1 z_2|}.$$

Therefore, $\theta \in \arg(z_1 z_2)$.
Conversely, we take any $\theta \in \arg(z_1 z_2)$. We consider $\theta_1 \in \arg z_1$ and we define $\theta_2 = \theta - \theta_1$. Then

$$\cos\theta_2 + i\sin\theta_2 = \tfrac{\cos\theta + i\sin\theta}{\cos\theta_1 + i\sin\theta_1} = \tfrac{z_1 z_2}{|z_1 z_2|} \Big/ \tfrac{z_1}{|z_1|} = \tfrac{z_2}{|z_2|}.$$

Therefore, $\theta_2 \in \arg z_2$ and $\theta = \theta_1 + \theta_2$. $\square$

We note that the equality $\operatorname{Arg}(z_1 z_2) = \operatorname{Arg} z_1 + \operatorname{Arg} z_2$ is not true in general.

Example 1.2.1. $\operatorname{Arg}(-1) + \operatorname{Arg}(-1) = \pi + \pi = 2\pi$, while $\operatorname{Arg}((-1)(-1)) = \operatorname{Arg} 1 = 0$.

The equalities $|z_1 z_2| = |z_1||z_2|$ and $\arg(z_1 z_2) = \arg z_1 + \arg z_2$ express the well known geometric rule: *when two complex numbers are multiplied, their distances from* 0 *are multiplied, and their angles are added.*

It is clear by now that for every $z \neq 0$ we may write

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

where $r = |z|$ and $\theta \in \arg z$. This is called **a polar representation** of z. There are infinitely many polar representations of z, one for each $\theta \in \arg z$. The polar representation with $\theta = \operatorname{Arg} z$ is called **principal polar representation** of z.

As in the case of the argument, *we do not define polar representation for the number* 0.

Exercises.

1.2.1. Which are all the possible values of $\operatorname{Arg}(z_1 z_2) - \operatorname{Arg} z_1 - \operatorname{Arg} z_2$?

1.2.2. Prove that $\arg(1/z) = \arg\overline{z} = -\arg z$ and $\arg(-z) = \pi + \arg z$, after you assign the proper meaning to these equalities.

1.2.3. Prove the following statement for any nonzero z, z_1, z_2. It is true that $z = z_1 z_2$ if and only if the triangle $T(0, 1, z_1)$ with vertices $0, 1, z_1$ is *similar* to the triangle $T(0, z_2, z)$ with vertices $0, z_2, z$ (0 corresponding to 0, 1 corresponding to z_2, and z_1 corresponding to z). This expresses the geometric visualization of the operation of multiplication in $\mathbb{C}$.

1.3 Stereographic projection and the sphere of Riemann.

Let

$$\mathbb{S}^2 = \{(\xi,\eta,\zeta) \in \mathbb{R}^3 \,|\, \xi^2+\eta^2+\zeta^2 = 1\}$$

be the unit sphere in $\mathbb{R}^3$. Through the usual identifications, we may consider $\mathbb{C} = \mathbb{R}^2$ as the set of points

$$z = x+iy = (x,y) = (x,y,0)$$

of $\mathbb{R}^3$.

A distinguished point of $\mathbb{S}^2$ is the north pole

$$N = (0,0,1).$$

Now we take any $z = x+iy \in \mathbb{C}$ and the line Nz in $\mathbb{R}^3$, which contains N and z. Clearly, this line intersects $\mathbb{S}^2$ at N. We shall see that there is a second point of intersection $A = (\xi,\eta,\zeta)$ of Nz and $\mathbb{S}^2$. That $A = (\xi,\eta,\zeta)$ belongs to Nz is equivalent to $\overrightarrow{NA} = t\,\overrightarrow{Nz}$ for some $t \in \mathbb{R}$. This is equivalent to

$$\begin{aligned}\xi - 0 &= t(x-0)\\ \eta - 0 &= t(y-0)\\ \zeta - 1 &= t(0-1)\end{aligned} \tag{1.1}$$

On the other hand, that $A = (\xi,\eta,\zeta)$ belongs to $\mathbb{S}^2$ is equivalent to

$$\xi^2+\eta^2+\zeta^2 = 1. \tag{1.2}$$

That $A = (\xi,\eta,\zeta)$ is a common point of Nz and $\mathbb{S}^2$ is equivalent to (ξ,η,ζ,t) being a solution of the system of the four equations (1.1) and (1.2). We easily solve this system and we find two distinct solutions: the point $N = (0,0,1)$, which we already know, and the point

$$A = \big(\tfrac{2x}{x^2+y^2+1}, \tfrac{2y}{x^2+y^2+1}, \tfrac{x^2+y^2-1}{x^2+y^2+1}\big).$$

Now we consider the mapping

$$\mathbb{C} \ni z = x+iy \mapsto A = (\xi,\eta,\zeta) = \big(\tfrac{2x}{x^2+y^2+1}, \tfrac{2y}{x^2+y^2+1}, \tfrac{x^2+y^2-1}{x^2+y^2+1}\big) \in \mathbb{S}^2 \setminus \{N\}$$

from $\mathbb{C}$ to $\mathbb{S}^2 \setminus \{N\}$. We check easily that this mapping is one-to-one and onto $\mathbb{S}^2 \setminus \{N\}$ and that the inverse mapping is

$$\mathbb{S}^2 \setminus \{N\} \ni A = (\xi,\eta,\zeta) \mapsto z = x+iy = \tfrac{\xi}{1-\zeta} + i\tfrac{\eta}{1-\zeta} \in \mathbb{C}.$$

The two mutually inverse mappings just defined between $\mathbb{C}$ and $\mathbb{S}^2 \setminus \{N\}$ are called **stereographic projections**. We write

$$\mathbb{C} \leftrightarrow \mathbb{S}^2 \setminus \{N\}$$

to denote the action of the two stereographic projections.

We shall see now that both stereographic projections are continuous. We take two points $z = x+iy$ and $z_0 = x_0+iy_0$ in $\mathbb{C}$. Let their images, through stereographic projection, be the points $A = (\xi,\eta,\zeta)$ and $A_0 = (\xi_0,\eta_0,\zeta_0)$ in $\mathbb{S}^2 \setminus \{N\}$. Using the formulas of stereographic projection and doing trivial algebraic manipulations, we can prove that the euclidean distance in $\mathbb{R}^3$ between A and A_0 equals

$$\begin{aligned}|A-A_0| &= \sqrt{(\xi-\xi_0)^2+(\eta-\eta_0)^2+(\zeta-\zeta_0)^2}\\ &= \dots\dots = \tfrac{2\sqrt{(x-x_0)^2+(y-y_0)^2}}{\sqrt{x^2+y^2+1}\,\sqrt{x_0^2+y_0^2+1}} = \tfrac{2|z-z_0|}{\sqrt{|z|^2+1}\,\sqrt{|z_0|^2+1}}.\end{aligned} \tag{1.3}$$

We also take $z = x + iy$ in $\mathbb{C}$ and let its image, through stereographic projection, be $A = (\xi, \eta, \zeta)$ in $\mathbb{S}^2 \setminus \{N\}$. We find that the euclidean distance in $\mathbb{R}^3$ between A and N equals

$$|A - N| = \sqrt{(\xi - 0)^2 + (\eta - 0)^2 + (\zeta - 1)^2} = \ldots\ldots = \tfrac{2}{\sqrt{x^2+y^2+1}} = \tfrac{2}{\sqrt{|z|^2+1}}. \tag{1.4}$$

If $z \to z_0$, then (1.3) implies that $A \to A_0$. Conversely, assume that $A \to A_0$. Then $A \not\to N$ and (1.4) shows that $|z|$ stays bounded. Hence (1.3) implies that $z \to z_0$. We conclude that both *stereographic projections are homeomorphisms between the metric spaces* $\mathbb{C}$ *and* $\mathbb{S}^2 \setminus \{N\}$.

We can continue the previous argument and examine the behaviour of z in $\mathbb{C}$ when its image A in $\mathbb{S}^2 \setminus \{N\}$ tends to the north pole N. Indeed, (1.4) shows that $A \to N$ if and only if $|z| \to +\infty$. In other words, $A \to N$ if and only if the euclidean distance of z from 0 becomes arbitrarily large.

Now, it is natural to introduce and attach to $\mathbb{C}$ an "ideal point", denoted ∞ and called **infinity**, whose "euclidean distance" from 0 is $+\infty$. We define the **extended complex plane** or the **sphere of Riemann** to be

$$\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}.$$

We also extend the previously defined stereographic projections $\mathbb{C} \leftrightarrow \mathbb{S}^2 \setminus \{N\}$ to be the **stereographic projections**

$$\widehat{\mathbb{C}} \leftrightarrow \mathbb{S}^2$$

which map each of $\infty \in \widehat{\mathbb{C}}$ and $N \in \mathbb{S}^2$ onto the other.

Thus, both stereographic projections $\widehat{\mathbb{C}} \leftrightarrow \mathbb{S}^2$ are bijective mappings between $\widehat{\mathbb{C}}$ and $\mathbb{S}^2$. We have seen that their restrictions $\mathbb{C} \leftrightarrow \mathbb{S}^2 \setminus \{N\}$ are homeomorphisms between the metric spaces $\mathbb{C}$ and $\mathbb{S}^2 \setminus \{N\}$. In order to examine the continuity properties of the extended stereographic projections, we have to equip the sets $\widehat{\mathbb{C}}$ and $\mathbb{S}^2$ with corresponding metrics. The metric on $\mathbb{S}^2 \setminus \{N\}$, i.e. the euclidean distance on $\mathbb{R}^3$, is also a metric on $\mathbb{S}^2$. But it is clear that the euclidean metric on $\mathbb{C}$ cannot be extended to become a metric on $\widehat{\mathbb{C}}$. The problem can be solved if we use the equalities (1.3) and (1.4) to transfer the metric on $\mathbb{S}^2$ to a metric on $\widehat{\mathbb{C}}$. If $z, z_0 \in \mathbb{C}$, we consider their images $A, A_0 \in \mathbb{S}^2 \setminus \{N\}$ and we define the *new distance* between z, z_0 to be equal to the euclidean distance in $\mathbb{R}^3$ between A, A_0 given by (1.3) in terms of z, z_0. If $z \in \mathbb{C}$ and $z_0 = \infty$, we consider their images $A \in \mathbb{S}^2 \setminus \{N\}$ and $A_0 = N$ and we define the *new distance* between z, z_0 to be equal to the euclidean distance in $\mathbb{R}^3$ between A, A_0 given by (1.4) in terms of z. The new distance between two points of $\mathbb{C}$ or between a point of $\mathbb{C}$ and ∞ is called **chordal distance**. In other words, we define the chordal distance $\chi(z_1, z_2)$ between z_1, z_2 in $\widehat{\mathbb{C}}$ to be the euclidean distance in $\mathbb{R}^3$ between their images, through stereographic projection, in $\mathbb{S}^2$. I.e.

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

Proposition 1.2. *The function* $\chi : \widehat{\mathbb{C}} \times \widehat{\mathbb{C}} \to \mathbb{R}$ *is a metric on* $\widehat{\mathbb{C}}$.

Proof. We must prove that chordal distance has the following basic properties:
(i) $\chi(z_1, z_2) \geq 0$ for every $z_1, z_2 \in \widehat{\mathbb{C}}$.
(ii) If $z_1, z_2 \in \widehat{\mathbb{C}}$, then: $\chi(z_1, z_2) = 0$ if and only if $z_1 = z_2$.
(iii) $\chi(z_1, z_2) = \chi(z_2, z_1)$ for every $z_1, z_2 \in \widehat{\mathbb{C}}$.
(iv) $\chi(z_1, z_3) \leq \chi(z_1, z_2) + \chi(z_2, z_3)$ for every $z_1, z_2, z_3 \in \widehat{\mathbb{C}}$.
The first three properties are obvious. The fourth, the triangle inequality, can be proved after many calculations using the formula of the chordal distance. But there is a better way. If we take the stereographic projections A_1, A_2, A_3 in $\mathbb{S}^2$ of z_1, z_2, z_3, then from the definition of the chordal distance we have $\chi(z_i, z_j) = |A_i - A_j|$ and we get

$$\chi(z_1, z_3) = |A_1 - A_3| \leq |A_1 - A_2| + |A_2 - A_3| = \chi(z_1, z_2) + \chi(z_2, z_3),$$

since the euclidean distance in $\mathbb{R}^3$ satisfies the triangle inequality. $\square$

The metric χ on $\widehat{\mathbb{C}}$ is called **chordal metric**.

We thus have a second way to measure distances in the complex plane: besides the euclidean distance $|z_1 - z_2|$ we also have the chordal distance $\chi(z_1, z_2)$.

Proposition 1.3. *$\widehat{\mathbb{C}}$ with the chordal metric and $\mathbb{S}^2$ with the euclidean metric of $\mathbb{R}^3$ are homeomorphic metric spaces.*

Proof. Stereographic projections are homeomorphisms between the two metric spaces. In fact they are more than that: they are **isometries**. Indeed, if $z_1, z_2 \in \widehat{\mathbb{C}}$ correspond to $A_1, A_2 \in \mathbb{S}^2$, then by the definition of the chordal metric we have $\chi(z_1, z_2) = |A_1 - A_2|$. I.e. stereographic projections preserve distances and so they are both continuous. $\square$

Proposition 1.4 describes the relation between the chordal metric and the euclidean metric in their common domain.

Proposition 1.4. *The chordal metric on $\mathbb{C}$ and the euclidean metric on $\mathbb{C}$ are equivalent.*

Proof. If $z, z_0 \in \mathbb{C}$, then $z \to z_0$ with respect to the euclidean distance if and only if $z \to z_0$ with respect to the chordal distance. To see this we consider the images $A, A_0 \in \mathbb{S}^2 \setminus \{N\}$ of z, z_0 under stereographic projection. We have proved already that $z \to z_0$ with respect to the euclidean distance in $\mathbb{C}$ if and only if $A \to A_0$ with respect to the euclidean distance in $\mathbb{R}^3$. But the euclidean distance between A, A_0 is *equal* to the chordal distance between z, z_0. Therefore,

$$|z - z_0| \to 0 \quad \Leftrightarrow \quad |A - A_0| \to 0 \quad \Leftrightarrow \quad \chi(z, z_0) \to 0.$$

Thus, the euclidean metric and the chordal metric on $\mathbb{C}$ are equivalent. $\square$

Proposition 1.5. *Let $z \in \mathbb{C}$. Then $z \to \infty$ in $\widehat{\mathbb{C}}$ if and only if $|z| \to +\infty$.*

Proof. This is obvious from $\chi(z, \infty) = 2/\sqrt{|z|^2 + 1}$. $\square$

We have introduced ∞ as the ideal point towards which a variable point z on the complex plane moves when its euclidean distance from 0 becomes arbitrarily large. *It is time to mention the difference with the ideal points $\pm\infty$ we attach to $\mathbb{R}$.* A variable point x on the real line moves away from 0 in exactly two specific directions: either to the left or to the right and then we say, respectively, that it moves *towards* $-\infty$ or *towards* $+\infty$. On the plane though there are no two uniquely specified directions. A point can move away from 0 either on arbitrary halflines (i.e. in infinitely many directions) or making an arbitrary "spiral-like movement" or in a completely arbitrary manner. Therefore, we may only say that the point moves *towards infinity*.

Now let us say a few things about neighborhoods of points in $\widehat{\mathbb{C}}$ with respect to the chordal metric. We start with the neighborhoods of ∞. If we denote $N_x(r)$ the r-neighborhood of a point x in the general metric space, then the r-neighborhood of ∞ in the metric space $(\widehat{\mathbb{C}}, \chi)$ is the set

$$\begin{aligned} N_\infty(r) &= \{z \in \widehat{\mathbb{C}} \,|\, \chi(z, \infty) < r\} = \{z \in \mathbb{C} \,|\, 2/\sqrt{|z|^2 + 1} < r\} \cup \{\infty\} \\ &= \begin{cases} \{z \in \mathbb{C} \,|\, |z| > \sqrt{(4/r^2) - 1}\} \cup \{\infty\}, & \text{if } 0 < r \leq 2 \\ \widehat{\mathbb{C}}, & \text{if } r > 2 \end{cases} \end{aligned}$$

We observe that the "small" neighborhoods of ∞, i.e. the neighborhoods $N_\infty(r)$ with $0 < r < 2$, are the complements of closed discs in $\mathbb{C}$ with center 0, together with ∞. To simplify notation we make the change of variable: $\frac{1}{s} = \sqrt{\frac{4}{r^2} - 1}$. If r increases in $(0, 2)$, then s increases in $(0, +\infty)$, and conversely. We call s**-neighborhood** of ∞ in $\widehat{\mathbb{C}}$ the set

$$D_\infty(s) = \left\{z \,\middle|\, |z| > \tfrac{1}{s}\right\} \cup \{\infty\},$$

i.e. the complement of the closed disc with center 0 and radius $\frac{1}{s}$, together with ∞.

We see that the neighborhoods of ∞ in $\widehat{\mathbb{C}}$ with respect to the chordal metric are of three kinds: the sets $D_\infty(s)$ with $s > 0$, the set $\widehat{\mathbb{C}} \setminus \{0\}$ (the case $r = 2$ or, equivalently, $s = +\infty$) and the whole set $\widehat{\mathbb{C}}$ (the case $r > 2$). Since in a metric space it is the "small" neighborhoods which actually characterize interior points, boundary points, limit points, limits of functions or sequences etc., in the case of $\widehat{\mathbb{C}}$ and ∞ we shall pay attention only to the neighborhoods of the form $D_\infty(s)$.

Now the following should be clear.

(i) The point ∞ is an interior point of $A \subseteq \widehat{\mathbb{C}}$ with respect to the chordal metric if and only if A contains, besides ∞, the complement of a closed disc in $\mathbb{C}$ with center 0.

(ii) The point ∞ is not a limit point of $A \subseteq \widehat{\mathbb{C}}$ with respect to the chordal metric if and only if A is contained in a closed disc with center 0 or, equivalently, A is a bounded set in $\mathbb{C}$ with respect to the euclidean metric.

(iii) If $\infty \notin A$, i.e. if $A \subseteq \mathbb{C}$, then we have the following equivalences: [∞ is a boundary point of A with respect to the chordal metric] $\Leftrightarrow$ [∞ is a limit point of A with respect to the chordal metric] $\Leftrightarrow$ [∞ is an accumulation point of A with respect to the chordal metric] $\Leftrightarrow$ [A is not bounded in $\mathbb{C}$ with respect to the euclidean metric].

Now we continue with the neighborhoods with respect to the chordal metric of a point $z_0 \in \mathbb{C}$. The r-neighborhood of $z_0 \in \mathbb{C}$ in $\widehat{\mathbb{C}}$ with respect to the chordal metric is the set

$$N_{z_0}(r) = \{z \in \widehat{\mathbb{C}} \,|\, \chi(z, z_0) < r\}.$$

This set does not have a simple form. Depending on the exact values of z_0 and r, it is an open disc or an open halfplane or the complement of a closed disc (together with ∞). Even when $N_{z_0}(r)$ is an open disc, z_0 is *not* its euclidean center. Look at exercise 1.3.2 for details. Since the chordal metric and the euclidean metric are equivalent in $\mathbb{C}$, we have the following relation between neighborhoods $N_{z_0}(r)$ with respect to the chordal metric and neighborhoods (i.e. the familiar discs) $D_{z_0}(r)$ with respect to the euclidean metric: for every $\epsilon > 0$ there is $\delta > 0$ so that $D_{z_0}(\delta) \subseteq N_{z_0}(\epsilon)$ and, conversely, for every $\epsilon > 0$ there is $\delta > 0$ so that $N_{z_0}(\delta) \subseteq D_{z_0}(\epsilon)$. From this we conclude easily that $z_0 \in \mathbb{C}$ is an interior point or a boundary point or a limit point of a set $A \subseteq \widehat{\mathbb{C}}$ with respect to the chordal metric if and only if it is, respectively, an interior point or a boundary point or a limit point of A with respect to the euclidean metric.

If $A \subseteq \mathbb{C}$ and we write A°, ∂A and $\overline{A}$ for the interior, the boundary and the closure of A with respect to the euclidean metric and $A^{\circ,\chi}$, $\partial_\chi A$ and $\overline{A}^\chi$ for the interior, the boundary and the closure of A with respect to the chordal metric, then we easily see that

$$A^{\circ,\chi} = A^\circ, \quad \partial_\chi A = \partial A, \quad \overline{A}^\chi = \overline{A}$$

for bounded $A \subseteq \mathbb{C}$, and

$$A^{\circ,\chi} = A^\circ, \quad \partial_\chi A = \partial A \cup \{\infty\}, \quad \overline{A}^\chi = \overline{A} \cup \{\infty\}$$

for unbounded $A \subseteq \mathbb{C}$. (Of course, when we say bounded or unbounded we mean with respect to the euclidean metric.)

For instance, if $A \subseteq \mathbb{C}$ is bounded, then it is open with respect to the chordal metric if and only if it is open with respect to the euclidean metric, and it is closed with respect to the chordal metric if and only if it is closed with respect to the euclidean metric. If $A \subseteq \mathbb{C}$ is not bounded, then again it is open with respect to the chordal metric if and only if it is open with respect to the euclidean metric, but, even if it is closed with respect to the euclidean metric, we have to attach ∞ to A to make it closed with respect to the chordal metric.

Regarding compactness, we know that $\mathbb{C}$ is not compact either with respect to the euclidean metric or with respect to the chordal metric. Indeed, $\mathbb{C}$ is not compact with respect to the euclidean metric, because it is not bounded. And then it is not compact with respect to the chordal metric, because the two metrics are equivalent in $\mathbb{C}$. But $\widehat{\mathbb{C}}$ *is compact* (with respect to the chordal metric, of course). Indeed, $\widehat{\mathbb{C}}$ is homeomorphic to $\mathbb{S}^2$, which is compact since it is a closed and bounded set in $\mathbb{R}^3$. Now, $\widehat{\mathbb{C}}$ is produced from $\mathbb{C}$ by the attachment to $\mathbb{C}$ of the single point ∞. This situation has a name in topology: we say that $\widehat{\mathbb{C}}$ is a **one-point compactification** of $\mathbb{C}$.

Based on the usual algebraic rules of limits, we may extend in the standard way the algebraic operations in the set $\widehat{\mathbb{C}}$:

$$z + \infty = \infty + z = \infty, \qquad -\infty = \infty, \qquad z - \infty = \infty - z = \infty,$$

$$z\,\infty = \infty\, z = \infty \quad \text{if } z \neq 0, \qquad \infty\,\infty = \infty,$$

$$\tfrac{1}{\infty} = 0, \qquad \tfrac{1}{0} = \infty, \qquad \tfrac{z}{\infty} = 0, \qquad \tfrac{\infty}{z} = \infty,$$

$$\overline{\infty} = \infty, \qquad |\infty| = +\infty.$$

For example, the rule $z_0 + \infty = \infty$ (when $z_0 \in \mathbb{C}$) can be based on the following argument. If $z \to z_0$ and $w \to \infty$ in $\widehat{\mathbb{C}}$, then $|z - z_0| \to 0$ and $|w| \to +\infty$ and then, by the triangle inequality, $|z + w| \geq |w| - |z - z_0| - |z_0| \to +\infty$. Hence $z + w \to \infty$ in $\widehat{\mathbb{C}}$. All other rules can be based on similar arguments.

The following are not defined:

$$\infty + \infty, \quad \infty - \infty, \quad 0\,\infty, \quad \infty\, 0, \quad \tfrac{\infty}{\infty}, \quad \tfrac{0}{0}.$$

They are called **indeterminate forms**.

For instance, regarding the case of $\infty + \infty$, one can easily find examples of points z, w such that $z \to \infty$ and $w \to \infty$ but such that $z + w$ has either no limit or any preassigned limit. The same is true in all other cases.

Observe the case of $\frac{1}{0} = \infty$. In $\mathbb{R}$ the expression $\frac{1}{0}$ is an indeterminate form, since when the real number x is small and > 0 then $\frac{1}{x}$ is large and > 0 and hence $\frac{1}{x}$ moves towards $+\infty$, and when x is small and < 0 then $\frac{1}{x}$ is large and < 0 and hence $\frac{1}{x}$ moves towards $-\infty$. But in $\mathbb{C}$, when z is small, i.e. when $|z|$ is small (and necessarily > 0), then the distance $|\frac{1}{z}| = \frac{1}{|z|}$ of $\frac{1}{z}$ from 0 is large and hence $\frac{1}{z}$ moves towards ∞ in $\widehat{\mathbb{C}}$. So we define $\frac{1}{0} = \infty$ in $\widehat{\mathbb{C}}$.

Example 1.3.1. Let us consider any polynomial function $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ with $a_n \neq 0$. The domain of definition of p is $\mathbb{C}$.
For every $z_0 \in \mathbb{C}$ we have

$$\lim_{z \to z_0} p(z) = p(z_0),$$

using the algebraic rules of limits and the trivial limits: $\lim_{z \to z_0} c = c$ and $\lim_{z \to z_0} z = z_0$. Therefore, p is continuous in $\mathbb{C}$.
If the degree of p is ≥ 1, i.e. $n \geq 1$ and $a_n \neq 0$, we write

$$p(z) = z^n \big(a_n + a_{n-1} \tfrac{1}{z} + \cdots + a_0 \tfrac{1}{z^n}\big)$$

and we get

$$\lim_{z \to \infty} p(z) = \infty.$$

Thus, if the degree of p is ≥ 1, we may define $p(\infty) = \infty$ and then $p : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ is continuous in $\widehat{\mathbb{C}}$.
If the degree of p is 0, then the function is constant: $p(z) = a_0$ for all z. Hence

$$\lim_{z \to \infty} p(z) = a_0.$$

In this case we may define $p(\infty) = a_0$ and again $p : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ is continuous in $\widehat{\mathbb{C}}$.

Example 1.3.2. Now we take a rational function $r(z) = \frac{p(z)}{q(z)} = \frac{a_n z^n + \cdots + a_1 z + a_0}{b_m z^m + \cdots + b_1 z + b_0}$ with $a_n, b_m \neq 0$.
The domain of definition of r is $\mathbb{C} \setminus \{z_1, \ldots, z_s\}$, where $z_1, \ldots, z_s$ are the roots of q.
If $z_0 \in \mathbb{C}$ and $q(z_0) \neq 0$, then using the algebraic rules of limits, we get:

$$\lim_{z \to z_0} r(z) = r(z_0).$$

Therefore r is continuous in its domain of definition.
Writing r in the form

$$r(z) = z^{n-m}(a_n + a_{n-1}\tfrac{1}{z} + \cdots + a_0\tfrac{1}{z^n})\big/(b_m + b_{m-1}\tfrac{1}{z} + \cdots + b_0\tfrac{1}{z^m}),$$

we can prove that

$$\lim_{z\to\infty} r(z) = \begin{cases} \infty, & \text{if } n > m \\ \frac{a_n}{b_n}, & \text{if } n = m \\ 0, & \text{if } n < m \end{cases}$$

Finally, let $z_0 \in \mathbb{C}$ and $q(z_0) = 0$, i.e. z_0 is one of the roots q. Then $z - z_0$ divides $q(z)$, and there is $k \geq 1$ and a polynomial $q_1(z)$ so that $q(z) = (z - z_0)^k q_1(z)$ and $q_1(z_0) \neq 0$. This means that the *multiplicity* of the root z_0 of $q(z)$ is k. There is also $l \geq 0$ and a polynomial $p_1(z)$ so that $p(z) = (z - z_0)^l p_1(z)$ and $p_1(z_0) \neq 0$. Indeed, if $p(z_0) = 0$, then $l \geq 1$ is the multiplicity of z_0 as a root of $p(z)$ and, if $p(z_0) \neq 0$, we take $l = 0$ (and we say that the multiplicity of z_0 as a root of $p(z)$ is zero) and $p_1(z) = p(z)$. So for every z different from the roots of $q(z)$ we have

$$r(z) = (z - z_0)^{l-k}\frac{p_1(z)}{q_1(z)}$$

and $p_1(z_0) \neq 0$, $q_1(z_0) \neq 0$. Now $\frac{p_1(z_0)}{q_1(z_0)}$ is neither ∞ nor 0, and hence

$$\lim_{z\to z_0} r(z) = \begin{cases} \infty, & \text{if } k > l \\ \frac{p_1(z_0)}{q_1(z_0)}, & \text{if } k = l \\ 0, & \text{if } k < l \end{cases}$$

Exactly as in the polynomial case, a rational function can be considered to be a function $r : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ continuous in $\widehat{\mathbb{C}}$. Indeed, as we just saw above, at every $z_0 \in \widehat{\mathbb{C}}$ a rational function r has a specific limit in $\widehat{\mathbb{C}}$. Now, if z_0 is in the usual domain of definition of r, then the limit of r at z_0 coincides with $r(z_0)$. Moreover, if z_0 is either ∞ or a root of the denominator of r, then we define $r(z_0)$ to be the limit of r at z_0.

Example 1.3.3. The sequence $((-2)^n)$ does not have a limit as a real sequence since its subsequences of the odd and the even indices have the different limits $-\infty$ and $+\infty$. But as a complex sequence $((-2)^n)$ tends to ∞, because $|(-2)^n| = 2^n \to +\infty$.

Example 1.3.4. Let us consider the geometric progression (z^n).
If $|z| < 1$, then $|z^n - 0| = |z|^n \to 0$ and hence $z^n \to 0$.
If $|z| > 1$, then $|z^n| = |z|^n \to +\infty$ and hence $z^n \to \infty$.
If $z = 1$, then $z^n = 1 \to 1$.
Finally, let $|z| = 1$, $z \neq 1$ and assume that $z^n \to w$. Since $|z^n| = |z|^n = 1$ for every n, we find that $|w| = 1$. From $z^n \to w$ we have $z = \frac{z^{n+1}}{z^n} \to \frac{w}{w} = 1$ and we arrive at a contradiction.
Thus:

$$z^n \begin{cases} \to 0, & \text{if } |z| < 1 \\ \to 1, & \text{if } z = 1 \\ \to \infty, & \text{if } |z| > 1 \\ \text{has no limit}, & \text{if } |z| = 1, z \neq 1 \end{cases}$$

Exercises.

1.3.1. Prove that $\chi(z_1, z_2) \leq 2$ for every $z_1, z_2 \in \widehat{\mathbb{C}}$. When does $\chi(z_1, z_2) = 2$ happen?

1.3.2. (i) Let l be any line in $\mathbb{C}$. We define $\widehat{l} = l \cup \{\infty\}$ and call it *line in* $\widehat{\mathbb{C}}$. We call *circle in* $\widehat{\mathbb{C}}$ every circle in $\mathbb{C}$. Prove that stereographic projection maps circles in $\widehat{\mathbb{C}}$ onto circles in $\mathbb{S}^2$ which do not contain N (and conversely) and lines in $\widehat{\mathbb{C}}$ onto circles in $\mathbb{S}^2$ which contain N (and conversely).
(ii) Find the images in $\mathbb{S}^2$ through stereographic projection of the following subsets (or collections of subsets) of $\widehat{\mathbb{C}}$:
(a) $\{z \,|\, |z| < 1\}$, $\{z \,|\, |z| = 1\}$, $\{z \,|\, |z| > 1\} \cup \{\infty\}$,
(b) $\{z \,|\, \mathrm{Re}\, z > 0\}$, $\{z \,|\, \mathrm{Re}\, z = 0\}$, $\{z \,|\, \mathrm{Re}\, z < 0\}$,
(c) the collection of lines containing a fixed point $\neq \infty$,
(d) the collection of circles with a fixed center,
(e) the collection of lines parallel to a fixed line,
(f) the collection of circles tangent to a fixed circle at a fixed point,
(g) the collection of circles containing two fixed points.
(iii) Let $z, w \in \widehat{\mathbb{C}}$ and let $A, B \in \mathbb{S}^2$ be their images through stereographic projection. If z, w are symmetric with respect to a line $\widehat{l}$ in $\widehat{\mathbb{C}}$ which contains 0, which is the relative position of A, B with respect to the image of $\widehat{l}$ in $\mathbb{S}^2$? If $w = \frac{1}{\overline{z}}$, which is the relative position of A, B in $\mathbb{S}^2$?
(iv) Consider a set of the form $P = \{z \in \widehat{\mathbb{C}} \,|\, \chi(z, z_0) = r\}$, where $z_0 \in \widehat{\mathbb{C}}$ and $r > 0$, i.e. a "circle" with respect to the chordal metric. If $z_0 = \infty$, prove that P is a circle in $\widehat{\mathbb{C}}$, i.e. in $\mathbb{C}$, and find its euclidean center and its euclidean radius. If $z_0 \in \mathbb{C}$, prove that P is either a circle in $\widehat{\mathbb{C}}$, i.e. in $\mathbb{C}$ -and in this case find its euclidean center and its euclidean radius- or a line in $\widehat{\mathbb{C}}$.
(v) If the lines $\widehat{l_1}, \widehat{l_2}$ have angle θ at their common point $z \in \mathbb{C}$, prove that their images through stereographic projection, i.e. two circles in $\mathbb{S}^2$ containing the image A of z and the north pole N, have the same angle θ at both A and N.

1.3.3. Let Σ be a collection of unbounded and connected subsets of $\mathbb{C}$. Prove that $(\bigcup_{A\in\Sigma} A) \cup \{\infty\}$ is a connected subset of $\widehat{\mathbb{C}}$.

Chapter 2

Series and curvilinear integrals.

2.1 Series of numbers.

A **series of complex numbers** or, simply, **complex series** is an expression

$$z_1 + z_2 + \cdots + z_n + \cdots \qquad \text{or} \qquad \sum_{n=1}^{+\infty} z_n.$$

If all z_n are real, we talk about a **series of real numbers** or **real series**. The $s_n = z_1 + \cdots + z_n$ are the **partial sums** of the series. We say that the series **converges** if the sequence (s_n) converges and then the limit s of (s_n) is called **sum** of the series and we write $\sum_{n=1}^{+\infty} z_n = s$. We say that the series **diverges** if (s_n) diverges. If (s_n) diverges to ∞, then we say that the series **diverges to** ∞ and that ∞ is the **sum** of the series and we write $\sum_{n=1}^{+\infty} z_n = \infty$.

We note that the sum of a complex series can be either a complex number or ∞. Only a real series can have sum equal to $+\infty$ or $-\infty$. Therefore, when we write $\sum_{n=1}^{+\infty} z_n = +\infty$ or $-\infty$, we accept that all z_n are real and that the series diverges to $+\infty$ or $-\infty$ as a real series. Of course, if a real series diverges to $+\infty$ or $-\infty$, then as a complex series it diverges to ∞.

Example 2.1.1. We have $\sum_{n=1}^{+\infty} c = 0$, if $c = 0$, and $\sum_{n=1}^{+\infty} c = \infty$, if $c \neq 0$.

Example 2.1.2. To examine the geometric series $\sum_{n=0}^{+\infty} z^n$, we use the formula $1 + z + \cdots + z^n = \frac{1-z^{n+1}}{1-z}$ for its partial sums, and we find that its sum is

$$\sum_{n=0}^{+\infty} z^n \begin{cases} = \frac{1}{1-z}, & \text{if } |z| < 1 \\ = \infty, & \text{if } |z| > 1 \text{ or } z = 1 \\ \text{does not exist}, & \text{if } |z| = 1,\ z \neq 1 \end{cases}$$

The usual simple algebraic rules, which hold for real series, hold also for complex series. We mention them without proofs. The proofs in the complex case are identical with the proofs in the real case.

Proposition 2.1. *If $\sum_{n=1}^{+\infty} z_n$ converges, then $z_n \to 0$.*

Proposition 2.2. *Provided that the right sides of the following formulas exist and that they are not indeterminate forms, we have*

$$\sum_{n=1}^{+\infty}(z_n + w_n) = \sum_{n=1}^{+\infty} z_n + \sum_{n=1}^{+\infty} w_n, \quad \sum_{n=1}^{+\infty} \lambda z_n = \lambda \sum_{n=1}^{+\infty} z_n, \quad \sum_{n=1}^{+\infty} \overline{z_n} = \overline{\sum_{n=1}^{+\infty} z_n}.$$

Moreover, if $z_n = x_n + iy_n$, then $\sum_{n=1}^{+\infty} z_n$ converges if and only if $\sum_{n=1}^{+\infty} x_n$ and $\sum_{n=1}^{+\infty} y_n$ converge, and

$$\sum_{n=1}^{+\infty} z_n = \sum_{n=1}^{+\infty} x_n + i \sum_{n=1}^{+\infty} y_n.$$

Regarding the *comparison theorems*, we may say that, since these are based on order relations which can be expressed only between real numbers, when we write $\sum_{n=1}^{+\infty} z_n \leq \sum_{n=1}^{+\infty} w_n$ as a consequence of $z_n \leq w_n$, we accept that all z_n, w_n are real and then we just apply the well known comparison theorems for real series.

Cauchy criterion. *The series* $\sum_{n=1}^{+\infty} z_n$ *converges if and only if for every* $\epsilon > 0$ *there is* n_0 *so that* $|\sum_{k=m+1}^{n} z_k| = |z_{m+1} + \cdots + z_n| < \epsilon$ *for every* m, n *with* $n > m \geq n_0$.

Proof. Let $s_n = z_1 + \cdots + z_n$. The series converges if and only if (s_n) converges or, equivalently, if and only if (s_n) is a Cauchy sequence. That (s_n) is a Cauchy sequence means that for every $\epsilon > 0$ there is n_0 so that $|z_{m+1} + \cdots + z_n| = |s_n - s_m| < \epsilon$ for every n, m with $n > m \geq n_0$. $\square$

We say that $\sum_{n=1}^{+\infty} z_n$ **converges absolutely** if the (real) series $\sum_{n=1}^{+\infty} |z_n|$ converges, i.e. if $\sum_{n=1}^{+\infty} |z_n| < +\infty$.

Criterion of absolute convergence. *If* $\sum_{n=1}^{+\infty} z_n$ *converges absolutely, then it converges and we have*

$$|\sum_{n=1}^{+\infty} z_n| \leq \sum_{n=1}^{+\infty} |z_n| < +\infty.$$

Proof. Let $\sum_{n=1}^{+\infty} |z_n|$ converge and take any $\epsilon > 0$. From the Cauchy criterion we have that there is n_0 so that $|z_{m+1}| + \cdots + |z_n| < \epsilon$ and hence $|z_{m+1} + \cdots + z_n| < \epsilon$ for every m, n with $n > m \geq n_0$. The Cauchy criterion, again, implies that $\sum_{n=1}^{+\infty} z_n$ converges.
Now we take the partial sums $s_n = z_1 + \cdots + z_n$ and $S_n = |z_1| + \cdots + |z_n|$. We have $|s_n| \leq S_n$ for all n and then we take the limit of this as $n \to +\infty$. $\square$

Ratio test of d' Alembert. *Let* $z_n \neq 0$ *for all* n.
(i) If $\overline{\lim} \left|\frac{z_{n+1}}{z_n}\right| < 1$, *then* $\sum_{n=1}^{+\infty} z_n$ *converges absolutely.*
(ii) If $\underline{\lim} \left|\frac{z_{n+1}}{z_n}\right| > 1$, *then* $\sum_{n=1}^{+\infty} z_n$ *diverges.*
(iii) If $\underline{\lim} \left|\frac{z_{n+1}}{z_n}\right| \leq 1 \leq \overline{\lim} \left|\frac{z_{n+1}}{z_n}\right|$, *then there is no general conclusion.*

Proof. (i) Take a so that $\overline{\lim} \left|\frac{z_{n+1}}{z_n}\right| < a < 1$. Then there is n_0 so that $\left|\frac{z_{n+1}}{z_n}\right| \leq a$ for every $n \geq n_0$. Now, for every $n \geq n_0$ we get

$$|z_n| = \left|\frac{z_n}{z_{n-1}}\right| \left|\frac{z_{n-1}}{z_{n-2}}\right| \cdots \left|\frac{z_{n_0+1}}{z_{n_0}}\right| |z_{n_0}| \leq a^{n-n_0} |z_{n_0}| = c\, a^n,$$

where $c = |z_{n_0}|/a^{n_0}$. Since $0 \leq a < 1$, the geometric series $\sum_{n=1}^{+\infty} a^n$ converges and, by comparison, $\sum_{n=1}^{+\infty} |z_n|$ also converges.
(ii) There is n_0 so that $\left|\frac{z_{n+1}}{z_n}\right| \geq 1$ for every $n \geq n_0$. Now, for every $n \geq n_0$ we have

$$|z_n| \geq |z_{n-1}| \geq \cdots \geq |z_{n_0}| > 0.$$

This implies that $z_n \not\to 0$ and so $\sum_{n=1}^{+\infty} z_n$ diverges.
(iii) For the series $\sum_{n=1}^{+\infty} \frac{1}{n}$ and $\sum_{n=1}^{+\infty} \frac{1}{n^2}$ we have that $\left|\frac{1/(n+1)}{1/n}\right| \to 1$ and $\left|\frac{1/(n+1)^2}{1/n^2}\right| \to 1$. The first series diverges and the second converges. $\square$

Root test of Cauchy. *(i) If* $\overline{\lim} \sqrt[n]{|z_n|} < 1$, *then* $\sum_{n=1}^{+\infty} z_n$ *converges absolutely.*
(ii) If $\overline{\lim} \sqrt[n]{|z_n|} > 1$, *then* $\sum_{n=1}^{+\infty} z_n$ *diverges.*
(iii) If $\overline{\lim} \sqrt[n]{|z_n|} = 1$, *then there is no general conclusion.*

Proof. (i) We consider any a such that $\overline{\lim} \sqrt[n]{|z_n|} < a < 1$. Then there is n_0 so that $\sqrt[n]{|z_n|} \leq a$ and hence $|z_n| \leq a^n$ for every $n \geq n_0$. Since $0 \leq a < 1$, the geometric series $\sum_{n=1}^{+\infty} a^n$ converges and, by comparison, $\sum_{n=1}^{+\infty} |z_n|$ also converges.
(ii) We have $\sqrt[n]{|z_n|} \geq 1$ for infinitely many n. Therefore, $|z_n| \geq 1$ for infinitely many n and hence $z_n \not\to 0$. Thus, $\sum_{n=1}^{+\infty} z_n$ diverges.
(iii) For the series $\sum_{n=1}^{+\infty} \frac{1}{n}$ and $\sum_{n=1}^{+\infty} \frac{1}{n^2}$ we have $\sqrt[n]{|1/n|} \to 1$ and $\sqrt[n]{|1/n^2|} \to 1$. The first series diverges and the second converges. $\square$

Applying the ratio test and the root test to specific series $\sum_{n=1}^{+\infty} z_n$, we find very often that the limits $\lim_{n\to+\infty} \left|\frac{z_{n+1}}{z_n}\right|$ and $\lim_{n\to+\infty} \sqrt[n]{|z_n|}$ exist. We know (and we used it in the proofs of parts (iii) of both tests) that in this case: $\underline{\lim} = \overline{\lim} = \lim$.

Example 2.1.3. To the series $\sum_{n=1}^{+\infty} \frac{z^n}{n!}$ we apply the ratio test. If $z = 0$, the series obviously converges absolutely. If $z \neq 0$, then $\left|\frac{z^{n+1}/(n+1)!}{z^n/n!}\right| = |z|/(n+1) \to 0 < 1$ and so the series converges absolutely for every z.
Now we apply the root test. We have $\sqrt[n]{|z^n/n!|} = |z|/\sqrt[n]{n!} \to 0 < 1$ and we arrive at the same conclusion as before.

Example 2.1.4. We consider $\sum_{n=1}^{+\infty} \frac{z^n}{n^2}$ and we apply the ratio test. If $z = 0$, the series obviously converges absolutely. If $z \neq 0$, then $\left|\frac{z^{n+1}/(n+1)^2}{z^n/n^2}\right| \to |z|$. Hence, if $0 < |z| < 1$, the series converges absolutely and, if $|z| > 1$, the series diverges.
Now we apply the root test. We have $\sqrt[n]{|z^n/n^2|} \to |z|$. Therefore, if $|z| < 1$, the series converges absolutely and, if $|z| > 1$, the series diverges.
If $|z| = 1$, none of the two tests applies. But we observe that $\sum_{n=1}^{+\infty} \left|\frac{z^n}{n^2}\right| = \sum_{n=1}^{+\infty} \frac{1}{n^2} < +\infty$ in this case, and $\sum_{n=1}^{+\infty} \frac{z^n}{n^2}$ converges absolutely.
Conclusion: $\sum_{n=1}^{+\infty} \frac{z^n}{n^2}$ converges absolutely if $|z| \leq 1$, and diverges if $|z| > 1$.

Lemma 2.1. *Let (a_n), (z_n) be two sequences and let $s_n = z_1 + \cdots + z_n$ for every n. Then*

$$\sum_{k=m+1}^{n} a_k z_k = \sum_{k=m+1}^{n} (a_k - a_{k+1}) s_k + a_{n+1} s_n - a_{m+1} s_m$$

for every n, m with $n > m$. This is the **summation by parts** *formula due to Abel.*

Proof. We have

$$\begin{aligned}\sum_{k=m+1}^{n} a_k z_k &= \sum_{k=m+1}^{n} a_k (s_k - s_{k-1}) = \sum_{k=m+1}^{n} a_k s_k - \sum_{k=m}^{n-1} a_{k+1} s_k \\ &= \sum_{k=m+1}^{n} (a_k - a_{k+1}) s_k + a_{n+1} s_n - a_{m+1} s_m\end{aligned}$$

and the proof is complete. □

Dirichlet test. *Let (a_n), (z_n) be two sequences and let $s_n = z_1 + \cdots + z_n$ for every n. If (a_n) is real and decreasing and $a_n \to 0$ and if (s_n) is bounded, then $\sum_{n=1}^{+\infty} a_n z_n$ converges.*

Proof. There is M so that $|s_n| \leq M$ for every n. Now, let $\epsilon > 0$. Since $a_n \to 0$, there is n_0 so that $0 \leq a_n < \frac{\epsilon}{2M+1}$ for every $n \geq n_0$. Then lemma 2.1 implies that, if $n_0 \leq m < n$,

$$\begin{aligned}\left|\sum_{k=m+1}^{n} a_k z_k\right| &\leq \sum_{k=m+1}^{n} (a_k - a_{k+1}) |s_k| + a_{n+1} |s_n| + a_{m+1} |s_m| \\ &\leq \sum_{k=m+1}^{n} (a_k - a_{k+1}) M + a_{n+1} M + a_{m+1} M = 2 a_{m+1} M < \epsilon.\end{aligned}$$

The criterion of Cauchy implies that $\sum_{n=1}^{+\infty} a_n z_n$ converges. □

Abel test. *Let (a_n), (z_n) be two sequences and let $s_n = z_1 + \cdots + z_n$ for every n. If (a_n) is real and decreasing and bounded below and if (s_n) converges, i.e. if $\sum_{n=1}^{+\infty} z_n$ converges, then $\sum_{n=1}^{+\infty} a_n z_n$ converges.*

Proof. Since (a_n) is real and decreasing and bounded below, there is a so that $a_n \to a$. We set $a'_n = a_n - a$ and then (a'_n) is real and decreasing and $a'_n \to 0$. We also have that (s_n) is bounded and so Dirichlet test implies that the series $\sum_{n=1}^{+\infty} a'_n z_n$ converges. Now, since $\sum_{n=1}^{+\infty} z_n$ also converges, we find that

$$\sum_{n=1}^{+\infty} a_n z_n = \sum_{n=1}^{+\infty} a'_n z_n + a \sum_{n=1}^{+\infty} z_n$$

and hence $\sum_{n=1}^{+\infty} a_n z_n$ converges. □

Example 2.1.5. If (a_n) is real and decreasing and $a_n \to 0$, then $\sum_{n=0}^{+\infty} a_n z^n$ converges for every z with $|z| \le 1$, $z \ne 1$.
Indeed, for the partial sums $s_n = 1 + z + z^2 + \cdots + z^n$ we have $|s_n| = \frac{|1-z^{n+1}|}{|1-z|} \le \frac{2}{|1-z|}$ and the Dirichlet test implies the convergence of $\sum_{n=0}^{+\infty} a_n z^n$.

Example 2.1.6. We consider $\sum_{n=1}^{+\infty} \frac{z^n}{n}$. As in example 2.1.4, the application of either the ratio test or the root test gives that the series converges absolutely if $|z| < 1$, and diverges if $|z| > 1$. If $|z| = 1$, none of the two tests applies. If $z = 1$, the series becomes $\sum_{n=1}^{+\infty} \frac{1}{n}$ and diverges. If $|z| = 1$, $z \ne 1$, then $\sum_{n=1}^{+\infty} \left|\frac{z^n}{n}\right| = \sum_{n=1}^{+\infty} \frac{1}{n} = +\infty$, and $\sum_{n=1}^{+\infty} \frac{z^n}{n}$ does not converge absolutely. But the series is a particular instance of the series in the previous example and hence converges if $|z| = 1$, $z \ne 1$. In general, when a series is convergent but not absolutely convergent we say that it is **conditionally convergent**.
Conclusion: $\sum_{n=1}^{+\infty} \frac{z^n}{n}$ converges absolutely if $|z| < 1$, diverges if $|z| > 1$ or $z = 1$, and converges conditionally if $|z| = 1$, $z \ne 1$.

Let $\sum_{n=0}^{+\infty} a_n$ and $\sum_{n=0}^{+\infty} b_n$ be two series. If $c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_1 + a_n b_0$ for every $n \ge 0$, then the series

$$\sum_{n=0}^{+\infty} c_n = \sum_{n=0}^{+\infty} (a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_1 + a_n b_0)$$

is called **Cauchy product** of the two series.

Proposition 2.3. *If one of the series $\sum_{n=0}^{+\infty} a_n$ and $\sum_{n=0}^{+\infty} b_n$ converges and the other converges absolutely, then their Cauchy product $\sum_{n=0}^{+\infty} c_n$ converges. Moreover, we have*

$$\sum_{n=0}^{+\infty} c_n = \sum_{n=0}^{+\infty} a_n \sum_{n=0}^{+\infty} b_n$$

for the sums of the three series.
If both $\sum_{n=0}^{+\infty} a_n$ and $\sum_{n=0}^{+\infty} b_n$ converge absolutely, then their Cauchy product $\sum_{n=0}^{+\infty} c_n$ converges absolutely.

Proof. First assume that both series converge absolutely. We have

$$|c_n| \le |a_0||b_n| + |a_1||b_{n-1}| + \cdots + |a_{n-1}||b_1| + |a_n||b_0|.$$

Hence, if $S = \sum_{n=0}^{+\infty} |a_n| < +\infty$ and $T = \sum_{n=0}^{+\infty} |b_n| < +\infty$, then

$$\sum_{n=0}^{N} |c_n| \le \sum_{n=0}^{N} \big(\sum_{k=0}^{n} |a_k||b_{n-k}| \big) = \sum_{k=0}^{N} |a_k| \big(\sum_{n=k}^{N} |b_{n-k}| \big) \le \sum_{k=0}^{N} |a_k| T \le ST$$

for every N. Thus, $\sum_{n=0}^{+\infty} |c_n| \le ST < +\infty$ and so $\sum_{n=0}^{+\infty} c_n$ converges absolutely.
Now, assume that $\sum_{n=0}^{+\infty} a_n$ converges absolutely, i.e. $S = \sum_{n=0}^{+\infty} |a_n| < +\infty$, and that $\sum_{n=0}^{+\infty} b_n$ converges and let

$$s = \sum_{n=0}^{+\infty} a_n, \quad t = \sum_{n=0}^{+\infty} b_n.$$

Moreover, let $s_n = a_0 + \cdots + a_n$, $t_n = b_0 + \cdots + b_n$ and $u_n = c_0 + \cdots + c_n$ be the partial sums of the three series and also $S_n = |a_0| + \cdots + |a_n|$. Then

$$\begin{aligned} u_N &= \sum_{n=0}^{N} c_n = \sum_{n=0}^{N} \big(\sum_{k=0}^{n} a_k b_{n-k} \big) = \sum_{k=0}^{N} a_k \big(\sum_{n=k}^{N} b_{n-k} \big) = \sum_{k=0}^{N} a_k \big(\sum_{m=0}^{N-k} b_m \big) \\ &= \sum_{k=0}^{N} a_k t_{N-k} \end{aligned}$$

and hence

$$s_N t_N - u_N = \sum_{k=0}^{N} a_k (t_N - t_{N-k}).$$

We take $p = [N/2]$ and we get

$$s_N t_N - u_N = \sum_{k=0}^{p} a_k (t_N - t_{N-k}) + \sum_{k=p+1}^{N} a_k (t_N - t_{N-k}). \tag{2.1}$$

If $0 \le k \le p$, then $N-k \ge N-p \ge p$ and hence

$$|t_N - t_{N-k}| \le \sup_{m,n\ge p} |t_m - t_n|. \tag{2.2}$$

If $p+1 \le k \le N$, then

$$|t_N - t_{N-k}| \le |t_N| + |t_{N-k}| \le 2\sup_{m\ge 1} |t_m| < +\infty. \tag{2.3}$$

Now, (2.1), (2.2) and (2.3) imply

$$\begin{aligned} |s_N t_N - u_N| &\le \textstyle\sum_{k=0}^{p} |a_k||t_N - t_{N-k}| + \sum_{k=p+1}^{N} |a_k||t_N - t_{N-k}| \\ &\le \textstyle\sup_{m,n\ge p} |t_m - t_n| \sum_{k=0}^{p} |a_k| + 2\sup_{m\ge 1} |t_m| \sum_{k=p+1}^{N} |a_k| \\ &\le \textstyle\sup_{m,n\ge p} |t_m - t_n|\, S + 2\sup_{m\ge 1} |t_m|\, (S - S_p). \end{aligned}$$

Now, $N \to +\infty$ implies $p \to +\infty$. Hence $S_p \to S$ and $\sup_{m,n\ge p} |t_m - t_n| \to 0$ by the Cauchy criterion for the convergent sequence (t_n). Therefore, $s_N t_N - u_N \to 0$ when $N \to +\infty$. We also have that $s_N \to s$ and $t_N \to t$ and we conclude that $u_N \to st$. □

Exercises.

2.1.1. Which of the following series converge?

$\sum_{n=1}^{+\infty}(\frac{1}{n} + \frac{i}{n^2})$, $\sum_{n=1}^{+\infty}(\frac{n}{2^n} + \frac{i}{n^3})$, $\sum_{n=1}^{+\infty} \frac{1+i^n}{n^2}$, $\sum_{n=1}^{+\infty} \frac{1}{2+i^n}$, $\sum_{n=1}^{+\infty} \frac{1}{n+i}$, $\sum_{n=1}^{+\infty} \frac{1}{n^2+in}$.

2.1.2. Find the sum of the series $\sum_{n=1}^{+\infty} n(-1)^{n-1}$ if we consider it as a complex series and also if we consider it as a real series.

2.1.3. (i) Apply the ratio test whenever possible:

$$\textstyle\sum_{n=1}^{+\infty} n^3 i^n,\quad \sum_{n=1}^{+\infty} \frac{n!}{i^n},\quad \sum_{n=1}^{+\infty} \frac{(1+i)^n}{n!},\quad \sum_{n=1}^{+\infty} \frac{(2i)^n n!}{n^n},\quad \sum_{n=1}^{+\infty} \frac{(2+i)^n n!}{n^n},$$

$$\textstyle\sum_{n=1}^{+\infty} \frac{e^n n!}{n^n},\quad \sum_{n=1}^{+\infty} \frac{(n!)^2}{(2n)!},\quad \sum_{n=1}^{+\infty} \frac{(4i)^n (n!)^2}{(2n)!},\quad \sum_{n=1}^{+\infty} \frac{(3+i)(6+i)(9+i)\cdots(3n+i)}{(3+4i)(3+8i)(3+12i)\cdots(3+4ni)}.$$

(ii) Apply the root test whenever possible:

$$\textstyle\sum_{n=1}^{+\infty} n^n i^n,\quad \sum_{n=1}^{+\infty}(\frac{n+i}{2n-i})^n,\quad \sum_{n=1}^{+\infty}(\frac{n+i}{n-i})^{2n},\quad \sum_{n=1}^{+\infty} \frac{n^3}{(1+2i)^n},$$

$$\textstyle\sum_{n=1}^{+\infty} n^3(1-i)^n,\quad \sum_{n=1}^{+\infty} \frac{(2+3i)^n}{n^n},\quad \sum_{n=1}^{+\infty} \frac{n+i}{(\sqrt[n]{n}+i)^n}.$$

2.1.4. If $\sum_{n=1}^{+\infty} |z_n| < +\infty$, prove that $\sum_{n=1}^{+\infty} z_n(\cos n\theta + i\sin n\theta)$ converges.

2.1.5. Let $z_n = x_n + iy_n$ for all n. Prove that $\sum_{n=1}^{+\infty} z_n$ converges absolutely if and only if $\sum_{n=1}^{+\infty} x_n$, $\sum_{n=1}^{+\infty} y_n$ converge absolutely.

2.1.6. Let $|a_n| r^n \le M n^k$ for all n. Prove that $\sum_{n=1}^{+\infty} a_n z^n$ converges for every z with $|z| < r$.

2.1.7. Find all z for which $\sum_{n=1}^{+\infty} \frac{z^n}{2+z^n}$ converges.

2.1.8. Let $|\operatorname{Arg} z_n| \le \theta_0 < \frac{\pi}{2}$ for every n. Prove that $\sum_{n=1}^{+\infty} z_n$ converges if and only if it converges absolutely. Prove that $\sum_{n=1}^{+\infty} z_n = \infty$ if and only if $\sum_{n=1}^{+\infty} |z_n| = +\infty$.

2.1.9. Find a series $\sum_{n=1}^{+\infty} z_n$ which converges and is such that $\sum_{n=1}^{+\infty} z_n^2$ diverges.

2.1.10. Check for every z the conditional convergence and the absolute convergence of the series:

$$\textstyle\sum_{n=1}^{+\infty} \frac{z^n}{n},\quad \sum_{n=2}^{+\infty} \frac{z^n}{n\log n},\quad \sum_{n=2}^{+\infty} \frac{z^n}{n\log^2 n},\quad \sum_{n=1}^{+\infty} z^n \sin\frac{1}{n},\quad \sum_{n=1}^{+\infty} z^n(1-\cos\frac{1}{n}).$$

2.1.11. For each $a > 0$ find all z for which the series $\sum_{n=1}^{+\infty} \frac{z^n}{n^a}$ converges.

2.1.12. (i) Let $s_n = z_1 + \cdots + z_n$ for all n. If $(a_{n+1}s_n)$ converges and if $\sum_{n=1}^{+\infty}(a_n - a_{n+1})s_n$ converges, prove that $\sum_{n=1}^{+\infty} a_n z_n$ converges. In particular: if (s_n) is bounded, if $a_n \to 0$ and if $\sum_{n=1}^{+\infty} |a_n - a_{n+1}| < +\infty$, prove that $\sum_{n=1}^{+\infty} a_n z_n$ converges.
What is the relation of all these with the tests of Dirichlet and Abel?
(ii) If the sequence (a_n) satisfies $\sum_{n=1}^{+\infty} |a_{n+1} - a_n| < +\infty$, we say that it is **of bounded variation**.
Prove that every sequence of bounded variation converges.
Prove that the set of all sequences of bounded variation is a linear space (over $\mathbb{C}$).
Prove that every *real* sequence which is monotone and bounded is of bounded variation.
For every $a \in \mathbb{R}$ we define $a_+ = (|a| + a)/2$ and $a_- = (|a| - a)/2$. Observe that $0 \leq a_+ \leq |a|$, $0 \leq a_- \leq |a|$, $|a| = a_+ + a_-$ and $a = a_+ - a_-$.
If (a_n) is a *real* sequence, then $\sum_{n=1}^{+\infty} |a_{n+1} - a_n| < +\infty$ implies $\sum_{n=1}^{+\infty}(a_{n+1} - a_n)_+ < +\infty$ and $\sum_{n=1}^{+\infty}(a_{n+1} - a_n)_- < +\infty$. Using this, prove that every *real* sequence of bounded variation with limit 0 is the difference of two decreasing sequences with limit 0.

2.1.13. Prove that the series $\sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{\sqrt{n}}$ converges but that the Cauchy product of this series with itself does not converge.

2.2 Curvilinear integrals.

We shall first extend the notion of integral of a *real* function over an interval to the notion of integral of a *complex* function over an interval.

Let f be a complex function defined in the interval $[a, b]$ and let $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$ be the real and imaginary parts of f. We say that f is **(Riemann) integrable** over $[a, b]$ if u, v are both (Riemann) integrable over $[a, b]$ and in this case we define the **(Riemann) integral** of f over $[a, b]$ to be

$$\int_a^b f(t)\, dt = \int_a^b u(t)\, dt + i \int_a^b v(t)\, dt. \tag{2.4}$$

Since the numbers $\int_a^b u(t)\, dt$ and $\int_a^b v(t)\, dt$ are real, we see that

$$\operatorname{Re} \int_a^b f(t)\, dt = \int_a^b \operatorname{Re} f(t)\, dt, \qquad \operatorname{Im} \int_a^b f(t)\, dt = \int_a^b \operatorname{Im} f(t)\, dt.$$

Now let us take any subdivision $\Delta = \{t_0, \ldots, t_n\}$ of $[a, b]$ and any choice $\Xi = \{\xi_1, \ldots, \xi_n\}$ of intermediate points $\xi_k \in [t_{k-1}, t_k]$ and the corresponding Riemann sum $\sum_{k=1}^n f(\xi_k)(t_k - t_{k-1})$. If $w(\Delta) = \max_{1 \leq k \leq n}(t_k - t_{k-1})$ is the width of the subdivision Δ, then we know that

$$\lim_{w(\Delta) \to 0} \sum_{k=1}^n u(\xi_k)(t_k - t_{k-1}) = \int_a^b u(t)\, dt$$

$$\lim_{w(\Delta) \to 0} \sum_{k=1}^n v(\xi_k)(t_k - t_{k-1}) = \int_a^b v(t)\, dt.$$

Multiplying the second relation with i, adding and using (2.4), we find

$$\lim_{w(\Delta) \to 0} \sum_{k=1}^n f(\xi_k)(t_k - t_{k-1}) = \int_a^b f(t)\, dt.$$

Example 2.2.1. If f is piecewise continuous in $[a, b]$, then $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$ are also piecewise continuous in $[a, b]$. Hence u, v are integrable, and so f is also integrable over $[a, b]$.

The following propositions are analogous to similar well known propositions about integrals of real functions and can be proved easily by the reader. One should decompose every complex function into its real and imaginary parts and use the analogous properties for real functions together with (2.4).

Proposition 2.4. *Let f_1, f_2 be integrable over $[a,b]$ and $\lambda_1, \lambda_2 \in \mathbb{C}$. Then $\lambda_1 f_1 + \lambda_2 f_2$ is integrable over $[a,b]$ and*

$$\int_a^b (\lambda_1 f_1(t) + \lambda_2 f_2(t))\, dt = \lambda_1 \int_a^b f_1(t)\, dt + \lambda_2 \int_a^b f_2(t)\, dt.$$

Proposition 2.5. *Let $a < b < c$. If f is integrable over $[a,b]$ and over $[b,c]$, then f is integrable over $[a,c]$ and*

$$\int_a^c f(t)\, dt = \int_a^b f(t)\, dt + \int_b^c f(t)\, dt.$$

Proposition 2.6. *If f_1, f_2 are integrable over $[a,b]$, then $f_1 f_2$ is integrable over $[a,b]$.*

The proof of the next proposition is not entirely trivial.

Proposition 2.7. *If f is integrable over $[a,b]$, then $|f|$ is integrable over $[a,b]$ and*

$$\left| \int_a^b f(t)\, dt \right| \leq \int_a^b |f(t)|\, dt.$$

Proof. Let $u = \operatorname{Re} f$, $v = \operatorname{Im} f$. Then u, v are integrable over $[a,b]$ hence $|f| = \sqrt{u^2 + v^2}$ is integrable over $[a,b]$. Now we have two cases.
If $\int_a^b f(t)\, dt = 0$, then $\left| \int_a^b f(t)\, dt \right| \leq \int_a^b |f(t)|\, dt$ is clearly true.
If $\int_a^b f(t)\, dt \neq 0$, then we take any element θ of the argument of the number $\int_a^b f(t)\, dt$ and we set $z = \cos\theta + i \sin\theta$. Now,

$$\left| \int_a^b f(t)\, dt \right| = \overline{z} \int_a^b f(t)\, dt = \int_a^b (\overline{z}\, f(t))\, dt.$$

The left side of this equality is real and hence its right side is also real and thus equal to its real part! Hence

$$\left| \int_a^b f(t)\, dt \right| = \operatorname{Re} \int_a^b (\overline{z}\, f(t))\, dt = \int_a^b \operatorname{Re}(\overline{z}\, f(t))\, dt \leq \int_a^b |\overline{z}\, f(t)|\, dt = \int_a^b |f(t)|\, dt$$

since $\operatorname{Re}(\overline{z}\, f(t)) \leq |\overline{z}\, f(t)|$ for every $t \in [a,b]$. ☐

We recall that every *continuous* complex function $\gamma : [a,b] \to \mathbb{C}$, where $[a,b]$ is any interval, is called *curve* in the complex plane.

The set of the values of a curve γ, i.e. the set $\gamma^* = \{\gamma(t) \,|\, t \in [a,b]\} \subseteq \mathbb{C}$ is the *trajectory* of the curve and it is a compact and connected subset of $\mathbb{C}$, since γ is continuous and $[a,b]$ is compact and connected. The points $\gamma(a)$ and $\gamma(b)$ are the endpoints, the initial and the final endpoint, respectively, of the curve.

The variable $t \in [a,b]$ is the *parameter* and $[a,b]$ is the *parametric interval* of the curve. When the parameter t increases in $[a,b]$, the variable point $\gamma(t)$ moves on the trajectory γ^* in a definite direction (from the initial to the final endpoint) which is the so-called *direction* of the curve. To be more precise, the sense of *direction* is understood as follows: if $a \leq t_1 < t_2 \leq b$, then we say that $\gamma(t_1)$ is *before* $\gamma(t_2)$ and that $\gamma(t_2)$ is *after* $\gamma(t_1)$ in the trajectory.

Finally,

$$z = \gamma(t), \qquad t \in [a,b],$$

is the so-called *parametric equation* of the curve γ.

If the endpoints of the curve γ coincide, i.e. $\gamma(a) = \gamma(b)$, then we say that the curve is *closed*.

If $\gamma(t) \in A$ for all $t \in [a,b]$, i.e. if $\gamma^* \subseteq A$, then we say that *the curve is in A*.

The term *curve* for the continuous function γ is justified by the fact that the shape of its trajectory γ^* is, usually, what in everyday language we call *curve in the plane*. Sometimes we use the term *curve* for the trajectory γ^* even though this is not typically correct. The reason is that there are cases of different curves γ_1, γ_2 with the same trajectory $\gamma_1^* = \gamma_2^*$.

Example 2.2.2. If $z_0, z_1 \in \mathbb{C}$, then the parametric equation

$$z = \gamma(t) = \tfrac{t-a}{b-a} z_1 + \tfrac{b-t}{b-a} z_0, \qquad t \in [a, b],$$

defines a curve γ whose trajectory γ^* is the linear segment $[z_0, z_1]$. Its initial and final endpoints are z_0 and z_1, respectively, and its direction is from z_0 to z_1. The *same* linear segment $[z_0, z_1]$ is the trajectory of another curve γ with parametric equation

$$z = \gamma(t) = t z_1 + (1 - t) z_0, \qquad t \in [0, 1].$$

Example 2.2.3. If $r > 0$, then the parametric equation

$$z = \gamma(t) = z_0 + r(\cos t + i \sin t), \qquad t \in [0, 2\pi],$$

defines a closed curve γ whose trajectory γ^* is the circle $C_{z_0}(r)$. The direction of this curve is the so-called *positive direction of rotation* around z_0: the counterclockwise rotation.
If we consider the curve γ with parametric equation $z = \gamma(t) = z_0 + r(\cos(2t) + i \sin(2t))$, $t \in [0, 2\pi]$, then we get a different curve. But the trajectories of the two curves coincide: the circle $C_{z_0}(r)$. The direction of the two curves is the same: the positive direction of rotation around z_0. But the first curve goes around z_0 only once, while the second curve goes around z_0 twice.

Let $\gamma : [a, b] \to \mathbb{C}$ be a curve and let $x = \operatorname{Re} \gamma$ and $y = \operatorname{Im} \gamma$ be the real and imaginary parts of γ, i.e. $\gamma(t) = x(t) + iy(t) = (x(t), y(t))$ for $t \in [a, b]$. If γ is differentiable at $t_0 \in [a, b]$ or, equivalently, if x, y are differentiable at t_0, then $\gamma'(t_0) = x'(t_0) + iy'(t_0) = (x'(t_0), y'(t_0))$ is the *tangent vector* of the trajectory γ^* at its point $\gamma(t_0)$. If $\gamma'(t_0) \neq 0$, then the vector $\gamma'(t_0)$ determines the *tangent line* of the trajectory γ^* at its point $\gamma(t_0)$ and the direction of $\gamma'(t_0)$ is the same as the direction of the curve. Strictly speaking, at its endpoints, $\gamma(a), \gamma(b)$, the curve can only have *tangent halflines*; not tangent lines. If $t_0 = a$ and $\gamma'(a) \neq 0$, then the vector $\gamma'(a)$ determines the *tangent halfline* of the trajectory at the endpoint $\gamma(a)$ with direction coinciding with the direction of the curve. If $t_0 = b$ and $\gamma'(b) \neq 0$, then the vector $-\gamma'(b)$ determines the *tangent halfline* of the trajectory at the endpoint $\gamma(b)$ with direction opposite to the direction of the curve. If at some $t_0 \in (a, b)$ the one-sided derivatives $\gamma'_-(t_0) \neq 0$ and $\gamma'_+(t_0) \neq 0$ exist but they are not equal, then the tangent halflines of the trajectory at its point $\gamma(t_0)$ may not be opposite and so there may be no tangent line of the trajectory at this point: one of the halflines is determined by $\gamma'_+(t_0)$ and the other by $-\gamma'_-(t_0)$.

We know that, if the curve $\gamma : [a, b] \to \mathbb{C}$ is *continuously differentiable* or *smooth*, i.e. if $\gamma' : [a, b] \to \mathbb{C}$ is continuous in $[a, b]$, then the *length* of the curve, denoted $l(\gamma)$, is equal to

$$l(\gamma) = \int_a^b |\gamma'(t)| \, dt. \tag{2.5}$$

Example 2.2.4. The curve γ with parametric equation $z = \gamma(t) = \frac{b-t}{b-a} z_0 + \frac{t-a}{b-a} z_1$, $t \in [a, b]$, has length

$$l(\gamma) = \int_a^b |\gamma'(t)| \, dt = \int_a^b \left| \tfrac{z_1 - z_0}{b-a} \right| dt = \left| \tfrac{z_1 - z_0}{b-a} \right| \int_a^b \, dt = |z_1 - z_0|.$$

Example 2.2.5. If $r > 0$ the curve γ with parametric equation $z = \gamma(t) = z_0 + r(\cos t + i \sin t)$, $t \in [0, 2\pi]$, has length

$$l(\gamma) = \int_a^b |\gamma'(t)| \, dt = \int_0^{2\pi} |r(-\sin t + i \cos t)| \, dt = \int_0^{2\pi} r \, dt = 2\pi r.$$

The same formula (2.5) gives the length of the curve γ if this is *piecewise continuously differentiable* or *piecewise smooth*, i.e. when there is a subdivision $a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b$ of the parametric interval $[a, b]$ so that the restriction of γ in every $[t_{k-1}, t_k]$ is continuously differentiable. (Strictly speaking, at the division points t_k the derivative of γ may not exist; the two one-sided derivatives should exist and be finite at these points.)

Now let $\gamma_1 : [a,b] \to \mathbb{C}$ be a curve. We consider any $\sigma : [c,d] \to [a,b]$ which is one-to-one in the interval $[c,d]$ and onto $[a,b]$, has continuous derivative in $[c,d]$ and has $\sigma'(s) > 0$ for every $s \in [c,d]$. Thus, σ is strictly increasing in $[c,d]$ and $\sigma(c) = a$, $\sigma(d) = b$. Every such σ is called *change of parameter*. Then $\gamma_2 = \gamma_1 \circ \sigma : [c,d] \to \mathbb{C}$ is continuous in $[c,d]$ and hence it is a new curve. We say that γ_2 is a *reparametrization* of γ_1: the parameter of γ_1 is $t \in [a,b]$ and the parameter of γ_2 is $s \in [c,d]$. The curves γ_1, γ_2 have the same trajectory, the same endpoints and the same direction. Since σ' is continuous and > 0, the two curves are simultaneously (piecewise) smooth and, in this case, their lengths are equal:

$$\begin{aligned} l(\gamma_2) &= \textstyle\int_c^d |\gamma_2'(s)|\, ds = \int_c^d |\gamma_1'(\sigma(s))||\sigma'(s)|\, ds = \int_c^d |\gamma_1'(\sigma(s))|\sigma'(s)\, ds \\ &= \textstyle\int_a^b |\gamma_1'(t)|\, dt = l(\gamma_1). \end{aligned}$$

We may define the following relation between curves: $\gamma_1 \sim \gamma_2$ if γ_2 is a reparametrization of γ_1. It is not difficult to prove that this relation between curves is an *equivalence relation*, i.e. it satisfies the following three properties:

(i) $\gamma \sim \gamma$.
(ii) $\gamma_1 \sim \gamma_2 \;\Rightarrow\; \gamma_2 \sim \gamma_1$.
(iii) $\gamma_1 \sim \gamma_2,\ \gamma_2 \sim \gamma_3 \;\Rightarrow\; \gamma_1 \sim \gamma_3$.

Indeed: (i) Let $\gamma : [a,b] \to \mathbb{C}$ be any curve. We consider the change of parameter $id : [a,b] \to [a,b]$, defined by $id(t) = t$, and then $\gamma = \gamma \circ id : [a,b] \to \mathbb{C}$. Thus, $\gamma \sim \gamma$. (ii) Let $\gamma_1 \sim \gamma_2$. Then $\gamma_2 = \gamma_1 \circ \sigma$ where $\sigma : [c,d] \to [a,b]$ is a change of parameter. But then $\sigma^{-1} : [a,b] \to [c,d]$ is also a change of parameter and $\gamma_1 = \gamma_2 \circ \sigma^{-1}$. Therefore $\gamma_2 \sim \gamma_1$. (iii) Let $\gamma_1 \sim \gamma_2$ and $\gamma_2 \sim \gamma_3$. Then $\gamma_2 = \gamma_1 \circ \sigma$ and $\gamma_3 = \gamma_2 \circ \tau$, where $\sigma : [c,d] \to [a,b]$ and $\tau : [e,f] \to [c,d]$ are changes of parameter. But then $\chi = \sigma \circ \tau : [e,f] \to [a,b]$ is a change of parameter and $\gamma_3 = \gamma_2 \circ \tau = (\gamma_1 \circ \sigma) \circ \tau = \gamma_1 \circ \chi$. Therefore $\gamma_1 \sim \gamma_3$.

It is of some value to note that if we have a curve γ with parametric interval $[a,b]$ and if we are given an arbitrary interval $[c,d]$, then there is a reparametrization of γ with parametric interval $[c,d]$ instead of $[a,b]$. We can do this if we can find an appropriate change of parameter $\sigma : [c,d] \to [a,b]$. There are many such σ, but a simple one is

$$t = \sigma(s) = \tfrac{d-s}{d-c}\, a + \tfrac{s-c}{d-c}\, b, \qquad s \in [c,d].$$

Therefore, if for some reason (and we shall presently see that there is such a reason) we do not want to distinguish between curves which are reparametrizations of each other, then the parametric interval of a curve is of no particular importance: we may consider a reparametrization of a given curve changing the given parametric interval to any other which we might prefer.

For every curve $\gamma : [a,b] \to \mathbb{C}$ we consider the curve $\neg\,\gamma : [a,b] \to \mathbb{C}$ given by

$$(\neg\,\gamma)(t) = \gamma(a+b-t), \qquad t \in [a,b].$$

Then $\neg\,\gamma$ is called *opposite* of γ. The curves γ and $\neg\,\gamma$ have the same trajectory but opposite directions. Also, the two curves are simultaneously (piecewise) smooth and, in this case, their lengths are equal:

$$l(\neg\,\gamma) = \textstyle\int_a^b |(\neg\,\gamma)'(t)|\, dt = \int_a^b |\gamma'(a+b-t)|\, dt = -\int_b^a |\gamma'(s)|\, ds = \int_a^b |\gamma'(s)|\, ds = l(\gamma).$$

If the curves $\gamma_1 : [a,b] \to \mathbb{C}$ and $\gamma_2 : [b,c] \to \mathbb{C}$ have $\gamma_1(b) = \gamma_2(b)$, then we say that γ_1, γ_2 (in this order) are *successive* and then we may define the curve $\gamma_1 \dot{+} \gamma_2 : [a,c] \to \mathbb{C}$ by

$$(\gamma_1 \dot{+} \gamma_2)(t) = \begin{cases} \gamma_1(t), & \text{if } a \le t \le b \\ \gamma_2(t), & \text{if } b \le t \le c \end{cases}$$

The curve $\gamma_1 \dot{+} \gamma_2$ is called *sum* of γ_1 and γ_2. If γ_1 and γ_2 are (piecewise) smooth, $\gamma_1 \dot{+} \gamma_2$ is also piecewise smooth. The trajectory $(\gamma_1 \dot{+} \gamma_2)^*$ is the union of the trajectories γ_1^* and γ_2^*.

Of course, the sum of two curves can be generalized to the sum of more than two curves provided that these are successive.

Example 2.2.6. Every polygonal line can be considered as the trajectory of a piecewise smooth curve. This curve is the sum of successive curves each of which has as its trajectory a corresponding linear segment of the polygonal line.

Through the operation of summation of successive curves, we may consider successive curves as one curve and, conversely, we may consider one curve as a sum of successive curves.

The length of the sum of successive piecewise smooth curves equals the sum of their lengths:

$$\begin{aligned} l(\gamma_1 \dot{+} \gamma_2) &= \int_a^c |(\gamma_1 \dot{+} \gamma_2)'(t)|dt = \int_a^b |(\gamma_1 \dot{+} \gamma_2)'(t)|dt + \int_b^c |(\gamma_1 \dot{+} \gamma_2)'(t)|dt \\ &= \int_a^b |\gamma_1'(t)|dt + \int_b^c |\gamma_2'(t)|dt = l(\gamma_1) + l(\gamma_2). \end{aligned}$$

Now we shall extend the notion of integral of a complex function over an *interval* to the notion of integral of a complex function over a *curve*. Let $\gamma : [a, b] \to \mathbb{C}$ be a piecewise smooth curve and let $f : \gamma^* \to \mathbb{C}$ be continuous in the trajectory $\gamma^* = \{\gamma(t) \,|\, t \in [a, b]\}$. Then $f \circ \gamma : [a, b] \to \mathbb{C}$ is continuous in $[a, b]$. Thus, $(f \circ \gamma)\gamma'$ is piecewise continuous in $[a, b]$ and hence integrable over $[a, b]$. We define the **curvilinear integral** of f over γ by

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

We shall usually write

$$\oint_\gamma f(z)\, dz$$

when γ is closed.

We remark that *whenever a curve γ is mentioned with respect either to its length $l(\gamma)$ or to a curvilinear integral of a function over γ we shall always assume that γ is piecewise smooth.*

Example 2.2.7. Let γ be the curve with parametric equation $z = \gamma(t) = (1-t)z_0 + tz_1$, $t \in [0, 1]$. The trajectory of γ is the linear segment $[z_0, z_1]$ having direction from z_0 to z_1. If f is continuous in $[z_0, z_1]$, then the curvilinear integral $\int_\gamma f(z)\, dz$ is denoted $\int_{[z_0,z_1]} f(z)\, dz$. I.e.

$$\int_{[z_0,z_1]} f(z)\, dz = \int_\gamma f(z)\, dz = (z_1 - z_0) \int_0^1 f((1-t)z_0 + tz_1)\, dt.$$

This is the *curvilinear integral of f over the linear segment $[z_0, z_1]$ from z_0 to z_1.*

Example 2.2.8. Let $r > 0$ and γ be the curve with parametric equation $z = \gamma(t) = z_0 + r(\cos t + i \sin t)$, $t \in [0, 2\pi]$. The trajectory of γ is the circle $C_{z_0}(r)$ with the positive direction of rotation around z_0. If f is continuous in the circle $C_{z_0}(r)$, then the curvilinear integral $\oint_\gamma f(z)\, dz$ is denoted $\oint_{C_{z_0}(r)} f(z)\, dz$. I.e.

$$\oint_{C_{z_0}(r)} f(z)\, dz = \oint_\gamma f(z)\, dz = \int_0^{2\pi} f\big(z_0 + r(\cos t + i \sin t)\big) r(-\sin t + i \cos t)\, dt.$$

This is the *curvilinear integral of f over the circle $C_{z_0}(r)$ with the positive direction of rotation.*

An important concrete instance of the previous example is the following.

Example 2.2.9. If $n \in \mathbb{Z}$, we know that $\int_0^{2\pi} \sin(nt)\, dt = 0$. Also, $\int_0^{2\pi} \cos(nt)\, dt = 2\pi$, if $n = 0$, and $\int_0^{2\pi} \cos(nt)\, dt = 0$, if $n \neq 0$. Therefore, if $n \in \mathbb{Z}$, we get

$$\begin{aligned} \oint_{C_{z_0}(r)} (z - z_0)^n\, dz &= \int_0^{2\pi} r^n (\cos t + i \sin t)^n r(-\sin t + i \cos t)\, dt \\ &= i r^{n+1} \int_0^{2\pi} (\cos t + i \sin t)^n (\cos t + i \sin t)\, dt \\ &= i r^{n+1} \int_0^{2\pi} \big(\cos((n+1)t) + i \sin((n+1)t) \big)\, dt \\ &= \begin{cases} 2\pi i, & \text{if } n = -1 \\ 0, & \text{if } n \neq -1 \end{cases} \end{aligned}$$

The following propositions are easy to prove.

Proposition 2.8. *If γ is a piecewise smooth curve, f_1, f_2 are continuous in γ^* and $\lambda_1, \lambda_2 \in \mathbb{C}$, then*

$$\int_\gamma (\lambda_1 f_1(z) + \lambda_2 f_2(z))\, dz = \lambda_1 \int_\gamma f_1(z)\, dz + \lambda_2 \int_\gamma f_2(z)\, dz.$$

Proof. An application of proposition 2.4 and of the definition of the curvilinear integral. □

We recall the notation for the *uniform norm* in A

$$\|f\|_A = \sup_{a\in A} |f(a)| = \sup\{|f(a)| \mid a \in A\}$$

of a bounded complex function $f : A \to \mathbb{C}$ defined in a nonempty set A.

Proposition 2.9. *If γ is a piecewise smooth curve and f is continuous in γ^*, then*

$$\big| \int_\gamma f(z)\, dz \big| \leq \sup_{z\in\gamma^*} |f(z)| l(\gamma) = \|f\|_{\gamma^*} l(\gamma).$$

Proof. If $\gamma : [a, b] \to \mathbb{C}$, then

$$\begin{aligned} \big| \int_\gamma f(z)\, dz \big| &= \big| \int_a^b f(\gamma(t))\gamma'(t)\, dt \big| \leq \int_a^b |f(\gamma(t))||\gamma'(t)|\, dt \leq \sup_{z\in\gamma^*} |f(z)| \int_a^b |\gamma'(t)|\, dt \\ &= \sup_{z\in\gamma^*} |f(z)| l(\gamma) = \|f\|_{\gamma^*} l(\gamma). \end{aligned}$$

The first inequality uses proposition 2.7. □

Proposition 2.10. *If γ is a piecewise smooth curve, f_n, ϕ are continuous in γ^* and $f_n \to f$ uniformly in γ^*, then*

$$\int_\gamma f_n(z)\phi(z)\, dz \to \int_\gamma f(z)\phi(z)\, dz.$$

Proof. Because of uniform convergence, f is continuous in γ^*. Therefore, the existence of the integrals $\int_\gamma f_n(z)\phi(z)\, dz$ and $\int_\gamma f(z)\phi(z)\, dz$ is assured. Now, proposition 2.9 implies

$$\begin{aligned} \big| \int_\gamma f_n(z)\phi(z)\, dz - \int_\gamma f(z)\phi(z)\, dz \big| &= \big| \int_\gamma (f_n(z) - f(z))\phi(z)\, dz \big| \\ &\leq \|(f_n - f)\phi\|_{\gamma^*} l(\gamma) \leq \|f_n - f\|_{\gamma^*} \|\phi\|_{\gamma^*} l(\gamma). \end{aligned}$$

Since $\|f_n - f\|_{\gamma^*} \to 0$, we get that $\int_\gamma f_n(z)\phi(z)\, dz \to \int_\gamma f(z)\phi(z)\, dz$. □

We may rewrite the result of proposition 2.10 in the form

$$\lim_{n\to+\infty} \int_\gamma f_n(z)\phi(z)\, dz = \int_\gamma \lim_{n\to+\infty} f_n(z)\phi(z)\, dz$$

of an interchange between the symbols $\lim_{n\to+\infty}$ and $\int_\gamma$. This interchange under the assumption of uniform convergence is the content of proposition 2.10.

Proposition 2.11. *If γ is a piecewise smooth curve, f_n, ϕ are continuous in γ^* and $\sum_{n=1}^{+\infty} f_n = s$ uniformly in γ^*, then*

$$\sum_{n=1}^{+\infty} \int_\gamma f_n(z)\phi(z)\, dz = \int_\gamma s(z)\phi(z)\, dz.$$

Proof. We consider the partial sums $s_n = f_1 + \cdots + f_n$ and we apply proposition 2.10 to them. Then

$$\sum_{k=1}^n \int_\gamma f_k(z)\phi(z)\, dz = \int_\gamma \sum_{k=1}^n f_k(z)\phi(z)\, dz = \int_\gamma s_n(z)\phi(z)\, dz \to \int_\gamma s(z)\phi(z)\, dz.$$

I.e. the series (of numbers) $\sum_{n=1}^{+\infty} \int_\gamma f_n(z)\phi(z)\, dz$ converges to (the number) $\int_\gamma s(z)\phi(z)\, dz$. □

As in the case of proposition 2.10, we may rewrite the result of proposition 2.11 in the form

$$\sum_{n=1}^{+\infty} \int_\gamma f_n(z)\phi(z)\,dz = \int_\gamma \sum_{n=1}^{+\infty} f_n(z)\phi(z)\,dz,$$

since $\sum_{n=1}^{+\infty} f_n(z) = s(z)$ for every $z \in \gamma^*$. Again, this interchange between the symbols $\sum_{n=1}^{+\infty}$ and $\int_\gamma$ under the assumption of uniform convergence is the content of proposition 2.11.

Proposition 2.12. *If each of the piecewise smooth curves γ_1, γ_2 is a reparametrization of the other and f is continuous in $\gamma_1^* = \gamma_2^*$, then*

$$\int_{\gamma_2} f(z)\,dz = \int_{\gamma_1} f(z)\,dz.$$

Proof. If $\gamma_1 : [a,b] \to \mathbb{C}$ and $\gamma_2 : [c,d] \to \mathbb{C}$, then there is a change of parameter $\sigma : [c,d] \to [a,b]$ so that $\gamma_2(s) = \gamma_1(\sigma(s))$ for all $s \in [c,d]$. Then

$$\begin{aligned}\int_{\gamma_2} f(z)\,dz &= \int_c^d f(\gamma_2(s))\gamma_2'(s)\,ds = \int_c^d f(\gamma_1(\sigma(s)))\gamma_1'(\sigma(s))\sigma'(s)\,ds \\ &= \int_a^b f(\gamma_1(t))\gamma_1'(t)\,dt = \int_{\gamma_1} f(z)\,dz\end{aligned}$$

after a change of parameter in the third integral. $\square$

At this point we observe that replacing a curve γ_1 with a reparametrization γ_2 of it does not alter certain objects related to the curve: its trajectory, its endpoints, its direction, its length, the number of times it covers its trajectory and, more important, the curvilinear integrals of continuous functions defined over its trajectory. Since in this course we shall use curves mostly to examine curvilinear integrals, we conclude that there is no reason to distinguish between a curve and its reparametrizations. Therefore, when we have a geometric object C which we would call, in everyday language, *curve in the plane*, e.g. a linear segment or a circle or a polygonal line, and a continuous function $f : C \to \mathbb{C}$, we can give a meaning to

$$\int_C f(z)\,dz$$

by specifying a piecewise continuously differentiable $\gamma : [a,b] \to \mathbb{C}$, i.e. a piecewise smooth curve, with trajectory γ^* coinciding with C, with endpoints coinciding with the endpoints of C and a specific assigned direction. The use of different curves, which are reparametrizations of the particular γ we have chosen, will not alter the value of the integral. In fact we have already seen two examples of this situation. One is the curvilinear integral $\int_{[z_0,z_1]} f(z)\,dz$ for which we use any parametric equation with trajectory equal to the linear segment $[z_0,z_1]$ and direction from z_0 to z_1. The simplest such parametric equation is $z = \gamma(t) = (1-t)z_0 + tz_1$, $t \in [0,1]$. The second example is the curvilinear integral $\oint_{C_{z_0}(r)} f(z)\,dz$ for which we use any parametric equation with trajectory equal to the circle $C_{z_0}(r)$ and which covers this circle once and in the positive direction of rotation around z_0. The simplest such parametric equation is $z = \gamma(t) = z_0 + r(\cos t + i\sin t)$, $t \in [0,2\pi]$.

Proposition 2.13. *Let γ_1, γ_2 be two successive piecewise smooth curves and let f be continuous in $\gamma_1^* \cup \gamma_2^*$. Then*

$$\int_{\gamma_1 \dot{+} \gamma_2} f(z)\,dz = \int_{\gamma_1} f(z)\,dz + \int_{\gamma_2} f(z)\,dz.$$

Proof. Let $\gamma_1 : [a,b] \to \mathbb{C}$ and $\gamma_2 : [b,c] \to \mathbb{C}$ with $\gamma_1(b) = \gamma_2(b)$. Then

$$\begin{aligned}\int_{\gamma_1 \dot{+} \gamma_2} f(z)\,dz &= \int_a^c f\big((\gamma_1 \dot{+} \gamma_2)(t)\big)(\gamma_1 \dot{+} \gamma_2)'(t)\,dt \\ &= \int_a^b f(\gamma_1(t))\gamma_1'(t)\,dt + \int_b^c f(\gamma_2(t))\gamma_2'(t)\,dt = \int_{\gamma_1} f(z)\,dz + \int_{\gamma_2} f(z)\,dz.\end{aligned}$$

The second equality uses proposition 2.5. $\square$

Proposition 2.14. *If γ is a piecewise smooth curve and f is continuous in γ^*, then*

$$\int_{\neg\gamma} f(z)\, dz = -\int_\gamma f(z)\, dz.$$

Proof. If $\gamma : [a,b] \to \mathbb{C}$, then

$$\begin{aligned}\int_{\neg\gamma} f(z)\, dz &= \int_a^b f((\neg\gamma)(t))(\neg\gamma)'(t)\, dt = -\int_a^b f(\gamma(a+b-t))\gamma'(a+b-t)\, dt \\ &= \int_b^a f(\gamma(s))\gamma'(s)\, ds = -\int_a^b f(\gamma(s))\gamma'(s)\, ds = -\int_\gamma f(z)\, dz.\end{aligned}$$

after a simple change of parameter in the third integral. □

Example 2.2.10. Let γ be the curve describing the linear segment $[z_0, z_1]$ from z_0 to z_1. Then $\neg\gamma$ describes the same segment from z_1 to z_0. Therefore, $\int_{[z_0,z_1]} f(z)\, dz = \int_\gamma f(z)\, dz$ and $\int_{[z_1,z_0]} f(z)\, dz = \int_{\neg\gamma} f(z)\, dz$. Hence

$$\int_{[z_1,z_0]} f(z)\, dz = -\int_{[z_0,z_1]} f(z)\, dz.$$

Before we leave this section, we should mention three variants of the notion of the curvilinear integral. Let $\gamma : [a,b] \to \mathbb{C}$ be a piecewise smooth curve and let $f : \gamma^* \to \mathbb{C}$ be continuous in the trajectory $\gamma^* = \{\gamma(t) \mid t \in [a,b]\}$. If $\gamma(t) = (x(t), y(t)) = x(t) + iy(t)$ for every $t \in [a,b]$, we define

$$\int_\gamma f(z)\, dx = \int_a^b f(\gamma(t))x'(t)\, dt, \qquad \int_\gamma f(z)\, dy = \int_a^b f(\gamma(t))y'(t)\, dt,$$

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

Trivially, we have

$$\int_\gamma f(z)\, dz = \int_\gamma f(z)\, dx + i\int_\gamma f(z)\, dy.$$

We leave to the reader the easy task to show that each of the three new integrals satisfies all properties of the original $\int_\gamma f(z)\, dz$, expressed in propositions 2.8 - 2.14. The only difference is with the integral $\int_\gamma f(z)\, |dz|$ which, regarding proposition 2.14, does *not* change its sign when we replace γ with $\neg\gamma$. Moreover, the basic inequality in proposition 2.9 takes the more precise form:

$$\left|\int_\gamma f(z)\, dz\right| \le \int_\gamma |f(z)|\, |dz| \le \|f\|_{\gamma^*} l(\gamma).$$

Indeed, observing the string of equalities/inequalities in the proof of proposition 2.9, we recognize $\int_\gamma |f(z)|\, |dz|$ as the third integral from the left. It is very common with beginning students to make the mistake: $|\int_\gamma f(z)\, dz| \le \int_\gamma |f(z)|\, dz$.

We should also say that

$$\int_\gamma |dz| = l(\gamma).$$

In calculus texts one usually sees the symbol ds instead of $|dz|$ for the infinitesimal length $|\gamma'(t)|\, dt$ over the curve.

Exercises.

2.2.1. Let $\sim$ be the relation of reparametrization.
(i) If $\gamma_1 \sim \gamma_2$, prove that $\neg\gamma_1 \sim \neg\gamma_2$.
(ii) If $\gamma_1 \sim \gamma_2$ and $\sigma_1 \sim \sigma_2$, prove that $\gamma_1 \dot{+} \sigma_1 \sim \gamma_2 \dot{+} \sigma_2$ (provided the two sums are defined).

2.2.2. Calculate $\int_\gamma |z|\, dz$, where γ is each of the following curves with initial endpoint $-i$ and final endpoint i.
(i) $\gamma(t) = it$, $t \in [-1, 1]$.
(ii) $\gamma(t) = \cos t + i \sin t$, $t \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.
(iii) $\gamma(t) = -\cos t + i \sin t$, $t \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

2.2.3. (i) If $n \in \mathbb{Z}$, $n \geq 0$, prove that $\int_\gamma z^n\,dz = \frac{z_1^{n+1} - z_0^{n+1}}{n+1}$, where z_0, z_1 are the initial and the final endpoint of the piecewise smooth γ.
(ii) Are there polynomials $p_n(z)$ so that $p_n(z) \to \frac{1}{z}$ uniformly in the circle $C_0(1)$? Think in terms of curvilinear integrals over the circle $C_0(1)$.

2.2.4. (i) Let f be continuous in the ring $\{z \,|\, 0 < |z| < r_0\}$ and $\lim_{r\to 0+} r\|f\|_{C_0(r)} = 0$. Prove that $\lim_{r\to 0+} \oint_{C_{z_0}(r)} f(z)\,dz = 0$.
(ii) Let f be continuous in $D_{z_0}(R)$. Prove that

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

2.2.5. Let $f : \Omega \to \mathbb{C}$ be continuous in the open set Ω and let $[a_n, b_n], [a, b] \subseteq \Omega$ for every n. If $a_n \to a$ and $b_n \to b$, prove that $\int_{[a_n,b_n]} f(z)\,dz \to \int_{[a,b]} f(z)\,dz$.

2.2.6. Let $f : \Omega \to \mathbb{C}$ be continuous in the open set Ω and γ be a piecewise smooth curve in Ω. Prove that for every $\epsilon > 0$ there is a polygonal curve σ in Ω so that $|\int_\sigma f(z)\,dz - \int_\gamma f(z)\,dz| < \epsilon$.

2.2.7. Prove that $|\int_a^b f(t)\,dt| = \int_a^b |f(t)|\,dt$ if and only if there is some halfline l with vertex 0 so that $f(t) \in l$ for every continuity point t of f.

2.2.8. Let $\gamma : [a, b] \to \mathbb{C}$ be a piecewise smooth curve and $f : \gamma^* \to \mathbb{C}$ be continuous in γ^*. Consider any subdivision $\Delta = \{t_0, \ldots, t_n\}$ of $[a, b]$ and any choice $\Xi = \{\xi_1, \ldots, \xi_n\}$ of intermediate points $\xi_k \in [t_{k-1}, t_k]$. Then we say that $\Delta^* = \{z_0, \ldots, z_n\}$, where $z_k = \gamma(t_k)$, is a subdivision of the trajectory γ^* and that $\Xi^* = \{\eta_1, \ldots, \eta_n\}$, where $\eta_k = \gamma(\xi_k)$, is a choice of intermediate points on the trajectory: η_k is between z_{k-1} and z_k on the trajectory. We say that $\sum_{k=1}^n f(z_k)(\eta_k - \eta_{k-1})$ is the corresponding Riemann sum. If $w(\Delta^*) = \max_{1\leq k\leq n} |z_k - z_{k-1}|$ is the width of the subdivision Δ^*, then prove that

$$\lim_{w(\Delta^*)\to 0} \sum_{k=1}^n f(z_k)(\eta_k - \eta_{k-1}) = \int_\gamma f(z)\,dz.$$

Chapter 3

Holomorphic functions.

3.1 Differentiability and holomorphy.

Let $f : A \to \mathbb{C}$ be a complex function defined in $A \subseteq \mathbb{C}$ and z_0 be an *interior* point of A. We say that f is **differentiable** at z_0 if $\lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0}$ exists and is a complex number. We call this limit **derivative** of f at z_0 and denote it

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

Example 3.1.1. The constant function c is differentiable at every point of $\mathbb{C}$ and its derivative is the constant function 0. Indeed, for every z_0 we have

$$\frac{dc}{dz}(z_0) = \lim_{z\to z_0} \frac{c-c}{z-z_0} = \lim_{z\to z_0} 0 = 0.$$

Example 3.1.2. The function z is differentiable at every point of $\mathbb{C}$ and its derivative is the constant function 1: for every z_0 we have

$$\frac{dz}{dz}(z_0) = \lim_{z\to z_0} \frac{z-z_0}{z-z_0} = \lim_{z\to z_0} 1 = 1.$$

Example 3.1.3. Let $f(z) = \overline{z}$. We shall prove that the $\lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0} = \lim_{z\to z_0} \frac{\overline{z}-\overline{z_0}}{z-z_0}$ does not exist, i.e. f is not differentiable at any z_0.
Let $z_0 = x_0 + iy_0$. The limit of $\frac{f(z)-f(z_0)}{z-z_0}$ when $z \to z_0$ on the horizontal line containing z_0 is

$$\lim_{x\to x_0} \frac{\overline{(x+iy_0)}-\overline{(x_0+iy_0)}}{(x+iy_0)-(x_0+iy_0)} = \lim_{x\to x_0} \frac{x-x_0}{x-x_0} = \lim_{x\to x_0} 1 = 1$$

and the limit of $\frac{f(z)-f(z_0)}{z-z_0}$ when $z \to z_0$ on the vertical line containing z_0 is

$$\lim_{y\to y_0} \frac{\overline{(x_0+iy)}-\overline{(x_0+iy_0)}}{(x_0+iy)-(x_0+iy_0)} = \lim_{y\to y_0} \frac{-iy+iy_0}{iy-iy_0} = \lim_{y\to y_0}(-1) = -1.$$

Since these two limits are different, the $\lim_{z\to z_0} \frac{\overline{z}-\overline{z_0}}{z-z_0}$ does not exist.

The proofs of the following four propositions are identical with the proofs of the well known analogous propositions for real functions of a real variable and we omit them.

Proposition 3.1. *If $f : A \to \mathbb{C}$ is differentiable at the interior point z_0 of $A \subseteq \mathbb{C}$, then f is continuous at z_0.*

Proposition 3.2. *If $f, g : A \to \mathbb{C}$ are differentiable at the interior point z_0 of $A \subseteq \mathbb{C}$, then $f+g, f-g, fg : A \to \mathbb{C}$ are also differentiable at z_0. Furthermore, if $g(z) \neq 0$ for all $z \in A$, then $\frac{f}{g} : A \to \mathbb{C}$ is differentiable at z_0. Finally,*

$$(f+g)'(z_0) = f'(z_0) + g'(z_0), \quad (f-g)'(z_0) = f'(z_0) - g'(z_0),$$

$$(fg)'(z_0) = f'(z_0)g(z_0) + f(z_0)g'(z_0), \quad (\tfrac{f}{g})'(z_0) = \frac{f'(z_0)g(z_0)-f(z_0)g'(z_0)}{g^2(z_0)}.$$

Proposition 3.3. *If $f : A \to B$ is differentiable at the interior point z_0 of $A \subseteq \mathbb{C}$ and $g : B \to \mathbb{C}$ is differentiable at the interior point $w_0 = f(z_0)$ of $B \subseteq \mathbb{C}$, then $g \circ f : A \to \mathbb{C}$ is differentiable at z_0. Also,*

$$(g \circ f)'(z_0) = g'(w_0)f'(z_0) = g'(f(z_0))f'(z_0).$$

Proposition 3.4. *Let $f : A \to B$ be one-to-one from $A \subseteq \mathbb{C}$ onto $B \subseteq \mathbb{C}$ and let $f^{-1} : B \to A$ be the inverse function. Let also z_0 be an interior point of A and $w_0 = f(z_0)$ be an interior point of B. If f is differentiable at z_0 and $f'(z_0) \neq 0$ and f^{-1} is continuous at w_0, then f^{-1} is differentiable at w_0 and*

$$(f^{-1})'(w_0) = 1/f'(z_0).$$

Example 3.1.4. Starting with the derivatives of the constant function c and of the function z and using the algebraic rules for derivatives, we get that every polynomial function is differentiable at every point of $\mathbb{C}$ and that its derivative is another polynomial function: if $p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$, then $p'(z) = a_1 + 2a_2 z + \cdots + n a_n z^{n-1}$.

Example 3.1.5. Every rational function is differentiable at every point of its domain of definition and its derivative is another rational function.

Example 3.1.6. If $h(z) = (z^2 - 3z + 2)^{15} - 3(z^2 - 3z + 2)^2$, then by the chain rule we get $h'(z) = 15(z^2 - 3z + 2)^{14}(2z - 3) - 6(z^2 - 3z + 2)(2z - 3)$.

Let f be a complex function defined in $A \subseteq \mathbb{C}$ and z_0 be an interior point of A. We say that f is **holomorphic** or **analytic** at z_0 if there is $r > 0$ so that $D_{z_0}(r) \subseteq A$ and f is differentiable at every point of $D_{z_0}(r)$.

The notion of holomorphy is stronger than the notion of differentiability: for a function to be holomorphic at a point it is necessary for it to be differentiable at this point *and* at all nearby points.

Example 3.1.7. Every polynomial function is holomorphic at every point of $\mathbb{C}$.

Example 3.1.8. Every rational function is holomorphic at every point of its domain of definition.

Example 3.1.9. Let $f(z) = |z|^2$. We have $\lim_{z\to 0} \frac{f(z)-f(0)}{z-0} = \lim_{z\to 0} \overline{z} = 0$ and so f is differentiable at 0 with $f'(0) = 0$.
We take any $z_0 \neq 0$ and we shall prove that the $\lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0} = \lim_{z\to z_0} \frac{|z|^2-|z_0|^2}{z-z_0}$ does not exist and therefore f is not differentiable at z_0. Indeed, let $z_0 = x_0 + iy_0$. The limit of $\frac{f(z)-f(z_0)}{z-z_0}$ when $z \to z_0$ on the horizontal line containing z_0 is

$$\lim_{x\to x_0} \frac{|x+iy_0|^2-|x_0+iy_0|^2}{(x+iy_0)-(x_0+iy_0)} = \lim_{x\to x_0} \frac{x^2-{x_0}^2}{x-x_0} = \lim_{x\to x_0}(x + x_0) = 2x_0$$

and the limit of $\frac{f(z)-f(z_0)}{z-z_0}$ when $z \to z_0$ on the vertical line containing z_0 is

$$\lim_{y\to y_0} \frac{|x_0+iy|^2-|x_0+iy_0|^2}{(x_0+iy)-(x_0+iy_0)} = \lim_{y\to y_0} \frac{y^2-{y_0}^2}{iy-iy_0} = -i\lim_{y\to y_0}(y + y_0) = -2iy_0.$$

Since $z_0 \neq 0$, these two limits are different and the $\lim_{z\to z_0} \frac{|z|^2-|z_0|^2}{z-z_0}$ does not exist.
We conclude that f is differentiable only at 0 and that it is nowhere holomorphic.

The set of points at which f is holomorphic is called **domain of holomorphy** of f.

Proposition 3.5. *If $B \subseteq \mathbb{C}$ is the set of the points at which the complex function f is differentiable, then the domain of holomorphy of f is the interior of B. In particular, the domain of holomorphy of f is an open set.*

Proof. Let U be the domain of holomorphy of f. If $z \in U$, there is $r > 0$ so that f is differentiable at every point of $D_z(r)$ and hence $D_z(r) \subseteq B$. Thus z is an interior point of B, i.e. $z \in B^\circ$.
Conversely, let $z \in B^\circ$. Then there is $r > 0$ so that $D_z(r) \subseteq B$, and so f is differentiable at every point of $D_z(r)$. Therefore f is holomorphic at z, i.e. $z \in U$. □

Example 3.1.10. The domain of holomorphy of any polynomial function is $\mathbb{C}$.

Example 3.1.11. The domain of holomorphy of any rational function is its domain of definition.

Example 3.1.12. The domain of holomorphy of both functions $f(z) = \overline{z}$ and $f(z) = |z|^2$ is the empty set.

Let $\Omega \subseteq \mathbb{C}$ be an open set. We say that the complex function f is holomorphic (or analytic) in Ω if it is holomorphic at every point of Ω or, equivalently, if Ω is a subset of the domain of holomorphy of f.

Clearly, the largest open set Ω in which f is holomorphic is its domain of holomorphy. It is also clear that *if f is differentiable at every point of an open set Ω, then f is holomorphic in Ω.*

Let the complex function f be defined in the neighborhood $D_\infty(r) = \{z \,|\, |z| > \frac{1}{r}\} \cup \{\infty\}$ of ∞. We consider the complex function g defined in $D_0(r) = \{w \,|\, |w| < r\}$ by

$$g(w) = f(1/w).$$

We say that f is differentiable or holomorphic at ∞ if g is differentiable or holomorphic, respectively, at 0.

We observe that $g(0) = f(\infty)$ and that the inverse functions $w = \frac{1}{z}$ and $z = \frac{1}{w}$ map each of the neighborhoods $D_\infty(r)$ and $D_0(r)$ onto the other. Now we shall see that the condition of differentiability of f at ∞, i.e. the differentiability of g at 0, can be translated into an equivalent condition in terms of f itself.

Proposition 3.6. *Let f be a complex function defined in $D_\infty(r)$. Then f is differentiable at ∞ if and only if*

$$\lim_{z\to\infty} z(f(z) - f(\infty)) \in \mathbb{C}.$$

Proof. Let $g(w) = f(\frac{1}{z})$ be the function considered in the above definition. Through the change of variable $w = \frac{1}{z}$, we have $\frac{g(w)-g(0)}{w-0} = z(f(z) - f(\infty))$. Thus, the existence of $\lim_{w\to 0} \frac{g(w)-g(0)}{w-0}$ is equivalent to the existence of $\lim_{z\to\infty} z(f(z) - f(\infty))$. In fact the two limits are equal. □

It is easy to see that differentiability of f at ∞ implies continuity of f at ∞.

Example 3.1.13. We shall check the differentiability (and hence holomorphy) at ∞ of polynomial and rational functions. We recall the notation and the results of examples 1.3.1 and 1.3.2.
A polynomial function p is continuous and complex valued at ∞ only if it is a constant $p(z) = a_0$ and $p(\infty) = a_0$. Then it is differentiable at ∞, since $\lim_{z\to\infty} z(p(z) - p(\infty)) = \lim_{z\to\infty} 0 = 0$.
A rational function $r(z) = \frac{a_n z^n + \cdots + a_1 z + a_0}{b_m z^m + \cdots + b_1 z + b_0}$ is continuous and complex valued at ∞ only if $n \leq m$, where n and m are the degrees of its numerator and denominator. If $n = m$, we set $r(\infty) = \frac{a_n}{b_n}$ and, after some algebraic manipulations, we get

$$\lim_{z\to\infty} z(r(z) - r(\infty)) = \lim_{z\to\infty} z\Big(\frac{a_n z^n + \cdots + a_1 z + a_0}{b_n z^n + \cdots + b_1 z + b_0} - \frac{a_n}{b_n}\Big) = \frac{a_{n-1} b_n - a_n b_{n-1}}{b_n^2}.$$

If $n < m$, we set $r(\infty) = 0$ and we get

$$\lim_{z\to\infty} z(r(z) - r(\infty)) = \lim_{z\to\infty} z\,\frac{a_n z^n + \cdots + a_1 z + a_0}{b_m z^m + \cdots + b_1 z + b_0} = \begin{cases} \frac{a_n}{b_{n+1}}, & \text{if } n + 1 = m \\ 0, & \text{if } n + 1 < m \end{cases}$$

Exercises.

3.1.1. Check the differentiability and holomorphy of the functions $\operatorname{Re} z$, $\operatorname{Im} z$ and $|z|$.

3.1.2. Let Ω be open and $f : \Omega \to \mathbb{C}$. We take $\Omega^* = \{z \,|\, \overline{z} \in \Omega\}$ and $f^* : \Omega^* \to \mathbb{C}$ given by $f^*(z) = \overline{f(\overline{z})}$ for every $z \in \Omega^*$. Prove that Ω^* is open and that, if f is differentiable at $z_0 \in \Omega$, then f^* is differentiable at $\overline{z_0} \in \Omega^*$.

3.1.3. Consider open sets U, V and $f : V \to U$, $g : U \to \mathbb{C}$, $h : V \to \mathbb{C}$ so that h is one-to-one and $h = g \circ f$. If h is differentiable at $w_0 \in V$, g is differentiable at $z_0 = f(w_0)$, $g'(z_0) \neq 0$ and f is continuous at w_0, prove that f is differentiable at w_0 and $f'(w_0) = \frac{h'(w_0)}{g'(z_0)}$.

3.1.4. (i) If p is a polynomial of degree n with roots $z_1, \dots, z_n$, prove

$$\frac{p'(z)}{p(z)} = \frac{1}{z-z_1} + \cdots + \frac{1}{z-z_n}$$

for $z \neq z_1, \dots, z_n$. Then prove that, if the roots of p are contained in a closed halfplane, then the roots of p' are contained in the same halfplane. Conclude that the roots of p' are contained in the smallest convex polygon which contains the roots of p.
(ii) For every a and every $n \in \mathbb{N}$, $n \geq 2$ prove that the equation $1 + z + az^n = 0$ has at least one root $z \in \overline{D}_0(2)$.

3.1.5. (i) Let $z_1, \dots, z_n$ be distinct and $q(z) = (z - z_1) \cdots (z - z_n)$. If the polynomial p has degree $< n$, prove

$$\frac{p(z)}{q(z)} = \sum_{k=1}^{n} \frac{p(z_k)}{q'(z_k)(z-z_k)}$$

for $z \neq z_1, \dots, z_n$.
(ii) Let $z_1, \dots, z_n$ be distinct. Prove that for every $c_1, \dots, c_n$ there is a unique polynomial p of degree $< n$ so that $p(z_k) = c_k$ for every $k = 1, \dots, n$.

3.1.6. Let f have continuous derivative in a neighborhood of z_0. Prove that $\frac{f(z_n')-f(z_n'')}{z_n'-z_n''} \to f'(z_0)$ if $z_n' \to z_0$, $z_n'' \to z_0$ and $z_n' \neq z_n''$ for every n.

3.2 The Cauchy-Riemann equations.

Now we shall relate the differentiability of f, as a complex function of $z = x + iy$, at some interior point $z_0 = x_0 + iy_0$ of its domain $A \subseteq \mathbb{C}$ with the partial derivatives of $u = \mathrm{Re}\, f$ and $v = \mathrm{Im}\, f$, as functions of (x, y) at the same point (x_0, y_0).

Theorem 3.1. *Let f be a complex function defined in $A \subseteq \mathbb{C}$, $z_0 = (x_0, y_0)$ be an interior point of A, and let u, v be the real and imaginary part of f. If f is differentiable at z_0, then u, v have partial derivatives with respect to x and y at (x_0, y_0), and*

$$\frac{\partial u}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0), \qquad \frac{\partial u}{\partial y}(x_0, y_0) = -\frac{\partial v}{\partial x}(x_0, y_0). \tag{3.1}$$

Proof. We assume

$$\lim_{z \to z_0} \frac{f(z)-f(z_0)}{z-z_0} = f'(z_0) = \mu + i\nu, \qquad \mu, \nu \in \mathbb{R}. \tag{3.2}$$

Since the limit of $\frac{f(z)-f(z_0)}{z-z_0}$ exists when z tends to z_0, the limits of the same expression when z tends to z_0 on the horizontal line containing z_0 as well as on the vertical line containing z_0 also exist and have the same value:

$$\lim_{x \to x_0} \frac{f(x,y_0)-f(x_0,y_0)}{x-x_0} = \mu + i\nu, \qquad \lim_{y \to y_0} \frac{f(x_0,y)-f(x_0,y_0)}{iy-iy_0} = \mu + i\nu. \tag{3.3}$$

From the first limit in (3.3) we get $\lim_{x \to x_0} \frac{u(x,y_0)+iv(x,y_0)-u(x_0,y_0)-iv(x_0,y_0)}{x-x_0} = \mu + i\nu$, and hence

$$\begin{aligned} \frac{\partial u}{\partial x}(x_0, y_0) &= \lim_{x \to x_0} \frac{u(x,y_0)-u(x_0,y_0)}{x-x_0} = \mu, \\ \frac{\partial v}{\partial x}(x_0, y_0) &= \lim_{x \to x_0} \frac{v(x,y_0)-v(x_0,y_0)}{x-x_0} = \nu. \end{aligned} \tag{3.4}$$

From the second limit in (3.3) we find $\lim_{y\to y_0} \frac{u(x_0,y)+iv(x_0,y)-u(x_0,y_0)-iv(x_0,y_0)}{iy-iy_0} = \mu + i\nu$, and hence

$$\begin{aligned} \tfrac{\partial v}{\partial y}(x_0, y_0) &= \lim_{y\to y_0} \tfrac{v(x_0,y)-v(x_0,y_0)}{y-y_0} = \mu, \\ \tfrac{\partial u}{\partial y}(x_0, y_0) &= \lim_{x\to x_0} \tfrac{u(x_0,y)-u(x_0,y_0)}{y-y_0} = -\nu. \end{aligned} \tag{3.5}$$

Comparing (3.4) and (3.5) we get (3.1). □

The equalities (3.1) are called **(system of) Cauchy-Riemann equations** at the point (x_0, y_0). We observe that, if f is differentiable at z_0, then (3.2), (3.4) and (3.5) imply

$$f'(z_0) = \tfrac{\partial u}{\partial x}(x_0, y_0) + i\tfrac{\partial v}{\partial x}(x_0, y_0) = \tfrac{\partial v}{\partial y}(x_0, y_0) - i\tfrac{\partial u}{\partial y}(x_0, y_0).$$

The next result is the converse of theorem 3.1 but with extra assumptions.

Theorem 3.2. *Let f be a complex function defined in $A \subseteq \mathbb{C}$, $z_0 = (x_0, y_0)$ be an interior point of A and let u, v be the real and imaginary part of f. If u, v have partial derivatives with respect to x and y at every point of some neighborhood of (x_0, y_0) and if these partial derivatives are continuous at (x_0, y_0) and if they satisfy the system of C-R equations at (x_0, y_0), then f is differentiable at z_0.*

Proof. Using the C-R equations, we define the real numbers μ and ν by:

$$\mu = \tfrac{\partial u}{\partial x}(x_0, y_0) = \tfrac{\partial v}{\partial y}(x_0, y_0), \qquad \nu = -\tfrac{\partial u}{\partial y}(x_0, y_0) = \tfrac{\partial v}{\partial x}(x_0, y_0). \tag{3.6}$$

Now take an arbitrary $\epsilon > 0$. Since $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$ are continuous at (x_0, y_0), there is $r > 0$ so that

$$\big|\tfrac{\partial u}{\partial x}(x, y) - \mu\big| < \tfrac{\epsilon}{4}, \quad \big|\tfrac{\partial u}{\partial y}(x, y) + \nu\big| < \tfrac{\epsilon}{4} \tag{3.7}$$

for every $(x, y) \in D_{(x_0,y_0)}(r)$. We take any $(x, y) \in D_{(x_0,y_0)}(r)$ and we write

$$u(x, y) - u(x_0, y_0) = u(x, y) - u(x_0, y) + u(x_0, y) - u(x_0, y_0). \tag{3.8}$$

By the mean value theorem, there is x' between x and x_0 so that

$$u(x, y) - u(x_0, y) = \tfrac{\partial u}{\partial x}(x', y)(x - x_0) \tag{3.9}$$

and y' between y and y_0 so that

$$u(x_0, y) - u(x_0, y_0) = \tfrac{\partial u}{\partial y}(x_0, y')(y - y_0). \tag{3.10}$$

The x', y' depend on x, y, but the points (x', y), (x_0, y') belong to $D_{(x_0,y_0)}(r)$. Therefore, (3.7) implies

$$\big|\tfrac{\partial u}{\partial x}(x', y) - \mu\big| < \tfrac{\epsilon}{4}, \qquad \big|\tfrac{\partial u}{\partial y}(x_0, y') + \nu\big| < \tfrac{\epsilon}{4}. \tag{3.11}$$

Combining (3.8), (3.9) and (3.10), we find

$$\begin{aligned} &u(x, y) - u(x_0, y_0) - \big(\mu(x - x_0) - \nu(y - y_0)\big) \\ &\qquad = \big(u(x, y) - u(x_0, y) - \mu(x - x_0)\big) + \big(u(x_0, y) - u(x_0, y_0) + \nu(y - y_0)\big) \\ &\qquad = \big(\tfrac{\partial u}{\partial x}(x', y) - \mu\big)(x - x_0) + \big(\tfrac{\partial u}{\partial y}(x_0, y') + \nu\big)(y - y_0) \end{aligned}$$

and, because of (3.11),

$$\begin{aligned} &\big|u(x, y) - u(x_0, y_0) - \big(\mu(x - x_0) - \nu(y - y_0)\big)\big| \\ &\qquad \le \big|\tfrac{\partial u}{\partial x}(x', y) - \mu\big||x - x_0| + \big|\tfrac{\partial u}{\partial y}(x_0, y') + \nu\big||y - y_0| \\ &\qquad < \tfrac{\epsilon}{4}|x - x_0| + \tfrac{\epsilon}{4}|y - y_0| < \tfrac{\epsilon}{2}\sqrt{(x - x_0)^2 + (y - y_0)^2}. \end{aligned} \tag{3.12}$$

In the same manner, for the function v we get

$$\big|v(x,y) - v(x_0,y_0) - \big(\nu(x-x_0) + \mu(y-y_0)\big)\big| < \tfrac{\epsilon}{2}\sqrt{(x-x_0)^2 + (y-y_0)^2}. \tag{3.13}$$

The inequalities (3.12) and (3.13) hold at every $(x,y) \in D_{(x_0,y_0)}(r)$.
We observe that the expressions inside the absolute values of the left sides of (3.12) and (3.13) are, respectively, the real and the imaginary part of the number

$$f(z) - f(z_0) - (\mu + i\nu)(z - z_0) = f(x,y) - f(x_0,y_0) - (\mu + i\nu)\big((x-x_0) + i(y-y_0)\big).$$

Therefore, (3.12) and (3.13) imply

$$|f(z) - f(z_0) - (\mu + i\nu)(z - z_0)| < \epsilon\sqrt{(x-x_0)^2 + (y-y_0)^2} = \epsilon|z - z_0|$$

for every $z \in D_{z_0}(r)$ and hence

$$\big|\tfrac{f(z)-f(z_0)}{z-z_0} - (\mu + i\nu)\big| < \epsilon$$

for every $z \in D_{z_0}(r)$, $z \neq z_0$. Thus, $\lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0} = \mu + i\nu$, and f is differentiable at z_0 with $f'(z_0) = \mu + i\nu$. □

Example 3.2.1. The real and the imaginary parts of the function $f(z) = z^2$ are $u(x,y) = x^2 - y^2$ and $v(x,y) = 2xy$. We find $\frac{\partial u}{\partial x}(x,y) = 2x$, $\frac{\partial u}{\partial y}(x,y) = -2y$, $\frac{\partial v}{\partial x}(x,y) = 2y$ and $\frac{\partial v}{\partial y}(x,y) = 2x$, and we see that the partial derivatives are continuous in the whole plane and they satisfy the C-R equations at every point. Theorem 3.2 implies that $f(z) = z^2$ is differentiable at every point and $f'(z) = \frac{\partial u}{\partial x}(x,y) + i\frac{\partial v}{\partial x}(x,y) = 2x + i2y = 2z$.

Example 3.2.2. We reconsider the function $f(z) = \overline{z}$ of example 3.1.3. Its real and imaginary parts are $u(x,y) = x$ and $v(x,y) = -y$. The partial derivatives $\frac{\partial u}{\partial x}(x,y) = 1$, $\frac{\partial u}{\partial y}(x,y) = 0$, $\frac{\partial v}{\partial x}(x,y) = 0$ and $\frac{\partial v}{\partial y}(x,y) = -1$ do not satisfy the C-R equations at any point (x,y). Theorem 3.1 implies that f is not differentiable at any point.

Example 3.2.3. We reconsider the function $f(z) = |z|^2$ of example 3.1.9. Its real and imaginary parts are $u(x,y) = x^2 + y^2$ and $v(x,y) = 0$. The partial derivatives are $\frac{\partial u}{\partial x}(x,y) = 2x$, $\frac{\partial u}{\partial y}(x,y) = 2y$, $\frac{\partial v}{\partial x}(x,y) = 0$ and $\frac{\partial v}{\partial y}(x,y) = 0$ and they satisfy the C-R equations only at the point $(0,0)$. Theorem 3.1 implies that f is not differentiable at any point besides, perhaps, the point $(0,0)$. Now, since the partial derivatives are continuous and satisfy the C-R equations at $(0,0)$, theorem 3.2 implies that f is differentiable at 0 and $f'(0) = \frac{\partial u}{\partial x}(0,0) + i\frac{\partial v}{\partial x}(0,0) = 0 + i0 = 0$.

Example 3.2.4. We shall see that the assumption of continuity of the partial derivatives of u, v at (x_0,y_0) in theorem 3.2 is crucial. We consider the function

$$f(z) = f(x,y) = \begin{cases} \frac{xy}{\sqrt{x^2+y^2}}, & \text{if } (x,y) \neq (0,0) \\ 0, & \text{if } (x,y) = (0,0) \end{cases}$$

Then its real and imaginary parts are

$$u(x,y) = \begin{cases} \frac{xy}{\sqrt{x^2+y^2}}, & \text{if } (x,y) \neq (0,0) \\ 0, & \text{if } (x,y) = (0,0) \end{cases} \qquad v(x,y) = 0.$$

It is clear that $\frac{\partial v}{\partial x}(x,y) = 0$ and $\frac{\partial v}{\partial y}(x,y) = 0$ and the partial derivatives of v are continuous at every (x,y). Moreover,

$$\frac{\partial u}{\partial x}(x,y) = \begin{cases} \frac{y^3}{\sqrt{(x^2+y^2)^3}}, & \text{if } (x,y) \neq (0,0) \\ 0, & \text{if } (x,y) = (0,0) \end{cases} \qquad \frac{\partial u}{\partial y}(x,y) = \begin{cases} \frac{x^3}{\sqrt{(x^2+y^2)^3}}, & \text{if } (x,y) \neq (0,0) \\ 0, & \text{if } (x,y) = (0,0) \end{cases}$$

The partial derivatives of u are continuous at every $(x,y) \neq (0,0)$ but they are *not* continuous at $(0,0)$. For instance, it is easy to see that the limit of $\frac{\partial u}{\partial x}(x,y) = \frac{y^3}{\sqrt{(x^2+y^2)^3}}$ when (x,y) tends to $(0,0)$ on the line with equation $y = x$ does not exist. Moreover, f is *not* differentiable at 0, *even though* u, v *do satisfy the C-R equations at* 0. In fact it is easy to see that the limit of $\frac{f(z)-f(0)}{z-0} = \frac{xy}{(x+iy)\sqrt{x^2+y^2}}$ when z tends to 0 on the line with equation $y = x$ does not exist.

The next proposition is a corollary of theorem 3.2. It is the form of theorem 3.2 in which this is usually applied.

Proposition 3.7. *Let f be a complex function defined in the open set $\Omega \subseteq \mathbb{C}$ and let u, v be the real and the imaginary part of f. If u, v have partial derivatives which are continuous and which satisfy the C-R equations at every point of Ω, then f is holomorphic in Ω.*

Proof. We take an arbitrary $z \in \Omega$ and a neighborhood of z which is contained in Ω. Theorem 3.2 implies that f is differentiable at z. Thus f is differentiable at every point of Ω and, since Ω is open, f is holomorphic in Ω. $\square$

An open and connected set Ω is called **region**.

Theorem 3.3. *Let f be holomorphic in the region $\Omega \subseteq \mathbb{C}$. If $f'(z) = 0$ for every $z \in \Omega$, then f is constant in Ω.*

First proof. Using $f' = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y}$, we find $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} = \frac{\partial u}{\partial y} = 0$ in Ω. We take any linear segment $[z_1, z_2]$ in Ω and its parametric equation $\gamma(t) = (1-t)z_1 + tz_2$, $t \in [0,1]$. By the mean value theorem, there is $t_0 \in (0,1)$ so that

$$\begin{aligned} u(z_2) - u(z_1) &= (u \circ \gamma)(1) - (u \circ \gamma)(0) = \tfrac{d(u \circ \gamma)}{dt}(t_0) \\ &= \tfrac{\partial u}{\partial x}(\gamma(t_0))(x_2 - x_1) + \tfrac{\partial u}{\partial y}(\gamma(t_0))(y_2 - y_1) = 0, \end{aligned}$$

where $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$. Thus, the values of u at the endpoints of any line segment in Ω are equal. Now we take arbitrary $z', z'' \in \Omega$. Then there is a polygonal line inside Ω which connects the two points z' and z''. The values of u at the endpoints of every line segment of the polygonal line are equal and hence $u(z') = u(z'')$. Therefore u is constant in Ω. Clearly, the same is true for the function v and hence for $f = u + iv$.

Second proof. We take arbitrary $z, w \in \Omega$. Since Ω is a region, there is a piecewise smooth curve $\gamma : [a,b] \to \Omega$ such that $\gamma(a) = z$, $\gamma(b) = w$. In fact we may choose γ to have a polygonal line in Ω as its trajectory. Then we have

$$f(w) - f(z) = (f \circ \gamma)(b) - (f \circ \gamma)(a) = \textstyle\int_a^b (f \circ \gamma)'(t)\, dt = \int_a^b f'(\gamma(t))\gamma'(t)\, dt = 0$$

because $f'(\gamma(t)) = 0$ for every $t \in [a,b]$. We conclude that $f(w) = f(z)$ for every $w, z \in \Omega$ and hence f is constant in Ω. $\square$

Let f be a complex function and let u, v be the real and imaginary part of f. If u, v have partial derivatives with respect to x, y at the point $z_0 = (x_0, y_0)$, it is trivial to prove that at the point z_0 we have

$$\tfrac{\partial f}{\partial x} = \tfrac{\partial u}{\partial x} + i\tfrac{\partial v}{\partial x}, \qquad \tfrac{\partial f}{\partial y} = \tfrac{\partial u}{\partial y} + i\tfrac{\partial v}{\partial y}. \tag{3.14}$$

We define the following differential operators:

$$\tfrac{\partial}{\partial z} = \tfrac{1}{2}\big(\tfrac{\partial}{\partial x} - i\tfrac{\partial}{\partial y}\big), \qquad \tfrac{\partial}{\partial \overline{z}} = \tfrac{1}{2}\big(\tfrac{\partial}{\partial x} + i\tfrac{\partial}{\partial y}\big). \tag{3.15}$$

Applying the differential operators $\frac{\partial}{\partial z}$ and $\frac{\partial}{\partial \overline{z}}$ to f and using (3.14), we have at the point z_0:

$$\begin{aligned} \tfrac{\partial f}{\partial z} &= \tfrac{1}{2}\big(\tfrac{\partial u}{\partial x} + \tfrac{\partial v}{\partial y}\big) + \tfrac{i}{2}\big(\tfrac{\partial v}{\partial x} - \tfrac{\partial u}{\partial y}\big), \\ \tfrac{\partial f}{\partial \overline{z}} &= \tfrac{1}{2}\big(\tfrac{\partial u}{\partial x} - \tfrac{\partial v}{\partial y}\big) + \tfrac{i}{2}\big(\tfrac{\partial v}{\partial x} + \tfrac{\partial u}{\partial y}\big). \end{aligned} \tag{3.16}$$

From the second of equations (3.16) we see that the system of C-R equations at the point z_0 is equivalent to the single equation

$$\frac{\partial f}{\partial \overline{z}} = 0$$

at z_0. Moreover, if the system of C-R equations is satisfied, then the first equation (3.16) implies

$$\frac{\partial f}{\partial z} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = f'$$

at z_0. We summarize.

Proposition 3.8. *If the complex function f is differentiable at z_0, then $\frac{\partial f}{\partial z}(z_0) = f'(z_0)$ and $\frac{\partial f}{\partial \overline{z}}(z_0) = 0$. Conversely, if $\frac{\partial f}{\partial z}$ and $\frac{\partial f}{\partial \overline{z}}$ exist in a neighborhood of the point z_0 and they are continuous at z_0 and if $\frac{\partial f}{\partial \overline{z}}(z_0) = 0$, then f is differentiable at z_0.*

Proof. Trivial. The converse is a restatement of theorem 3.2. Indeed, (3.16) implies that the existence or the continuity of $\frac{\partial f}{\partial z}$, $\frac{\partial f}{\partial \overline{z}}$ at a point is equivalent to the existence or the continuity, respectively, of $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, $\frac{\partial v}{\partial y}$ at the point. □

Sometimes a complex function f is given to us through an expression $f(x, y)$ as a function of two real variables and we are interested in finding an expression $f(z)$ of the function in terms of the single complex variable z. We then write $x = \frac{z+\overline{z}}{2}$, $y = \frac{z-\overline{z}}{2i}$ and hence

$$f(x, y) = f\left(\frac{z+\overline{z}}{2}, \frac{z-\overline{z}}{2i}\right). \tag{3.17}$$

In general, even after performing various algebraic simplifications we end up with an expression in terms of *both* variables z and $\overline{z}$. In order to end up with the occurence of z only, it is reasonable to impose the condition that the derivative of $f(x, y)$ with respect to $\overline{z}$ vanishes. From (3.17) and a formal chain rule we get

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

This is exactly the second differential operator (3.15) applied to f and we saw that the condition $\frac{\partial f}{\partial \overline{z}} = 0$ is equivalent to the system of C-R equations. We conclude that *the function $f(x, y)$ is a function of the single variable z if and only if its real and imaginary parts satisfy the C-R equations.*

Exercises.

3.2.1. Solve exercise 3.1.1 under the light of C-R equations.

3.2.2. (i) Prove that $F(x, y) = \sqrt{|xy|}$ satisfies the C-R equations at 0 but that it is not differentiable at 0.
(ii) Prove that the function with $G(x, y) = \frac{x^2y}{x^4+y^2}$ if $(x, y) \neq (0, 0)$ and with $G(0, 0) = 0$ satisfies the C-R equations at 0, that $\frac{G(z)-G(0)}{z-0}$ has a limit when $z \to 0$ on every line which contains 0, but that G is not differentiable at 0.

3.2.3. Let $f = u + iv$ be a complex function and $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, $\frac{\partial v}{\partial y}$ exist in a neighborhood of z_0 and be continuous at z_0.
(i) If $\lim_{z\to z_0} \operatorname{Re} \frac{f(z)-f(z_0)}{z-z_0}$ exists and is a real number, prove that f is differentiable at z_0.
(ii) If $\lim_{z\to z_0} \left|\frac{f(z)-f(z_0)}{z-z_0}\right|$ exists and is a real number, prove that either f is differentiable at z_0 or $\overline{f}$ is differentiable at z_0.

3.2.4. Let $f = u + iv$ be holomorphic in the region $\Omega \subseteq \mathbb{C}$.
(i) If either u or v is constant in Ω, prove that f is constant in Ω.
(ii) More generally, if for some line l it is true that $f(z) \in l$ for every $z \in \Omega$, prove that f is constant in Ω.
(iii) Consider (ii) with a circle C instead of a line l.

3.2.5. This exercise juxtaposes the notion of differentiability of a function of two real variables, which we learn in multivariable calculus, and the notion of differentiability of a function of one complex variable, which we learn in complex analysis: to distinguish between them we call the first $\mathbb{R}$-differentiability and the second $\mathbb{C}$-differentiability.
We recall from multivariable calculus that a *real* function u defined in $A \subseteq \mathbb{R}^2$ is $\mathbb{R}$-differentiable at the interior point (x_0, y_0) of A if there are $a, b \in \mathbb{R}$ so that

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

In this case we have that $\frac{\partial u}{\partial x}(x_0, y_0) = a$ and $\frac{\partial u}{\partial y}(x_0, y_0) = b$.
We also recall that a *vector* function $f = (u, v)$ defined in $A \subseteq \mathbb{R}^2$ is $\mathbb{R}$-differentiable at the interior point (x_0, y_0) of A if its real components u and v are both $\mathbb{R}$-differentiable at (x_0, y_0), i.e. if there are $a, b, c, d \in \mathbb{R}$ so that

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

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

In this case we have that $\frac{\partial u}{\partial x}(x_0, y_0) = a$, $\frac{\partial u}{\partial y}(x_0, y_0) = b$, $\frac{\partial v}{\partial x}(x_0, y_0) = c$, $\frac{\partial v}{\partial y}(x_0, y_0) = d$ and that the $\mathbb{R}$-derivative of f is the 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.
Prove that $f = (u, v) = u + iv$ is $\mathbb{C}$-differentiable at $z_0 = (x_0, y_0)$, i.e. that the $\lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0}$ exists and is a complex number, if and only if f is $\mathbb{R}$-differentiable at $z_0 = (x_0, y_0)$ and its $\mathbb{R}$-derivative is an antisymmetric matrix: $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. In this case the $\mathbb{C}$-derivative and the $\mathbb{R}$-derivative of f are related by $f'(z_0) = a + ib$.

3.2.6. Consider the functions $z^n, \overline{z}^n, |z|^2$ and, using the differential operator $\frac{\partial}{\partial \overline{z}}$, examine whether they are functions of z only or, equivalently, whether they are holomorphic.

3.2.7. Let f be a complex function. If $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist in a neighborhood of the point z_0 and are continuous at z_0, prove that

$$\lim_{r\to 0+} \frac{1}{2\pi i r^2} \oint_{C_{z_0}(r)} f(z)\, dz = \frac{\partial f}{\partial \overline{z}}(z_0).$$

3.3 Conformality.

Let the complex function f be continuous in $A \subseteq \mathbb{C}$ and $\gamma : [a, b] \to A$ be a curve. Thus the trajectory of γ is contained in the domain of definition of f. We define the function

$$f(\gamma) = f \circ \gamma : [a, b] \to \mathbb{C},$$

which is continuous in $[a, b]$. Then $f(\gamma)$ is a curve and we call it **image of γ through** f.

Now we also consider an interior point z of A and we assume that f is differentiable at z and $f(z) = w$, $f'(z) \neq 0$. We also take any curve $\gamma : [a, b] \to A$ with $\gamma(a) = z$. Then γ has z as its initial point and its trajectory is contained in A. We also assume that γ is differentiable at a and that $\gamma'(a) \neq 0$, i.e. that γ has a non-zero tangent vector at the point z. The image curve $f(\gamma) : [a, b] \to \mathbb{C}$ has $f(\gamma)(a) = (f \circ \gamma)(a) = f(\gamma(a)) = f(z) = w$ as its initial point and its tangent vector at w is $f(\gamma)'(a) = (f \circ \gamma)'(a) = f'(\gamma(a))\gamma'(a) = f'(z)\gamma'(a) \neq 0$. From this equality we have two conclusions. The first is that

$$|f(\gamma)'(a)| = |f'(z)||\gamma'(a)|.$$

Thus, the length of the tangent vector of $f(\gamma)$ at its initial point w equals the length of the tangent vector of γ at its initial point z multiplied with the factor $|f'(z)| > 0$. We express this as:
f multiplies the lengths of tangent vectors at z with the factor $|f'(z)| > 0$ or, in other words, *f expands the tangent vectors at z by the factor $|f'(z)| > 0$.*
The second conclusion is that

$$\arg f(\gamma)'(a) = \arg f'(z) + \arg \gamma'(a). \tag{3.18}$$

Thus, we find the angle of the tangent vector of $f(\gamma)$ at its initial point w by adding the angle of $f'(z)$ to the angle of the tangent vector of γ at its initial point z. We express this as:
f increases the angles of the tangent vectors at z by the angle of $f'(z)$ or, in other words, *f rotates the tangent vectors at z by the angle of $f'(z)$.*

We observe that the expansion and the rotation of the tangent vectors at z is uniform over all these vectors: *independently of their directions and their lengths, all these tangent vectors are expanded by the same factor $|f'(z)|$ and they are rotated by the same angle* $\arg f'(z)$. Since, any two of these tangent vectors are rotated by the same angle, we conclude that their relative angles remain unchanged! Indeed, let us consider two of the above curves, γ_1 and γ_2. Then the angle between their tangent vectors at z is $\arg \gamma_2'(a) - \arg \gamma_1'(a)$ and the angle between the tangent vectors of $f(\gamma_1)$ and $f(\gamma_2)$ at w is $\arg f(\gamma_2)'(a) - \arg f(\gamma_1)'(a)$. From (3.18) for γ_1 and γ_2 we get

$$\arg f(\gamma_2)'(a) - \arg f(\gamma_1)'(a) = \arg \gamma_2'(a) - \arg \gamma_1'(a).$$

Therefore, the angle between the tangent vectors of $f(\gamma_1)$ and $f(\gamma_2)$ at w is the same as the angle between the tangent vectors of γ_1 and γ_2 at z. We say:
f preserves the angle between tangent vectors at z.

This last property of f is called **conformality** of f at z and holds, as we just saw, under the assumption that f is differentiable at z and $f'(z) \neq 0$.

Exercises.

3.3.1. Consider the holomorphic function $w = f(z) = az + b$ with $a \neq 0$.
(i) Prove that f is one-to-one from $\mathbb{C}$ onto $\mathbb{C}$.
(ii) Prove that f maps lines and circles onto lines and circles, respectively.
(iii) Consider two lines with equations $kx + ly = m$ and $k'x + l'y = m'$. Which is the condition for the two lines to intersect? Under this condition, find their intersection point and the angle of the two lines at this point. Then find the equations of the images of the two lines through f and find their intersection point and their angle at this point. Confirm the conformality of f.

3.3.2. Consider the holomorphic function $w = z^2$.
(i) With any fixed u_0, v_0, consider the hyperbolas with equations $x^2 - y^2 = u_0$ and $2xy = v_0$ on the z-plane ($z = x + iy$). Do they intersect and at which points? Find the angle of the two hyperbolas at each of their common points.
(ii) With any fixed $x_0, y_0 \neq 0$, consider the parabolas with equations $u = \frac{1}{4{y_0}^2}v^2 - {y_0}^2$ and $u = -\frac{1}{4{x_0}^2}v^2 + {x_0}^2$ on the w-plane ($w = u + iv$). Do they intersect and at which points? Find the angle of the two parabolas at each of their common points.

3.3.3. Let f be holomorphic in the open set $U \subseteq \mathbb{C}$ so that f' is continuous in U, let γ be a piecewise smooth curve in U and $\Gamma = f(\gamma)$ be the image of γ through f. If the complex function ϕ is continuous in Γ^*, prove that

$$\textstyle\int_\Gamma \phi(w)\,dw = \int_\gamma \phi(f(z))f'(z)\,dz.$$

Chapter 4

Examples of holomorphic functions.

4.1 Linear fractional transformations.

Every rational function of the form

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

is called **linear fractional transformation**. We assume that $ad - bc \neq 0$. It is easy to show that $ad - bc \neq 0$ if and only if the function T is not constant.

In order to have the full picture of the definition of a linear fractional transformation T, we have to say something about the values of T at the roots of the denominator and at ∞. There are two cases. If $c = 0$, then because of $ad - bc \neq 0$ we have $ad \neq 0$ and then $T(z) = \frac{a}{d}z + \frac{b}{d}$ for all $z \in \mathbb{C}$. Since $\frac{a}{d} \neq 0$, we have that $T(\infty) = \infty$. Thus

$$T(z) = \begin{cases} \frac{a}{d}z + \frac{b}{d}, & \text{if } z \in \mathbb{C} \\ \infty, & \text{if } z = \infty \end{cases} \qquad \text{if } c = 0. \tag{4.1}$$

If $c \neq 0$, then the denominator has $z = -\frac{d}{c}$ as its root, which, because of $ad - bc \neq 0$, is not a root of the numerator. Hence $T(-\frac{d}{c}) = \infty$. Also $T(\infty) = \frac{a}{c}$. Thus

$$T(z) = \begin{cases} \frac{az+b}{cz+d}, & \text{if } z \in \mathbb{C},\ z \neq -\frac{d}{c} \\ \infty, & \text{if } z = -\frac{d}{c} \\ \frac{a}{c}, & \text{if } z = \infty \end{cases} \qquad \text{if } c \neq 0. \tag{4.2}$$

We conclude that every linear fractional transformation (l.f.t.) is a function $T : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ and, even though we write $T(z) = \frac{az+b}{cz+d}$, we must have in mind the full formulas (4.1) and (4.2).

It is very easy to see that every l.f.t. is one-to-one from $\widehat{\mathbb{C}}$ onto $\widehat{\mathbb{C}}$. The formula of the inverse l.f.t. of T is

$$T^{-1}(z) = \frac{dz-b}{-cz+a}.$$

The identity function $id(z) = z$ is clearly a l.f.t. with $a = d = 1$, $b = c = 0$, and we easily see that the composition of two l.f.t. is another l.f.t. Indeed, if $T(z) = \frac{az+b}{cz+d}$ and $S(z) = \frac{a'z+b'}{c'z+d'}$, then

$$(S \circ T)(z) = \frac{a'T(z)+b'}{c'T(z)+d'} = \frac{a'\frac{az+b}{cz+d}+b'}{c'\frac{az+b}{cz+d}+d'} = \frac{(a'a+b'c)z+(a'b+b'd)}{(c'a+d'c)z+(c'b+d'd)}.$$

Observe that $(a'a + b'c)(c'b + d'd) - (a'b + b'd)(c'a + d'c) = (a'd' - b'c')(ad - bc) \neq 0$.

Thus, the set of all l.f.t. is a group with the binary operation of composition. The neutral element of this group is the identity function.

Since a l.f.t. is a rational function, it is continuous in $\widehat{\mathbb{C}}$, and, as a particular instance of example 3.1.13, it is holomorphic in $\widehat{\mathbb{C}}$ except at the point at which it takes the value ∞.

Now we shall make a comment on an interesting relation between circles and lines. We observe that the equations of circles and lines can be unified in the following manner: if $\alpha, \beta, \gamma \in \mathbb{R}$, $w \in \mathbb{C}$, $w \neq 0$, $\alpha^2 + \beta^2 \neq 0$ and $\beta^2 |w|^2 > 4\alpha\gamma$, then the equation

$$\alpha |z|^2 + \beta \operatorname{Re}(\overline{w} z) + \gamma = 0$$

is the equation of a line, if $\alpha = 0$, and the equation of a circle, if $\alpha \neq 0$. In fact, if $\alpha = 0$, then $\beta \neq 0$ and the equation becomes $\operatorname{Re}(\overline{w} z) = -\frac{\gamma}{\beta}$ and this is the equation of a line. If $\alpha \neq 0$, the equation becomes $|z + \frac{\beta}{2\alpha} w|^2 = \frac{\beta^2 |w|^2 - 4\alpha\gamma}{4\alpha^2}$. This is the equation of the circle with center $-\frac{\beta}{2\alpha} w$ and radius $\frac{\sqrt{\beta^2 |w|^2 - 4\alpha\gamma}}{2|\alpha|}$. Conversely, every circle and every line have equations of this form. If, for instance, we take the equation $\operatorname{Re}(\overline{w} z) = c$ of a line, with $w \in \mathbb{C}$, $w \neq 0$, and $c \in \mathbb{R}$, we may write it in the form $\alpha |z|^2 + \beta \operatorname{Re}(\overline{w} z) + \gamma = 0$ by taking $\alpha = 0$, $\beta = 1$ and $\gamma = -c$. If we take the equation $|z - z_0| = r$ of a circle with $z_0 \in \mathbb{C}$ and $r > 0$, we may write it as $|z|^2 - 2\operatorname{Re}(\overline{z_0} z) + |z_0|^2 = r^2$. This becomes $\alpha |z|^2 + \beta \operatorname{Re}(\overline{w} z) + \gamma = 0$ by taking $\alpha = 1$, $\gamma = |z_0|^2 - r^2$ and: $\beta = -2$ and $w = z_0$, in case $z_0 \neq 0$, or $\beta = 0$ and $w = 1$, in case $z_0 = 0$. In all cases the choices of the parameters satisfy the restrictions: $\alpha, \beta, \gamma \in \mathbb{R}$, $w \in \mathbb{C}$, $w \neq 0$, $\alpha^2 + \beta^2 \neq 0$ and $\beta^2 |w|^2 > 4\alpha\gamma$.

This consideration of the equations of a line and a circle as special cases of one equation permits us to unify the notions of circle and line into the single notion of **generalized circle** in $\mathbb{C}$. If we attach the point ∞ to any line (and leave circles unchanged), then we are talking about generalized circles in $\widehat{\mathbb{C}}$. Look at exercise 1.3.2 for another interesting unification of the notions of circle and line: *generalized circles in* $\widehat{\mathbb{C}}$ *are the images of circles in* $\mathbb{S}^2$ *through stereographic projection*.

Now, an important property of every l.f.t. is that it maps generalized circles in $\widehat{\mathbb{C}}$ onto generalized circles in $\widehat{\mathbb{C}}$. To prove it we consider three special cases first.

Example 4.1.1. Every function $T(z) = z + b$ is a l.f.t. with $a = 1$, $c = 0$, $d = 1$ and, for an obvious reason, it is called **translation** by b.
Every such T is holomorphic in $\mathbb{C}$, one-to-one from $\mathbb{C}$ onto $\mathbb{C}$ and satisfies $T(\infty) = \infty$. It is trivial to prove that T maps lines in $\widehat{\mathbb{C}}$ onto lines in $\widehat{\mathbb{C}}$ and circles in $\mathbb{C}$ onto circles in $\mathbb{C}$.

Example 4.1.2. Every function $T(z) = az$ with $a \neq 0$ is a l.f.t. with $b = c = 0$, $d = 1$ and it is called **homothety** with center 0.
Every such T rotates points around 0 by the fixed angle $\arg a$. Indeed, if $w = T(z) = az$, then $\arg w = \arg z + \arg a$. Moreover, T multiplies distances between points by the fixed factor $|a|$. Indeed, if $w_1 = T(z_1) = az_1$ and $w_2 = T(z_2) = az_2$, then $|w_1 - w_2| = |a||z_1 - z_2|$.
Also T is holomorphic in $\mathbb{C}$, one-to-one from $\mathbb{C}$ onto $\mathbb{C}$, satisfies $T(\infty) = \infty$ and it is easy to prove that T maps lines in $\widehat{\mathbb{C}}$ onto lines in $\widehat{\mathbb{C}}$ and circles in $\mathbb{C}$ onto circles in $\mathbb{C}$.

Example 4.1.3. The function $T(z) = \frac{1}{z}$ is a l.f.t. with $a = d = 0$, $c = b = 1$ and it is called **inversion** with respect to the circle $\mathbb{T} = C_0(1)$.
The inversion T is holomorphic in $\widehat{\mathbb{C}} \setminus \{0\}$, one-to-one from $\widehat{\mathbb{C}} \setminus \{0, \infty\}$ onto $\widehat{\mathbb{C}} \setminus \{0, \infty\}$ and satisfies $T(0) = \infty$ and $T(\infty) = 0$. Moreover, it is easy to show that T maps (i) lines in $\widehat{\mathbb{C}}$ which do not contain 0 onto circles in $\mathbb{C}$ which contain 0, (ii) lines in $\widehat{\mathbb{C}}$ which contain 0 onto lines in $\widehat{\mathbb{C}}$ which contain 0, (iii) circles in $\mathbb{C}$ which contain 0 onto lines in $\widehat{\mathbb{C}}$ which do not contain 0 and (iv) circles in $\mathbb{C}$ which do not contain 0 onto circles in $\mathbb{C}$ which do not contain 0.

Lemma 4.1. *Every l.f.t. is a composition of finitely many translations, homotheties and inversions.*

Proof. Let $T(z) = \frac{az+b}{cz+d}$.
If $c = 0$, then $T(z) = a'z + b'$, where $a' = \frac{a}{d} \neq 0$ and $b' = \frac{b}{d}$. If we consider the homothety $T_1(z) = a'z$ and the translation $T_2(z) = z + b'$, then $T = T_2 \circ T_1$.
If $c \neq 0$, then

$$T(z) = \frac{\frac{a}{c}(cz+d) + (b - \frac{ad}{c})}{cz+d} = \frac{a}{c} + \frac{bc-ad}{c^2} \frac{1}{z + \frac{d}{c}}.$$

If we consider the translation $T_1(z) = z + \frac{d}{c}$, the inversion $T_2(z) = \frac{1}{z}$, the homothety $T_3(z) = \frac{bc-ad}{c^2} z$ and the translation $T_4(z) = z + \frac{a}{c}$, then $T = T_4 \circ T_3 \circ T_2 \circ T_1$. $\square$

Proposition 4.1. *Every l.f.t. maps generalized circles in $\widehat{\mathbb{C}}$ onto generalized circles in $\widehat{\mathbb{C}}$.*

Proof. A corollary of lemma 4.1 and of the examples 4.1.1, 4.1.2 and 4.1.3. $\square$

Proposition 4.2. *Take the distinct $z_1, z_2, z_3 \in \widehat{\mathbb{C}}$ and the distinct $w_1, w_2, w_3 \in \widehat{\mathbb{C}}$. Then there is a unique l.f.t. T so that $T(z_j) = w_j$ for $j = 1, 2, 3$.*

Proof. We consider the l.f.t. S which, depending on whether one of z_1, z_2, z_3 is ∞ or not, has the formula

$$S(z) = \begin{cases} \frac{z_2-z_3}{z_2-z_1}\frac{z-z_1}{z-z_3}, & \text{if } z_1 \neq \infty, z_2 \neq \infty, z_3 \neq \infty \\ \frac{z-z_1}{z_2-z_1}, & \text{if } z_3 = \infty \\ \frac{z-z_1}{z-z_3}, & \text{if } z_2 = \infty \\ \frac{z_2-z_3}{z-z_3}, & \text{if } z_1 = \infty \end{cases}$$

The l.f.t. S has values: $S(z_1) = 0$, $S(z_2) = 1$, $S(z_3) = \infty$.
There is a similar l.f.t. R with values: $R(w_1) = 0$, $R(w_2) = 1$, $R(w_3) = \infty$.
Then the l.f.t. $T = R^{-1} \circ S$ has values: $T(z_1) = w_1, T(z_2) = w_2, T(z_3) = w_3$.
To prove the uniqueness of T with $T(z_1) = w_1, T(z_2) = w_2, T(z_3) = w_3$ we consider the previous l.f.t S, R and then the l.f.t. $Q = R \circ T \circ S^{-1}$ has values: $Q(0) = 0$, $Q(1) = 1$, $Q(\infty) = \infty$. Since $Q(\infty) = \infty$, we get that Q has the form $Q(z) = az + b$ with $a \neq 0$. Now from $Q(0) = 0$, $Q(1) = 1$ we find $a = 1$, $b = 0$ and hence Q is the identity l.f.t. id with $id(z) = z$. Thus $R \circ T \circ S^{-1} = id$ and hence $T = R^{-1} \circ S$. $\square$

When we apply the previous results we should bear in mind that every three distinct points in $\widehat{\mathbb{C}}$ belong to a unique generalized circle in $\widehat{\mathbb{C}}$.

Example 4.1.4. The l.f.t. which maps the triple $i, 2, 1$ onto the triple $0, 1, \infty$ is

$$w = T(z) = \tfrac{2-1}{2-i}\tfrac{z-i}{z-1} = \tfrac{2+i}{5}\tfrac{z-i}{z-1} = \tfrac{(2+i)z+(1-2i)}{5z-5}.$$

The points $i, 2, 1$ in the z-plane are not co-linear and hence belong to a circle A. The points $0, 1, \infty$ belong to the line $B = \mathbb{R} \cup \{\infty\}$ in $\widehat{\mathbb{C}}$. Now, T maps the circle A in the z-plane onto some generalized circle $T(A)$ in the w-plane. Since A contains $i, 2, 1$, $T(A)$ must contain the images of $i, 2, 1$, i.e. $0, 1, \infty$. Thus $T(A) = B$.
If we want to determine the circle $A = C_{z_0}(r)$ which contains $i, 2, 1$, we have to find z_0, r so that $i, 2, 1$ satisfy the equation $|z - z_0| = r$: we just solve a system of three equations in three real unknowns: x_0, y_0, r. But there is a second and probably easier way to find the equation of A. Indeed, w belongs to $\mathbb{R}$ if and only if $\operatorname{Im} w = 0$ and, using simple algebra, we see that this is equivalent to $|z - \frac{3}{2}(1 + i)|^2 = \frac{5}{2}$, $z \neq 1$. Since $z = 1$ is mapped onto $w = \infty$, we have that w belongs to B if and only if z belongs to the circle $C_{3(1+i)/2}\big(\sqrt{5/2}\big)$. We conclude that $A = C_{3(1+i)/2}\big(\sqrt{5/2}\big)$.

Exercises.

4.1.1. Find l.f.t. T so that $T(1) = i, T(i) = 0, T(-1) = -i$. Find $T(\mathbb{T})$ and $T(\mathbb{D})$.

4.1.2. Find l.f.t. T so that $T(\mathbb{D}) = \{z \mid \operatorname{Im} z > 0\}, T(i) = 1, T(1) = 0, T(a) = -1$, where $a \in \mathbb{T}$. Can a be an arbitrary point of $\mathbb{T}$?

4.1.3. (i) Let $T_1(z) = \frac{a_1z+b_1}{c_1z+d_1}$ and $T_2(z) = \frac{a_2z+b_2}{c_2z+d_2}$. Prove that T_1, T_2 are the same function if and only if there is $\lambda \neq 0$ so that $a_2 = \lambda a_1$, $b_2 = \lambda b_1$, $c_2 = \lambda c_1$, $d_2 = \lambda d_1$.
(ii) Prove that every l.f.t. T can take the form $T(z) = \frac{az+b}{cz+d}$ with $ad - bc = 1$.

4.1.4. Let A be a generalized circle of the z-plane $\widehat{\mathbb{C}}$ and B be a generalized circle of the w-plane $\widehat{\mathbb{C}}$. Then, in an obvious way, A splits $\widehat{\mathbb{C}}$ into two disjoint sets A_+ and A_- and, similarly, B splits $\widehat{\mathbb{C}}$ into two disjoint sets B_+ and B_-. Now, let T be a l.f.t. and let $T(A) = B$. Assume that $z_0 \in A_+$ and $w_0 = T(z_0) \in B_+$. Prove that $T(A_+) = B_+$ and $T(A_-) = B_-$.

4.1.5. A point $z \in \widehat{\mathbb{C}}$ is called **fixed point** of the l.f.t. T if $T(z) = z$. If the l.f.t. T is not the identity (in which case T has infinitely many fixed points), prove that T has either one or two fixed points in $\widehat{\mathbb{C}}$. In each case, which are the images through T of the generalized circles which contain its fixed points?
Apply the above to each of $T(z) = z + 2$, $T(z) = 2z - 1$, $T(z) = \frac{z-1}{z+1}$, $T(z) = \frac{3z-4}{z-1}$.

4.1.6. (i) The points $a, b \in \widehat{\mathbb{C}}$ are called **symmetric** with respect to $C_{z_0}(r)$ if either $a = z_0$, $b = \infty$ or $a = \infty$, $b = z_0$ or $a, b \in \mathbb{C}$ are on the same halfline with vertex z_0 and $|a - z_0||b - z_0| = r^2$. Observe that either a, b coincide with one and the same point of $C_{z_0}(r)$ or a, b are on different sides of $C_{z_0}(r)$. Given $a \in \widehat{\mathbb{C}} \setminus \{z_0, \infty\}$, describe a geometric construction "with ruler and compass" of its symmetric point, $b \in \widehat{\mathbb{C}} \setminus \{z_0, \infty\}$, with respect to $C_{z_0}(r)$. Prove that a, b are symmetric with respect to $C_{z_0}(r)$ if and only if

$$b = z_0 + \tfrac{r^2}{\overline{a} - \overline{z_0}}.$$

(ii) The points $a, b \in \widehat{\mathbb{C}}$ are called **symmetric** with respect to the line $\widehat{l} = l \cup \{\infty\}$ in $\widehat{\mathbb{C}}$ if either $a = b = \infty$ or $a, b \in \mathbb{C}$ are symmetric with respect to l. Prove that a, b are symmetric with respect to $\widehat{l}$ if and only if

$$b = z_1 + \tfrac{z_2 - z_1}{\overline{z_2} - \overline{z_1}}(\overline{a} - \overline{z_1}),$$

where z_1, z_2 are any two distinct fixed points of the line l.
(iii) We take a l.f.t. $w = T(z)$ and generalized circles A in the z-plane $\widehat{\mathbb{C}}$ and B in the w-plane $\widehat{\mathbb{C}}$. Prove that, if T maps A onto B, then T maps symmetric points with respect to A onto symmetric points with respect to B.
(iv) Find l.f.t. T so that $T(C_0(1)) = C_i(3)$, $T(i) = 3 + i$, $T(\frac{1}{2}) = 0$.

4.1.7. The l.f.t. $w = T(z)$ is called **real** if it maps the real line (with ∞) in the z-plane $\widehat{\mathbb{C}}$ onto the real line (with ∞) in the w-plane $\widehat{\mathbb{C}}$.
(i) Prove that the l.f.t. T is real if and only if there are $a, b, c, d \in \mathbb{R}$ with $ad - bc \neq 0$ so that $T(z) = \frac{az+b}{cz+d}$.
(ii) If the l.f.t. T is real and $T(z) = \frac{az+b}{cz+d}$, with $a, b, c, d \in \mathbb{R}$, $ad - bc \neq 0$, we define $\operatorname{sign} T$ to be the sign of $ad - bc$. Using exercise 4.1.3(i), prove that $\operatorname{sign} T$ is well defined.
(iii) Prove that, if the l.f.t. T is real, then T^{-1} is real, and that, if the l.f.t. S, T are real, then $S \circ T$ is real. Also prove that

$$\operatorname{sign} T^{-1} = \operatorname{sign} T, \quad \operatorname{sign}(S \circ T) = \operatorname{sign} S \, \operatorname{sign} T.$$

(iv) Take a real l.f.t. T. Prove that T maps the upper halfplane onto the upper halfplane (and the lower onto the lower) if and only if $\operatorname{sign} T = +1$ and that T maps the upper halfplane onto the lower halfplane (and the lower onto the upper) if and only if $\operatorname{sign} T = -1$.

4.1.8. (i) Let $z_0 \in \mathbb{D}$ and $|\lambda| = 1$ and consider the l.f.t.

$$T(z) = \lambda \tfrac{z - z_0}{1 - \overline{z_0} z}.$$

Prove that $T(\mathbb{T}) = \mathbb{T}$ and $T(z_0) = 0$. Find $T(\mathbb{D})$.
(ii) Let $z_0 \in \mathbb{D}$ and let T be a l.f.t. such that $T(\mathbb{T}) = \mathbb{T}$ and $T(z_0) = 0$. Prove that there is λ with $|\lambda| = 1$ so that $T(z) = \lambda \frac{z - z_0}{1 - \overline{z_0} z}$.
(iii) Let $a, b \in \mathbb{D}$ and let T be a l.f.t. such that $T(\mathbb{T}) = \mathbb{T}$ and $T(a) = b$. Prove that there is λ with $|\lambda| = 1$ so that $\frac{T(z) - b}{1 - \overline{b} T(z)} = \lambda \frac{z - b}{1 - \overline{a} z}$.

4.1.9. Consider $\mathbb{H}_+ = \{z \mid \text{Im}\, z > 0\}$.
(i) Let $z_0 \in \mathbb{H}_+$ and $|\lambda| = 1$ and consider the l.f.t.

$$T(z) = \lambda \tfrac{z-z_0}{z-\overline{z_0}}.$$

Prove that $T(\mathbb{R} \cup \{\infty\}) = \mathbb{T}$ and $T(z_0) = 0$. Find $T(\mathbb{H}_+)$.
(ii) Let $z_0 \in \mathbb{H}_+$ and let T be a l.f.t. such that $T(\mathbb{R} \cup \{\infty\}) = \mathbb{T}$ and $T(z_0) = 0$. Prove that there is λ with $|\lambda| = 1$ so that $T(z) = \lambda \frac{z-z_0}{z-\overline{z_0}}$.

4.1.10. Consider distinct $z_1, z_2, z_3, z_4 \in \widehat{\mathbb{C}}$. We define the **double ratio** of z_1, z_2, z_3, z_4 (in this order) to be

$$(z_1, z_2, z_3, z_4) = \begin{cases} \frac{z_1-z_3}{z_1-z_4}\frac{z_2-z_4}{z_2-z_3}, & \text{if } z_1 \neq \infty, z_2 \neq \infty, z_3 \neq \infty, z_4 \neq \infty \\ \frac{z_2-z_4}{z_2-z_3}, & \text{if } z_1 = \infty \\ \frac{z_1-z_3}{z_1-z_4}, & \text{if } z_2 = \infty \\ \frac{z_2-z_4}{z_1-z_4}, & \text{if } z_3 = \infty \\ \frac{z_1-z_3}{z_2-z_3}, & \text{if } z_4 = \infty \end{cases}$$

(i) Prove that
$$(T(z_1), T(z_2), T(z_3), T(z_4)) = (z_1, z_2, z_3, z_4)$$
for every l.f.t. T and every distinct $z_1, z_2, z_3, z_4 \in \widehat{\mathbb{C}}$.
(ii) Prove that the distinct $z_1, z_2, z_3, z_4 \in \widehat{\mathbb{C}}$ belong to the same generalized circle if and only if $(z_1, z_2, z_3, z_4) \in \mathbb{R} \setminus \{0\}$.
(iii) If $(z_1, z_2, z_3, z_4) = \lambda$, find all values (depending on λ) which result from this double ratio after all rearrangements of z_1, z_2, z_3, z_4.

4.1.11. Prove that the group of all l.f.t. is simple, i.e. that its only normal subgroups are itself and $\{I\}$, where I is the identity l.f.t.

4.2 The exponential function.

We define the **exponential function** $\exp : \mathbb{C} \to \mathbb{C}$ by

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

for every $z = x + iy$.

If $z \in \mathbb{R}$, i.e. $z = x + i0$, then $\exp z = e^x(\cos 0 + i \sin 0) = e^x = e^z$. This implies that we may use the symbol e^z instead of $\exp z$ without the danger of contradiction, in the case that z is real, between the symbol e^z as we just defined it and the symbol e^z as we know it from infinitesimal calculus. Therefore, we define

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

for every $z = x + iy$.

Since $z = x + iy$ implies $|e^z| = |e^x||\cos y + i \sin y| = e^x$, we have that

$$|e^z| = e^{\text{Re}\, z}.$$

From $e^z = e^x(\cos y + i \sin y)$ and $|e^z| = e^x$ we get $e^z = |e^z|(\cos y + i \sin y)$. So y is one of the elements of $\arg e^z$ and hence

$$\arg e^z = \{\text{Im}\, z + k2\pi \mid k \in \mathbb{Z}\}.$$

We have the basic equality

$$e^{z_1}e^{z_2} = e^{z_1+z_2}.$$

Indeed, $e^{x_1}(\cos y_1 + i\sin y_1)e^{x_2}(\cos y_2 + i\sin y_2) = e^{x_1+x_2}(\cos(y_1+y_2) + i\sin(y_1+y_2))$ from the addition formulas of cos and sin.

If $z_2 - z_1 = k2\pi i$ for some $k \in \mathbb{Z}$, then $e^{z_2} = e^{z_1}e^{k2\pi i} = e^{z_1}(\cos(k2\pi) + i\sin(k2\pi)) = e^{z_1}$. Conversely, assume $e^{z_2} = e^{z_1}$ and let $z_2 - z_1 = x + iy$. Then $e^x(\cos y + i\sin y) = e^{z_2 - z_1} = \frac{e^{z_2}}{e^{z_1}} = 1$ and hence $e^x = 1$, $\cos y = 1$ and $\sin y = 0$. Therefore, $x = 0$ and $y = k2\pi$ for some $k \in \mathbb{Z}$. Thus, $z_2 - z_1 = k2\pi i$ with $k \in \mathbb{Z}$. We proved that

$$e^{z_2} = e^{z_1} \quad \Leftrightarrow \quad z_2 - z_1 = k2\pi i \text{ for some } k \in \mathbb{Z}.$$

For all $z = x + iy$ we have $|e^z| = e^x > 0$ and hence

$$e^z \neq 0.$$

On the other hand, if we take any $w \neq 0$ and if we use the notation

$$\ln : (0, +\infty) \to \mathbb{R}$$

for the well known logarithmic function from infinitesimal calculus, then the solutions of the equation $e^z = w$ are described as follows:

$$e^z = w \quad \Leftrightarrow \quad z = \ln|w| + iy \text{ for some } y \in \arg w.$$

Indeed, if we write $z = x + iy$, then the equality $w = e^z$ becomes $w = e^x(\cos y + i\sin y)$ and it just means that its right side is one of the polar representations of w. Hence, $w = e^z$ if and only if $e^x = |w|$ and y is a value of $\arg w$. Now, $e^x = |w|$ is equivalent to $x = \ln|w|$. Therefore, the equation $e^z = w$ has these infinitely many solutions: $z = \ln|w| + iy$ where y is any value of $\arg w$. All these solutions have the same real part, $x = \ln|w|$, and their imaginary parts are the elements of $\arg w$.

From what we said already, it is clear that the exponential function is onto $\mathbb{C} \setminus \{0\}$ but not one-to-one in $\mathbb{C}$. In fact the exponential function is *infinity-to-one* since there are infinitely many values of z corresponding to the same value of $w \neq 0$.

Based on the equality $e^{iy} = \cos y + i\sin y$, we may write the polar representations of any $z \neq 0$ in an equivalent form:

$$z = r(\cos\theta + i\sin\theta) = re^{i\theta},$$

where $r = |z|$ and $\theta \in \arg z$. The second form is simpler and we shall use it extensively in the rest of the course. For instance, we may rewrite the examples 2.2.8 and 2.2.9 as follows.

Example 4.2.1. Using the parametric equation $z = \gamma(t) = z_0 + re^{it}$, $t \in [0, 2\pi]$, for the circle $C_{z_0}(r)$, we have

$$\oint_{C_{z_0}(r)} f(z)\,dz = \oint_\gamma f(z)\,dz = \int_0^{2\pi} f(z_0 + re^{it})ire^{it}\,dt.$$

Example 4.2.2. If $n \in \mathbb{Z}$, we have $\int_0^{2\pi} e^{int}\,dt = 2\pi$, if $n = 0$, and $\int_0^{2\pi} e^{int}\,dt = 0$, if $n \neq 0$. Therefore, if $n \in \mathbb{Z}$, we get

$$\oint_{C_{z_0}(r)} (z - z_0)^n\,dz = \int_0^{2\pi} r^n e^{int} ire^{it}\,dt = ir^{n+1}\int_0^{2\pi} e^{i(n+1)t}\,dt = \begin{cases} 2\pi i, & \text{if } n = -1 \\ 0, & \text{if } n \neq -1 \end{cases}$$

The real and imaginary parts of e^z are $u(x,y) = e^x\cos y$ and $v(x,y) = e^x\sin y$. Therefore, u, v have partial derivatives $\frac{\partial u}{\partial x}(x,y) = e^x\cos y$, $\frac{\partial u}{\partial y} = -e^x\sin y$, $\frac{\partial v}{\partial x} = e^x\sin y$, $\frac{\partial v}{\partial y} = e^x\cos y$, which are continuous and satisfy the system of C-R equations in $\mathbb{C}$ and hence e^z is holomorphic

in $\mathbb{C}$. To calculate the derivative of e^z we write $\frac{\partial u}{\partial x}(x,y) + i\frac{\partial v}{\partial x}(x,y) = e^x\cos y + ie^x\sin y = e^x(\cos y + i\sin y)$ and hence
$$\frac{d\,e^z}{dz} = e^z.$$

We shall now examine the mapping properties of the function $w = e^z$. We write $z = x + iy$ and $w = u + iv$.

If $z = x + iy$ varies on the horizontal line h_y in the z-plane which intersects the y-axis at the fixed point iy, then $w = e^z = e^x(\cos y + i\sin y)$ varies on the halfline r_y in the w-plane with vertex 0 (without 0) which forms angle y with the positive u-semiaxis. Also, if z varies on the horizontal line h_y from left to right, i.e. when x increases from $-\infty$ to $+\infty$, then $w = e^z$ varies on the halfline r_y from 0 to ∞. If y increases by $\Delta y > 0$, i.e. if the horizontal line h_y moves upward, then the corresponding halfline r_y rotates in the positive direction around 0 by the angle Δy. The two horizontal lines h_y and $h_{y+2\pi}$ are mapped onto the same halfline $r_y = r_{y+2\pi}$.

If the point $z = x + iy$ varies on the vertical line v_x in the z-plane which intersects the x-axis at the fixed point x, then $w = e^z = e^x(\cos y + i\sin y)$ varies on the circle $C_0(e^x)$, call it c_x, in the w-plane. Also, if z moves upward on the vertical line v_x, i.e. if y increases from $-\infty$ to $+\infty$, then $w = e^z$ rotates on the circle c_x infinitely many times in the positive direction. If y increases over an interval of length 2π, then $w = e^z$ describes the whole circle c_x once in the positive direction. If x increases by $\Delta x > 0$, i.e. if the vertical line v_x moves to the right, then the circle c_x with radius e^x becomes the circle $c_{x+\Delta x}$ with radius $e^{x+\Delta x} = e^x e^{\Delta x}$.

We may combine the above results. For instance, if we consider the open rectangle
$$\Pi = \{x + iy \,|\, x_1 < x < x_2, y_1 < y < y_2\}$$
in the z-plane with sides parallel to the two coordinate axes, then Π is the intersection of the open horizontal zone between the lines h_{y_1} and h_{y_2} and the open vertical zone between the lines v_{x_1} and v_{y_2}. If $y_2 - y_1 < 2\pi$, then Π is mapped onto the open "circular rectangle"
$$R = \{re^{i\theta} \,|\, e^{x_1} < r < e^{x_2}, y_1 < \theta < y_2\},$$
in the w-plane, which is the intersection of the angular region between the halflines r_{y_1} and r_{y_2} and the open ring between the circles c_{x_1} and c_{x_2}. If $y_2 - y_1 = 2\pi$, then the "circular rectangle" R is the open ring between the circles c_{x_1} and c_{x_2} without its linear segment which belongs to the halfline $r_{y_1} = r_{y_2}$. Of course, in this case, if Π includes at least one of its horizontal sides, then its image R is the whole open ring between the circles c_{x_1} and c_{x_2}.

Starting from $e^{iy} = \cos y + i\sin y$ and $e^{-iy} = \cos(-y) + i\sin(-y) = \cos y - i\sin y$, we easily find that $\cos y = \frac{1}{2}(e^{iy} + e^{-iy})$ and $\sin y = \frac{1}{2i}(e^{iy} - e^{-iy})$ for every $y \in \mathbb{R}$. Now we extend the trigonometric functions **cosine** and **sine** from $\mathbb{R}$ to $\mathbb{C}$ by defining
$$\cos z = \tfrac{1}{2}(e^{iz} + e^{-iz}), \quad \sin z = \tfrac{1}{2i}(e^{iz} - e^{-iz})$$
for every $z \in \mathbb{C}$. It is clear from the holomorphy of the exponential function that cos and sin are holomorphic in $\mathbb{C}$ and that
$$\frac{d\cos z}{dz} = -\sin z, \qquad \frac{d\sin z}{dz} = \cos z.$$

It is also easy to show that cos and sin are 2π-periodic.

Now we extend the **tangent** and the **cotangent** from $\mathbb{R}$ to $\mathbb{C}$ by defining
$$\tan z = \frac{\sin z}{\cos z} = \frac{e^{iz} - e^{-iz}}{e^{iz} + e^{-iz}}, \qquad \cot z = \frac{\cos z}{\sin z} = \frac{e^{iz} + e^{-iz}}{e^{iz} - e^{-iz}}$$
for every $z \in \mathbb{C}$. It is easy to see that the solutions of $\cos z = 0$ are $z = \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$, and the solutions of $\sin z = 0$ are $z = k\pi$, $k \in \mathbb{Z}$. Therefore, tan is defined and holomorphic in the open set $\mathbb{C} \setminus \{\frac{\pi}{2} + k\pi \,|\, k \in \mathbb{Z}\}$ and cot is defined and holomorphic in the open set $\mathbb{C} \setminus \{k\pi \,|\, k \in \mathbb{Z}\}$. Both functions are π-periodic.

Exercises.

4.2.1. Prove that $\overline{e^z} = e^{\overline{z}}$ for all z.

4.2.2. Prove that $|e^z - 1| \leq e^{|z|} - 1 \leq |z|e^{|z|}$.

4.2.3. Let $z \to \infty$ on any halfline. Depending on the halfline, study the existence of the $\lim e^z$ in $\widehat{\mathbb{C}}$. Which characteristic of the halfline determines the existence and the value of the limit?

4.2.4. Find the images through the exponential function of:

$$\{x + iy \,|\, a < x < b, \theta < y < \theta + \pi\}, \quad \{x + iy \,|\, a < x < b, \theta < y < \theta + 2\pi\},$$
$$\{x + iy \,|\, x < b, \theta < y < \theta + \pi\}, \quad \{x + iy \,|\, x < b, \theta < y < \theta + 2\pi\},$$
$$\{x + iy \,|\, a < x, \theta < y < \theta + \pi\}, \quad \{x + iy \,|\, a < x, \theta < y < \theta + 2\pi\}.$$

4.2.5. Every horizontal and every vertical line in the z-plane are perpendicular. Also, every halfline with vertex 0 and every circle with center 0 in the w-plane are perpendicular. How do these facts relate to the conformality of the function $w = e^z$?

4.2.6. Prove that
(i) $\sin^2 z + \cos^2 z = 1$.
(ii) $\sin(z + w) = \sin z \cos w + \cos z \sin w$, $\cos(z + w) = \cos z \cos w - \sin z \sin w$.
(iii) $|\cos(x+iy)|^2 = \cos^2 x + \sinh^2 y$, $|\sin(x+iy)|^2 = \sin^2 x + \sinh^2 y$ where $\sinh y = \frac{1}{2}(e^y - e^{-y})$.

4.2.7. Study the function $w = \sin z$ in the vertical zone $\{x + iy \,|\, -\frac{\pi}{2} < x < \frac{\pi}{2}\}$ and the function $w = \cos z$ in the vertical zone $\{x + iy \,|\, 0 < x < \pi\}$. Examine the images through these functions of the various horizontal linear segments (of length π) and the various vertical lines inside these two vertical zones.

4.3 Branches of the logarithmic function.

In the last section we proved, for every $w \neq 0$, the equivalence

$$e^z = w \quad \Leftrightarrow \quad z = \ln|w| + iy \text{ for some } y \in \arg w.$$

For every $w \neq 0$ we consider the set

$$\log w = \{\ln|w| + iy \,|\, y \in \arg w\}$$

and we call it **logarithm** of w. So the elements of $\log w$ are the solutions of $e^z = w$, i.e.

$$e^z = w \quad \Leftrightarrow \quad z \in \log w.$$

If we take $y = \operatorname{Arg} w$, then we get the particular element

$$\operatorname{Log} w = \ln|w| + i \operatorname{Arg} w$$

of $\log w$ and this is called **principal logarithm** of w.

If $r = |w|$ and if θ is any of the values of the argument of w, i.e. if $w = r(\cos\theta + i\sin\theta) = re^{i\theta}$ is any of the polar representations of w, then the values of $\arg w$ are the numbers $\theta + k2\pi$, $k \in \mathbb{Z}$. Hence the values of $\log w$ are the numbers $\ln r + i(\theta + k2\pi)$, $k \in \mathbb{Z}$.

Example 4.3.1. (i) $\operatorname{Log} 1 = 0$ and $\log 1 = \{i2k\pi \,|\, k \in \mathbb{Z}\}$.
(ii) $\operatorname{Log}(-1) = i\pi$ and $\log(-1) = \{i(2k+1)\pi \,|\, k \in \mathbb{Z}\}$.
(iii) $\operatorname{Log} i = i\frac{\pi}{2}$ and $\log i = \{i(2k + \frac{1}{2})\pi \,|\, k \in \mathbb{Z}\}$.
(iv) $\operatorname{Log}(-3i) = \ln 3 - i\frac{\pi}{2}$ and $\log(-3i) = \{\ln 3 + i(2k - \frac{1}{2})\pi \,|\, k \in \mathbb{Z}\}$.
(v) $\operatorname{Log}(1 + i) = \ln\sqrt{2} + i\frac{\pi}{4}$ and $\log(1 + i) = \{\ln\sqrt{2} + i(2k + \frac{1}{4})\pi \,|\, k \in \mathbb{Z}\}$.
(vi) $\operatorname{Log}(1 - i\sqrt{3}) = \ln 2 - i\frac{\pi}{3}$ and $\log(1 - i\sqrt{3}) = \{\ln 2 + i(2k - \frac{1}{3})\pi \,|\, k \in \mathbb{Z}\}$.

For any fixed $w \neq 0$ the set $\log w$ has infinitely many elements, and any two of them differ by an integral multiple of $i2\pi$. All elements of $\log w$ have the same real part $x = \ln|w|$, and hence they are on the same vertical line v_x with equation $x = \ln|w|$, and the vertical differences between them are the integral multiples of 2π. Therefore, every vertical segment of the line v_x, which has length 2π and includes only one of its endpoints, contains exactly one element of $\log w$. Moreover, every horizontal zone, which has vertical width 2π and includes only one of its boundary lines (either the upper or the lower one), contains exactly one element of $\log w$ for every $w \neq 0$. More precisely, if we consider any θ_0 and the horizontal zone

$$Z_{\theta_0} = \{x + iy \,|\, \theta_0 < y \leq \theta_0 + 2\pi\} \quad \text{or} \quad Z_{\theta_0} = \{x + iy \,|\, \theta_0 \leq y < \theta_0 + 2\pi\},$$

then Z_{θ_0} contains exactly one element of $\log w$: the one with imaginary part y equal to the (unique) $\theta \in \arg w$ satisfying $\theta_0 < \theta \leq \theta_0 + 2\pi$ or $\theta_0 \leq \theta < \theta_0 + 2\pi$, respectively. For instance, if we consider the special zone determined by $\theta_0 = -\pi$ which contains its upper boundary line, i.e.

$$Z_{-\pi} = \{x + iy \,|\, -\pi < y \leq \pi\},$$

then, for every $w \neq 0$, the unique element of $\log w$ which is contained in this zone is the principal logarithm $\operatorname{Log} w$.

Proposition 4.3. *For all $w_1, w_2 \neq 0$ we have*

$$\log(w_1 w_2) = \log w_1 + \log w_2.$$

By this we mean that the sum of any element of $\log w_1$ *and any element of* $\log w_2$ *is an element of* $\log(w_1 w_2)$ *and, conversely, any element of* $\log(w_1 w_2)$ *is the sum of an element of* $\log w_1$ *and an element of* $\log w_2$.

Proof. A corollary of proposition 1.1 and of the equality $\ln|w_1 w_2| = \ln|w_1| + \ln|w_2|$. □

It is already clear that the exponential function $w = \exp z = e^z$ from $\mathbb{C}$ onto $\mathbb{C} \setminus \{0\}$ is not one-to-one. Therefore, there is no inverse of the exponential function. If we want to produce some kind of inverse of the exponential function, we may take any w in the range $\mathbb{C} \setminus \{0\}$ of the function and select one value of z out of the infinitely many in $\mathbb{C}$ which satisfy the $e^z = w$. There are many instances of this method at a more elementary level. Let us consider for instance the function $y = x^2$ from $(-\infty, +\infty)$ onto $[0, +\infty)$, which is not one-to-one in $(-\infty, +\infty)$. We take any $y \in [0, +\infty)$ (the range of $y = x^2$) and find one x such that $x^2 = y$. There are exactly two such x: $x = \sqrt{y}$ and $x = -\sqrt{y}$. Therefore, one might say that we have only two choices for the inverse function: the choice $x = \sqrt{y}$ for every $y \in [0, +\infty)$ and the choice $x = -\sqrt{y}$ for every $y \in [0, +\infty)$. *But this is not correct.* We may choose $x = \sqrt{y}$ for some $y \in [0, +\infty)$ and $x = -\sqrt{y}$ for the remaining $y \in [0, +\infty)$. It is obvious that there are infinitely many such inverse functions, depending on the particular choice we make between $x = \sqrt{y}$ and $x = -\sqrt{y}$ for each value of y. Nevertheless, there is a criterion which reduces the number of our inverse functions to *exactly two*: the criterion of *continuity!* We observe that the last function, with the double formula, is not continuous. On the contrary, the function $x = \sqrt{y}$ for every $y \in [0, +\infty)$ and the function $x = -\sqrt{y}$ for every $y \in [0, +\infty)$ are both continuous. To prove that these are the only continuous inverse functions is a simple exercise in real analysis. Indeed, assume that there is some continuous inverse function $x = f(y)$ of $y = x^2$ defined in $[0, +\infty)$ (the range of $y = x^2$). I.e. $f : [0, +\infty) \to \mathbb{R}$ is continuous in $[0, +\infty)$ and $f(y)^2 = y$ for every $y \in [0, +\infty)$. Let there be $y_1, y_2 > 0$ with $y_1 \neq y_2$ such that $f(y_1) = \sqrt{y_1}$ and $f(y_2) = -\sqrt{y_2}$. Since f is continuous in the interval between y_1, y_2 and its values at the endpoints are opposite, there is some y in this interval so that: $f(y) = 0$. This is impossible, because $y > 0$ and either $f(y) = \sqrt{y} > 0$ or $f(y) = -\sqrt{y} < 0$. Therefore, there are no such $y_1, y_2 > 0$ and hence we have exactly two cases: either $f(y) = \sqrt{y}$ for every $y \geq 0$ or $f(y) = -\sqrt{y}$ for every $y \geq 0$. We may say that there are exactly two *continuous branches of the square root* in $[0, +\infty)$: the branch $x = \sqrt{y}$ and the branch $x = -\sqrt{y}$.

Now let us go back to the determination of possible inverses of the exponential function.

Let $A \subseteq \mathbb{C} \setminus \{0\}$. We say that the function f is a **continuous branch of log** in A if f is continuous in A and for every $w \in A$ we have that $f(w)$ is an element of $\log w$ or, equivalently,

$$e^{f(w)} = w$$

for every $w \in A$.

Proposition 4.4 gives many useful examples of continuous branches of the logarithm.

Proposition 4.4. *Let $\theta_0 \in \mathbb{R}$. We consider the set*

$$A_{\theta_0} = \{re^{i\theta} \,|\, 0 < r < +\infty, \theta_0 < \theta < \theta_0 + 2\pi\}$$

in the w-plane (i.e. $\mathbb{C}$ without the halfline with vertex 0 which forms angle θ_0 with the positive u-semiaxis, where $w = u + iv$) and the open horizontal zone

$$Z_{\theta_0} = \{x + iy \,|\, -\infty < x < +\infty, \theta_0 < y < \theta_0 + 2\pi\}$$

in the z-plane. We define the function $f : A_{\theta_0} \to Z_{\theta_0}$ as follows: for every $w \in A_{\theta_0}$ we take $f(w)$ to be the unique element of $\log w$ *in the zone Z_{θ_0}. Then f is continuous in A_{θ_0} and so it is a continuous branch of* $\log$ *in A_{θ_0}.*

Proof. Assume that f is not continuous at some w in A_{θ_0}. Then there is a sequence (w_n) in A_{θ_0} so that $w_n \to w$ and $f(w_n) \not\to f(w)$. This implies that there is $\delta > 0$ so that $|f(w_n) - f(w)| \geq \delta > 0$ for infinitely many n. These infinitely many n define a subsequence of (w_n). Now we ignore the rest of the sequence (w_n) and concentrate on the specific subsequence. For simplicity we rename the subsequence and call it (w_n) again. Therefore, we have a sequence (w_n) in A_{θ_0} such that

$$w_n \to w \qquad \text{and} \qquad |f(w_n) - f(w)| \geq \delta > 0 \tag{4.3}$$

for every n. We set $z = f(w) \in Z_{\theta_0}$ and $z_n = f(w_n) \in Z_{\theta_0}$ for every n. Then $e^z = w$ and $e^{z_n} = w_n$ for every n and (4.3) becomes

$$e^{z_n} \to e^z \qquad \text{and} \qquad |z_n - z| \geq \delta > 0 \tag{4.4}$$

for every n. The real parts of the z_n are equal to $\ln|w_n|$ and, since $\ln|w_n| \to \ln|w|$, the real parts of the z_n are bounded. Moreover, since $z_n \in Z_{\theta_0}$, the imaginary parts of the z_n are also bounded. Therefore, the sequence (z_n) is bounded and the Bolzano-Weierstrass theorem implies that there is a subsequence (z_{n_k}) so that $z_{n_k} \to z'$ for some z'. Since all z_{n_k} belong to Z_{θ_0}, we see that z' belongs to the closed zone $\overline{Z}_{\theta_0} = \{x + iy \,|\, -\infty < x < +\infty, \theta_0 \leq y \leq \theta_0 + 2\pi\}$. Taking the limit in (4.4), we get that $e^{z'} = e^z$ and $|z' - z| \geq \delta$. Therefore, z' and z differ by a non-zero integral multiple of $i2\pi$. But this is impossible, because z belongs to the open zone Z_{θ_0} and z' belongs to the closed zone $\overline{Z}_{\theta_0}$.
Thus f is continuous at every w in A_{θ_0}. $\square$

Our study of the mapping properties of the exponential function in the previous section gives the following information about the mapping properties of the continuous branch $f : A_{\theta_0} \to Z_{\theta_0}$ of log, which is defined in proposition 4.4: f maps the halflines in A_{θ_0} with vertex 0 (without 0) onto the horizontal lines in Z_{θ_0} and the circles with center 0 (without their point on the halfline which is excluded from A_{θ_0}) onto the vertical segments of Z_{θ_0}.

Choosing any real θ_0, we have defined a continuous branch of log in the subset A_{θ_0} of the w-plane, whose range is the zone Z_{θ_0} of the z-plane. If, instead of θ_0, we consider $\theta_0 + k2\pi$ with any $k \in \mathbb{Z}$, then the domain $A = A_{\theta_0 + k2\pi}$ *remains the same* but the range, i.e. the zone $Z_{\theta_0 + k2\pi}$, moves vertically by $k2\pi$. The various zones $Z_{\theta_0+k2\pi}$ are successive and cover the whole z-plane (except for their boundary lines with equations $y = \theta_0 + k2\pi$). We summarize:

If we exclude from the w-plane a halfline with vertex 0, then in the remaining open set A there are infinitely many continuous branches of log *defined. Each of them maps A onto some open horizontal zone of the z-plane of width 2π. These various open zones, which correspond to the various continuous branches of* log *(in the same set A), are mutually disjoint, successive and cover the z-plane (except for their boundary lines). Of course, if we change the original halfline which determines the set A, then the corresponding zones and the corresponding continuous branches of* log *also change.*

Example 4.3.2. One particular example of a continuous branch of log is defined when we choose $\theta_0 = -\pi$. Then the set $A_{-\pi} = \{re^{i\theta} \,|\, 0 < r < +\infty, -\pi < \theta < \pi\}$ is the w-plane without the negative u-semiaxis (where $w = u + iv$) and the range of the branch is the zone $Z_{-\pi} = \{x + iy \,|\, -\infty < x < +\infty, -\pi < y < \pi\}$. It is obvious that this branch is the function which maps every $w \in A_{-\pi}$ onto the principal value $z = \mathrm{Log}\, w$ of $\log w$. I.e. we get the so-called **principal branch of log**

$$\mathrm{Log} : A_{-\pi} \to Z_{-\pi}.$$

We must keep in mind that in the same set $A_{-\pi}$ of the w-plane, besides the principal branch, there are infinitely many other continuous branches of log defined. Each of them maps $A_{-\pi}$ in a corresponding zone $Z_{-\pi+k2\pi}$, with $k \in \mathbb{Z}$, which is $Z_{-\pi}$ moved vertically by $k2\pi$. This branch results from the principal branch Log by adding the constant $ik2\pi$ and its formula is $\mathrm{Log} + i2k\pi$.

Now, we introduce a slight generalization of the notion of the branch of log, i.e. we define the notion of the branch of $\log g$, where g is a more general function than the identity $g(w) = w$.

Let $A \subseteq \mathbb{C}$ and $g : A \to \mathbb{C} \setminus \{0\}$ be continuous in A. We say that the function f is a **continuous branch of log g** in A if f is continuous in A and for every $w \in A$ we have that $f(w)$ is an element of $\log g(w)$ or, equivalently,

$$e^{f(w)} = g(w)$$

for every $w \in A$.

Example 4.3.3. Let $g : A \to \mathbb{C} \setminus \{0\}$ be continuous in $A \subseteq \mathbb{C}$. If there is a continuous branch h of log in $g(A)$, then $f = h \circ g$ is a continuous branch of $\log g$ in A.
Indeed, $f = h \circ g$ is continuous in A and, since $e^{h(z)} = z$ for every $z \in g(A)$, we also have

$$e^{f(w)} = e^{h(g(w))} = g(w)$$

for every $w \in A$.
This is a standard way to produce continuous branches of $\log g$ when we know continuous branches of log in the range of g.
For instance, if $g(w) = w - w_0$ and $A = \mathbb{C} \setminus l$, where l is a halfline with vertex w_0, then $g(A) = \mathbb{C} \setminus l'$, where l' is the halfline with vertex 0 which is parallel to l. We know that there are infinitely many branches of log defined in $g(A)$ and hence there are infinitely many branches of $\log(w - w_0)$ defined in A.

Proposition 4.5. *Let $g : A \to \mathbb{C} \setminus \{0\}$ be continuous in $A \subseteq \mathbb{C}$ and let f be any continuous branch of* $\log g$ *in A. If w_0 is an interior point of A and g is differentiable at w_0, then f is differentiable at w_0 and $f'(w_0) = \frac{g'(w_0)}{g(w_0)}$. Hence, if g is holomorphic in the interior of A, then f is also holomorphic in the interior of A.*

Proof. We set $z_0 = f(w_0)$ and $z = f(w)$ for every $w \in A$. Then $e^{z_0} = g(w_0)$ and $e^z = g(w)$. Since f is continuous, $w \to w_0$ implies $z \to z_0$. Therefore, using the derivative of the exponential function at z_0, we see that

$$\tfrac{f(w)-f(w_0)}{w-w_0} = \tfrac{z-z_0}{e^z-e^{z_0}} \, \tfrac{g(w)-g(w_0)}{w-w_0} \to \tfrac{g'(w_0)}{e^{z_0}} = \tfrac{g'(w_0)}{g(w_0)}$$

when $w \to w_0$. Thus f is differentiable at w_0 and $f'(w_0) = \frac{g'(w_0)}{g(w_0)}$. $\square$

Therefore, if $g : A \to \mathbb{C} \setminus \{0\}$ is holomorphic in the open set A, *every continuous branch of* $\log g$ *can be called* **holomorphic branch** *of* $\log g$ *in* A.

Example 4.3.4. We have defined infinitely many continuous branches of log in the open set which results when we exclude any halfline with vertex 0 from the w-plane. All these branches are *holomorphic branches of* log. In particular the principal branch $\mathrm{Log} : A_{-\pi} \to Z_{-\pi}$ is holomorphic in $A_{-\pi}$.

Proposition 4.6. *Let* $g : A \to \mathbb{C} \setminus \{0\}$ *be continuous in* $A \subseteq \mathbb{C}$.
(i) If f_1 *is a continuous branch of* $\log g$ *in* A *and* $f_2 - f_1 = ik2\pi$ *in* A*, where* k *is a fixed integer, then* f_2 *is also a continuous branch of* $\log g$ *in* A.
(ii) If, morever, A *is connected and* f_1, f_2 *are continuous branches of* $\log g$ *in* A*, then* $f_2 - f_1 = ik2\pi$ *in* A*, where* k *is a fixed integer. In particular, if* $f_1(w_0) = f_2(w_0)$ *for some* $w_0 \in A$*, then* $f_1 = f_2$ *in* A.

Proof. (i) The continuity of f_1 in A implies the continuity of f_2 in A. We also have $e^{f_1(w)} = g(w)$ for every $w \in A$ and hence $e^{f_2(w)} = e^{f_1(w)+ik2\pi} = e^{f_1(w)}e^{ik2\pi} = g(w)$ for every $w \in A$. Therefore, f_2 is a continuous branch of $\log g$ in A.
(ii) We consider the function $k = \frac{1}{i2\pi}(f_2 - f_1)$. Since for every $w \in A$ both $f_2(w)$ and $f_1(w)$ are elements of $\log g(w)$, we have that $k(w)$ is an integer. Also, since both f_1, f_2 are continuous in A, k is continuous in A. Now, k is a continuous real function in the connected set A, and hence it has the intermediate value property. But since its only values are integers, it is constant in A. So there is a fixed integer k so that $\frac{1}{i2\pi}(f_2 - f_1) = k$ or, equivalently, $f_2 - f_1 = ik2\pi$ in A.
If $f_2(w_0) = f_1(w_0)$ for some $w_0 \in A$, then the integer k is 0 and we get that $f_2 = f_1$ in A. □

Thus, *if we know one continuous branch of* $\log g$ *in the connected set* A*, then we find every other possible continuous branch of* $\log g$ *in* A *by adding to the known branch an arbitrary constant of the form* $ik2\pi$ *with* $k \in \mathbb{Z}$.

Example 4.3.5. Let $A = A_{-\pi}$ be the w-plane without the negative u-semiaxis (where $w = u+iv$). We want to find a continuous branch of log in A having value $z = 0$ when $w = 1$.
We already know that the principal branch Log of the logarithm has value $z = \mathrm{Log}\, 1 = 0$ at $w = 1$. Since A is connected, there is no other such continuous branch of log in A.
Now, in the same set $A = A_{-\pi}$ we want to find a continuous branch of log taking the value $z = i4\pi$ at $w = 1$.
Since A is connected the branch we are looking for has the form $\mathrm{Log} + ik2\pi$ for some fixed integer k. We try $w = 1$ in this equality and get $k = 2$.

Example 4.3.6. Let $A = A_0 = \{re^{i\theta} \,|\, 0 < r < +\infty, 0 < \theta < 2\pi\}$ be the w-plane without the positive u-semiaxis (where $w = u + iv$). We want to find a continuous branch of log in A taking the value $z = i(\frac{\pi}{2} + 4\pi)$ at $w = i$.
We consider the horizontal zones in the z-plane which correspond to the set A: to each $k \in \mathbb{Z}$ corresponds the zone $Z_{0+k2\pi} = \{x + iy \,|\, -\infty < x < +\infty, k2\pi < y < 2\pi + k2\pi\}$. Now we choose the particular zone which contains the value $z = i(\frac{\pi}{2} + 4\pi)$. This zone corresponds to $k = 2$ and it is $Z_{4\pi} = \{x + iy \,|\, -\infty < x < +\infty, 4\pi < y < 6\pi\}$. Then a continuous branch f of log which maps A onto $Z_{4\pi}$ is given by $f(w) = \ln r + i\theta$, where $r = |w|$ and θ is the unique value of $\arg w$ which is contained in the interval $(4\pi, 6\pi)$. Since A is connected, there is no other such continuous branch of log in A.

Exercises.

4.3.1. Let $z \neq 0$. Prove that the only element of $\exp(\log z)$ is z and that the elements of $\log(\exp z)$ are $z + k2\pi i$, $k \in \mathbb{Z}$.

4.3.2. Let $0 < r_1 < r_2$. Find $\mathrm{Log}(A)$, if $A = \{w \,|\, r_1 \leq |w| \leq r_2\} \setminus [-r_2, -r_1]$.

4.3.3. Find the inverse images through the exponential function $w = e^z$ of the following sets:

$$\{re^{i\theta} \,|\, 1 < r < 3, -\tfrac{\pi}{2} < \theta < \tfrac{\pi}{2}\}, \quad \{re^{i\theta} \,|\, 1 < r, -\tfrac{\pi}{2} < \theta < \tfrac{\pi}{2}\},$$

$$\{re^{i\theta} \,|\, r < 1, -\tfrac{\pi}{2} < \theta < \tfrac{\pi}{2}\}, \quad \{re^{i\theta} \,|\, 1 < r < 3, -\pi < \theta < \pi\},$$

$$\{re^{i\theta} \,|\, 1 < r < 3, 0 < \theta < 2\pi\}, \quad \{re^{i\theta} \,|\, 1 < r < 3, -\tfrac{\pi}{2} < \theta < \tfrac{3\pi}{2}\}.$$

In which of these sets is the principal branch of log defined and which other continuous branches of log are defined in these sets? Which continuous branches of log are defined in the remaining sets? In any case write the formulas of the continuous branches of log as well as the image of each set through the corresponding continuous branches of log.

4.3.4. Work on the following in both cases: $\theta_0 = -\pi$ and $\theta_0 = 0$.
Consider A_{θ_0}, i.e. the w-plane without the halfline with vertex 0 which forms angle θ_0 with the positive u-semiaxis. Consider also θ_1, θ_2 with $\theta_0 < \theta_1 < \theta_2 < \theta_0 + 2\pi$ as well as r_1, r_2 with $0 < r_1 < r_2 < +\infty$. Draw the set $P = \{w = re^{i\theta} \,|\, r_1 < r < r_2, \theta_1 < \theta < \theta_2\}$ and its images through the various continuous branches of log in A_{θ_0}.

4.3.5. Let $P = \{re^{i\theta} \,|\, 1 < r < 2, -\frac{3\pi}{4} < \theta < \frac{3\pi}{4}\}$, $Q = \{w = re^{i\theta} \,|\, 1 < r < 2, \frac{\pi}{4} < \theta < \frac{7\pi}{4}\}$. We know that there is a continuous branch f of log in P and a continuous branch g of log in Q. Is it possible for f and g to coincide in $P \cap Q$?

4.3.6. Look back at exercise 1.2.1 and find all the possible values of $\mathrm{Log}(z_1 z_2) - \mathrm{Log}\, z_1 - \mathrm{Log}\, z_2$.

4.3.7. Prove that there is no continuous branch of log defined in any circle $C_0(r)$ and hence in any set A which contains such a circle.

4.3.8. Define $w^a = e^{a\, \mathrm{Log}\, w}$ for every $w \in D_1(1)$, and for every z prove that

$$\lim_{x \to +\infty} \left(1 + \tfrac{z}{x}\right)^x = e^z.$$

4.3.9. Let $A \subseteq \mathbb{C} \setminus \{0\}$. If A is connected and if f_1, f_2 are two different continuous branches of log in A, prove that $f_1(A) \cap f_2(A) = \emptyset$. (Observe how this result is confirmed by the special case of A being $\mathbb{C}$ without a halfline with vertex 0 in which case the various continuous branches of log in A map A onto disjoint horizontal zones.)

4.3.10. Let $a, b \in \mathbb{R}$ with $a < b$. Discuss the geometric meaning of the number

$$\mathrm{Arg}\, \tfrac{z-b}{z-a} = \mathrm{Im}(\mathrm{Log}\, \tfrac{z-b}{z-a})$$

for z in $\mathbb{H}_+ = \{z \,|\, \mathrm{Im}\, z > 0\}$. How does this number vary when z varies in $\mathbb{H}_+$? Find the geometric locus of the z in $\mathbb{H}_+$ for which $\mathrm{Arg}\, \frac{z-b}{z-a} = c$ is constant, $0 \leq c \leq \pi$.

4.4 Powers and branches of roots.

If $n \in \mathbb{N}$, $n \geq 2$, the function

$$w = z^n$$

is holomorphic in the z-plane $\mathbb{C}$ and we shall examine some mapping properties of this function.

We work with polar representations:

$$z = re^{i\theta}, \qquad w = r^n e^{in\theta}.$$

If $\theta \in \mathbb{R}$ is constant and r varies in $(0, +\infty)$, i.e. if z moves on the halfline r_θ in the z-plane with vertex 0 (without 0) which forms angle θ with the positive x-semiaxis, then $w = z^n$ moves on the halfline r_ϕ in the w-plane with vertex 0 (without 0) which forms angle $\phi = n\theta$ with the positive u-semiaxis. Also, if z moves on the halfline r_θ from 0 to ∞, then $w = z^n$ moves on the halfline r_ϕ from 0 to ∞. If θ increases by $\Delta\theta > 0$, i.e. if the halfline r_θ turns in the positive direction by an angle $\Delta\theta$, then the corresponding halfline r_ϕ turns in the positive direction by an angle $\Delta\phi = n\Delta\theta$. The two halflines r_θ and $r_{\theta+\frac{2\pi}{n}}$ are mapped onto the same halfline $r_\phi = r_{\phi+2\pi}$.

If $r \in (0, +\infty)$ is constant and θ varies in $\mathbb{R}$, i.e. if the point z moves on the circle $C_0(r)$ in the z-plane, then $w = z^n$ moves on the circle $C_0(r^n)$ in the w-plane. Also, if z rotates once on $C_0(r)$ in the positive direction, i.e. if θ increases in an interval of length 2π, then $w = z^n$ rotates n times on $C_0(r^n)$ in the positive direction. If θ increases in an interval of length $\frac{2\pi}{n}$, then $w = z^n$ rotates once on $C_0(r^n)$ in the positive direction. If r increases, i.e. if the circle $C_0(r)$ expands, then the corresponding circle $C_0(r^n)$ also expands.

In the following as well as in the whole course, we shall use the symbol $\sqrt[n]{x}$ only to denote the unique nonnegative n-th root of a nonnegative real number x.

If $n \in \mathbb{N}$, $n \geq 2$ and if we take any polar representation $w = Re^{i\Theta}$ of $w \neq 0$, then the equation $z^n = w$ has n solutions which are described as follows:

$$z^n = w = Re^{i\Theta} \quad \Leftrightarrow \quad z = \sqrt[n]{R}\, e^{i(\frac{\Theta}{n}+k\frac{2\pi}{n})} \text{ for some } k = 0, 1, \dots, n-1. \tag{4.5}$$

Indeed, if we write $z = re^{i\theta}$, then the equality $z^n = w$ becomes $r^n e^{in\theta} = Re^{i\Theta}$ and this is equivalent to $r^n = R$ and $n\theta = \Theta + k2\pi$ for some $k \in \mathbb{Z}$. Solving for r and θ, we find the solutions $z = \sqrt[n]{R}\, e^{i(\frac{\Theta}{n}+k\frac{2\pi}{n})}$, $k \in \mathbb{Z}$. It is trivial to see that two of these solutions are the same if and only if the corresponding values of k differ by a multiple of n and hence there are n distinct solutions corresponding to the values $0, 1, \dots, n-1$ of k. We easily see that the solutions of $z^n = w$ are the vertices of a regular n-gon inscribed in the circle $C_0(\sqrt[n]{R})$.

The set of the solutions of $z^n = w$, which appear in the right side of (4.5), is called **n-th root** of w and it is denoted $w^{\frac{1}{n}}$ or $w^{1/n}$, i.e.

$$w^{\frac{1}{n}} = \big\{ \sqrt[n]{R}\, e^{i(\frac{\Theta}{n}+k\frac{2\pi}{n})} \,\big|\, k = 0, 1, \dots, n-1 \big\},$$

where $w = Re^{i\Theta}$ is any polar representation of w. Thus, we have the equivalence

$$z^n = w \quad \Leftrightarrow \quad z \in w^{\frac{1}{n}}.$$

Of course, if $w = 0$, then the equation $z^n = w$ has the unique solution $z = 0$ and then we define $0^{\frac{1}{n}} = \{0\}$.

Example 4.4.1. The n-th root of 1 is called **n-th root of unity**.
Since $1 = 1e^{i0}$, the elements of the n-th root of unity are the numbers $e^{ik\frac{2\pi}{n}}$, $k = 0, 1, \dots, n-1$. Obviously, one of them is 1 and, if we denote $e^{i\frac{2\pi}{n}}$ by the symbol ω_n, we find that the elements of the n-th root of unity are the numbers

$$1, \omega_n, \omega_n^2, \dots, \omega_n^{n-1}.$$

This ω_n is called **principal n-th root of unity**.

We saw that, if $w \neq 0$, then $w^{\frac{1}{n}}$ has exactly n elements which are on the vertices of a regular n-gon inscribed in the circle $C_0(\sqrt[n]{|w|})$ of the z-plane. Therefore, every arc of this circle with central angle $\frac{2\pi}{n}$, which includes only one of its endpoints, contains exactly one of the elements of $w^{\frac{1}{n}}$. Thus, every angular set in the z-plane with vertex 0 and angle $\frac{2\pi}{n}$, which includes only one of its boundary halflines, contains, for every $w \neq 0$, exactly one element of $w^{\frac{1}{n}}$. In particular, if we consider any θ_0 and the angular set

$$A_{\theta_0} = \big\{ re^{i\theta} \,\big|\, r > 0, \theta_0 < \theta \leq \theta_0 + \tfrac{2\pi}{n} \big\} \quad \text{or} \quad A_{\theta_0} = \big\{ re^{i\theta} \,\big|\, r > 0, \theta_0 \leq \theta < \theta_0 + \tfrac{2\pi}{n} \big\},$$

then A_{θ_0} contains exactly one element of $w^{\frac{1}{n}}$.

Clearly, the function $w = z^n$ from $\mathbb{C} \setminus \{0\}$ onto $\mathbb{C} \setminus \{0\}$ is n-to-one and has no inverse. So we shall define *branches* of an inverse of $w = z^n$.

Let $A \subseteq \mathbb{C} \setminus \{0\}$. We say that the function f is a **continuous branch of $\mathbf{w}^{\frac{1}{n}}$** in A if f is continuous in A and for every $w \in A$ we have that $f(w)$ is an element of $w^{\frac{1}{n}}$ or, equivalently,

$$f(w)^n = w$$

for every $w \in A$.

Proposition 4.7 gives many examples of continuous branches of $w^{\frac{1}{n}}$.

Proposition 4.7. *Let $\phi_0 \in \mathbb{R}$. We consider the set*

$$A_{\phi_0} = \{se^{i\phi} \,|\, s > 0, \phi_0 < \phi < \phi_0 + 2\pi\}$$

in the w-plane (i.e. $\mathbb{C}$ without the halfline with vertex 0 which forms angle ϕ_0 with the positive u-semiaxis, where $w = u + iv$) and the angular region

$$B_{\phi_0/n} = \left\{re^{i\theta} \,\middle|\, r > 0, \tfrac{\phi_0}{n} < \theta < \tfrac{\phi_0}{n} + \tfrac{2\pi}{n}\right\}$$

in the z-plane. We define the function $f : A_{\phi_0} \to B_{\phi_0/n}$ as follows: for every $w \in A_{\phi_0}$ we take $f(w)$ to be the unique element of $w^{\frac{1}{n}}$ in the angular region $B_{\phi_0/n}$. Then f is continuous in A_{ϕ_0} and so it is a continuous branch of $w^{\frac{1}{n}}$ in A_{ϕ_0}.

Proof. Assume that f is not continuous at some w in A_{ϕ_0}. Then there is a sequence (w_k) in A_{ϕ_0} so that $w_k \to w$ and $f(w_k) \not\to f(w)$. Then there is $\delta > 0$ so that $|f(w_k) - f(w)| \geq \delta > 0$ for infinitely many k. These infinitely many k define a subsequence of (w_k). Now we ignore the rest of the sequence (w_k) and concentrate on the specific subsequence. For simplicity we rename the subsequence and call it (w_k) again. Therefore, we have a sequence (w_k) in A_{ϕ_0} such that

$$w_k \to w \qquad \text{and} \qquad |f(w_k) - f(w)| \geq \delta > 0 \tag{4.6}$$

for every k. We set $z = f(w) \in B_{\phi_0/n}$ and $z_k = f(w_k) \in B_{\phi_0/n}$ for every k. Then $z^n = w$ and $z_k^n = w_k$ for every k and (4.6) becomes

$$z_k^n \to z^n \qquad \text{and} \qquad |z_k - z| \geq \delta > 0 \tag{4.7}$$

for every k. Since $|z_k|^n \to |z|^n$ and hence $|z_k| \to |z|$, we get that the sequence (z_k) is bounded and the Bolzano-Weierstrass theorem implies that there is a subsequence (z_{k_m}) so that $z_{k_m} \to z'$ for some z'. Since all z_{k_m} belong to $B_{\phi_0/n}$, we have that z' belongs to the closed angular region $\overline{B}_{\phi_0/n} = \{z = re^{i\theta} \,|\, r \geq 0, \frac{\phi_0}{n} \leq \theta \leq \frac{\phi_0}{n} + \frac{2\pi}{n}\}$. Taking the limit in (4.7), we get $z'^n = z^n$ and $|z' - z| \geq \delta$. This is impossible, because z belongs to $B_{\phi_0/n}$ and z' belongs to $\overline{B}_{\phi_0/n}$.
Thus f is continuous at every w in A_{ϕ_0}. □

From the mapping properties of the function $w = z^n$ we get the following for the mapping properties of the continuous branch $f : A_{\phi_0} \to B_{\phi_0/n}$ of $w^{\frac{1}{n}}$, which is defined in proposition 4.7. The function f maps the halflines in A_{ϕ_0} with vertex 0 (without 0) onto the halflines in $B_{\phi_0/n}$ with vertex 0 (without 0) and the circular arcs in A_{ϕ_0} with center 0 onto the circular arcs in $B_{\phi_0/n}$ with center 0.

Choosing any real ϕ_0, we have defined a continuous branch of $w^{\frac{1}{n}}$ in the subset A_{ϕ_0} of the w-plane, whose range is the angular region $B_{\phi_0/n}$ of the z-plane. If, instead of ϕ_0, we consider $\phi_0 + k2\pi$ with any $k = 0, 1, \ldots, n-1$, then the set $A = A_{\phi_0 + k2\pi}$ *remains the same* but the range, i.e. the angular region $B_{(\phi_0 + k2\pi)/n}$, rotates by an angle $k\frac{2\pi}{n}$. The n angular regions $B_{(\phi_0 + k2\pi)/n}$ with $k = 0, 1, \ldots, n-1$ are successive and cover the z-plane (except for their n boundary halflines with vertex 0). We summarize:

If we exclude from the w-plane any halfline with vertex 0, then in the remaining open set A there are n continuous branches of $w^{\frac{1}{n}}$ defined. Each of them maps A onto some open angular region of the z-plane with vertex 0 and angle $\frac{2\pi}{n}$. These various angular regions, which correspond to the various continuous branches of $w^{\frac{1}{n}}$ (in the same set A), are mutually disjoint, successive and cover the z-plane (except for their boundary halflines). Of course, if we change the original halfline which determines the set A, then the corresponding angular regions and the corresponding branches of $w^{\frac{1}{n}}$ also change.

Example 4.4.2. We get a concrete example of a continuous branch of $w^{\frac{1}{n}}$ when we take $\phi_0 = -\pi$. Then the set $A_{-\pi} = \{se^{i\phi} \mid s > 0, -\pi < \phi < \pi\}$ is the w-plane without the negative u-semiaxis (where $w = u + iv$) and the range of the continuous branch of $w^{\frac{1}{n}}$ is the angular region $B_{-\pi/n} = \{re^{i\theta} \mid r > 0, -\frac{\pi}{n} < \theta < \frac{\pi}{n}\}$. The value of this branch at every $w \in A_{-\pi}$ is given by

$$z = \sqrt[n]{s}\, e^{i\frac{\phi}{n}},$$

where $w = se^{i\phi}$ is the polar representation of w with $-\pi < \phi < \pi$. Clearly,

$$z = \sqrt[n]{|w|}\, e^{i\frac{\operatorname{Arg} w}{n}} = e^{\frac{\operatorname{Log} w}{n}}.$$

On the same set $A_{-\pi}$ of the w-plane, besides the above continuous branch of $w^{\frac{1}{n}}$, we may define n continuous branches of $w^{\frac{1}{n}}$. Each of them maps $A_{-\pi}$ onto a corresponding angular region $B_{(-\pi+k2\pi)/n}$ with $k = 0, 1, \ldots, n-1$, which results by rotating $B_{-\pi/n}$ in the positive direction by the angle $k\frac{2\pi}{n}$. This branch results from the original branch by multiplication by the constant $e^{ik\frac{2\pi}{n}}$ and its value at every $w \in A_{-\pi}$ is given by

$$z = \sqrt[n]{s}\, e^{i(\frac{\phi}{n}+k\frac{2\pi}{n})},$$

where $w = se^{i\phi}$ is the polar representation of w with $-\pi < \phi < \pi$.

Now we introduce a generalization of the notion of continuous branch of $w^{\frac{1}{n}}$. We define the notion of continuous branch of $g^{\frac{1}{n}}$, where g is a more general function than $g(w) = w$.

Let $A \subseteq \mathbb{C}$ and $g : A \to \mathbb{C} \setminus \{0\}$ be continuous in A. We say that the function f is a **continuous branch of $\mathbf{g}^{\frac{1}{n}}$** in A if f is continuous in A and for every $w \in A$ we have that $f(w)$ is an element of $g(w)^{\frac{1}{n}}$ or, equivalently,

$$f(w)^n = g(w)$$

for every $w \in A$.

Example 4.4.3. Let $g : A \to \mathbb{C} \setminus \{0\}$ be continuous in $A \subseteq \mathbb{C}$. If there is a continuous branch h of $w^{\frac{1}{n}}$ in $g(A)$, then $f = h \circ g$ is a continuous branch of $g^{\frac{1}{n}}$ in A.
Indeed, $f = h \circ g$ is continuous in A and, since $h(z)^n = z$ for every $z \in g(A)$, we have that $f(w)^n = h(g(w))^n = g(w)$ for every $w \in A$.

Example 4.4.4. Let $g : A \to \mathbb{C} \setminus \{0\}$ be continuous in $A \subseteq \mathbb{C}$. If there is a continuous branch h of $\log g$ in A, then $f = e^{\frac{1}{n}h}$ is a continuous branch of $g^{\frac{1}{n}}$ in A.
Indeed, $f = e^{\frac{1}{n}h}$ is continuous in A and, since $e^{h(w)} = g(w)$ for every $w \in A$, we get that $f(w)^n = e^{h(w)} = g(w)$ for every $w \in A$.
This is a standard way to produce continuous branches of $g^{\frac{1}{n}}$ when we know continuous branches of $\log g$.

Proposition 4.8. *Let $g : A \to \mathbb{C} \setminus \{0\}$ be continuous in $A \subseteq \mathbb{C}$ and f be any continuous branch of $g^{\frac{1}{n}}$ in A. If w_0 is an interior point of A and g is differentiable at w_0, then f is differentiable at w_0 and $f'(w_0) = \frac{g'(w_0)f(w_0)}{ng(w_0)}$. Hence, if g is holomorphic in the interior of A, then f is also holomorphic in the interior of A.*

Proof. We set $z_0 = f(w_0)$ and $z = f(w)$ for every $w \in A$. Then $z_0^n = g(w_0)$ and $z^n = g(w)$. Since f is continuous, $w \to w_0$ implies $z \to z_0$. Therefore, using the derivative of the exponential function at z_0, we see that

$$\tfrac{f(w)-f(w_0)}{w-w_0} = \tfrac{z-z_0}{z^n-z_0^n}\,\tfrac{g(w)-g(w_0)}{w-w_0} \to \tfrac{g'(w_0)}{nz_0^{n-1}} = \tfrac{g'(w_0)f(w_0)}{ng(w_0)}$$

when $w \to w_0$. Thus f is differentiable at w_0 and $f'(w_0) = \frac{g'(w_0)f(w_0)}{ng(w_0)}$. □

Therefore, if $g : A \to \mathbb{C} \setminus \{0\}$ is holomorphic in the open set A, *every continuous branch of* $g^{\frac{1}{n}}$ *can be called* **holomorphic branch** *of* $g^{\frac{1}{n}}$ *in* A.

Example 4.4.5. We have defined n distinct continuous branches of $w^{\frac{1}{n}}$ in the open set A which results when we exclude any halfline with vertex 0 from the w-plane. All these branches are holomorphic branches of $w^{\frac{1}{n}}$ in A.

Proposition 4.9. *Let* $g : A \to \mathbb{C} \setminus \{0\}$ *be continuous in* $A \subseteq \mathbb{C}$. *Let also* $\omega_n = e^{i\frac{2\pi}{n}}$ *be the principal* n*-th root of unity.*
(i) If f_1 *is a continuous branch of* $g^{\frac{1}{n}}$ *in* A *and* $\frac{f_2}{f_1} = \omega_n^k$ *in* A*, where* $k = 0, 1, \dots, n-1$ *is fixed, then* f_2 *is also a continuous branch of* $g^{\frac{1}{n}}$ *in* A.
(ii) If, moreover, A *is connected and* f_1, f_2 *are continuous branches of* $g^{\frac{1}{n}}$ *in* A*, then* $\frac{f_2}{f_1} = \omega_n^k$ *in* A*, where* $k = 0, 1, \dots, n-1$ *is fixed. In particular, if* $f_1(w_0) = f_2(w_0)$ *for some* $w_0 \in A$*, then* $f_1 = f_2$ *in* A.

Proof. (i) The continuity of f_1 in A implies the continuity of f_2 in A. We also have $f_1(w)^n = g(w)$ for every $w \in A$ and hence $f_2(w)^n = f_1(w)^n(\omega_n^k)^n = g(w)(\omega_n^n)^k = g(w)$ for every $w \in A$. Thus, f_2 is a continuous branch of $g^{\frac{1}{n}}$ in A.
(ii) For each $w \in A$ the numbers $f_2(w), f_1(w)$ are elements of $g(w)^{\frac{1}{n}}$. Hence $(\frac{f_2(w)}{f_1(w)})^n = \frac{g(w)}{g(w)} = 1$ and so $\frac{f_2}{f_1} : A \to \{1, \omega_n, \dots, \omega_n^{n-1}\}$. Now, the function $\frac{f_2}{f_1}$ is continuous in A and A is connected, hence the set $\frac{f_2}{f_1}(A)$ is also connected. Since $\frac{f_2}{f_1}(A) \subseteq \{1, \omega_n, \dots, \omega_n^{n-1}\}$, the set $\frac{f_2}{f_1}(A)$ contains only one point. I.e. $\frac{f_2}{f_1}$ is constant in A and hence $\frac{f_2}{f_1} = \omega_n^k$ in A, where $k = 0, 1, \dots, n-1$ is fixed.
In case $f_2(w_0) = f_1(w_0)$, then the integer k is 1 and we get $f_2 = f_1$ in A. □

Thus, *if we know one continuous branch of* $g^{\frac{1}{n}}$ *in the connected set A, then we can find every other of the* n *possible continuous branches of* $g^{\frac{1}{n}}$ *in A by multiplying the known branch with any constant* n*-th root of unity.*

Example 4.4.6. Let $A_{-\pi} = \{se^{i\phi} \,|\, s > 0, -\pi < \phi < \pi\}$ be the w-plane without the negative u-semiaxis (where $w = u + iv$). We want to find a continuous branch of the square root $w^{\frac{1}{2}}$ in $A_{-\pi}$ taking the value $z = 1$ at $w = 1$.
From the example 4.4.2 we already know the continuous branch of the square root which maps $A_{-\pi}$ onto the angular region $B_{-\pi/2} = \{re^{i\theta} \,|\, r > 0, -\frac{\pi}{2} < \theta < \frac{\pi}{2}\}$, i.e. onto the right halfplane of the z-plane: the value of this branch at every $w \in A_{-\pi}$ is given by

$$z = \sqrt{s}\, e^{i\frac{\phi}{2}},$$

where $w = se^{i\phi}$ is the polar representation of w with $-\pi < \phi < \pi$. Since $A_{-\pi}$ is connected, there is no other continuous branch of the square root in $A_{-\pi}$ taking the value $z = 1$ at $w = 1$.

Example 4.4.7. Consider $A_{-\pi} = \{se^{i\phi} \,|\, s > 0, -\pi < \phi < \pi\}$ again. Now we want to find a continuous branch of the square root $w^{\frac{1}{2}}$ in A taking the value $z = -1$ at $w = 1$.
In the previous example we found one continuous branch of the square root in A. Since A is

connected, there are exactly two continuous branches of the square root in A. We consider the principal square root of 1, i.e. $\omega_2 = e^{i\frac{2\pi}{2}} = e^{i\pi} = -1$. (Trivial: the square roots of 1 are the solutions of $z^2 = 1$, i.e. the numbers $1, -1$.) Then the value of the second continuous branch of the square root at every $w \in A_{-\pi}$ is given by

$$z = \sqrt{s}\, e^{i\frac{\phi}{2}} \omega_2 = -\sqrt{s}\, e^{i\frac{\phi}{2}},$$

where $w = se^{i\phi}$ is the polar representation of w with $-\pi < \phi < \pi$. This branch of the square root is the opposite of the branch in example 4.4.6 and maps $A_{-\pi}$ onto the angular region $B_{(-\pi+2\pi)/2} = B_{\pi/2} = \{re^{i\theta} \,|\, r > 0, \frac{\pi}{2} < \theta < \frac{3\pi}{2}\}$, i.e. onto the left halfplane of the z-plane.

Exercises.

4.4.1. Describe the sets

$$(-1)^{\frac{1}{2}}, \quad (-1)^{\frac{1}{3}}, \quad (-1)^{\frac{1}{4}}, \quad i^{\frac{1}{2}}, \quad i^{\frac{1}{3}}, \quad i^{\frac{1}{4}}, \quad \left(\tfrac{1-i\sqrt{3}}{2}\right)^{\frac{1}{2}}, \quad \left(\tfrac{1-i\sqrt{3}}{2}\right)^{\frac{1}{3}}, \quad \left(\tfrac{1-i\sqrt{3}}{2}\right)^{\frac{1}{4}}.$$

4.4.2. (i) Find the elements of $\log(i^2)$ and of $2\log i$ and observe that the two sets are different.
(ii) Prove that for every $w \neq 0$ and every $n \in \mathbb{N}$ the sets $\log(w^{\frac{1}{n}})$ and $\frac{1}{n}\log w$ are equal.

4.4.3. If $w \neq 0$, prove that $w^{\frac{1}{n}} = \{e^{\frac{\zeta}{n}} \,|\, \zeta \in \log w\}$.

4.4.4. Let $w \neq 0$ and z be any of the elements of $w^{\frac{1}{n}}$. Prove that the elements of $w^{\frac{1}{n}}$ are the numbers $z, z\omega_n, z\omega_n^2, \ldots, z\omega_n^{n-1}$.

4.4.5. The set $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ is a group under multiplication. Let $n \in \mathbb{N}$, $n \geq 2$.
(i) Prove that the n-th root of unity, i.e. the set $\{1, \omega_n, \omega_n^2, \ldots, \omega_n^{n-1}\}$, is a subgroup of $\mathbb{C}^*$.
(ii) Let $z = \omega_n^k$ be any of the elements of the n-th root of unity and $\langle z\rangle = \{z^m \,|\, m \in \mathbb{Z}\}$ be the group generated by z. Prove that z is a generator of $\{1, \omega_n, \omega_n^2, \ldots, \omega_n^{n-1}\}$ or, equivalently, $\langle z\rangle = \{1, \omega_n, \omega_n^2, \ldots, \omega_n^{n-1}\}$ if and only if $\gcd\{k, n\} = 1$.
(iii) Prove that $\{1, \omega_n, \omega_n^2, \ldots, \omega_n^{n-1}\}$ has no subgroups other than $\{1\}$ and itself if and only if n is a prime number.

4.4.6. Look at exercise 3.3.2. Consider the curves on the z-plane with equations $x^2 - y^2 = \alpha$ and $2xy = \beta$. If the two curves intersect at a point (x_0, y_0), find in two ways their angle at this point.

4.4.7. Prove that there is no continuous branch of $w^{\frac{1}{n}}$ in any circle $C_0(r)$ and hence in any set A which contains such a circle.

4.4.8. Consider the sets:

$$\{re^{i\theta} \,|\, 0 < r < +\infty, -\tfrac{\pi}{2} < \theta < \tfrac{3\pi}{2}\}, \quad \{re^{i\theta} \,|\, 0 < r < +\infty, -\pi < \theta < \pi\},$$

$$\{re^{i\theta} \,|\, 0 < r < +\infty, 0 < \theta < 2\pi\}, \quad \{re^{i\theta} \,|\, 0 < r < +\infty, \tfrac{\pi}{2} < \theta < \tfrac{5\pi}{2}\}.$$

In each of these sets write the formulas of the continuous branches of the square root, of the cube root and of the sixth root.

4.4.9. (i) Considering a holomorphic branch of $(w+1)^{\frac{1}{2}}$ in $\mathbb{C} \setminus (-\infty, -1]$ and a holomorphic branch of $(w-1)^{\frac{1}{2}}$ in $\mathbb{C} \setminus [1, +\infty)$, prove that there is a holomorphic branch of $(w^2-1)^{\frac{1}{2}}$ in $\Omega = \mathbb{C} \setminus ((-\infty, -1] \cup [1, +\infty))$.
(ii) Considering a holomorphic branch of $(w+1)^{\frac{1}{2}}$ in $\mathbb{C} \setminus (-\infty, -1]$ and a holomorphic branch of $(w-1)^{\frac{1}{2}}$ in $\mathbb{C} \setminus (-\infty, 1]$, prove that there is a holomorphic branch of $(w^2-1)^{\frac{1}{2}}$ in $\Omega' = \mathbb{C} \setminus [-1, 1]$. (This is not as trivial as (i).)
(iii) Prove that there is no continuous branch of $(w^2-1)^{\frac{1}{2}}$ in any circle which surrounds one of the points ± 1 but not the other.

4.4.10. Prove that we can define a holomorphic branch f of $(1-w)^{\frac{1}{2}}+(1+w)^{\frac{1}{2}}$ in the region A which results when we exclude from $\mathbb{C}$ two non-intersecting halflines, one with vertex $+1$ and another with vertex -1. Prove that every such f satisfies $f(w)^4-4f(w)^2+4w^2=0$ for every $w\in A$. How many such branches f exist in A?

4.4.11. (i) Let $w\neq 0$ and $a\in\mathbb{Z}$. Prove that $\{e^{az}\,|\,z\in\log w\}$ has only one element, namely w^a.
(ii) Generalizing (i), let $w\neq 0$ and $a\notin\mathbb{Z}$. We define

$$w^a=\{e^{az}\,|\,z\in\log w\}$$

and this set may have more than one elements. When does w^a have finitely many elements and when does it have infinitely many elements?
(iii) Describe the sets $(\frac{1-i\sqrt{3}}{2})^{\frac{1}{2}}$, $i^{\frac{1}{4}}$, 2^i, $i^{\sqrt{2}}$ and draw their elements.
(iv) Prove that the elements of w^{a+b} are also elements of w^aw^b, and that the elements of w^{ab} are also elements of $(w^a)^b$.
(v) Let f be a continuous branch of log in $A\subseteq\mathbb{C}\setminus\{0\}$. Prove that $g=e^{af}$ is a continuous branch of w^a in A and that g is differentiable at every interior point w_0 of A and $g'(w_0)=\frac{ag(w_0)}{w_0}$.
(vi) Prove that there is a unique holomorphic branch f of $(1-w)^i=e^{i\log(1-w)}$ in $\mathbb{D}$ so that $f(0)=1$. Then prove that there are $c_1,c_2>0$ so that $c_1<|f(w)|<c_2$ for every $w\in\mathbb{D}$. Find the best such c_1,c_2.

4.4.12. We define

$$\arccos w=\{z\,|\,\cos z=w\},\quad \arcsin w=\{z\,|\,\sin z=w\},\quad \arctan w=\{z\,|\,\tan z=w\}.$$

(i) Prove that the three sets are non-empty, except in the case of $\arctan(\pm i)$.
(ii) Express arccos, arcsin and arctan in terms of log.
(iii) It should be clear from exercise 4.2.7 that sin is one-to-one from $\{x+iy\,|\,-\frac{\pi}{2}<x<\frac{\pi}{2}\}$ onto $\Omega=\mathbb{C}\setminus((-\infty,-1]\cup[1,+\infty))$. Prove that the inverse function g_0 is a continuous branch of arcsin in Ω, i.e. g_0 is continuous in Ω and $\sin g_0(w)=w$ for every $w\in\Omega$. Describe all continuous branches g of arcsin in Ω and prove that they are holomorphic in Ω with

$$g'(w)=\tfrac{1}{(1-w^2)^{1/2}}$$

for every $w\in\Omega$, where at the denominator appears a specific holomorphic branch of $(1-w^2)^{\frac{1}{2}}$ in Ω (see exercise 4.4.9).
(iv) From exercise 4.2.7 again, it is clear that cos is one-to-one from $\{x+iy\,|\,0<x<\pi\}$ onto $\Omega=\mathbb{C}\setminus((-\infty,-1]\cup[1,+\infty))$. Prove that the inverse function h_0 is a continuous branch of arccos in Ω, i.e. h_0 is continuous in Ω and $\cos h_0(w)=w$ for every $w\in\Omega$. Describe all continuous branches h of arccos in Ω and prove that they are holomorphic in Ω with

$$h'(w)=-\tfrac{1}{(1-w^2)^{1/2}}$$

for every $w\in\Omega$, where at the denominator appears a specific holomorphic branch of $(1-w^2)^{\frac{1}{2}}$ in Ω.
(v) Prove that tan is one-to-one from $\{x+iy\,|\,-\frac{\pi}{2}<x<\frac{\pi}{2}\}$ onto $U=\mathbb{C}\setminus\{iv\,|\,v\leq -1 \text{ or } 1\leq v\}$. Prove that the inverse function k_0 is a continuous branch of arctan in U, i.e. k_0 is continuous in U and $\tan k_0(w)=w$ for every $w\in U$. Describe all continuous branches k of arctan in U and prove that they are holomorphic in U with

$$k'(w)=\tfrac{1}{1+w^2}$$

for every $w\in U$.

4.4.13. Considering appropriate continuous branches of $w^{\frac{1}{2}}$, evaluate $\int_\gamma\frac{1}{w^{1/2}}\,dw$ for both curves $\gamma_1(t)=e^{it}$, $t\in[0,\pi]$, and $\gamma_2(t)=e^{-it}$, $t\in[0,\pi]$.

4.5 Functions defined by curvilinear integrals.

4.5.1 Indefinite integrals.

Let the complex functions f, F be defined in the region $\Omega \subseteq \mathbb{C}$. We say that F is a **primitive** of f in Ω if $F'(z) = f(z)$ for every $z \in \Omega$.

Proposition 4.10. *Let the complex function f be continuous in the region $\Omega \subseteq \mathbb{C}$. Then the following are equivalent.*
(i) $\oint_\gamma f(z)\,dz = 0$ for every closed piecewise smooth curve γ in Ω.
(ii) $\int_{\gamma_1} f(z)\,dz = \int_{\gamma_2} f(z)\,dz$ for every two piecewise smooth curves γ_1, γ_2 in Ω with the same endpoints.
(iii) There is a primitive of f in Ω.

Proof. (iii)$\Rightarrow$(i) Let F be any primitive of f in Ω. We take an arbitrary piecewise smooth curve $\gamma : [a,b] \to \Omega$ with $\gamma(a) = \gamma(b)$. Then

$$\begin{aligned}\oint_\gamma f(z)\,dz = \oint_\gamma F'(z)\,dz &= \int_a^b F'(\gamma(t))\gamma'(t)\,dt = \int_a^b (F\circ\gamma)'(t)\,dt \\ &= (F\circ\gamma)(b) - (F\circ\gamma)(a) = F(\gamma(b)) - F(\gamma(a)) = 0.\end{aligned}$$

(i)$\Rightarrow$(ii) Assume that the piecewise smooth curves γ_1, γ_2 in Ω have the same endpoints. Then the piecewise smooth curve $\gamma = \gamma_1 \dot{+} (\neg\,\gamma_2)$ is a closed curve in Ω and then

$$\int_{\gamma_1} f(z)\,dz - \int_{\gamma_2} f(z)\,dz = \int_{\gamma_1} f(z)\,dz + \int_{\neg\,\gamma_2} f(z)\,dz = \oint_\gamma f(z)\,dz = 0.$$

(ii)$\Rightarrow$(iii) We consider an arbitrary fixed $z_0 \in \Omega$. Then for every $z \in \Omega$ there is at least one piecewise smooth curve γ in Ω with initial point z_0 and final point z. We define the function $F : \Omega \to \mathbb{C}$ by

$$F(z) = \int_\gamma f(\zeta)\,d\zeta. \tag{4.8}$$

This formula defines $F(z)$ uniquely, since the value of the curvilinear integral depends only on the point z and not on the particular piecewise smooth curve γ which we use to join z_0 to z.
Now we shall prove that F is a primitive of f in Ω. We take an arbitrary $z \in \Omega$ and a disc $D_z(r) \subseteq \Omega$. We also take a piecewise smooth curve γ in Ω with initial point z_0 and final point z. Then the value of $F(z)$ is given by (4.8). Now we consider any $w \in D_z(r)$ and the curve $\gamma \dot{+} [z,w]$. This curve is in Ω, it is piecewise smooth and has initial point z_0 and final point w. Therefore,

$$F(w) = \int_{\gamma\dot{+}[z,w]} f(\zeta)\,d\zeta = \int_\gamma f(\zeta)\,d\zeta + \int_{[z,w]} f(\zeta)\,d\zeta. \tag{4.9}$$

From (4.8) and (4.9) we get

$$F(w) - F(z) - f(z)(w-z) = \int_{[z,w]} f(\zeta)\,d\zeta - f(z)\int_{[z,w]} d\zeta = \int_{[z,w]} (f(\zeta) - f(z))\,d\zeta. \tag{4.10}$$

Now, since f is continuous, for every $\epsilon > 0$ there is $\delta > 0$ so that $|f(\zeta) - f(z)| < \epsilon$ for every $\zeta \in \Omega$ with $|\zeta - z| < \delta$. Taking $w \in D_z(r)$ with $|w - z| < \delta$ we automatically have $|\zeta - z| < \delta$ for every $\zeta \in [z,w]$ and (4.10) implies

$$|F(w) - F(z) - f(z)(w-z)| \leq \epsilon |w - z|.$$

Therefore, $\left|\frac{F(w)-F(z)}{w-z} - f(z)\right| \leq \epsilon$ for every w with $0 < |w-z| < \delta$ and hence $F'(z) = f(z)$. $\square$

Let the complex function f be continuous in the region $\Omega \subseteq \mathbb{C}$. If either one of the equivalent conditions (i), (ii) of proposition 4.10 is satisfied, then as we saw in the proof of (ii)$\Rightarrow$(iii) of proposition 4.10, we may choose a fixed point $z_0 \in \Omega$ and define $F(z) = \int_\gamma f(\zeta)\,d\zeta$ for every

$z \in \Omega$, where γ is an arbitrary piecewise smooth curve in Ω with initial point z_0 and final point z. Now, any function F of the form

$$F(z) = \int_\gamma f(\zeta)\, d\zeta + c,$$

where γ is any piecewise smooth curve in Ω with fixed (but otherwise arbitrary) initial point $z_0 \in \Omega$ and final point $z \in \Omega$ and where c is an arbitrary constant, is called **indefinite integral** of f in Ω.

The crucial condition for the existence of an indefinite integral is (ii) (or its equivalent (i)) of proposition 4.10. As soon as this is satisfied, then by changing the *base point* $z_0 \in \Omega$ or the constant c we get different indefinite integrals F.

In the proof of proposition 4.10 we saw that every indefinite integral of f is a primitive of f. The converse is also true. Indeed, let F be any primitive of f in the region Ω, i.e. let $F'(z) = f(z)$ for every $z \in \Omega$. Proposition 4.10 implies that condition (ii) is satisfied and, if we take any piecewise smooth curve $\gamma : [a, b] \to \Omega$ with initial point a fixed $z_0 \in \Omega$ and final point $z \in \Omega$, then

$$\begin{aligned} \int_\gamma f(\zeta)\, d\zeta = \int_\gamma F'(\zeta)\, d\zeta = \int_a^b F'(\gamma(t))\gamma'(t)\, dt = \int_a^b (F \circ \gamma)'(t)\, dt \\ = (F \circ \gamma)(b) - (F \circ \gamma)(a) = F(z) - F(z_0). \end{aligned} \tag{4.11}$$

Thus, F has the form $F(z) = \int_\gamma f(\zeta)\, d\zeta + F(z_0)$ and hence it is an indefinite integral of f in Ω.

We summarize. *Let the complex function f be continuous in the region $\Omega \subseteq \mathbb{C}$. Then the notion of primitive of f in Ω coincides with the notion of indefinite integral of f in Ω. Moreover, the existence of a primitive or, equivalently, of an indefinite integral of f in Ω is equivalent to the validity of condition (ii) (or (i)) of proposition 4.10.*

Regarding the number of possible primitives of f in Ω we may easily see that, if there is at least one primitive F of f in Ω, then all others are of the form $F + c$ for an arbitrary constant c. Indeed, it is obvious that $F + c$ is a primitive of f in Ω. Conversely, if G is a primitive of f in Ω, then we have $(G - F)'(z) = G'(z) - F'(z) = f(z) - f(z) = 0$ for every $z \in \Omega$. Now, theorem 3.3 implies that $G - F$ is a constant in Ω.

Since it is useful for calculations of curvilinear integrals, we state relation (4.11) as a separate proposition.

Proposition 4.11. *Let F be a primitive of the continuous function f in the region $\Omega \subseteq \mathbb{C}$. Then for every piecewise smooth curve γ in Ω with initial endpoint z_1 and final endpoint z_2 we have*

$$\int_\gamma f(z)\, dz = F(z_2) - F(z_1).$$

Example 4.5.1. Every polynomial function $p(z) = a_0 + a_1 z + \cdots + a_n z^n$ has the primitive $a_0 z + \frac{a_1}{2} z^2 + \cdots + \frac{a_n}{n+1} z^{n+1}$ in $\mathbb{C}$. Therefore, we have

$$\oint_\gamma p(z)\, dz = 0$$

for every closed piecewise smooth curve γ.
In particular,

$$\oint_\gamma (z - z_0)^n\, dz = 0 \qquad \text{if } n \in \mathbb{Z}, n \geq 0,$$

for every closed piecewise smooth curve γ. A very special case of this, with the circle $C_{z_0}(r)$, we saw in examples 2.2.9 and 4.2.2.

Example 4.5.2. The exponential function e^z has the primitive e^z in $\mathbb{C}$. Hence

$$\oint_\gamma e^z\, dz = 0$$

for every closed piecewise smooth curve γ.

Example 4.5.3. If $n \in \mathbb{N}$, $n \geq 2$, the function $\frac{1}{(z-z_0)^n}$ has the primitive $-\frac{1}{(n-1)(z-z_0)^{n-1}}$ in $\mathbb{C} \setminus \{z_0\}$. Therefore,

$$\oint_\gamma \frac{1}{(z-z_0)^n}\, dz = 0 \qquad \text{if } n \in \mathbb{N}, n \geq 2,$$

for every closed piecewise smooth curve γ in $\mathbb{C} \setminus \{z_0\}$. A very special case of this, with the circle $C_{z_0}(r)$, we saw in examples 2.2.9 and 4.2.2.

Example 4.5.4. The function $\frac{1}{z-z_0}$ (the case $n = 1$ of the previous example) has no primitive in $\mathbb{C} \setminus \{z_0\}$ or even in any open ring $D_{z_0}(r_1, r_2) = \{z \,|\, r_1 < |z - z_0| < r_2\}$.
Indeed, if $\frac{1}{z-z_0}$ had a primitive in $D_{z_0}(r_1, r_2)$, then we would have $\oint_\gamma \frac{1}{z-z_0}\, dz = 0$ for every closed piecewise smooth curve γ in $D_{z_0}(r_1, r_2)$. Now, if we take a radius r so that $r_1 < r < r_2$ and the curve $\gamma : [0, 2\pi] \to D_{z_0}(r_1, r_2)$ with parametric equation $\gamma(t) = z_0 + re^{it}$, then we have

$$\oint_\gamma \frac{1}{z-z_0}\, dz = \oint_{C_{z_0}(r)} \frac{1}{z-z_0}\, dz = \int_0^{2\pi} \frac{1}{re^{it}} rie^{it}\, dt = 2\pi i \neq 0.$$

In fact, we did exactly the same calculation in example 4.2.2.

The following result is important.

Theorem 4.1. *Let $g : \Omega \to \mathbb{C} \setminus \{0\}$ be holomorphic in the region $\Omega \subseteq \mathbb{C}$ and let g' be continuous in Ω. Then a holomorphic branch of* $\log g$ *exists in Ω if and only if*

$$\oint_\gamma \frac{g'(z)}{g(z)}\, dz = 0$$

for every closed piecewise smooth curve γ in Ω.

Proof. Assume that there is a holomorphic branch of $\log g$ in Ω, i.e. there is F holomorphic in Ω so that $e^{F(z)} = g(z)$ for every $z \in \Omega$. Then $F'(z)e^{F(z)} = g'(z)$ for every $z \in \Omega$ and hence $F'(z) = \frac{g'(z)}{g(z)}$ for every $z \in \Omega$. Therefore, F is a primitive of $\frac{g'}{g}$ in Ω and thus, $\oint_\gamma \frac{g'(z)}{g(z)}\, dz = 0$ for every closed piecewise smooth curve γ in Ω.
Conversely, assume $\oint_\gamma \frac{g'(z)}{g(z)}\, dz = 0$ for every closed piecewise smooth curve γ in Ω. Then $\frac{g'}{g}$ has a primitive, say F, in Ω. Now, we have $\frac{d}{dz}(g(z)e^{-F(z)}) = g'(z)e^{-F(z)} - g(z)F'(z)e^{-F(z)} = 0$ for every $z \in \Omega$. This implies that, for some constant c, we have $g(z)e^{-F(z)} = c$ for every $z \in \Omega$. Since $c \neq 0$, there is a constant d so that $e^d = c$ and we finally get that $e^{F(z)+d} = g(z)$ for every $z \in \Omega$. Now the function $F + d$ is a holomorphic branch of $\log g$ in Ω. □

In the next chapter we shall prove that for every holomorphic g the derivative g' is automatically continuous. Therefore, *a posteriori*, the assumption in theorem 4.1 that g' is continuous is unnecessary.

Example 4.5.5. If the region $\Omega \subseteq \mathbb{C} \setminus \{z_0\}$ contains a circle $C_{z_0}(r)$, then there is no holomorphic branch of $\log(z - z_0)$ in Ω. In fact, example 4.5.4 shows that $\oint_{C_{z_0}(r)} \frac{1}{z-z_0}\, dz \neq 0$.

Example 4.5.6. Let $g : \Omega \to \mathbb{C} \setminus \{0\}$ be holomorphic in the region $\Omega \subseteq \mathbb{C}$, let g' be continuous in Ω and suppose that there is a halfline with vertex 0 so that $g(\Omega) \subseteq \mathbb{C} \setminus l$.
We know that a holomorphic branch of log exists in $\mathbb{C} \setminus l$ and now example 4.3.3 says that a holomorphic branch of $\log g$ exists in Ω. From theorem 4.1 we also get that $\oint_\gamma \frac{g'(z)}{g(z)}\, dz = 0$ for every closed piecewise smooth curve γ in Ω.

4.5.2 Integrals with parameter.

Lemma 4.2. *Let $n \in \mathbb{N}$ and γ be any piecewise smooth curve. If the complex function ϕ is continuous in the trajectory γ^*, we define*

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

for every $z \notin \gamma^*$. Then f is holomorphic in the open set $\mathbb{C} \setminus \gamma^*$ and

$$f'(z) = n \int_\gamma \frac{\phi(\zeta)}{(\zeta-z)^{n+1}}\, d\zeta$$

for every $z \notin \gamma^*$.

Proof. We take any $z \in \mathbb{C} \setminus \gamma^*$. Since $\mathbb{C} \setminus \gamma^*$ is open, there is $\delta > 0$ so that $D_z(\delta) \subseteq \mathbb{C} \setminus \gamma^*$. We consider the smaller circle $D_z(\frac{\delta}{2})$ and we have $|\zeta - w| \geq \frac{\delta}{2}$ for every $\zeta \in \gamma^*$ and every $w \in D_z(\frac{\delta}{2})$. Now for every $w \in D_z(\frac{\delta}{2})$ we get

$$\frac{f(w)-f(z)}{w-z} - n \int_\gamma \frac{\phi(\zeta)}{(\zeta-z)^{n+1}}\, d\zeta = \int_\gamma \Big(\frac{\frac{1}{(\zeta-w)^n} - \frac{1}{(\zeta-z)^n}}{w-z} - \frac{n}{(\zeta-z)^{n+1}} \Big)\, \phi(\zeta)\, d\zeta. \tag{4.12}$$

To simplify the notation, we temporarily set $a = \zeta - w$ and $b = \zeta - z$, and, to estimate the parenthesis in (4.12), we use the algebraic identity

$$\frac{\frac{1}{a^n} - \frac{1}{b^n}}{b-a} - \frac{n}{b^{n+1}} = (b-a)\Big(\frac{1}{a^n b^2} + \frac{2}{a^{n-1}b^3} + \cdots + \frac{n-1}{a^2 b^n} + \frac{n}{ab^{n+1}}\Big).$$

We have that $|a| \geq \frac{\delta}{2}$ and $|b| \geq \frac{\delta}{2}$ for every $\zeta \in \gamma^*$ and $w \in D_z(\frac{\delta}{2})$ and hence

$$\begin{aligned} \Big| \frac{\frac{1}{a^n} - \frac{1}{b^n}}{b-a} - \frac{n}{b^{n+1}} \Big| &\leq |b-a| \Big(\frac{1}{|a|^n |b|^2} + \cdots + \frac{n}{|a||b|^{n+1}} \Big) \\ &\leq |w-z| \frac{1+2+\cdots+(n-1)+n}{(\delta/2)^{n+2}} \leq |w-z|\, \frac{n^2 2^{n+2}}{\delta^{n+2}}. \end{aligned} \tag{4.13}$$

Now, (4.12) and (4.13) imply

$$\Big| \frac{f(w)-f(z)}{w-z} - n \int_\gamma \frac{\phi(\zeta)}{(\zeta-z)^{n+1}}\, d\zeta \Big| \leq |w-z|\, \frac{n^2 2^{n+2}}{\delta^{n+2}}\, \|\phi\|_{\gamma^*} l(\gamma)$$

for every $w \in D_z(\frac{\delta}{2})$. Therefore, $\lim_{w \to z} \frac{f(w)-f(z)}{w-z} = n \int_\gamma \frac{\phi(\zeta)}{(\zeta-z)^{n+1}}\, d\zeta$ and f is differentiable at z with $f'(z) = n \int_\gamma \frac{\phi(\zeta)}{(\zeta-z)^{n+1}}\, d\zeta$. □

Observe that lemma 4.2 justifies the change of order of the operations of integration and differentiation with respect to the parameter z:

$$f'(z) = \frac{d}{dz} f(z) = \frac{d}{dz} \int_\gamma \frac{\phi(\zeta)}{(\zeta-z)^n}\, d\zeta = \int_\gamma \frac{d}{dz}\Big(\frac{\phi(\zeta)}{(\zeta-z)^n}\Big)\, d\zeta = n \int_\gamma \frac{\phi(\zeta)}{(\zeta-z)^{n+1}}\, d\zeta.$$

Proposition 4.12. *Let γ be any piecewise smooth curve and the complex function ϕ be continuous in the trajectory γ^*. Then the function $f(z) = \int_\gamma \frac{\phi(\zeta)}{\zeta - z}\, d\zeta$ is infinitely many times differentiable in the open set $\mathbb{C} \setminus \gamma^*$ and*

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

for every $z \notin \gamma^*$.

Proof. Successive applications of lemma 4.2. □

Exercises.

4.5.1. Let f, g be holomorphic in the region $\Omega \subseteq \mathbb{C}$ and let f', g' be continuous in Ω.
(i) If $|f(z) - 1| < 1$ for every $z \in \Omega$, prove that $\oint_\gamma \frac{f'(z)}{f(z)}\, dz = 0$ for every closed piecewise smooth curve γ in Ω.
(ii) If $|f(z) - g(z)| < |g(z)|$ for every $z \in \Omega$, prove that $\oint_\gamma \frac{f'(z)}{f(z)}\, dz = \oint_\gamma \frac{g'(z)}{g(z)}\, dz$ for every closed piecewise smooth curve γ in Ω.

4.5.2. Let γ be a piecewise smooth curve and the complex function ϕ be continuous in γ^*. We know that the function $f(z) = \int_\gamma \frac{\phi(\zeta)}{\zeta - z}\, d\zeta$ is holomorphic in $\mathbb{C} \setminus \gamma^*$. Prove that f is holomorphic at ∞.

4.5.3. Let the complex function ϕ be continuous in $\mathbb{R}$ and let $\int_{-\infty}^{+\infty} \frac{|\phi(t)|}{1+|t|}\,dt < +\infty$. Prove that the function $f(z) = \int_{-\infty}^{+\infty} \frac{\phi(t)}{t-z}\,dt$ is holomorphic in $\mathbb{C} \setminus \mathbb{R}$.

4.5.4. Let the complex function ϕ be continuous in $\mathbb{R}$ and $\int_{-\infty}^{+\infty} |\phi(t)|e^{M|t|}\,dt < +\infty$ for every $M > 0$. Prove that the function $f(z) = \int_{-\infty}^{+\infty} \phi(t)e^{tz}\,dt$ is holomorphic in $\mathbb{C}$.

4.5.5. Find the domains of holomorphy of the following functions of z:

$$\int_0^1 \frac{1}{1+tz}\,dt, \quad \int_{-1}^1 \frac{e^{tz}}{1+t^2}\,dt, \quad \int_0^{+\infty} \frac{e^{tz}}{1+t^2}\,dt, \quad \int_0^{+\infty} e^{-tz^2}\,dt.$$

4.6 Functions defined by power series.

Every series of the form

$$\sum_{n=0}^{+\infty} a_n(z-z_0)^n = a_0 + a_1(z-z_0) + a_2(z-z_0)^2 + \cdots$$

is called **power series** with **center** z_0 and **coefficients** a_n. The $R \in [0,+\infty]$ defined by

$$R = \frac{1}{\overline{\lim}\ \sqrt[n]{|a_n|}}$$

is called **radius of convergence** of the power series. (Of course we understand that $R = 0$ if $\overline{\lim}\ \sqrt[n]{|a_n|} = +\infty$ and $R = +\infty$ if $\overline{\lim}\ \sqrt[n]{|a_n|} = 0$.)

Proposition 4.13. *Let $\sum_{n=0}^{+\infty} a_n(z-z_0)^n$ be a power series with radius of convergence R.*
If $R = 0$, then the series converges only at z_0. If $R > 0$, then:
(i) The power series converges absolutely at every $z \in D_{z_0}(R)$.
(ii) The power series diverges at every $z \notin \overline{D}_{z_0}(R)$.
(iii) The power series converges uniformly in every closed disc $\overline{D}_{z_0}(r)$ with $r < R$.
(iv) The sum

$$s(z) = \sum_{n=0}^{+\infty} a_n(z-z_0)^n, \qquad z \in D_{z_0}(R),$$

is holomorphic in $D_{z_0}(R)$. The derivative of s in $D_{z_0}(R)$ is the sum of the power series which results from $\sum_{n=0}^{+\infty} a_n(z-z_0)^n$ by formal termwise differentiation. I.e.

$$s'(z) = \sum_{n=1}^{+\infty} na_n(z-z_0)^{n-1}, \qquad z \in D_{z_0}(R).$$

Proof. If $z = z_0$, then the power series consists only of its constant term a_0 and hence converges. If $z \neq z_0$, then by the definition of R we get $\overline{\lim}\ \sqrt[n]{|a_n(z-z_0)^n|} = \overline{\lim}\ \sqrt[n]{|a_n|}\,|z-z_0| = \frac{|z-z_0|}{R}$. The root test of Cauchy for general series implies that the power series converges absolutely if $|z-z_0| < R$ and diverges if $|z-z_0| > R$ and this is the content of (i) and (ii).
(iii) Let $0 < r < R$. We take any R' with $r < R' < R$. Then $\overline{\lim}\ \sqrt[n]{|a_n|} < \frac{1}{R'}$ and so there is n_0 so that $\sqrt[n]{|a_n|} \leq \frac{1}{R'}$ for every $n \geq n_0$. Then for every $z \in \overline{D}_{z_0}(r)$ we have $|a_n(z-z_0)^n| = |a_n|\,|z-z_0|^n \leq (\frac{r}{R'})^n$ for every $n \geq n_0$. Since $\frac{r}{R'} < 1$, we have $\sum_{n=0}^{+\infty}(\frac{r}{R'})^n < +\infty$ and the test of Weierstrass implies that the power series $\sum_{n=0}^{+\infty} a_n(z-z_0)^n$ converges uniformly in $\overline{D}_{z_0}(r)$.
(iv) Besides $\sum_{n=0}^{+\infty} a_n(z-z_0)^n$, we also consider the power series $\sum_{n=1}^{+\infty} na_n(z-z_0)^{n-1}$. The second power series results from the first by formal termwise differentiation. We shall prove that the second series converges at every $z \in D_{z_0}(R)$ and that its sum is the derivative of the sum of the first series at every $z \in D_{z_0}(R)$.
We have $\overline{\lim}\ \sqrt[n]{|na_n|} = \overline{\lim}\ \sqrt[n]{n}\,\sqrt[n]{|a_n|} = \overline{\lim}\ \sqrt[n]{|a_n|}$ and the radius of convergence of the series $\sum_{n=1}^{+\infty} na_n(z-z_0)^n$ is also R. Thus, $\sum_{n=1}^{+\infty} na_n(z-z_0)^{n-1}$ converges at every $z \in D_{z_0}(R)$. We define

$$s(z) = \sum_{n=0}^{+\infty} a_n(z-z_0)^n, \quad t(z) = \sum_{n=1}^{+\infty} na_n(z-z_0)^{n-1}, \qquad z \in D_{z_0}(R).$$

Now at every $z, w \in D_{z_0}(R)$ we have

$$s(w) - s(z) = \sum_{n=0}^{+\infty} a_n((w - z_0)^n - (z - z_0)^n).$$

For simplicity, we shall set temporarily $a = z - z_0$ and $b = w - z_0$ and then we have

$$\begin{aligned} \tfrac{s(w)-s(z)}{w-z} - t(z) &= \sum_{n=2}^{+\infty} a_n(b^{n-1} + b^{n-2}a + \cdots + ba^{n-2} + a^{n-1} - na^{n-1}) \\ &= (w - z) \sum_{n=2}^{+\infty} a_n(b^{n-2} + 2b^{n-3}a + \cdots + (n-2)ba^{n-3} + (n-1)a^{n-2}). \end{aligned} \tag{4.14}$$

We fix $z \in D_{z_0}(R)$ and $\delta = \frac{R-|z-z_0|}{2} > 0$. We also set $R_1 = |z-z_0|+\delta = R-\delta$. If $w \in D_z(\delta)$, then $|b| \leq R_1$ and $|a| \leq R_1$ and (4.14) implies

$$\left|\tfrac{s(w)-s(z)}{w-z} - t(z)\right| \leq |w - z| \sum_{n=2}^{+\infty} n^2|a_n|R_1^{n-2}.$$

Since $\overline{\lim}\, \sqrt[n]{|n^2 a_n R_1^n|} = \frac{R_1}{R} < 1$, the last sum is a finite number independent of $w \in D_z(\delta)$. Therefore, $\lim_{w\to z} \frac{s(w)-s(z)}{w-z} = t(z)$ and s is differentiable at z with $s'(z) = t(z)$. ☐

If R is the radius of convergence of $\sum_{n=0}^{+\infty} a_n(z - z_0)^n$, then the open disc $D_{z_0}(R)$ is called **disc of convergence** of the power series.

We saw that, if $0 < R \leq +\infty$, the sum s of the power series is a holomorphic function in $D_{z_0}(R)$. In fact the derivative s' is the sum of the power series we get by formal termwise differentiation of the original power series. We saw that the differentiated power series has the same disc of convergence as the original series and hence we may repeat our arguments: the function s' is holomorphic in $D_{z_0}(R)$ and its derivative, i.e. the second derivative of s, is the sum of the power series which we get by a second formal termwise differentiation of the original power series. We conclude that *the function s is infinitely many times differentiable in the disc of convergence $D_{z_0}(R)$ and*

$$s^{(k)}(z) = \sum_{n=k}^{+\infty} n(n-1)\cdots(n-k+1)a_n(z - z_0)^{n-k}, \qquad z \in D_{z_0}(R).$$

Example 4.6.1. For the power series $\sum_{n=1}^{+\infty} \frac{z^n}{n}$ we get $\overline{\lim}\, \sqrt[n]{|1/n|} = 1$, and hence $R = 1$. The disc of convergence is $\mathbb{D}$. If s is the function defined by the power series in $\mathbb{D}$, then

$$s'(z) = \sum_{n=1}^{+\infty} z^{n-1} = \tfrac{1}{1-z}$$

for every $z \in \mathbb{D}$. We observe that $-\operatorname{Log}(1 - z)$ is defined and is holomorphic in $\mathbb{D}$. Its derivative is $\frac{1}{1-z}$ and its value at 0 is 0. Since the functions $s(z)$ and $-\operatorname{Log}(1 - z)$ have the same derivative in the region $\mathbb{D}$ and the same value at 0, we conclude that

$$\sum_{n=1}^{+\infty} \tfrac{z^n}{n} = -\operatorname{Log}(1 - z)$$

for every $z \in \mathbb{D}$. We shall come back to this identity when we study the Taylor series of the function $-\operatorname{Log}(1 - z)$ in $\mathbb{D}$.

Example 4.6.2. For $\sum_{n=1}^{+\infty} \frac{z^n}{n^2}$ we get $\overline{\lim}\, \sqrt[n]{|1/n^2|} = 1$, and hence $R = 1$. The disc of convergence is $\mathbb{D}$.

Example 4.6.3. For $\sum_{n=0}^{+\infty} \frac{z^n}{n!}$ we have $\overline{\lim}\, \sqrt[n]{|1/n!|} = 0$ and hence $R = +\infty$. The disc of convergence is $\mathbb{C}$. If s is the function defined by the power series in $\mathbb{C}$, then

$$s'(z) = \sum_{n=1}^{+\infty} \tfrac{z^{n-1}}{(n-1)!} = \sum_{n=0}^{+\infty} \tfrac{z^n}{n!} = s(z)$$

for every z. Now we have that $\frac{d}{dz}(e^{-z}s(z)) = -e^{-z}s(z) + e^{-z}s'(z) = 0$ for every z. Since the value of $e^{-z}s(z)$ at 0 is 1, we find that $e^{-z}s(z) = 1$ for every z and thus

$$\sum_{n=0}^{+\infty} \tfrac{z^n}{n!} = e^z$$

for every z. We shall reprove this identity later, when we study the Taylor series of the function e^z. On the other hand, since the series $\sum_{n=0}^{+\infty}\frac{z^n}{n!}$ and $\sum_{n=0}^{+\infty}\frac{w^n}{n!}$ converge absolutely, proposition 2.3 implies that

$$e^z e^w = \sum_{n=0}^{+\infty}\frac{z^n}{n!}\sum_{n=0}^{+\infty}\frac{w^n}{n!} = \sum_{n=0}^{+\infty}\Big(\sum_{k=0}^{n}\frac{z^k}{k!}\frac{w^{n-k}}{(n-k)!}\Big) = \sum_{n=0}^{+\infty}\frac{1}{n!}\Big(\sum_{k=0}^{n}\binom{n}{k}z^k w^{n-k}\Big)$$
$$= \sum_{n=0}^{+\infty}\frac{(z+w)^n}{n!} = e^{z+w}.$$

This provides us with a second proof of the identity $e^z e^w = e^{z+w}$.

Example 4.6.4. For $\sum_{n=1}^{+\infty} n! z^n$ we have $\overline{\lim}\ \sqrt[n]{n!} = +\infty$, and hence $R = 0$. The power series converges only at 0.

Every series of the form

$$\sum_{-\infty}^{n=-1} a_n(z-z_0)^n = \cdots + \frac{a_{-3}}{(z-z_0)^3} + \frac{a_{-2}}{(z-z_0)^2} + \frac{a_{-1}}{z-z_0}$$

is called **power series of second type** with **center** z_0 and **coefficients** a_n. The $R \in [0,+\infty]$ defined by

$$R = \overline{\lim}\ \sqrt[m]{|a_{-m}|}$$

is called **radius of convergence** of the power series.

The usual power series of the form $\sum_{n=0}^{+\infty} a_n(z-z_0)^n$ are also called *power series of first type*, to distinguish them from the power series of second type.

We observe that a power series of second type has no meaning at z_0, in the same way that any power series of first type (with $a_n \neq 0$ for at least one $n \geq 1$) has no meaning at ∞. On the other hand, if $z = \infty$, then a power series of second type becomes $\sum_{-\infty}^{n=-1} 0 = 0$ and hence converges with sum 0.

From now on in these notes we shall use the notations

$$D_{z_0}(R,+\infty) = \{z \,|\, R < |z-z_0|\}, \qquad \overline{D}_{z_0}(R,+\infty) = \{z \,|\, R \leq |z-z_0|\}$$

for the open and the closed unbounded ring with center z_0 and internal radius R. We also use

$$D_{z_0}(R_1,R_2) = \{z \,|\, R_1 < |z-z_0| < R_2\}, \qquad \overline{D}_{z_0}(R_1,R_2) = \{z \,|\, R_1 \leq |z-z_0| \leq R_2\}$$

to denote the open and the closed bounded ring with center z_0, internal radius R_1 and external radius R_2.

Proposition 4.14. *Let $\sum_{-\infty}^{n=-1} a_n(z-z_0)^n$ be a power series of second type with radius of convergence R.*
If $R = +\infty$, then the series converges only at ∞. If $R < +\infty$, then
(i) The power series converges absolutely at every $z \in D_{z_0}(R,+\infty) \cup \{\infty\}$.
(ii) The power series diverges at every $z \notin \overline{D}_{z_0}(R,+\infty)$.
(iii) The power series converges uniformly in every $\overline{D}_{z_0}(r,+\infty) \cup \{\infty\}$ with $r > R$.
(iv) The sum

$$s(z) = \sum_{-\infty}^{n=-1} a_n(z-z_0)^n, \qquad z \in D_{z_0}(R,+\infty) \cup \{\infty\},$$

is holomorphic in $D_{z_0}(R,+\infty) \cup \{\infty\}$. The derivative of s in $D_{z_0}(R,+\infty) \cup \{\infty\}$ is the sum of the power series which results from $\sum_{-\infty}^{n=-1} a_n(z-z_0)^n$ by formal termwise differentiation. I.e.

$$s'(z) = \sum_{-\infty}^{n=-1} n a_n(z-z_0)^{n-1}, \qquad z \in D_{z_0}(R,+\infty) \cup \{\infty\}.$$

Proof. The easiest way is to reduce a power series of second type to a power series of first type with the simple change of variable $w = \frac{1}{z-z_0}$. Then the power series $\sum_{-\infty}^{n=-1} a_n(z-z_0)^n$ takes the form

$$\sum_{-\infty}^{n=-1} a_n w^{-n} = \sum_{m=1}^{+\infty} a_{-m} w^m$$

of a power series of first type with center 0. We also observe that z varies in the unbounded ring $D_{z_0}(R,+\infty)$ if and only if w varies in the punctured disc $D_0(\frac{1}{R})\setminus\{0\}$. Also, z varies in the unbounded ring $\overline{D}_{z_0}(r,+\infty)$ if and only if w varies in the punctured disc $\overline{D}_0(\frac{1}{r})\setminus\{0\}$. Now we can use everything we know about the series $\sum_{m=1}^{+\infty}a_{-m}w^m$ from proposition 4.13 to get the corresponding results about the series $\sum_{-\infty}^{n=-1}a_n(z-z_0)^n$. For example, the differentiability of $\sum_{-\infty}^{n=-1}a_n(z-z_0)^n$ results from the differentiability of $\sum_{m=1}^{+\infty}a_{-m}w^m$ and the differentiability of the function $w=\frac{1}{z-z_0}$. We leave all the details to the reader. We shall only say a few things about the differentiability of $s(z)=\sum_{-\infty}^{n=-1}a_n(z-z_0)^n$ at ∞, using again the transformed power series $s_*(w)=\sum_{m=1}^{+\infty}a_{-m}w^m$. Since $s(\infty)=0$ and $s_*(0)=0$, we have

$$\lim_{z\to\infty}z(s(z)-s(\infty))=\lim_{z\to\infty}zs(z)=\lim_{w\to0}(1+z_0w)\,\tfrac{s_*(w)}{w}=s_*'(0)=a_{-1}.$$

Therefore, s is differentiable at ∞ □

If R is the radius of convergence of $\sum_{-\infty}^{n=-1}a_n(z-z_0)^n$, then the open ring $D_{z_0}(R,+\infty)$ is called **ring of convergence** of the power series. In fact the series converges in $D_{z_0}(R,+\infty)\cup\{\infty\}$, which is an open set in $\widehat{\mathbb{C}}$ with respect to the chordal metric.

If $0\le R<+\infty$, we saw that the sum s of the power series is a holomorphic function in $D_{z_0}(R,+\infty)\cup\{\infty\}$. In fact the derivative s' is the sum of the power series we get by formal termwise differentiation of the original power series. The differentiated power series converges in the same set $D_{z_0}(R,+\infty)\cup\{\infty\}$. Therefore, we may repeat our arguments: the function s' is holomorphic in $D_{z_0}(R,+\infty)\cup\{\infty\}$ and its derivative, i.e. the second derivative of s, is the sum of the power series which we get by a second formal termwise differentiation of the original power series. We conclude that *the function $s(z)$ is infinitely many times differentiable in $D_{z_0}(R,+\infty)\cup\{\infty\}$ and*

$$s^{(k)}(z)=\sum_{-\infty}^{n=-1}n(n-1)\cdots(n-k+1)a_n(z-z_0)^{n-k},\qquad z\in D_{z_0}(R,+\infty)\cup\{\infty\}.$$

Example 4.6.5. $\sum_{-\infty}^{n=-1}\frac{z^n}{-n}=\sum_{m=1}^{+\infty}\frac{1}{mz^m}$ converges in $D_0(1,+\infty)\cup\{\infty\}=\widehat{\mathbb{C}}\setminus\overline{\mathbb{D}}$.

Example 4.6.6. $\sum_{-\infty}^{n=-1}\frac{z^n}{n^2}=\sum_{m=1}^{+\infty}\frac{1}{m^2z^m}$ converges in $D_0(1,+\infty)\cup\{\infty\}=\widehat{\mathbb{C}}\setminus\overline{\mathbb{D}}$.

Example 4.6.7. $\sum_{-\infty}^{n=-1}\frac{z^n}{(-n)!}=\sum_{m=1}^{+\infty}\frac{1}{m!z^m}$ converges in $D_0(0,+\infty)\cup\{\infty\}=\widehat{\mathbb{C}}\setminus\{0\}$.

Example 4.6.8. $\sum_{-\infty}^{n=-1}(-n)!z^n=\sum_{m=1}^{+\infty}\frac{m!}{z^m}$ converges only at ∞.

Finally, we consider a series of the form

$$\sum_{-\infty}^{+\infty}a_n(z-z_0)^n=\cdots+\tfrac{a_{-2}}{(z-z_0)^2}+\tfrac{a_{-1}}{z-z_0}+a_0+a_1(z-z_0)+a_2(z-z_0)^2+\cdots$$

which consists of a power series of first type and a power series of second type. We assume that $a_n\neq0$ for at least one $n<0$ and for at least one $n>0$. Then the original series is called **power series of third type** with **center** z_0 and **coefficients** a_n. The radius of convergence R_1 of $\sum_{-\infty}^{n=-1}a_n(z-z_0)^n$ and the radius of convergence R_2 of $\sum_{n=0}^{+\infty}a_n(z-z_0)^n$ are called **radii of convergence** of our power series. We say that $\sum_{-\infty}^{+\infty}a_n(z-z_0)^n$ converges at z if both $\sum_{-\infty}^{n=-1}a_n(z-z_0)^n$ and $\sum_{n=0}^{+\infty}a_n(z-z_0)^n$ converge at z, and we say that $\sum_{-\infty}^{+\infty}a_n(z-z_0)^n$ diverges at z in all other cases.

A power series of third type with center z_0 has no meaning at the points z_0 and ∞.

A power series of third type is a combination of a power series of first type and a power series of second type. Therefore, we expect that the properties of a power series of this new type are a combination of properties of power series of the two previous types. Indeed, the next result is a direct combination of propositions 4.13 and 4.14 and we omit the proof.

Proposition 4.15. *Let $\sum_{-\infty}^{+\infty} a_n(z-z_0)^n$ be a power series of third type with radii of convergence R_1, R_2.*
If $R_2 \le R_1$, then the series diverges at every z, except in the case $0 < R_1 = R_2 = R < +\infty$ and then it may converge only at some $z \in C_{z_0}(R)$. If $R_1 < R_2$, then
(i) The power series converges absolutely at every $z \in D_{z_0}(R_1, R_2)$.
(ii) The power series diverges at every $z \notin \overline{D}_{z_0}(R_1, R_2)$.
(iii) The power series converges uniformly in every $\overline{D}_{z_0}(r_1, r_2)$ with $R_1 < r_1 < r_2 < R_2$.
(iv) The sum

$$s(z) = \sum_{-\infty}^{+\infty} a_n(z-z_0)^n, \qquad z \in D_{z_0}(R_1, R_2),$$

is holomorphic in $D_{z_0}(R_1, R_2)$. The derivative of s in $D_{z_0}(R_1, R_2)$ is the sum of the power series which results from $\sum_{-\infty}^{+\infty} a_n(z-z_0)^n$ by formal termwise differentiation. I.e.

$$s'(z) = \sum_{-\infty}^{+\infty} n a_n(z-z_0)^{n-1}, \qquad z \in D_{z_0}(R_1, R_2).$$

If $R_1 < R_2$, then $D_{z_0}(R_1, R_2)$ is called **ring of convergence** of $\sum_{-\infty}^{+\infty} a_n(z-z_0)^n$ and the function s defined by the power series is infinitely many times differentiable in $D_{z_0}(R_1, R_2)$.

Example 4.6.9. We consider $\sum_{-\infty}^{n=-1} \frac{2^n}{-n} z^n + 1 + \sum_{n=1}^{+\infty} \frac{1}{n^2} z^n$.
Then $\sum_{-\infty}^{n=-1} \frac{2^n}{-n} z^n$ has radius of convergence $\frac{1}{2}$ and $1 + \sum_{n=1}^{+\infty} \frac{1}{n^2} z^n$ has radius of convergence 1. Therefore, $D_0(\frac{1}{2}, 1)$ is the ring of convergence of $\sum_{-\infty}^{n=-1} \frac{2^n}{-n} z^n + 1 + \sum_{n=1}^{+\infty} \frac{1}{n^2} z^n$.

Exercises.

4.6.1. Find the discs of convergence of the following power series:

$$\sum_{n=0}^{+\infty} n^{13} z^n, \quad \sum_{n=1}^{+\infty} \frac{1}{n^5} z^n, \quad \sum_{n=1}^{+\infty} \frac{1}{n^n} z^n, \quad \sum_{n=1}^{+\infty} n^{\ln n} z^n,$$

$$\sum_{n=1}^{+\infty} \ln^n n \, z^n, \quad \sum_{n=1}^{+\infty} \frac{n!}{n^n} z^n, \quad \sum_{n=1}^{+\infty} \frac{(n!)^2}{n^n} z^n, \quad \sum_{n=0}^{+\infty} \frac{(n!)^2}{(2n)!} z^n.$$

4.6.2. Find the rings of convergence of the following power series:

$$\sum_{-\infty}^{n=-1} n^3 z^n, \quad \sum_{-\infty}^{n=-1} \frac{1}{n^2} z^n, \quad \sum_{-\infty}^{n=-1} \frac{1}{2^n} z^n, \quad \sum_{-\infty}^{n=-1} 3^n z^n, \quad \sum_{-\infty}^{n=-1} \frac{1}{(-n)! n^n} z^n.$$

4.6.3. Find the ring of convergence and the sum of $\sum_{-\infty}^{n=-1} (-1)^n z^n + \sum_{n=1}^{+\infty} (\frac{1}{2i})^{n+1} z^n$.

4.6.4. (i) Using the geometric series $\sum_{n=0}^{+\infty} z^n$, write $\frac{1}{1-z}$ as a power series with disc of convergence $D_0(1)$ and as power series with ring of convergence $D_0(1, +\infty)$.
(ii) Write $\frac{1}{(z-3)(z-4)}$ as a power series with disc of convergence $D_0(3)$, as a power series with ring of convergence $D_0(3, 4)$, and as a power series with ring of convergence $D_0(4, +\infty)$.

4.6.5. If $m \in \mathbb{N}$, using the geometric series $\sum_{n=0}^{+\infty} z^n$, write $\frac{1}{(1-z)^m}$ as a power series $\sum_{n=0}^{+\infty} a_n z^n$, and determine its disc of convergence.

4.6.6. Find the radius of convergence of

$$1 + \sum_{n=1}^{+\infty} \frac{a(a+1)\cdots(a+n-1)b(b+1)\cdots(b+n-1)}{1\cdot 2\cdots n\cdot c(c+1)\cdots(c+n-1)} z^n,$$

where $c \neq 0, -1, -2, \ldots$. This power series is called **hypergeometric series** with parameters a, b, c. Prove that the function $w = F(z; a, b, c)$, which is defined by the hypergeometric series in its disc of convergence, is a solution of the differential equation

$$z(1-z)w'' + (c - (a+b+1)z)w' - abw = 0.$$

4.6.7. (i) Prove that, if two power series of the type $\sum_{n=0}^{+\infty} a_n(z-z_0)^n$ with positive radii of convergence define the same function in the intersection of their discs of convergence (with common center z_0), then the two series coincide, i.e. they have the same coefficients a_n.
(ii) Prove a result analogous to (i) for two power series of the type $\sum_{-\infty}^{n=-1} a_n(z-z_0)^n$.

4.6.8. Let $0 < R < +\infty$.
(i) If $\sum_{n=0}^{+\infty} a_n(z-z_0)^n$ converges absolutely for some $z \in C_{z_0}(R)$, prove that it converges absolutely for every $z \in \overline{D}_{z_0}(R)$.
(ii) If $\sum_{n=0}^{+\infty} a_n(z-z_0)^n$ converges for some $z \in C_{z_0}(R)$, prove that it converges absolutely for every $z \in D_{z_0}(R)$.

4.6.9. Let R', R'' and R be the radii of convergence of $\sum_{n=0}^{+\infty} a'_n(z-z_0)^n$, $\sum_{n=0}^{+\infty} a''_n(z-z_0)^n$ and $\sum_{n=0}^{+\infty}(a'_n + a''_n)(z-z_0)^n$, respectively. If $R' \neq R''$, prove that $R = \min\{R', R''\}$. If $R' = R''$, prove that $R \geq R' = R''$.

4.6.10. Let $c_n = a_0b_n + a_1b_{n-1} + \cdots + a_{n-1}b_1 + a_nb_0$ for every $n \geq 0$. If the power series $\sum_{n=0}^{+\infty} a_n(z-z_0)^n$ and $\sum_{n=0}^{+\infty} b_n(z-z_0)^n$ converge in the disc $D_{z_0}(R)$, prove that the power series $\sum_{n=0}^{+\infty} c_n(z-z_0)^n$ also converges in $D_{z_0}(R)$ and that

$$\sum_{n=0}^{+\infty} c_n(z-z_0)^n = \sum_{n=0}^{+\infty} a_n(z-z_0)^n \sum_{n=0}^{+\infty} b_n(z-z_0)^n$$

for every $z \in D_{z_0}(R)$.

4.6.11. Let R be the radius of convergence of $\sum_{n=1}^{+\infty} a_n(z-z_0)^n$. If $0 < R < +\infty$, find the radii of convergence of $\sum_{n=1}^{+\infty} n^k a_n(z-z_0)^n$, $\sum_{n=1}^{+\infty} n! a_n(z-z_0)^n$ and $\sum_{n=1}^{+\infty} \frac{a_n}{n!}(z-z_0)^n$.

4.6.12. Let $k \in \mathbb{N}$, $k \geq 2$. Find the z for which $\sum_{n=1}^{+\infty} \frac{z^{kn}}{n}$ converges.

4.6.13. Find the z for which $\sum_{n=1}^{+\infty} z^{n!}$ converges.

4.6.14. Let $0 < b < 1$. Find the ring of convergence of $\sum_{n=-\infty}^{+\infty} b^{n^2} z^n$.

4.6.15. If $s(z) = \sum_{n=0}^{+\infty} a_n(z-z_0)^n$ for every $z \in D_{z_0}(R)$ and $|a_1| \geq \sum_{n=2}^{+\infty} n|a_n|r^{n-1}$ for some r with $0 < r \leq R$, prove that s is one-to-one in $D_{z_0}(r)$. Conclude that, if $a_1 \neq 0$ and $r > 0$ is small enough, then s is one-to-one in $D_{z_0}(r)$.

4.6.16. Consider the power series $\frac{z^3}{1} - \frac{z^{2\cdot 3}}{1} + \frac{z^{3^2}}{2} - \frac{z^{2\cdot 3^2}}{2} + \cdots + + \frac{z^{3^n}}{n} - \frac{z^{2\cdot 3^n}}{n} + \cdots$. Prove that the radius of convergence of this power series is 1 and that the set of $z \in C_0(1)$ for which the power series converges as well as the set of $z \in C_0(1)$ for which the power series diverges are both dense in $C_0(1)$.

4.6.17. (i) Let $s(z) = \sum_{n=0}^{+\infty} a_n z^n$ for every $z \in D_0(R)$. Use the summation by parts formula of lemma 2.1 to prove that, if the series converges for some $\zeta \in C_0(R)$, then the series $\sum_{n=0}^{+\infty} a_n r^n \zeta^n$ converges uniformly as a series of functions of r in the interval $[0,1]$. Apply this to prove that in this case we have that $\lim_{r\to 1-} s(r\zeta) = s(\zeta)$.
(ii) Use the series in example 4.6.1 to prove that

$$\sum_{n=1}^{+\infty} \frac{1}{n}\cos(n\theta) = -\ln\left(2\sin\frac{\theta}{2}\right), \quad \sum_{n=1}^{+\infty} \frac{1}{n}\sin(n\theta) = \frac{\pi-\theta}{2}$$

for every $\theta \in (0, 2\pi)$.

Chapter 5

Local behaviour and basic properties of holomorphic functions.

5.1 The theorem of Cauchy for triangles.

Let Δ be a closed triangular region. We write $\oint_{\partial\Delta} f(z)\,dz$ to denote the curvilinear integral over a piecewise smooth curve γ with trajectory $\gamma^* = \partial\Delta$ which describes the triangle $\partial\Delta$ once and in the positive direction. For instance, if z_1, z_2, z_2 are the vertices of the triangle in the order which agrees with the positive direction of $\partial\Delta$, then a valid curve is $\gamma = [z_1, z_2] \dot{+} [z_2, z_3] \dot{+} [z_3, z_1]$. Hence,

$$\oint_{\partial\Delta} f(z)\,dz = \int_{[z_1,z_2]} f(z)\,dz + \int_{[z_2,z_3]} f(z)\,dz + \int_{[z_3,z_1]} f(z)\,dz.$$

Of course there are analogous statements for integrals $\oint_{\partial R} f(z)\,dz$, when R is a closed rectangular region or, more generally, a closed convex polygonal region.

The theorem of Cauchy-Goursat. *If f is holomorphic in an open set Ω which contains the closed triangular region Δ, then*

$$\oint_{\partial\Delta} f(z)\,dz = 0.$$

Proof. We write $I = \oint_{\partial\Delta} f(z)\,dz$, and we have to show that $I = 0$.
Let $\Delta = \Delta(z_1, z_2, z_3)$ be the given closed triangular region with vertices z_1, z_2, z_3 written in the order which agrees with the positive direction of $\partial\Delta$. We take the points w_3, w_1, w_2, which are the midpoints of the linear segments $[z_1, z_2]$, $[z_2, z_3]$, $[z_3, z_1]$, respectively. Then the closed triangular region $\Delta(z_1, z_2, z_3)$ splits into the four closed triangular regions

$$\Delta^{(1)} = \Delta(z_1, w_3, w_2),\ \ \Delta^{(2)} = \Delta(w_3, z_2, w_1),\ \ \Delta^{(3)} = \Delta(w_1, z_3, w_2),\ \ \Delta^{(4)} = \Delta(w_3, w_1, w_2)$$

and we define the corresponding curvilinear integrals:

$$I^{(1)} = \oint_{\partial\Delta^{(1)}} f(z)\,dz,\ \ I^{(2)} = \oint_{\partial\Delta^{(2)}} f(z)\,dz,\ \ I^{(3)} = \oint_{\partial\Delta^{(3)}} f(z)\,dz,\ \ I^{(4)} = \oint_{\partial\Delta^{(4)}} f(z)\,dz.$$

We analyse each of the four integrals into three integrals over the three linear segments of the corresponding triangle, we add the resulting twelve integrals and we observe the cancellations which occur between integrals over pairs of linear segments with opposite directions. We end up with six integrals over six successive linear segments which add up to give the three linear segments of the original triangle $\partial\Delta$. The result is

$$I = I^{(1)} + I^{(2)} + I^{(3)} + I^{(4)}.$$

This implies

$$|I| \le |I^{(1)}| + |I^{(2)}| + |I^{(3)}| + |I^{(4)}|$$

and hence $|I^{(j)}| \geq \frac{1}{4}|I|$ for at least one j. Now we take the corresponding closed triangular region $\Delta^{(j)}$ and, for simplicity, we denote it Δ_1. We also denote I_1 the corresponding integral $I^{(j)}$. We have proved that there is a closed triangular region Δ_1 contained in the original Δ such that, if $I = \oint_{\partial\Delta} f(z)\,dz$ and $I_1 = \oint_{\partial\Delta_1} f(z)\,dz$, then $|I_1| \geq \frac{1}{4}|I|$. We also observe that $\operatorname{diam}\Delta_1 = \frac{1}{2}\operatorname{diam}\Delta$. We may continue inductively and produce a sequence of closed triangular regions Δ_n and the corresponding sequence of curvilinear integrals

$$I_n = \oint_{\partial\Delta_n} f(z)\,dz$$

so that:
(i) $\Delta \supseteq \Delta_1 \supseteq \cdots \supseteq \Delta_n \supseteq \Delta_{n+1} \supseteq \cdots$,
(ii) $|I_n| \geq \frac{1}{4^n}|I|$,
(iii) $\operatorname{diam}\Delta_n = \frac{1}{2^n}\operatorname{diam}\Delta$.
Now, (i), (iii) imply that there is a (unique) point z contained in all Δ_n. In particular, $z \in \Delta$ and hence f is differentiable at z. If we take an arbitrary $\epsilon > 0$, then there is $\delta > 0$ so that $|\frac{f(\zeta)-f(z)}{\zeta - z} - f'(z)| < \epsilon$ for every ζ with $0 < |\zeta - z| < \delta$. Thus,

$$|f(\zeta) - f(z) - f'(z)(\zeta - z)| \leq \epsilon|\zeta - z| \tag{5.1}$$

for every ζ with $|\zeta - z| < \delta$. Because of (iii), there is some large n so that $\operatorname{diam}\Delta_n < \delta$. Since $z \in \Delta_n$ and $\operatorname{diam}\Delta_n < \delta$, we get $|\zeta - z| \leq \operatorname{diam}\Delta_n < \delta$ for every $\zeta \in \partial\Delta_n \subseteq \Delta_n$ and now (5.1) and (iii) imply

$$|f(\zeta) - f(z) - f'(z)(\zeta - z)| \leq \epsilon|\zeta - z| \leq \epsilon \operatorname{diam}\Delta_n = \frac{\epsilon}{2^n}\operatorname{diam}\Delta$$

for every $\zeta \in \partial\Delta_n$. Therefore,

$$\big|\oint_{\partial\Delta_n}(f(\zeta) - f(z) - f'(z)(\zeta - z))\,d\zeta\big| \leq \frac{\epsilon}{2^n}\operatorname{diam}\Delta\, l(\partial\Delta_n) \leq \frac{3\epsilon}{4^n}(\operatorname{diam}\Delta)^2. \tag{5.2}$$

Since $f(z) + f'(z)(\zeta - z)$ is a polynomial function of ζ, we get

$$\oint_{\partial\Delta_n}(f(z) + f'(z)(\zeta - z))\,d\zeta = 0$$

from example 4.5.1, and (5.2) becomes

$$|I_n| = |\oint_{\partial\Delta_n} f(\zeta)\,d\zeta| \leq \frac{3\epsilon}{4^n}(\operatorname{diam}\Delta)^2.$$

Finally, (ii) implies

$$|I| \leq 3\epsilon(\operatorname{diam}\Delta)^2$$

and since $\epsilon > 0$ is arbitrary, we conclude that $I = 0$. □

5.2 Primitives and the theorem of Cauchy in convex regions.

Proposition 5.1. *If f is holomorphic in the convex region Ω, then f has a primitive in Ω.*

Proof. We fix $z_0 \in \Omega$. Then for every $z \in \Omega$ the linear segment $[z_0, z]$ is contained in Ω and we define $F(z) = \int_{[z_0,z]} f(\zeta)\,d\zeta$. We shall prove that F is a primitive of f in Ω. We take arbitrary $z, w \in \Omega$ and consider the closed triangular region Δ with vertices z_0, z, w. Since Ω is convex, Δ is contained in Ω and the Cauchy-Goursat theorem implies $\oint_{\partial\Delta} f(z)\,dz = 0$, i.e.

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

Therefore $F(w) - F(z) = \int_{[z,w]} f(\zeta)\,d\zeta$ and hence

$$F(w) - F(z) - f(z)(w - z) = \int_{[z,w]} f(\zeta)\,d\zeta - f(z)\int_{[z,w]} d\zeta = \int_{[z,w]}(f(\zeta) - f(z))\,d\zeta. \tag{5.3}$$

Since f is continuous, for every $\epsilon > 0$ there is $\delta > 0$ so that $|f(\zeta) - f(z)| < \epsilon$ for every $\zeta \in \Omega$ with $|\zeta - z| < \delta$. Taking $w \in \Omega$ with $|w - z| < \delta$ we automatically have $|\zeta - z| < \delta$ for every $\zeta \in [z, w]$ and (5.3) implies

$$|F(w) - F(z) - f(z)(w - z)| \leq \epsilon |w - z|.$$

Therefore, $\left|\frac{F(w)-F(z)}{w-z} - f(z)\right| \leq \epsilon$ for every w with $0 < |w-z| < \delta$ and hence $F'(z) = f(z)$. $\square$

The theorem of Cauchy in convex regions. *If f is holomorphic in the convex region Ω, then*

$$\oint_\gamma f(z)\, dz = 0$$

for every closed piecewise smooth curve γ in Ω.

Proof. Direct from propositions 4.10 and 5.1. $\square$

Now we shall decribe a very useful technique to handle curvilinear integrals of holomorphic functions. Every closed piecewise smooth curve γ we shall refer to will be *visually simple*, for instance a circle or a triangle or a rectangle, and we shall be able to distinguish between the points *inside* γ and the points *outside* γ. We assume that γ surrounds every point inside it once and in the positive direction and that it does not surround the points outside it. The points inside γ form the *region inside* γ and the points outside γ form the *region outside* γ. Then γ^* is the common boundary of the region inside γ and the region outside γ. We shall concentrate on two characteristic cases.

First case. Let f be holomorphic in the open set Ω and let γ be a closed piecewise smooth curve in Ω. We want to evaluate $\oint_\gamma f(z)\, dz$.

If Ω is convex, then $\oint_\gamma f(z)\, dz = 0$. So let us assume that Ω is not convex. To continue, we assume that the region inside γ, call it D, is contained in Ω, and hence f is holomorphic in D as well as in $\partial D = \gamma^*$. Now our technique is the following. We split D into specific disjoint open sets $E_1, \dots, E_m$ so that their boundaries $\partial E_1, \dots, \partial E_m$ are trajectories of closed piecewise smooth curves $\sigma_1, \dots, \sigma_m$, so that $\overline{D} = \overline{E}_1 \cup \cdots \cup \overline{E}_m$ and, finally, so that, when we analyse in an appropriate way each of $\sigma_1, \dots, \sigma_m$ in successive subcurves and drop those subcurves which appear as pairs of opposite curves, the remaining subcurves can be summed up to give the original curve γ. The result is:

$$\oint_\gamma f(z)\, dz = \oint_{\sigma_1} f(z)\, dz + \cdots + \oint_{\sigma_m} f(z)\, dz.$$

In fact we applied this technique in the proof of the theorem of Cauchy-Goursat.

Now, if the various $E_1, \dots, E_m$ can be chosen so that each $\overline{E}_1, \dots, \overline{E}_m$ is contained in a corresponding *convex* open subset of Ω, then we conclude that

$$\oint_\gamma f(z)\, dz = \oint_{\sigma_1} f(z)\, dz + \cdots + \oint_{\sigma_m} f(z)\, dz = 0 + \cdots + 0 = 0.$$

Second case. Let f be holomorphic in the open set Ω and let $\gamma, \gamma_1, \dots, \gamma_n$ be $n+1$ closed piecewise smooth curves in Ω. We want to relate $\oint_\gamma f(z)\, dz$, $\oint_{\gamma_1} f(z)\, dz, \dots, \oint_{\gamma_n} f(z)\, dz$.

We assume that the regions inside $\gamma_1, \dots, \gamma_n$ are disjoint and that they are all contained in the region inside γ. Let us call D the *intermediate* region, i.e. the set consisting of the points which are inside γ and outside every $\gamma_1, \dots, \gamma_n$, i.e. the intersection of the region inside γ and the regions ouside $\gamma_1, \dots, \gamma_n$. We further assume that D is a subset of Ω, and hence f is holomorphic in D as well as in $\partial D = \gamma^* \cup \gamma_1^* \cup \cdots \cup \gamma_n^*$. Now, here is the technique. We split D into specific disjoint open sets $E_1, \dots, E_m$ so that their boundaries $\partial E_1, \dots, \partial E_m$ are trajectories of closed piecewise smooth curves $\sigma_1, \dots, \sigma_m$, so that $\overline{E} = \overline{E}_1 \cup \cdots \cup \overline{E}_m$ and, finally, so that, when we analyse in an appropriate way each of $\sigma_1, \dots, \sigma_m$ in successive subcurves and drop those subcurves which appear as pairs of opposite curves, the remaining subcurves can be summed up to give γ as well as the opposites of $\gamma_1, \dots, \gamma_n$. The result is:

$$\oint_\gamma f(z)\,dz - \oint_{\gamma_1} f(z)\,dz - \cdots - \oint_{\gamma_n} f(z)\,dz = \oint_{\sigma_1} f(z)\,dz + \cdots + \oint_{\sigma_m} f(z)\,dz.$$

If the various $E_1, \dots, E_m$ can be chosen so that each $\overline{E}_1, \dots, \overline{E}_m$ is contained in a corresponding *convex* open subset of Ω, then $\oint_{\sigma_1} f(z)\,dz + \cdots + \oint_{\sigma_m} f(z)\,dz = 0 + \cdots + 0 = 0$ and hence

$$\oint_\gamma f(z)\,dz = \oint_{\gamma_1} f(z)\,dz + \cdots + \oint_{\gamma_n} f(z)\,dz.$$

Corollary 5.1. *Let $C, C_1, \dots, C_n$ be $n+1$ circles and let $D, D_1, \dots, D_n$ be the corresponding open discs. Assume that $\overline{D_1}, \dots, \overline{D_n}$ are disjoint and that they are all contained in D. Consider also the closed region $M = \overline{D} \setminus (D_1 \cup \cdots \cup D_n)$. If $f : \Omega \to \mathbb{C}$ is holomorphic in an open set Ω which contains M, then*

$$\oint_C f(z)\,dz = \oint_{C_1} f(z)\,dz + \cdots + \oint_{C_n} f(z)\,dz.$$

Instead of circles we may consider rectangles or triangles or any combination of the three shapes.

Exercises.

5.2.1. Let γ_R be the closed piecewise smooth curve which is the sum of the linear segment $[0, R]$, the arc of the circle $C_0(R)$ from R to $Re^{i\frac{\pi}{4}}$ in the positive direction and the linear segment $[Re^{i\frac{\pi}{4}}, 0]$. Also, let σ_R be the curve wich describes only the above arc from R to $Re^{i\frac{\pi}{4}}$.
(i) Prove that $\int_{\sigma_R} e^{-z^2}\,dz \to 0$ when $R \to +\infty$.
(ii) Using γ_R appropriately together with the formula $\int_0^{+\infty} e^{-x^2}\,dt = \frac{\sqrt{\pi}}{2}$, prove the formulas for the so-called Fresnel integrals:

$$\int_0^{+\infty} \sin(x^2)\,dx = \int_0^{+\infty} \cos(x^2)\,dx = \frac{\sqrt{\pi}}{2\sqrt{2}}.$$

5.2.2. Let $y, R > 0$ and $\gamma_{R,y}$ be the closed piecewise smooth curve which is the sum of the linear segments $[-R, R]$, $[R, R+iy]$, $[R+iy, -R+iy]$ and $[-R+iy, -R]$.
(i) If $y > 0$ is constant, prove that $\int_{[R,R+iy]} e^{-z^2}\,dz \to 0$ and $\int_{[-R+iy,-R]} e^{-z^2}\,dz \to 0$ when $R \to +\infty$.
(ii) Using $\gamma_{R,y}$ appropriately, prove that $\int_{-\infty}^{+\infty} e^{-(x+iy)^2}\,dx$ does not depend on $y \in [0, +\infty)$.
(iii) Using the formula $\int_0^{+\infty} e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}$, prove that

$$\int_{-\infty}^{+\infty} e^{-x^2} \cos(2xy)\,dx = \sqrt{\pi}e^{-y^2}$$

for every $y \geq 0$ (and hence for every $y \leq 0$ also). This identity is very important in harmonic analysis.

5.3 Cauchy's formulas for circles and infinite differentiability.

Cauchy's formula for circles. *If f is holomorphic in an open set Ω containing the closed disc $\overline{D}_{z_0}(R)$, then*

$$f(z) = \frac{1}{2\pi i} \oint_{C_{z_0}(R)} \frac{f(\zeta)}{\zeta - z}\,d\zeta$$

for every $z \in D_{z_0}(R)$.

Proof. Let $z \in D_{z_0}(R)$. We consider any open disc $D_z(r)$ with $r < R - |z - z_0|$. Then $\overline{D}_z(r) \subseteq D_{z_0}(R)$ and the function $\frac{f(\zeta)}{\zeta - z}$ is holomorphic in the open set $\Omega \setminus \{z\}$ which contains the closed region between the circles $C_z(r)$ and $C_{z_0}(R)$. Corollary 5.1 implies

$$\oint_{C_{z_0}(R)} \frac{f(\zeta)}{\zeta - z}\,d\zeta = \oint_{C_z(r)} \frac{f(\zeta)}{\zeta - z}\,d\zeta. \tag{5.4}$$

Now, we have $\oint_{C_z(r)} \frac{1}{\zeta - z}\, d\zeta = \int_0^{2\pi} \frac{1}{re^{it}}\, ire^{it}\, dt = 2\pi i$ and hence

$$\oint_{C_z(r)} \frac{f(\zeta)}{\zeta - z}\, d\zeta - 2\pi i f(z) = \oint_{C_z(r)} \frac{f(\zeta) - f(z)}{\zeta - z}\, d\zeta. \tag{5.5}$$

We take $\epsilon > 0$. Since f is continuous at z, there is $\delta > 0$ so that $|f(\zeta) - f(z)| < \epsilon$ for every $\zeta \in \Omega$ with $|\zeta - z| < \delta$. Therefore, if $r < \delta$, (5.5) implies

$$\Big| \oint_{C_z(r)} \frac{f(\zeta)}{\zeta - z}\, d\zeta - 2\pi i f(z) \Big| \leq \frac{\epsilon}{r}\, 2\pi r = 2\pi\epsilon.$$

Since ϵ is arbitrary, we conclude that

$$\lim_{r \to 0} \oint_{C_z(r)} \frac{f(\zeta)}{\zeta - z}\, d\zeta = 2\pi i f(z).$$

Now, letting $r \to 0$ in (5.4), we get $\oint_{C_{z_0}(R)} \frac{f(\zeta)}{\zeta - z}\, d\zeta = 2\pi i f(z)$. $\square$

A particular instance of the formula of Cauchy is when $z = z_0$, the center of the circle $C_{z_0}(R)$. Using the parametric equation $\zeta = z_0 + Re^{it}$, $t \in [0, 2\pi]$, we get

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + Re^{it})\, dt$$

and this is called **mean value property** of the holomorphic function f.

Cauchy's formula for derivatives and circles. *If f is holomorphic in an open set Ω containing the closed disc $\overline{D}_{z_0}(R)$, then f is infinitely many times differentiable at every $z \in D_{z_0}(R)$ and*

$$f^{(n)}(z) = \frac{n!}{2\pi i} \oint_{C_{z_0}(R)} \frac{f(\zeta)}{(\zeta - z)^{n+1}}\, d\zeta$$

for every $z \in D_{z_0}(R)$ and every $n \in \mathbb{N}$.

Proof. Proposition 4.12 says that $\frac{1}{2\pi i} \oint_{C_{z_0}(R)} \frac{f(\zeta)}{\zeta - z}\, d\zeta$ is an infinitely many times differentiable function of z in the disc $D_{z_0}(R)$. On the other hand, Cauchy's formula says that this function coincides with the function f in the same disc. Therefore f is infinitely many times differentiable in $D_{z_0}(R)$. Moreover, the derivatives of f are the same as the derivatives of $\frac{1}{2\pi i} \oint_{C_{z_0}(R)} \frac{f(\zeta)}{\zeta - z}\, d\zeta$ and these are given by the formulas in proposition 4.12. $\square$

Example 5.3.1. Let $n \in \mathbb{N}$. Then

$$\oint_{C_{z_0}(R)} \frac{1}{(\zeta - z)^n}\, d\zeta = 0$$

for every $z \notin \overline{D}_{z_0}(R)$. To see this we observe that the circle $C_{z_0}(R)$ is contained in a slightly larger open disc $D_{z_0}(R')$ which does not contain z: it is enough to take $R < R' < |z - z_0|$. Then the disc $D_{z_0}(R')$ is a convex region and $\frac{1}{(\zeta - z)^n}$ is a holomorphic function of ζ in $D_{z_0}(R')$. Now the result is an application of the theorem of Cauchy in convex regions.
On the other hand, for every $z \in D_{z_0}(R)$ we have

$$\oint_{C_{z_0}(R)} \frac{1}{(\zeta - z)^n}\, d\zeta = \begin{cases} 2\pi i, & \text{if } n = 1, \\ 0, & \text{if } n \geq 2 \end{cases}$$

This is a simple application of Cauchy's formula (for a function and its derivatives) to the constant function 1. The special case $z = z_0$ we have already seen in examples 2.2.9 and 4.2.2 and the general case (for $n \geq 2$) in example 4.5.3.

Theorem 5.1. *If f is holomorphic in the open set Ω, then f is infinitely many times differentiable in Ω.*

Proof. Let $z_0 \in \Omega$. We take a closed disc $\overline{D}_{z_0}(R) \subseteq \Omega$ and then f is infinitely many times differentiable in $D_{z_0}(R)$ and hence at z_0. □

It is time to recall the remark after theorem 4.1. The assumption of continuity of the derivative in theorem 4.1 is superfluous. The same we may say for the hypothesis in example 4.5.6 and in exercises 3.3.3 and 4.5.1.

Cauchy's estimates. *If f is holomorphic in an open set containing the closed disc $\overline{D}_{z_0}(R)$ and if $|f(\zeta)| \leq M$ for every $\zeta \in C_{z_0}(R)$, then*

$$|f^{(n)}(z_0)| \leq \tfrac{n!M}{R^n}$$

for every $n \in \mathbb{N}$.

Proof. Direct application of Cauchy's formulas. □

Exercises.

5.3.1. Evaluate $\oint_{C_0(r)} \frac{z^2+1}{z(z^2+4)}\,dz$ for $0 < r < 2$ and for $2 < r < +\infty$.

5.3.2. If $n \in \mathbb{N}$, evaluate

$$\oint_{C_0(1)} \tfrac{e^z}{z^n}\,dz, \quad \int_0^{2\pi} e^{\cos\theta} \sin(n\theta - \sin\theta)\,d\theta, \quad \int_0^{2\pi} e^{\cos\theta} \cos(n\theta - \sin\theta)\,d\theta.$$

5.3.3. If $n \in \mathbb{N}$, evaluate

$$\oint_{C_0(1)} \tfrac{e^{iz}}{z^n}\,dz, \quad \oint_{C_0(1)} \tfrac{\sin z}{z^n}\,dz, \quad \oint_{C_0(1)} \tfrac{e^z - e^{-z}}{z^n}\,dz, \quad \oint_{C_1(\frac{1}{2})} \tfrac{\operatorname{Log} z}{(z-1)^n}\,dz.$$

5.3.4. Let f be holomorphic in $\mathbb{C}$ and let $|f(z)| \leq A + M|z|^n$ for every z. Prove that $f^{(n+1)}(z) = 0$ for every z and that f is a polynomial function of degree $\leq n$.

5.3.5. Let the complex function f be continuous in $\overline{D}_{z_0}(R)$ and holomorphic in $D_{z_0}(R)$. Prove that $f(z) = \frac{1}{2\pi i} \oint_{C_{z_0}(R)} \frac{f(\zeta)}{\zeta - z}\,d\zeta$ for every $z \in D_{z_0}(R)$.

5.3.6. Let f be holomorphic in an open set containing the closed disc $\overline{D}_{z_0}(R)$ and let $0 < r < R$. If $|f(z)| \leq M$ for every $z \in C_{z_0}(R)$, find an upper bound for $|f^{(n)}|$ in $\overline{D}_{z_0}(r)$, which depends only on n, r, R, M and not on f or z_0.

5.3.7. Let f be holomorphic in $D_{z_0}(R)$. If $|f(z)| \leq \frac{1}{R - |z - z_0|}$ for every $z \in D_{z_0}(R)$, find the smallest possible upper bound for $|f^{(n)}(z_0)|$, which depends only on n, R and not on f or z_0.

5.3.8. Let f be holomorphic in $\mathbb{D}$ with $\iint_{\mathbb{D}} |f(z)|\,dxdy < +\infty$ $(z = x + iy)$. Prove that

$$f(w) = \tfrac{1}{\pi} \iint_{\mathbb{D}} \tfrac{f(z)}{(1 - \overline{z}w)^2}\,dxdy \qquad (z = x + iy)$$

for every $w \in \mathbb{D}$.

5.3.9. Let f be holomorphic in $D_{z_0}(R)$.
(i) Using the mean value property, prove that

$$f(z_0) = \tfrac{1}{\pi R^2} \iint_{D_{z_0}(R)} f(z)\,dxdy \qquad (z = x + iy).$$

(ii) If $1 \leq p < +\infty$, prove that

$$|f(z_0)|^p \leq \tfrac{1}{\pi R^2} \iint_{D_{z_0}(R)} |f(z)|^p\,dxdy \qquad (z = x + iy).$$

5.3.10. Prove that

$$\int_0^{2\pi} \ln|1 - ae^{i\theta}|\,d\theta = \begin{cases} 2\pi \ln|a|, & \text{if } |a| \geq 1 \\ 0, & \text{if } |a| \leq 1 \end{cases}$$

5.4 Morera's theorem.

Theorem 5.1 and proposition 4.10 imply the following corollary. *If the complex function f is continuous in the region $\Omega \subseteq \mathbb{C}$ and if $\oint_\gamma f(z)\,dz = 0$ for every closed piecewise smooth curve γ in Ω, then f is holomorphic in Ω.* Indeed, since $\oint_\gamma f(z)\,dz = 0$ for every closed piecewise smooth curve γ in Ω, we get that f has a primitive, say F, in Ω. This means that $F' = f$ in Ω and hence F is holomorphic in Ω. Therefore, F is infinitely many times differentiable in Ω and then f is also infinitely many times differentiable in Ω. In particular, f is holomorphic in Ω.

The next theorem proves the same result with weaker assumptions.

The theorem of Morera. *If the complex function f is continuous in the open set $\Omega \subseteq \mathbb{C}$ and if $\oint_{\partial\Delta} f(z)\,dz = 0$ for every closed triangular region Δ in Ω, then f is holomorphic in Ω.*

Proof. Let $z_0 \in \Omega$. We consider a disc $D_{z_0}(R) \subseteq \Omega$. This disc is a convex set and we have that $\oint_{\partial\Delta} f(z)\,dz = 0$ for every closed triangular region Δ in $D_{z_0}(R)$. Then the *proof* of proposition 5.1 applies, and we get that f has a primitive, say F, in $D_{z_0}(R)$. This means that $F' = f$ in $D_{z_0}(R)$ and hence F is holomorphic in $D_{z_0}(R)$. Therefore, F is infinitely many times differentiable in $D_{z_0}(R)$ and f is also infinitely many times differentiable in $D_{z_0}(R)$. In particular, f is holomorphic in $D_{z_0}(R)$ and hence at z_0. $\square$

Exercises.

5.4.1. If the complex function f is continuous in the open set Ω and holomorphic in $\Omega \setminus l$, where l is a line, prove that f is holomorphic in Ω.

5.5 Liouville's theorem. The fundamental theorem of algebra.

The theorem of Liouville. *If f is holomorphic and bounded in $\mathbb{C}$, then f is constant in $\mathbb{C}$.*

Proof. There is $M \geq 0$ so that $|f(z)| \leq M$ for every z. We take any z_0 and apply Cauchy's estimate for $n = 1$ with an arbitrary circle $C_{z_0}(R)$ and we find that $|f'(z_0)| \leq \frac{M}{R}$. Letting $R \to +\infty$, we get $f'(z_0) = 0$. Since z_0 is arbitrary, we conclude that f is constant. $\square$

Fundamental theorem of algebra. *Every polynomial of degree ≥ 1 has at least one root in $\mathbb{C}$.*

Proof. Let p be a polynomial of degree ≥ 1 and assume that p has no root in $\mathbb{C}$.
We consider the function $f = \frac{1}{p}$, which is holomorphic in $\mathbb{C}$, and we see easily that it is also bounded in $\mathbb{C}$. Indeed, since $\lim_{z\to\infty} p(z) = \infty$, we have $\lim_{z\to\infty} f(z) = 0$, and hence there is $R > 0$ so that $|f(z)| \leq 1$ for every z with $|z| > R$. Since $|f|$ is continuous in the compact disc $\overline{D}_0(R)$, there is $M' \geq 0$ so that $|f(z)| \leq M'$ for every z with $|z| \leq R$. Taking $M = \max\{M', 1\}$, we have that $|f(z)| \leq M$ for every z and hence f is bounded.
Liouville's theorem implies that f and hence p is constant and we arrive at a contradiction. $\square$

Having proved that a polynomial p has a root z_1, we may prove in a purely algebraic way that $z - z_1$ is a factor of p, i.e. there is a polynomial p_1 so that $p(z) = (z - z_1)p_1(z)$ for every z. Continuing inductively, we conclude that, if $n \geq 1$ is the degree of p, there are $z_1, \dots, z_n$ so that

$$p(z) = c(z - z_1)\cdots(z - z_n) \qquad \text{for every } z$$

where c is a constant. Thus, *every polynomial p of degree $n \geq 1$ has exactly n roots in $\mathbb{C}$.*

Exercises.

5.5.1. If $f : \mathbb{C} \to \mathbb{C}$ is holomorphic in $\mathbb{C}$ and $\mathrm{Re}\, f$ is bounded in $\mathbb{C}$, prove that f is constant in $\mathbb{C}$.

5.5.2. We say that z, w are symmetric with respect to $\mathbb{T}$ if either $z = 0, w = \infty$ or $z = \infty, w = 0$ or $z, w \in \mathbb{C}$, $z = \frac{1}{\overline{w}}$.
Let p, q be two polynomials with no common root and so that $|p(z)| = |q(z)|$ for every $z \in \mathbb{T}$. Prove that, if $a \in \mathbb{C} \setminus \{0\}$ is a root of p of multiplicity k, then $b = \frac{1}{\overline{a}}$ is a root of q of multiplicity k and conversely. I.e. the roots of p and the roots of q form pairs of points symmetric with respect to $\mathbb{T}$. (In particular, p and q have the same degree.)

5.6 Taylor series and Laurent series.

Proposition 5.2. *Let f be holomorphic in the open set Ω, $z_0 \in \Omega$ and let $D_{z_0}(R)$ be the largest disc with center z_0 which is contained in Ω. Then there is a unique power series $\sum_{n=0}^{+\infty} a_n(z-z_0)^n$ so that*

$$f(z) = \sum_{n=0}^{+\infty} a_n(z - z_0)^n$$

for every $z \in D_{z_0}(R)$. The coefficients are given by

$$a_n = \frac{f^{(n)}(z_0)}{n!} = \frac{1}{2\pi i} \oint_{C_{z_0}(r)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta$$

for $0 < r < R$.

Proof. We take $z \in D_{z_0}(R)$, and then $|z - z_0| < R$. If $|z - z_0| < r < R$, then $z \in D_{z_0}(r)$ and, according to the formula of Cauchy, we have

$$f(z) = \frac{1}{2\pi i} \oint_{C_{z_0}(r)} \frac{f(\zeta)}{\zeta-z}\, d\zeta. \tag{5.6}$$

Now for every $\zeta \in C_{z_0}(r)$ we have $\left|\frac{z-z_0}{\zeta-z_0}\right| = \frac{|z-z_0|}{r} < 1$ and hence

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

The test of Weierstrass implies that $\sum_{n=0}^{+\infty}\left(\frac{z-z_0}{\zeta-z_0}\right)^n$ converges, as a series of functions of ζ, uniformly in $C_{z_0}(r)$. Indeed, $\left|\frac{z-z_0}{\zeta-z_0}\right|^n = \left(\frac{|z-z_0|}{r}\right)^n$ for every $\zeta \in C_{z_0}(r)$ and $\sum_{n=0}^{+\infty}\left(\frac{|z-z_0|}{r}\right)^n < +\infty$. So from (5.6) we have that

$$f(z) = \sum_{n=0}^{+\infty} \frac{1}{2\pi i} \oint_{C_{z_0}(r)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta\, (z - z_0)^n. \tag{5.7}$$

Now, we observe that the radius r has been chosen to satisfy the inequality $|z - z_0| < r < R$ and so the integrals $\frac{1}{2\pi i} \int_{C_{z_0}(r)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta$ depend *a priori* on z. But there are two reasons that these integrals actually do not depend on the value of r in the interval $(0, R)$ and hence on z. The first reason is that from the formulas of Cauchy we get

$$\frac{1}{2\pi i} \oint_{C_{z_0}(r)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta = \frac{f^{(n)}(z_0)}{n!}$$

when $0 < r < R$. The second reason is that $\frac{f(\zeta)}{(\zeta-z_0)^{n+1}}$ is holomorphic in $D_{z_0}(R) \setminus \{z_0\}$, and because of corollary 5.1, we have

$$\frac{1}{2\pi i} \oint_{C_{z_0}(r_1)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta = \frac{1}{2\pi i} \oint_{C_{z_0}(r_2)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta$$

when $0 < r_1 < r_2 < R$. We conclude from (5.7) that $f(z) = \sum_{n=0}^{+\infty} a_n(z - z_0)^n$ for every $z \in D_{z_0}(R)$, where $a_n = \frac{f^{(n)}(z_0)}{n!} = \frac{1}{2\pi i} \oint_{C_{z_0}(r)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta$ for $0 < r < R$.

Regarding uniqueness, assume that $f(z) = \sum_{n=0}^{+\infty} b_n(z-z_0)^n$ for every $z \in D_{z_0}(R)$. Then, if $0 < r < R$, the series $\sum_{n=0}^{+\infty} b_n(z-z_0)^n$ converges uniformly in $C_{z_0}(r)$ and we get

$$\begin{aligned} 2\pi i a_k &= \oint_{C_{z_0}(r)} \tfrac{f(\zeta)}{(\zeta-z_0)^{k+1}}\, d\zeta = \oint_{C_{z_0}(r)} \tfrac{1}{(\zeta-z_0)^{k+1}} \sum_{n=0}^{+\infty} b_n(\zeta-z_0)^n\, d\zeta \\ &= \sum_{n=0}^{+\infty} b_n \oint_{C_{z_0}(r)} (\zeta-z_0)^{n-k-1}\, d\zeta = 2\pi i b_k. \end{aligned}$$

The last equality uses the calculation in example 4.2.2. □

The power series provided by proposition 5.2 is called **Taylor series** of f in the disc $D_{z_0}(R)$, the largest open disc with center z_0 which is contained in the domain of holomorphy of f.

Example 5.6.1. The function $f(z) = \frac{1}{1-z}$ is holomorphic in $\mathbb{C} \setminus \{1\}$ and the largest open disc with center 0 which is contained in $\mathbb{C} \setminus \{1\}$ is $D_0(1)$. To find the Taylor series of f in $D_0(1)$ we calculate the derivatives $f^{(n)}(z) = \frac{n!}{(1-z)^{n+1}}$ for every $n \geq 0$. Thus, $a_n = \frac{f^{(n)}(0)}{n!} = 1$ for every $n \geq 0$ and the Taylor series of f is $\sum_{n=0}^{+\infty} z^n$. I.e. $\frac{1}{1-z} = \sum_{n=0}^{+\infty} z^n$ for every $z \in D_0(1)$. Of course, this is already known.

Example 5.6.2. The function $f(z) = \frac{1}{1+z^2} = \frac{1}{(z+i)(z-i)}$ is holomorphic in the open set $\mathbb{C} \setminus \{i, -i\}$ and the largest open disc with center 0 which is contained in $\mathbb{C} \setminus \{i, -i\}$ is $D_0(1)$. To find the Taylor series of f in $D_0(1)$ we calculate the derivatives of f. We write $f(z) = -\frac{1}{2i}\big(\frac{1}{i-z} + \frac{1}{i+z}\big)$ and get

$$f^{(n)}(z) = -\tfrac{1}{2i}\big(\tfrac{n!}{(i-z)^{n+1}} + (-1)^n \tfrac{n!}{(i+z)^{n+1}}\big)$$

for every $n \geq 0$. Hence $a_n = \frac{f^{(n)}(0)}{n!} = \frac{1+(-1)^n}{2i^n}$ for every $n \geq 0$. Thus, $a_n = 0$, if n is odd, and $a_n = \frac{1}{i^n} = (-1)^{\frac{n}{2}}$, if n is even. So the Taylor series of f is $\sum_{k=0}^{+\infty}(-1)^k z^{2k}$. I.e. $\frac{1}{1+z^2} = \sum_{k=0}^{+\infty}(-1)^k z^{2k}$ for every $z \in D_0(1)$.
We may find the same formula if we use the Taylor series of $\frac{1}{1-z}$, i.e. $\frac{1}{1-z} = \sum_{n=0}^{+\infty} z^n$. We replace z with $-z^2$ and find $\frac{1}{1+z^2} = \sum_{n=0}^{+\infty}(-1)^n z^{2n}$. From the moment that we have found *some* power series which coincides with our function in $D_0(1)$, then, because of uniqueness, *this* is the Taylor series of our function.

Example 5.6.3. The exponential function $f(z) = e^z$ is holomorphic in $\mathbb{C}$ and the largest open disc with center 0 which is contained in $\mathbb{C}$ is $D_0(+\infty) = \mathbb{C}$. The derivatives of f are $f^{(n)}(z) = e^z$ for every $n \geq 0$ and the coefficients of the Taylor series of f are $a_n = \frac{f^{(n)}(0)}{n!} = \frac{1}{n!}$ for every $n \geq 0$. Thus, the Taylor series of f is $\sum_{n=0}^{+\infty} \frac{1}{n!} z^n$ and we have

$$e^z = \sum_{n=0}^{+\infty} \tfrac{1}{n!} z^n$$

for every z. We have proven this identity differently in example 4.6.3.

Example 5.6.4. The function $f(z) = \cos z$ is holomorphic in $\mathbb{C}$ and the largest open disc with center 0 which is contained in $\mathbb{C}$ is $D_0(+\infty) = \mathbb{C}$. The derivatives of f are $f^{(n)}(z) = (-1)^{\frac{n}{2}} \cos z$ for even n and $f^{(n)}(z) = (-1)^{\frac{n+1}{2}} \sin z$ for odd n. Therefore, the coefficients of the Taylor series are $a_n = \frac{f^{(n)}(0)}{n!} = \frac{(-1)^{\frac{n}{2}}}{n!}$ for even n and $a_n = \frac{f^{(n)}(0)}{n!} = 0$ for odd n. Thus, the Taylor series of f is $\sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k)!} z^{2k}$ and we have

$$\cos z = \sum_{k=0}^{+\infty} \tfrac{(-1)^k}{(2k)!} z^{2k}$$

for every z. In the same manner we can prove that

$$\sin z = \sum_{k=1}^{+\infty} \tfrac{(-1)^{k-1}}{(2k-1)!} z^{2k-1}$$

for every z.
Another way to find the Taylor series of cos and sin is through the definitions of the two functions and the Taylor series of e^z. For instance:

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

The power series we found coincides with the function cos in the largest open disc with center 0 which is contained in the domain of holomorphy of cos and, because of uniqueness, this is the Taylor series of cos.

Example 5.6.5. The function $f(z) = -\operatorname{Log}(1-z)$ is defined and holomorphic in $\mathbb{C}\setminus[1,+\infty)$. The largest disc with center 0 in $\mathbb{C}\setminus[1,+\infty)$ is $\mathbb{D}$. The derivatives of f are $f^{(n)}(z) = \frac{(n-1)!}{(1-z)^n}$ for every $n \geq 1$. Thus, $a_0 = 0$ and $a_n = \frac{f^{(n)}(0)}{n!} = \frac{1}{n}$ for every $n \geq 1$ and the Taylor series of f is $\sum_{n=1}^{+\infty}\frac{z^n}{n}$. I.e.

$$-\operatorname{Log}(1-z) = \sum_{n=1}^{+\infty}\tfrac{z^n}{n}$$

for every $z \in \mathbb{D}$. We found the same result in example 4.6.1.

Proposition 5.3. *Let f be holomorphic in the open set Ω and let $D_{z_0}(R_1,R_2)$ be a largest open ring with center z_0 which is contained in Ω. Then there is a unique power series $\sum_{-\infty}^{+\infty}a_n(z-z_0)^n$ so that*

$$f(z) = \sum_{-\infty}^{+\infty}a_n(z-z_0)^n$$

for every $z \in D_{z_0}(R_1,R_2)$. The coefficients are given by

$$a_n = \tfrac{1}{2\pi i}\oint_{C_{z_0}(r)}\tfrac{f(\zeta)}{(\zeta-z_0)^{n+1}}\,d\zeta$$

for $R_1 < r < R_2$.

Proof. We take $z \in D_{z_0}(R_1,R_2)$, and then $R_1 < |z-z_0| < R_2$. We choose any r_1, r_2 so that $R_1 < r_1 < |z-z_0| < r_2 < R_2$. Then $z \in D_{z_0}(r_1,r_2)$ and

$$f(z) = \tfrac{1}{2\pi i}\oint_{C_{z_0}(r_2)}\tfrac{f(\zeta)}{\zeta-z}\,d\zeta - \tfrac{1}{2\pi i}\oint_{C_{z_0}(r_1)}\tfrac{f(\zeta)}{\zeta-z}\,d\zeta. \tag{5.8}$$

To prove (5.8), we consider an open disc $D_z(r)$ with $r < \min\{r_2-|z-z_0|, |z-z_0|-r_1\}$. Then $\overline{D}_z(r) \subseteq D_{z_0}(r_1,r_2)$ and we apply corollary 5.1 to $\frac{f(\zeta)}{\zeta-z}$, which is a holomorphic function of ζ in $D_{z_0}(R_1,R_2)\setminus\{z\}$. We get

$$\oint_{C_{z_0}(r_2)}\tfrac{f(\zeta)}{\zeta-z}\,dz - \oint_{C_{z_0}(r_1)}\tfrac{f(\zeta)}{\zeta-z}\,dz = \oint_{C_z(r)}\tfrac{f(\zeta)}{\zeta-z}\,d\zeta.$$

Now as in the proof of Cauchy's formula for circles, we have

$$\lim_{r\to0}\oint_{C_z(r)}\tfrac{f(\zeta)}{\zeta-z}\,d\zeta = 2\pi i f(z)$$

and the proof of (5.8) is complete.
For every $\zeta \in C_{z_0}(r_2)$ we have

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

because $\left|\frac{z-z_0}{\zeta-z_0}\right| = \frac{|z-z_0|}{r_2} < 1$. Similarly, for every $\zeta \in C_{z_0}(r_1)$ we have

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

because $\left|\frac{\zeta-z_0}{z-z_0}\right| = \frac{r_1}{|z-z_0|} < 1$. Exactly as in the proof of proposition 5.2, we see that these two series of functions converge uniformly and (5.8) implies

$$f(z) = \sum_{n=0}^{+\infty} \frac{1}{2\pi i} \oint_{C_{z_0}(r_2)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta\, (z-z_0)^n$$
$$+ \sum_{n=0}^{+\infty} \frac{1}{2\pi i} \oint_{C_{z_0}(r_1)} f(\zeta)(\zeta-z_0)^n\, d\zeta\, \frac{1}{(z-z_0)^{n+1}}.$$

In the last series we change $n+1$ to $-n$ and get

$$\begin{aligned} f(z) &= \sum_{n=0}^{+\infty} \frac{1}{2\pi i} \oint_{C_{z_0}(r_2)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta\, (z-z_0)^n \\ &\quad + \sum_{-\infty}^{n=-1} \frac{1}{2\pi i} \oint_{C_{z_0}(r_1)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta\, (z-z_0)^n. \end{aligned} \tag{5.9}$$

Now, $\frac{f(\zeta)}{(\zeta-z_0)^{n+1}}$ is holomorphic in $D_{z_0}(R_1, R_2)$ and another application of corollary 5.1 implies that

$$\oint_{C_{z_0}(r_1)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta = \oint_{C_{z_0}(r_2)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta$$

for $R_1 < r_1 < r_2 < R_2$. Therefore the coefficients of both series in (5.9) do not depend on the values of r_1, r_2, and we replace both radii with any r with $R_1 < r < R_2$. We conclude that $f(z) = \sum_{-\infty}^{+\infty} a_n(z-z_0)^n$ for every $z \in D_{z_0}(R_1, R_2)$, where $a_n = \frac{1}{2\pi i} \oint_{C_{z_0}(r)} \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta$ for $R_1 < r < R_2$.
Regarding uniqueness, assume that $f(z) = \sum_{-\infty}^{+\infty} b_n(z-z_0)^n$ for every $z \in D_{z_0}(R_1, R_2)$. We take any r with $R_1 < r < R_2$, and then $\sum_{-\infty}^{+\infty} b_n(z-z_0)^n$ converges uniformly in $C_{z_0}(r)$. Then

$$\begin{aligned} 2\pi i a_k &= \oint_{C_{z_0}(r)} \frac{f(\zeta)}{(\zeta-z_0)^{k+1}}\, d\zeta = \oint_{C_{z_0}(r)} \frac{1}{(\zeta-z_0)^{k+1}} \sum_{-\infty}^{+\infty} b_n(\zeta-z_0)^n\, d\zeta \\ &= \sum_{-\infty}^{+\infty} b_n \oint_{C_{z_0}(r)} (\zeta-z_0)^{n-k-1}\, d\zeta = 2\pi i b_k \end{aligned}$$

and we get that $b_k = a_k$ for every k. $\square$

The power series given by proposition 5.3 is called **Laurent series** of f in $D_{z_0}(R_1, R_2)$, a largest open ring with center z_0 which is contained in the domain of holomorphy of f.

Example 5.6.6. The function $f(z) = \frac{1}{z}$ is holomorphic in $\mathbb{C} \setminus \{0\}$. The ring $D_0(0, +\infty) = \mathbb{C} \setminus \{0\}$ is the largest open ring with center 0 which is contained in $\mathbb{C} \setminus \{0\}$. To find the Laurent series of f in $D_0(0, +\infty)$ we evaluate the coefficients a_n. We take any r with $0 < r < +\infty$, and then we have

$$a_n = \frac{1}{2\pi i} \oint_{C_0(r)} \frac{1/\zeta}{\zeta^{n+1}}\, d\zeta = \frac{1}{2\pi i} \oint_{C_0(r)} \frac{1}{\zeta^{n+2}}\, d\zeta$$

for every n. If $n \neq -1$, then $a_n = 0$ and, if $n = -1$, then $a_{-1} = 1$. Therefore, the Laurent series of f in $D_0(0, +\infty)$ is $\sum_{-\infty}^{+\infty} a_n z^n = z^{-1}$ and hence we have the obvious identity $\frac{1}{z} = z^{-1}$ for every $z \in D_0(0, +\infty)$.

In the following examples we shall use the uniqueness of the Laurent series to find the Laurent series of certain functions without evaluating integrals: we find in an indirect way a power series which coincides with the function in a specific ring and then, because of uniqueness, this is the Laurent series of the function in the ring.

Example 5.6.7. The function $f(z) = \frac{1}{1-z}$ is holomorphic in the open set $\mathbb{C} \setminus \{1\}$. We have seen that the largest open disc with center 0 which is contained in $\mathbb{C} \setminus \{1\}$ is $D_0(1)$ and that the Taylor series of f in this disc is $\sum_{n=0}^{+\infty} z^n$.
Another largest open ring with center 0 which is contained in $\mathbb{C} \setminus \{1\}$ is $D_0(1, +\infty)$. To find the Laurent series of f in this ring, we may evaluate the coefficients a_n using their formulas with the integrals. But we can do something simpler. If $z \in D_0(1, +\infty)$, then $\left|\frac{1}{z}\right| < 1$ and hence

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

Because of uniqueness, the Laurent series of f in $D_0(1, +\infty)$ is $-\sum_{-\infty}^{n=-1} z^n$.

Example 5.6.8. The function $f(z)=\frac{1}{(z-1)(z-2)}$ is holomorphic in $\mathbb{C}\setminus\{1,2\}$. There is a largest open disc and two largest open rings with center 0 which are contained in $\mathbb{C}\setminus\{1,2\}$: the disc $D_0(1)$ and the rings $D_0(1,2)$ and $D_0(2,+\infty)$. To find the corresponding Taylor and Laurent series we write f as a sum of simple fractions: $f(z)=\frac{1}{z-2}-\frac{1}{z-1}$.
If $z\in D_0(1)$, then $|z|<1$ and $\left|\frac{z}{2}\right|<1$, and hence

$$f(z)=-\frac{1}{2}\frac{1}{1-\frac{z}{2}}+\frac{1}{1-z}=-\frac{1}{2}\sum_{n=0}^{+\infty}(\tfrac{z}{2})^n+\sum_{n=0}^{+\infty}z^n=\sum_{n=0}^{+\infty}(1-\tfrac{1}{2^{n+1}})z^n.$$

Therefore, the Taylor series of f in $D_0(1)$ is $\sum_{n=0}^{+\infty}(1-\frac{1}{2^{n+1}})z^n$.
If $z\in D_0(1,2)$, then $\left|\frac{1}{z}\right|<1$ and $\left|\frac{z}{2}\right|<1$, and hence

$$f(z)=-\frac{1}{2}\frac{1}{1-\frac{z}{2}}-\frac{1}{z}\frac{1}{1-\frac{1}{z}}=-\frac{1}{2}\sum_{n=0}^{+\infty}(\tfrac{z}{2})^n-\frac{1}{z}\sum_{n=0}^{+\infty}(\tfrac{1}{z})^n=-\sum_{-\infty}^{n=-1}z^n-\sum_{n=0}^{+\infty}\tfrac{1}{2^{n+1}}\,z^n.$$

Therefore, the Laurent series of f in $D_0(1,2)$ is $-\sum_{-\infty}^{n=-1}z^n-\sum_{n=0}^{+\infty}\frac{1}{2^{n+1}}\,z^n$.
If $z\in D_0(2,+\infty)$, then $\left|\frac{1}{z}\right|<1$ and $\left|\frac{2}{z}\right|<1$, and hence

$$f(z)=\frac{1}{z}\frac{1}{1-\frac{2}{z}}-\frac{1}{z}\frac{1}{1-\frac{1}{z}}=\frac{1}{z}\sum_{n=0}^{+\infty}(\tfrac{2}{z})^n-\frac{1}{z}\sum_{n=1}^{+\infty}(\tfrac{1}{z})^n=\sum_{-\infty}^{n=-2}(\tfrac{1}{2^{n+1}}-1)z^n.$$

Therefore, the Laurent series of f in $D_0(2,+\infty)$ is $\sum_{-\infty}^{n=-2}(\frac{1}{2^{n+1}}-1)z^n$.

Example 5.6.9. The function $f(z)=e^{\frac{1}{z}}$ is holomorphic in $\mathbb{C}\setminus\{0\}$. Then $D_0(0,+\infty)=\mathbb{C}\setminus\{0\}$ is the only largest open ring with center 0 which is contained in $\mathbb{C}\setminus\{0\}$. We find the Laurent series of f in $D_0(0,+\infty)$ using the Taylor series of e^z in $\mathbb{C}$. In the identity $e^z=\sum_{n=0}^{+\infty}\frac{1}{n!}z^n$ we replace z with $\frac{1}{z}$ and we find

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

for every $z\neq 0$. Therefore, the Laurent series of f in $D_0(0,+\infty)$ is $\sum_{-\infty}^{n=-1}\frac{1}{(-n)!}z^n+1$.

Exercises.

5.6.1. Let $0<|a|<|b|$. Find the three Laurent series with center 0, the two Laurent series with center a and the two Laurent series with center b of the function $\frac{z}{(z-a)(z-b)}$.

5.6.2. Find the Taylor series of $\frac{1}{1+z^2}$ with center any $a\in\mathbb{R}$.

5.6.3. Find the Taylor series with center 1 of the holomorphic branch of $z^{\frac{1}{2}}$ with value 1 at 1.

5.6.4. Let f be holomorphic in $D_{z_0}(R)$ and let $\sum_{n=0}^{+\infty}a_n(z-z_0)^n$ be the Taylor series of f.
(i) Prove that, if $0\leq r<R$, then

$$\frac{1}{2\pi}\int_0^{2\pi}|f(z_0+re^{it})|^2\,dt=\sum_{n=0}^{+\infty}|a_n|^2r^{2n}.$$

(ii) If $|f(z)|\leq M$ for every $z\in D_{z_0}(R)$, prove that $\sum_{n=0}^{+\infty}|a_n|^2R^{2n}\leq M^2$.
(iii) If g is also holomorphic in $D_{z_0}(R)$ with Taylor series $\sum_{n=0}^{+\infty}b_n(z-z_0)^n$, prove that, if $0\leq r<R$, then

$$\frac{1}{2\pi}\int_0^{2\pi}f(z_0+re^{it})\,\overline{g(z_0+re^{it})}\,dt=\sum_{n=0}^{+\infty}a_n\overline{b_n}r^{2n}.$$

5.6.5. Let f be holomorphic in $D_{z_0}(R_1,R_2)$. Prove that there are functions f_1,f_2 so that f_2 is holomorphic in $D_{z_0}(R_2)$ and f_1 is holomorphic in $D_{z_0}(R_1,+\infty)\cup\{\infty\}$ and so that $f=f_1+f_2$ in $D_{z_0}(R_1,R_2)$. Prove that, if f is bounded in $D_{z_0}(R_1,R_2)$, then f_1,f_2 are bounded in $D_{z_0}(R_1,R_2)$.

5.6.6. Let f be holomorphic in $D_0(R,+\infty)$. Prove that f is holomorphic also at ∞ if and only if the Laurent series of f in $D_0(R,+\infty)$ is of the form $\sum_{-\infty}^{n=-1}a_nz^n+a_0$. Observe that $f(\infty)=a_0$.

5.6.7. Prove that

$$\tfrac{1}{\cos z} = 1 + \sum_{k=1}^{+\infty} \tfrac{E_{2k}}{(2k)!} z^{2k}$$

for $|z| < \frac{\pi}{2}$, where the numbers E_{2k} satisfy the recursive relations

$$E_{2n} - \tbinom{2n}{2n-2} E_{2n-2} + \tbinom{2n}{2n-4} E_{2n-4} - \cdots + (-1)^{n-1} \tbinom{2n}{2} E_2 + (-1)^n = 0.$$

Evaluate E_2, E_4, E_6, E_8. The numbers E_{2k} are called **Euler constants**.

5.6.8. Let f be holomorphic in the horizontal zone $\Omega = \{x + iy \,|\, A < y < B\}$ and periodic with period 1, i.e. $f(z+1) = f(z)$ for every $z \in \Omega$.
(i) Prove that there are c_n so that

$$f(z) = \sum_{-\infty}^{+\infty} c_n e^{2\pi i n z}$$

for every $z \in \Omega$ and find formulas for the coefficients c_n.
(ii) Prove that the series in (i) converges uniformly in every smaller zone $\{x + iy \,|\, a < y < b\}$ with $A < a < b < B$.

5.6.9. (i) Prove that

$$e^{\frac{w}{2}(z - \frac{1}{z})} = b_0(w) + \sum_{n=1}^{+\infty} b_n(w) \big(z^n + \tfrac{(-1)^n}{z^n}\big)$$

for every $z \neq 0$, where

$$b_n(w) = \tfrac{1}{\pi} \int_0^{\pi} \cos(nt - w \sin t)\, dt$$

for $n \in \mathbb{N}_0$.
(ii) If $m, n \in \mathbb{N}_0$, prove that

$$\tfrac{1}{2\pi i} \int_{C_0(1)} \tfrac{(z^2 \pm 1)^m}{z^{m+n+1}}\, dz = \begin{cases} \frac{(\pm 1)^p (n+2p)!}{p!(n+p)!} & \text{if } m = n + 2p,\, p \in \mathbb{N}_0 \\ 0, & \text{otherwise} \end{cases}$$

(iii) The function $b_n(w)$ is called Bessel function of the first kind. Find the Taylor series of $b_n(w)$ with center 0.

5.6.10. $f : I \to \mathbb{C}$ is called **real analytic** in the open interval I in $\mathbb{R}$ if for every $t_0 \in I$ there are $\epsilon > 0$ and $a_n \in \mathbb{C}$, $n \in \mathbb{N}_0$, so that $(t_0 - \epsilon, t_0 + \epsilon) \subseteq I$ and $f(t) = \sum_{n=0}^{+\infty} a_n (t - t_0)^n$ for every $t \in (t_0 - \epsilon, t_0 + \epsilon)$.
Prove that, if f is real analytic in I, then there is an open set $\Omega \subseteq \mathbb{C}$ so that $I \subseteq \Omega$ and so that f can be extended as a function $f : \Omega \to \mathbb{C}$ holomorphic in Ω.

5.7 Roots and the principle of identity.

Let f be holomorphic in the open set Ω and $z_0 \in \Omega$. We consider the largest open disc $D_{z_0}(R)$ which is contained in Ω and the Taylor series of f in this disc. Then

$$f(z) = \sum_{n=0}^{+\infty} a_n (z - z_0)^n = a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \cdots$$

for every $z \in D_{z_0}(R)$.

We assume that z_0 is a root of f or, equivalently, that $a_0 = 0$ and we distinguish between two cases.

First case: $a_n = 0$ *for every* n.

Then, obviously, $f(z) = 0$ for every $z \in D_{z_0}(R)$, i.e. f is identically 0 in $D_{z_0}(R)$. Because of the formulas for a_n, the condition $a_n = 0$ for every n is equivalent to $f^{(n)}(z_0) = 0$ for every n.

Second case: $a_n \neq 0$ *for at least one* n.

We consider the smallest $n \geq 1$ with $a_n \neq 0$ and let this be N. I.e. $a_0 = a_1 = \ldots = a_{N-1} = 0$ and $a_N \neq 0$. This is equivalent to

$$f(z_0) = f^{(1)}(z_0) = \ldots = f^{(N-1)}(z_0) = 0, \quad f^{(N)}(z_0) \neq 0.$$

Then we have

$$f(z) = (z - z_0)^N \sum_{n=0}^{+\infty} a_{N+n}(z - z_0)^n$$

for every $z \in D_{z_0}(R)$. The power series $\sum_{n=0}^{+\infty} a_{N+n}(z - z_0)^n$ converges in the disc $D_{z_0}(R)$ and defines a function g holomorphic in $D_{z_0}(R)$. Then

$$f(z) = (z - z_0)^N g(z)$$

for every $z \in D_{z_0}(R)$, and thus $g(z) = \frac{f(z)}{(z-z_0)^N}$ for every $z \in D_{z_0}(R) \setminus \{z_0\}$. We observe that $\frac{f(z)}{(z-z_0)^N}$ is a holomorphic function in $\Omega \setminus \{z_0\}$ and not only in $D_{z_0}(R) \setminus \{z_0\}$. Therefore, we may consider g as defined in $\Omega \setminus \{z_0\}$ with the same formula: $g(z) = \frac{f(z)}{(z-z_0)^N}$. We also recall that g is defined, through its power series, at z_0 and it is holomorphic in $D_{z_0}(R) \subseteq \Omega$. In fact its value at z_0 is $g(z_0) = a_N = \frac{f^{(N)}(z_0)}{N!}$. Thus, the formula of g, as a function holomorphic in Ω, can be written:

$$g(z) = \begin{cases} \frac{f(z)}{(z-z_0)^N}, & \text{if } z \in \Omega \setminus \{z_0\} \\ a_N = \frac{f^{(N)}(z_0)}{N!}, & \text{if } z = z_0 \end{cases}$$

Since $g(z_0) = a_N \neq 0$ and since g is continuous at z_0, there is r with $0 < r \leq R$ so that $g(z) \neq 0$ for every $z \in D_{z_0}(r)$, and hence $f(z) \neq 0$ for every $z \in D_{z_0}(r) \setminus \{z_0\}$.

Let f be holomorphic in the open set Ω, $z_0 \in \Omega$ and $\sum_{n=0}^{+\infty} a_n(z - z_0)^n$ be the Taylor series of f at z_0. Then we have three cases. If $a_n = 0$ for every n, then we say that z_0 is a root of f of **infinite multiplicity**. If $a_0 = a_1 = \ldots = a_{N-1} = 0$ and $a_N \neq 0$ for some $N \geq 1$, then we say that z_0 is a root of f of **multiplicity N**. Finally, if $f(z_0) = a_0 \neq 0$, we say that z_0 is a root of f of **multiplicity 0**.

We saw that, if z_0 is a root of f of infinite multiplicity, then f is identically 0 in the largest disc with center z_0 which is contained in the domain of holomorphy of f. If z_0 is a root of f of finite multiplicity, then there is some disc $D_{z_0}(r)$ which contains no other root of f besides z_0 and hence we say that the root z_0 is **isolated**. Moreover, if the multiplicity of z_0 is N, then the function $g(z) = \frac{f(z)}{(z-z_0)^N}$, which is holomorphic in $\Omega \setminus \{z_0\}$, can be defined at z_0 as $g(z_0) = a_N = \frac{f^{(N)}(z_0)}{N!}$ and then it is holomorphic in Ω. In other words, *we can factorize $(z - z_0)^N$ from $f(z)$, i.e. we can write $f(z) = (z - z_0)^N g(z)$ with a function g holomorphic in Ω*. This is a striking generalization of the analogous factorization for polynomials: is z_0 is a root of the polynomial $p(z)$ of multiplicity N, then we can write $p(z) = (z - z_0)^N q(z)$, where $q(z)$ is another polynomial.

Example 5.7.1. The function $e^{z^3} - 1$ is holomorphic in $\mathbb{C}$ and its Taylor series with center 0 is $\sum_{n=1}^{+\infty} \frac{1}{n!} z^{3n}$. Therefore,

$$e^{z^3} - 1 = z^3 \sum_{n=1}^{+\infty} \frac{1}{n!} z^{3(n-1)} = z^3 \sum_{n=0}^{+\infty} \frac{1}{(n+1)!} z^{3n} = z^3 g(z)$$

for every z, where g is the function defined by the power series $\sum_{n=0}^{+\infty} \frac{1}{(n+1)!} z^{3n}$. Now g is holomorphic in $\mathbb{C}$ with $g(0) = 1 \neq 0$, hence 0 is a root of $e^{z^3} - 1$ of multiplicity 3.

Lemma 5.1. *If f is holomorphic in the region Ω and if $z_0 \in \Omega$ is a root of f of infinite multiplicity, then f is identically 0 in Ω.*

Proof. f is identically 0 in some disc with center z_0. We define

$$B = \{z \in \Omega \mid f \text{ is identically } 0 \text{ in some disc with center } z\}$$

and the complementary set $C = \Omega \setminus B$. Obviously, $B \cup C = \Omega$ and $B \neq \emptyset$, since $z_0 \in B$.
If $z \in B$, then f is identically 0 in some disc $D_z(r)$, and if we take any $w \in D_z(r)$, then f is identically 0 in some small disc $D_w(r') \subseteq D_z(r)$. Thus every $w \in D_z(r)$ belongs to B, i.e. $D_z(r) \subseteq B$ and z is not a limit point of C.
Now, let $z \in C$. Then f is identically 0 in no disc with center z, and hence z is not a root of infinite multiplicity of f. Therefore, there is a disc $D_z(r)$ in which the only possible root of f is its center z. Then this disc contains no $w \in B$ and z is not a limit point of B.
Thus, none of B, C contains a limit point of the other. Since $B \neq \emptyset$, we must have $C = \emptyset$, otherwise B, C would form a decomposition of Ω. Hence $\Omega = B$ and f is identically 0 in Ω. □

Principle of identity. *If f is holomorphic in the region Ω and if the roots of f have an accumulation point in Ω, then f is identically 0 in Ω.*

Proof. Suppose that there is a sequence (z_n) of roots of f so that $z_n \to z$ with $z \in \Omega$ and $z_n \neq z$ for every n. Since f is continuous at z and $z_n \to z$, we have $0 = f(z_n) \to f(z)$ and hence $f(z) = 0$. If z is a root of finite multiplicity of f, then there would be some disc $D_z(r)$ in which the only root of f is its center z. This is wrong, since $D_z(r)$ contains, after some index, all roots z_n and these are different from z. Therefore, z is a root of infinite multiplicity of f, and lemma 5.1 implies that f is identically 0 in Ω. □

Lemma 5.1 and the principle of identity can be stated for a non-connected open set Ω. Then the result of lemma 5.1 holds in the connected component of Ω which contains the root of infinite multiplicity z_0 and the result of the principle of identity holds in the connected component of Ω which contains the accumulation point of the roots of f.

Instead of speaking only about the roots of f, i.e. the solutions of the equation $f(z) = 0$, we may state our results for the solutions of the equation $f(z) = w$ for any fixed w. The results are the same as before. We just consider the function $g(z) = f(z) - w$, and then the solutions of $f(z) = w$ are the same as the roots of g. For instance, if z_0 is a solution of $f(z) = w$ of infinite multiplicity, then f is constant w in some disc $D_{z_0}(R)$ and, if z_0 is a solution of $f(z) = w$ of finite multiplicity N, then in some disc $D_{z_0}(r)$ the function f takes the value w only at the center z_0. Then lemma 5.1 says that, if f is holomorphic in the region Ω and z_0 is a solution of $f(z) = w$ of infinite multiplicity, then f is constant w in Ω. And the principle of identity says that, if f is holomorphic in the region Ω and the solutions of $f(z) = w$ have an accumulation point in Ω, then f is constant w in Ω.

The principle of identity has another equivalent form.

Principle of identity. *If f is holomorphic in the region Ω and if some compact $K \subseteq \Omega$ contains infinitely many roots of f, then f is identically 0 in Ω.*

Proof. Let us assume the previous principle of identity and let us suppose that some compact $K \subseteq \Omega$ contains infinitely many roots of f. Then there is a sequence (z_n) of roots of f in K with distinct terms. Since K is compact, there is a subsequence (z_{n_k}) so that $z_{n_k} \to z$ for some $z \in K$. But then $z \in \Omega$ is an accumulation point of roots of f and hence f is identically 0 in Ω.
Conversely, let us assume the present form of the principle of identity and let us suppose that the roots of f have an accumulation point in Ω. Then there is a sequence (z_n) of roots of f so that $z_n \to z$ with $z \in \Omega$ and $z_n \neq z$ for every n. Then the set $\{z_n \,|\, n \in \mathbb{N}\} \cup \{z\}$ is a compact $\subseteq \Omega$ and contains infinitely many roots of f. So f is identically 0 in Ω. □

Example 5.7.2. Assume that there is f holomorphic in $\mathbb{C}$ so that $f(\frac{1}{n}) = \frac{n}{n+1}$ for every $n \in \mathbb{N}$.
We write $f(\frac{1}{n}) = \frac{1}{1+\frac{1}{n}}$ and compare the functions $f(z)$ and $\frac{1}{1+z}$. Both are holomorphic in $\mathbb{C} \setminus \{-1\}$ and their difference $f(z) - \frac{1}{1+z}$ has roots at the points $\frac{1}{n}$ which have 0 as their accumulation point. Since $0 \in \mathbb{C} \setminus \{-1\}$ and $\mathbb{C} \setminus \{-1\}$ is connected, we have that $f(z) - \frac{1}{1+z}$ is identically 0 in this set, i.e. $f(z) = \frac{1}{1+z}$ for every $z \neq -1$. Since we assume that f is holomorphic at -1, we get $\lim_{z \to -1} \frac{1}{1+z} = \lim_{z \to -1} f(z) = f(-1)$ and we arrive at a contradiction.

Example 5.7.3. Assume that there is some f holomorphic in $\mathbb{C}\setminus\{0\}$ so that $f(x)=\sqrt{x}$ for every $x\in(0,+\infty)$ or even for every x in some subinterval (a,b) of $(0,+\infty)$.
We consider the continuous branch g of $z^{\frac{1}{2}}$ in the region $\Omega=\mathbb{C}\setminus(-\infty,0]$ which has value 1 at $z=1$. The function g is given by

$$g(z)=\sqrt{r}\,e^{i\frac{\theta}{2}},$$

where $z=re^{i\theta}$ is the polar representation of $z\in\Omega$ with $-\pi<\theta<\pi$. So $f(x)=\sqrt{x}=g(x)$ for every $x\in(a,b)$. Hence $f-g$ is holomorphic in the region Ω and has roots at all points of (a,b). We conclude that $f-g$ is identically 0 in Ω. I.e.

$$f(z)=\sqrt{r}\,e^{i\frac{\theta}{2}},$$

where $z=re^{i\theta}$ is the polar representation of $z\in\Omega$ with $-\pi<\theta<\pi$. Since f is holomorphic in $\mathbb{C}\setminus\{0\}$, it is continuous at every point of $(-\infty,0)$, e.g. at -1.
We take points $z=re^{i\theta}$ converging to -1 from the upper halfplane. This means that $r\to 1$ and $\theta\to\pi-$. Then we have

$$f(-1)=\lim_{r\to 1,\theta\to\pi-}\sqrt{r}\,e^{i\frac{\theta}{2}}=e^{i\frac{\pi}{2}}=i.$$

Now we take points $z=re^{i\theta}$ converging to -1 from the lower halfplane. This means that $r\to 1$ and $\theta\to-\pi+$. Then we have

$$f(-1)=\lim_{r\to 1,\theta\to-\pi+}\sqrt{r}\,e^{i\frac{\theta}{2}}=e^{-i\frac{\pi}{2}}=-i.$$

We arrive at a contradiction.

Exercises.

5.7.1. Let f be holomorphic in the disc $D_{z_0}(R)$ and let z_0 be a root of multiplicity $N\geq 1$ of f. If F is a primitive of f in $D_{z_0}(R)$ and $F(z_0)=w_0$, which is the multiplicity of z_0 as a solution of $F(z)=w_0$?

5.7.2. Is there any f holomorphic in $\mathbb{C}$ which satisfies one of the following?
(i) $f(\frac{1}{n})=(-1)^n$ for every $n\in\mathbb{N}$.
(ii) $f(\frac{1}{n})=\frac{1+(-1)^n}{n}$ for every $n\in\mathbb{N}$.
(iii) $f(\frac{1}{2k})=f(\frac{1}{2k+1})=\frac{1}{k}$ for every $k\in\mathbb{N}$.

5.7.3. Is there any f holomorphic in $\mathbb{C}\setminus\{0\}$ so that $f(x)=|x|$ for every $x\in\mathbb{R}\setminus\{0\}$?

5.7.4. Let f,g be holomorphic in the region Ω and $0\in\Omega$. If f, g have no root in Ω and if $f'(\frac{1}{n})/f(\frac{1}{n})=g'(\frac{1}{n})/g(\frac{1}{n})$ for every $n\in\mathbb{N}$, what do you conclude about f,g?

5.7.5. Let f,g be holomorphic in the region Ω. If $fg=0$ in Ω, prove that either $f=0$ in Ω or $g=0$ in Ω.

5.7.6. Let f,g be holomorphic in the region Ω. If $\overline{f}\,g$ is holomorphic in Ω, prove that either $g=0$ in Ω or f is constant in Ω.

5.7.7. (i) Let the region Ω be symmetric with respect to $\mathbb{R}$, i.e. $\overline{z}\in\Omega$ for every $z\in\Omega$. If $\Omega\neq\emptyset$, prove that $\Omega\cap\mathbb{R}\neq\emptyset$. Let also f be holomorphic in Ω and assume that $f(z)\in\mathbb{R}$ for every $z\in\Omega\cap\mathbb{R}$. Prove that $f(\overline{z})=\overline{f(z)}$ for every $z\in\Omega$.
(ii) Let the region $\Omega\subseteq\mathbb{C}\setminus\{0\}$ be symmetric with respect to $\mathbb{T}$, i.e. $\frac{1}{\overline{z}}\in\Omega$ for every $z\in\Omega$. If $\Omega\neq\emptyset$, prove that $\Omega\cap\mathbb{T}\neq\emptyset$. Let also f be holomorphic in Ω and assume that $f(z)\in\mathbb{T}$ for every $z\in\Omega\cap\mathbb{T}$. Prove that $f(\frac{1}{\overline{z}})=\frac{1}{\overline{f(z)}}$ for every $z\in\Omega$.
(iii) Let f be holomorphic in $\mathbb{C}$ and let $f(z)\in\mathbb{T}$ for every $z\in\mathbb{T}$. Prove that there is $c\in\mathbb{T}$ and $n\in\mathbb{N}_0$ so that $f(z)=cz^n$ for every z.

5.7.8. Many of the results of this section hold also for the point ∞.
(i) Let $\Omega \subseteq \mathbb{C}$ be an open set containing some ring $D_0(R, +\infty)$ and let f be holomorphic in $\Omega \cup \{\infty\}$. Then, according to exercice 5.6.6, the Laurent series of f in $D_0(R, +\infty)$ is of the form $\sum_{-\infty}^{n=-1} a_n z^n + a_0$ and also $f(\infty) = a_0$.
If $a_n = 0$ for every $n \leq 0$, we say that ∞ is a root of f of multiplicity $+\infty$, and in this case prove that f is identically 0 in the connected component of Ω which contains $D_0(R, +\infty)$.
If $a_0 = a_{-1} = \ldots = a_{-N+1} = 0$ and $a_{-N} \neq 0$, we say that ∞ is a root of f of multiplicity N, and in this case prove that ∞ is an isolated root of f, i.e. there is some $r \geq R$ so that f has no root in $D_0(r, +\infty)$.
Of course, if $a_0 \neq 0$, we say that ∞ is a root of f of multiplicity 0.
If ∞ is an accumulation point of roots of f, prove that f is identically 0 in the connected component of Ω which contains $D_0(R, +\infty)$.
Prove that ∞ is a root of f of multiplicity N if and only if 0 is a root of g of multiplicity N, where g is defined by $g(w) = f(\frac{1}{w})$.
(ii) Let $r = \frac{p}{q}$ be a rational function and let n be the degree of the polynomial p and m be the degree of the polynomial q. If $n \leq m$, prove that ∞ is a root of r of multiplicity $m - n$.

5.8 Isolated singularities.

We say that z_0 is an **isolated singularity** of f if there is some disc $D_{z_0}(R)$ so that f is holomorphic in $D_{z_0}(R) \setminus \{z_0\}$. Then f has a Laurent series in $D_{z_0}(0, R) = D_{z_0}(R) \setminus \{z_0\}$. I.e.

$$f(z) = \sum_{-\infty}^{+\infty} a_n (z - z_0)^n$$

for every $z \in D_{z_0}(R) \setminus \{z_0\}$.

Now we have three cases. If $a_n = 0$ for every $n < 0$, then we say that z_0 is a **removable singularity** of f. If $a_n \neq 0$ for at least one $n < 0$ and there are only finitely many $n < 0$ such that $a_n \neq 0$, then we say that z_0 is a **pole** of f. Finally, if $a_n \neq 0$ for infinitely many $n < 0$, then we say that z_0 is an **essential singularity** of f.

Let us start with the case of a removable singularity z_0. Then

$$f(z) = \sum_{n=0}^{+\infty} a_n (z - z_0)^n$$

for every $z \in D_{z_0}(R) \setminus \{z_0\}$. The power series $\sum_{n=0}^{+\infty} a_n (z - z_0)^n$ converges at every $z \in D_{z_0}(R)$ and defines a holomorphic function in $D_{z_0}(R)$ with value a_0 at z_0. The function f may not be defined at z_0 or it may be defined at z_0 with a value $f(z_0)$ either equal to a_0 or not equal to a_0. Now, in any case, we define (or redefine) f at z_0 to be $f(z_0) = a_0$. Then we have $f(z) = \sum_{n=0}^{+\infty} a_n (z - z_0)^n$ for every $z \in D_{z_0}(R)$ and f becomes holomorphic in $D_{z_0}(R)$.

We summarize. *If $z_0 \in \Omega$ is a removable singularity of f, then f can be defined (or redefined) appropriately at z_0 so that it becomes holomorphic in a disc with center z_0. The Laurent series of f at z_0 reduces to a power series of first type and this power series is the Taylor series of the (extended) f in a disc with center z_0.*

Here is a useful test to decide if an isolated singularity is removable without calculating the Laurent series of the function.

Riemann's criterion. *Let z_0 be an isolated singularity of f. If*

$$\lim_{z \to z_0} (z - z_0) f(z) = 0,$$

then z_0 is a removable singularity of f.

Proof. Let $f(z) = \sum_{-\infty}^{+\infty} a_n (z - z_0)^n$ for every $z \in D_{z_0}(R) \setminus \{z_0\}$. We take any $\epsilon > 0$ and then there is $\delta > 0$ so that $|z - z_0||f(z)| \leq \epsilon$ for every $z \in D_{z_0}(R)$ with $0 < |z - z_0| < \delta$. Now, we consider any r with $0 < r < \min\{\delta, R, 1\}$ and any $n < 0$. Then we have

$$|a_n| = \left|\frac{1}{2\pi i} \oint_{C_{z_0}(r)} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} \, d\zeta\right| \leq \frac{1}{2\pi} \frac{\epsilon}{r^{n+2}} \, 2\pi r = \epsilon r^{-n-1} = \epsilon r^{|n|-1} \leq \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we get $a_n = 0$ for every $n < 0$ and z_0 is a removable singularity of f. $\square$

In the case of an isolated singularity z_0 for f, sometimes we know that the $\lim_{z\to z_0} f(z)$ exists and it is finite or that f is bounded close to z_0. In both cases we have that $\lim_{z\to z_0}(z-z_0)f(z) = 0$ is satisfied and we conclude that z_0 is a removable singularity of f.

Example 5.8.1. The function $f(z) = \frac{z^2-3z+2}{z-2}$ is holomorphic in $\mathbb{C}\setminus\{2\}$. Since $\lim_{z\to 2} f(z) = 1$, the point 2 is a removable singularity of f. If we define $f(2) = 1$, then f, now defined in $\mathbb{C}$, is holomorphic in $\mathbb{C}$. In fact, the extended f is the simple function $z - 1$ in $\mathbb{C}$.

Now we consider the case of a pole z_0 of f. Let $\sum_{-\infty}^{+\infty} a_n(z-z_0)^n$ be the Laurent series of f in the ring $D_{z_0}(R)\setminus\{z_0\}$ and then there is a largest $m \ge 1$ so that $a_{-m} \neq 0$. Let N be this largest m. Then we say that z_0 is a pole of f of **order N** or of **multiplicity N** and we have

$$f(z) = \frac{a_{-N}}{(z-z_0)^N} + \cdots + \frac{a_{-1}}{z-z_0} + \sum_{n=0}^{+\infty} a_n(z-z_0)^n$$

for every $z \in D_{z_0}(R)\setminus\{z_0\}$ with $a_{-N} \neq 0$. We may write this as

$$f(z) = \frac{1}{(z-z_0)^N}\sum_{n=0}^{+\infty} a_{n-N}(z-z_0)^n$$

for every $z \in D_{z_0}(R)\setminus\{z_0\}$. Since the power series $\sum_{n=0}^{+\infty} a_{n-N}(z-z_0)^n$ converges in the disc $D_{z_0}(R)$, it defines a function g holomorphic in $D_{z_0}(R)$ and we have

$$f(z) = \frac{g(z)}{(z-z_0)^N}$$

for every $z \in D_{z_0}(R)\setminus\{z_0\}$. Observe that $g(z_0) = a_{-N} \neq 0$.

It is easy to prove the converse. Suppose there is a g holomorphic in some disc $D_{z_0}(R)$ so that $g(z_0) \neq 0$ and $f(z) = \frac{g(z)}{(z-z_0)^N}$ for every $z \in D_{z_0}(R)\setminus\{z_0\}$. Let $\sum_{n=0}^{+\infty} b_n(z-z_0)^n$ be the Taylor series of g and then we have

$$f(z) = \frac{b_0}{(z-z_0)^N} + \cdots + \frac{b_{N-1}}{z-z_0} + \sum_{n=0}^{+\infty} b_{n+N}(z-z_0)^n$$

for $z \in D_{z_0}(R)\setminus\{z_0\}$. The last power series is the Laurent series of f in $D_{z_0}(R)\setminus\{z_0\}$ and since $b_0 = g(z_0) \neq 0$, we have that z_0 is a pole of f of order N.

Here are some more comments. Since $g(z_0) \neq 0$ and g is continuous at z_0, we have that g does not vanish at any point of some disc $D_{z_0}(r)$ with $0 < r \le R$. Then $h(z) = \frac{1}{g(z)}$ is holomorphic in $D_{z_0}(r)$ and $\frac{1}{f(z)} = (z-z_0)^N h(z)$ for every $z \in D_{z_0}(r)\setminus\{z_0\}$. Therefore, z_0 is a removable singularity of $\frac{1}{f}$. Moreover, if we define $\frac{1}{f}$ to take the value 0 at z_0, then we have $\frac{1}{f}(z) = (z-z_0)^N h(z)$ for every $z \in D_{z_0}(r)$ and, since $h(z_0) \neq 0$, then z_0 is a root of the extended $\frac{1}{f}$ of multiplicity N. It is easy to prove in a similar way the converse, and we conclude that z_0 *is a pole of f of order N if and only if it is a root of $\frac{1}{f}$ of mutiplicity N.*

Example 5.8.2. Many times we meet functions of the form $f = \frac{p}{q}$, where p, q are holomorphic in a neighborhood of z_0. For instance, if p, q are polynomials, then f is a rational function.
Let z_0 be a root of p and q of multiplicity $M \ge 0$ and $N \ge 0$, respectively. In this case we saw that there are holomorphic functions p_1 and q_1 in a neighborhood $D_{z_0}(R)$ of z_0 so that

$$p(z) = (z-z_0)^M p_1(z), \quad q(z) = (z-z_0)^N q_1(z)$$

for every $z \in D_{z_0}(R)$ and also $p_1(z_0) \neq 0$ and $q_1(z_0) \neq 0$. (Of course we consider the case that none of p, q is identically 0.) Then there is r with $0 < r \le R$ so that $p_1(z) \neq 0$ and $q_1(z) \neq 0$ for every $z \in D_{z_0}(r)$, and then we have

$$f(z) = \frac{p(z)}{q(z)} = (z-z_0)^{M-N}\frac{p_1(z)}{q_1(z)} = (z-z_0)^{M-N} g(z)$$

for every $z \in D_{z_0}(r) \setminus \{z_0\}$, where the function $g = \frac{p_1}{q_1}$ is holomorphic in $D_{z_0}(r)$ and $g(z_0) = \frac{p_1(z_0)}{q_1(z_0)} \neq 0$. Now we have two cases. If $M \geq N$, then z_0 is a removable singularity of f, and f (after we extend it appropriately at z_0) is holomorphic at z_0 and z_0 is a root of f of multiplicity $M - N$. If $M < N$, then z_0 is a pole of order $N - M$ of f.

Here are some concrete instances of this example.

Example 5.8.3. The function $f(z) = \frac{z^2-3z+2}{(z-2)^2}$ is holomorphic in $\mathbb{C} \setminus \{2\}$.
Since $z^2 - 3z + 2 = (z - 2)(z - 1)$, we have $f(z) = \frac{z-1}{z-2}$ for $z \neq 2$. The function $g(z) = z - 1$ is holomorphic in $\mathbb{C}$ and $g(2) = 1 \neq 0$. Therefore, 2 is a pole of f of order 1.

Example 5.8.4. The function $f(z) = \frac{e^z-1}{z^3}$ is holomorphic in $\mathbb{C} \setminus \{0\}$.
The Taylor series of $e^z - 1$ with center 0 is $z + \frac{1}{2!}\, z^2 + \frac{1}{3!}\, z^3 + \cdots$. Hence $e^z - 1 = zg(z)$ with $g(z) = 1 + \frac{1}{2!}\, z + \frac{1}{3!}\, z^2 + \cdots$. The function g is holomorphic in $\mathbb{C}$ and $g(0) = 1 \neq 0$ and we have $f(z) = \frac{g(z)}{z^2}$ for $z \neq 0$. Therefore, 0 is a pole of f of order 2.

Example 5.8.5. The function $\cot z = \frac{\cos z}{\sin z}$ is holomorphic in $\mathbb{C} \setminus \{k\pi \mid k \in \mathbb{Z}\}$.
The points $k\pi$, $k \in \mathbb{Z}$, are isolated singularities of $\cot z$ and we shall prove that they are all poles of order 1. We fix $k \in \mathbb{Z}$. The Taylor series of $\sin z$ with center $k\pi$ results from the Taylor series of $\sin z$ with center 0, as follows

$$\begin{aligned}
\sin z &= \sin((z - k\pi) + k\pi) = \cos k\pi \sin(z - k\pi) = (-1)^k \sin(z - k\pi) \\
&= (-1)^k\big((z - k\pi) - \tfrac{1}{3!}(z - k\pi)^3 + \cdots\big) = (-1)^k(z - k\pi) - \tfrac{(-1)^k}{3!}(z - k\pi)^3 + \cdots .
\end{aligned}$$

Therefore, $\sin z = (z - k\pi)q_1(z)$ for every z, where the function q_1 is holomorphic in $\mathbb{C}$ with $q_1(k\pi) = (-1)^k$. Hence,

$$\cot z = \tfrac{\cos z}{(z-k\pi)q_1(z)} = \tfrac{g(z)}{z-k\pi}$$

with $g(z) = \frac{\cos z}{q_1(z)}$ and g is holomorphic in the disc $D_{k\pi}(\pi)$ and $g(k\pi) = \frac{\cos k\pi}{q_1(k\pi)} = 1$. Therefore, $k\pi$ is a pole of $\cot z$ of order 1.
(Observe that $D_{k\pi}(\pi)$ is the largest open disc with center $k\pi$ which is contained in the domain of holomorphy of g because it is the largest open disc with center $k\pi$ which does not contain any root of q_1. This is true because $q_1(z) = \frac{\sin z}{z-k\pi}$ vanishes at every $l\pi$ with $l \in \mathbb{Z}$, $l \neq k$.)
The Laurent series of $\cot z$ in $D_{k\pi}(0, \pi)$ is

$$\cot z = \tfrac{1}{z-k\pi} + g'(k\pi) + \tfrac{1}{2}g''(k\pi)(z - k\pi) + \cdots .$$

For the determination of poles there is a criterion similar to the criterion of Riemann for removable singularities.

Proposition 5.4. *Let z_0 be an isolated singularity of f. Then z_0 is a pole of f if and only if* $\lim_{z\to z_0} f(z) = \infty$.

Proof. There is a disc $D_{z_0}(R)$ so that f is holomorphic in $D_{z_0}(R) \setminus \{z_0\}$.
If z_0 is a pole of order N of f, then we saw that there is a function g holomorphic in $D_{z_0}(R)$ so that $g(z_0) \neq 0$ and $f(z) = \frac{g(z)}{(z-z_0)^N}$ for every $z \in D_{z_0}(R) \setminus \{z_0\}$. This implies $\lim_{z\to z_0} f(z) = \infty$.
Conversely, let $\lim_{z\to z_0} f(z) = \infty$. Then there is r with $0 < r \leq R$ so that $f(z) \neq 0$ for every $z \in D_{z_0}(r) \setminus \{z_0\}$ and then the function $h = \frac{1}{f}$ is holomorphic in $D_{z_0}(r) \setminus \{z_0\}$. Since $\lim_{z\to z_0} h(z) = \lim_{z\to z_0} \frac{1}{f(z)} = 0$, the criterion of Riemann implies that z_0 is a removable singularity of h. Therefore, we may define h appropriately at z_0 so that it becomes holomorphic in $D_{z_0}(r)$: we set $h(z_0) = \lim_{z\to z_0} h(z) = 0$. It is clear that z_0 is the only root of (the extended) h in $D_{z_0}(r)$ and, if N is the multiplicity of this root, then $h(z) = (z - z_0)^N h_1(z)$ where h_1 is holomorphic in $D_{z_0}(r)$ and has no root in $D_{z_0}(r)$. Thus, the function $g = \frac{1}{h_1}$ is holomorphic in $D_{z_0}(r)$ and, clearly, has no root in $D_{z_0}(r)$. Now we have altogether that $f(z) = \frac{g(z)}{(z-z_0)^N}$ for every $z \in D_{z_0}(r) \setminus \{z_0\}$ with $g(z_0) \neq 0$, and so z_0 is a pole of f of order N. $\square$

There is one more test for the case of a pole which also determines the exact order of the pole.

Proposition 5.5. *Let z_0 be an isolated singularity of f. Then z_0 is a pole of f of order $N \geq 1$ if and only if the* $\lim_{z\to z_0}(z-z_0)^N f(z)$ *exists and it is finite and* $\neq 0$.

Proof. If z_0 is a pole of f of order N, then we repeat the beginning of the proof of proposition 5.4 and we get that $\lim_{z\to z_0}(z-z_0)^N f(z) = \lim_{z\to z_0} g(z) = g(z_0) \neq 0$.
Conversely, let $\lim_{z\to z_0}(z-z_0)^N f(z)$ be finite and $\neq 0$. Riemann's criterion implies that the function $g(z) = (z-z_0)^N f(z)$, which is holomorphic in some ring $D_{z_0}(R)\setminus\{z_0\}$, can be extended at z_0 by setting $g(z_0) = \lim_{z\to z_0} g(z) = \lim_{z\to z_0}(z-z_0)^N f(z) \neq 0$, and the extended g is holomorphic in $D_{z_0}(R)$. Therefore, there is a g holomorphic in $D_{z_0}(R)$ with $g(z_0) \neq 0$ so that $f(z) = \frac{g(z)}{(z-z_0)^N}$ for every $z \in D_{z_0}(R) \setminus \{z_0\}$ and z_0 is a pole of f of order N. □

Finally, for the case of an essential singularity we have the following result.

Proposition 5.6. *Let z_0 be an isolated singularity of f. Then z_0 is an essential singularity of f if and only if the* $\lim_{z\to z_0} f(z)$ *does not exist.*

Proof. By the criterion of Riemann, z_0 is a removable singularity if and only if the $\lim_{z\to z_0} f(z)$ exists and it is finite. Proposition 5.4 says that z_0 is a pole if and only if $\lim_{z\to z_0} f(z) = \infty$. □

Example 5.8.6. In example 5.6.9 we saw that $\sum_{-\infty}^{n=-1} \frac{1}{(-n)!} z^n + 1$ is the Laurent series of $e^{\frac{1}{z}}$ in $D_0(0,+\infty)$. Hence 0 is an essential singularity of $e^{\frac{1}{z}}$.
Therefore, the $\lim_{z\to 0} e^{\frac{1}{z}}$ does not exist. We can see this without proving first that 0 is an essential singularity of $e^{\frac{1}{z}}$. In fact, proving that the $\lim_{z\to 0} e^{\frac{1}{z}}$ does not exist is another way to see that 0 is an essential singularity of $e^{\frac{1}{z}}$. Indeed, if $z = x$ tends to 0 on the positive x-semiaxis, then $|e^{\frac{1}{z}}| = e^{\frac{1}{x}} \to +\infty$, and hence $e^{\frac{1}{z}} \to \infty$. If $z = x$ tends to 0 on the negative x-semiaxis, then $|e^{\frac{1}{z}}| = e^{\frac{1}{x}} \to 0$, and hence $e^{\frac{1}{z}} \to 0$. Thus, the $\lim_{z\to 0} e^{\frac{1}{z}}$ does not exist.

Let z_0 be an isolated singularity of f and let $\sum_{-\infty}^{+\infty} a_n(z-z_0)^n$ be the Laurent series of f in the ring $D_{z_0}(0,R) = D_{z_0}(R)\setminus\{z_0\}$. Then $\sum_{-\infty}^{n=-1} a_n(z-z_0)^n$ is called the **singular part** of the Laurent series of f or, simply, the singular part of f at z_0. Also, $\sum_{n=0}^{+\infty} a_n(z-z_0)^n$ is called the **regular part** of the Laurent series of f or, simply, the regular part of f at z_0.

We have seen that in the case of a removable singularity z_0 the singular part of f at z_0 is zero and the Laurent series of f at z_0 consists only of its regular part. In the case of a pole z_0 of f of order N the singular part at z_0 is a finite sum of the form $\sum_{n=1}^{N} \frac{a_{-n}}{(z-z_0)^n}$ with $a_{-N} \neq 0$. In this case the singular part is a *rational function* whose denominator is $(z-z_0)^N$. In the case of an essential singularity z_0 the singular part at z_0 has infinitely many terms.

If we subtract from f its singular part at its singularity z_0, then we get

$$f(z) - \textstyle\sum_{-\infty}^{n=-1} a_n(z-z_0)^n = \sum_{n=0}^{+\infty} a_n(z-z_0)^n,$$

which is a power series of first type and hence converges in the disc $D_{z_0}(R)$, including the center z_0. Therefore, z_0 is a removable singularity of the function $F(z) = f(z) - \sum_{-\infty}^{n=-1} a_n(z-z_0)^n$ and if we define F to have value $F(z_0) = a_0$ at z_0, then this function is holomorphic in $D_{z_0}(R)$.

We shall now establish the well known *analysis of a rational function into a sum of simple fractions*.

Proposition 5.7. *Let $r = \frac{p}{q}$ be a rational function. We assume that the polynomials p, q have no common roots (and hence no common factors), that the degree of p is n, the degree of q is m and that $z_1, \dots, z_k$ are the roots of q with corresponding multiplicities $m_1, \dots, m_k$. Then*

$$r(z) = p_1\left(\tfrac{1}{z-z_1}\right) + \cdots + p_k\left(\tfrac{1}{z-z_k}\right) + p_0(z),$$

where $p_1, \dots, p_k$ are polynomials without constant terms and of degrees $m_1, \dots, m_k$, respectively, and p_0 is either the null polynomial, if $n < m$, or a polynomial of degree $n - m$, if $n \geq m$.

Proof. We saw in example 5.8.2 that each z_j is a pole of r of degree m_j. Then the singular part of r at z_j has the form $\sum_{l=1}^{m_j} \frac{a_{-l}}{(z-z_0)^l}$ with $a_{-m_j} \neq 0$. This can be written

$$\sum_{l=1}^{m_j} \frac{a_{-l}}{(z-z_0)^l} = p_j\big(\tfrac{1}{z-z_j}\big),$$

where p_j is the polynomial $p_j(z) = \sum_{l=1}^{m_j} a_{-l} z^l$ without constant term and of degree m_j. We subtract from r all its singular parts at $z_1, \dots, z_k$ and we form the function

$$p_0(z) = r(z) - \big(p_1(\tfrac{1}{z-z_1}) + \cdots + p_k(\tfrac{1}{z-z_k})\big).$$

This function is a rational function defined in $\mathbb{C} \setminus \{z_1, \dots, z_k\}$ and its only *possible* poles are the points $z_1, \dots, z_k$. We observe, though, that every z_j is a removable singularity of $r(z) - p_j(\frac{1}{z-z_j})$ and that each of $p_1(\frac{1}{z-z_1}), \dots, p_k(\frac{1}{z-z_k})$, besides $p_j(\frac{1}{z-z_j})$, is holomorphic at z_j. Thus, every z_j is a removable singularity of p_0. In other words, the rational function p_0 has *no poles* and hence it is a polynomial. Now, we have the identity

$$r(z) = p_1(\tfrac{1}{z-z_1}) + \cdots + p_k(\tfrac{1}{z-z_k}) + p_0(z)$$

and we consider two cases. If $n < m$, then $\lim_{z\to\infty} r(z) = 0$ and, since $\lim_{z\to\infty} p_j(\frac{1}{z-z_j}) = 0$ for every j, we have that $\lim_{z\to\infty} p_0(z) = 0$. Thus, p_0 is the null polynomial. If $n \geq m$, then $c = \lim_{z\to\infty} \frac{r(z)}{z^{n-m}}$ is a complex number $\neq 0$. Since $\lim_{z\to\infty} p_j(\frac{1}{z-z_j})/z^{n-m} = 0$ for every j, we have that $\lim_{z\to\infty} \frac{p_0(z)}{z^{n-m}} = c \neq 0$. Thus the polynomial p_0 has degree $n-m$. $\square$

Exercises.

5.8.1. Is 0 an isolated singularity of $\frac{1}{\sin(1/z)}$?

5.8.2. Find the isolated (non-removable) singularities of:

$$\frac{1}{z^2+5z+6}, \quad \frac{1}{(z^2-1)^2}, \quad \frac{e^z-1}{z}, \quad \frac{e^z-1}{z^3}, \quad \frac{z^2}{\sin z}, \quad \frac{1}{\sin z}, \quad \tan z, \quad \frac{1}{\sin^2 z}, \quad e^z + e^{1/z}, \quad \frac{1}{e^z-1}.$$

Which of the singularities are poles and what is their order?

5.8.3. Find the initial four terms of the Laurent series at 0 of the functions:

$$\cot z, \quad \frac{1}{\sin z}, \quad \frac{z}{\sin^2 z}, \quad \frac{1}{e^z-1}.$$

5.8.4. Prove that an isolated singularity of f cannot be a pole of e^f.

5.8.5. Let z_0 be an isolated singularity of f, which is not constant in any neighborhood of z_0. If there is $s \in \mathbb{R}$ so that $\lim_{z\to z_0} |z-z_0|^s |f(z)| \in [0, +\infty]$, prove that z_0 is either a removable singularity or a pole of f and that there is $m \in \mathbb{Z}$ so that

$$\lim_{z\to z_0} |z-z_0|^s |f(z)| \begin{cases} = 0, & \text{if } s > m \\ = +\infty, & \text{if } s < m \\ \in (0, +\infty), & \text{if } s = m \end{cases}$$

5.8.6. Let f be holomorphic in $\mathbb{C} \setminus \{0\}$ so that $\lim_{z\to 0} \frac{f(z)}{\sqrt{|z|}} = 0$ and $\lim_{z\to\infty} \frac{f(z)}{|z|\sqrt{|z|}} = 0$. What do you conclude about f?

5.8.7. Let f be holomorphic in $D_{z_0}(R) \setminus \{z_0\}$ and let either $\operatorname{Re} f$ or $\operatorname{Im} f$ be bounded either from above or from below in $D_{z_0}(R) \setminus \{z_0\}$. Prove that z_0 is a removable singularity of f.

5.8.8. Let f be holomorphic in $D_0(R) \setminus \{z_0\}$, where $R > 1$ and $|z_0| = 1$, and let z_0 be a pole of f. If $f(z) = \sum_{n=0}^{+\infty} a_n z^n$ is the Taylor series of f in $D_0(1)$, prove that $\frac{a_n}{a_{n+1}} \to z_0$.

5.8.9. Let Ω be a region so that every point of Ω is either a point of holomorphy or an isolated singularity of f. If the roots of f have an accumulation point in Ω, which is not an essential singularity of f, prove that f is identically 0 in Ω.

5.8.10. (i) Let z_0 be an essential singularity of f and let $w \in \mathbb{C}$. Prove that for every $r > 0$ the function $\frac{1}{f-w}$ *is not* bounded in $D_{z_0}(r) \setminus \{z_0\}$.
(ii) Prove the **Casorati-Weierstrass theorem**. If z_0 is an essential singularity of f, then for every w there is a sequence (z_n) with $z_n \to z_0$ and $z_n \neq z_0$ for every n so that $f(z_n) \to w$.

5.8.11. (i) Prove that every $2k\pi i$, $k \in \mathbb{Z}$, is a pole of $\frac{1}{e^z-1}$ of order 1.
(ii) Prove that

$$\tfrac{1}{e^z-1} = \tfrac{1}{z} - \tfrac{1}{2} + \textstyle\sum_{k=1}^{+\infty}(-1)^{k-1}\frac{B_k}{(2k)!}z^{2k-1}$$

for $|z| < 2\pi$, where the numbers B_k satisfy the recursive relations

$$\tfrac{1}{(2k+1)!} - \tfrac{1}{2(2k)!} + \textstyle\sum_{\nu=1}^{k}\frac{(-1)^{\nu-1}B_\nu}{(2\nu)!(2k-2\nu+1)!} = 0$$

for $k \geq 1$. Evaluate B_1, B_2, B_3. The numbers B_k are called **Bernoulli constants**.

5.8.12. Look at exercises 5.6.6 and 5.7.8. We shall extend what we said in this section to the case of the point ∞.
(i) We say that ∞ is an isolated singularity of f if f is holomorphic in some ring $D_0(R, +\infty)$. Let $\sum_{-\infty}^{+\infty} a_n z^n$ be the Laurent series of f in this ring. If $a_n = 0$ for every $n \geq 1$, then we say that ∞ is a removable singularity of f. If $a_n \neq 0$ for at least one $n \geq 1$ and for only finitely many $n \geq 1$, then we say that ∞ is a pole of f. Finally, if $a_n \neq 0$ for infinitely many $n \geq 1$, then we say that ∞ is an essential singularity of f.
Prove that ∞ is a removable singularity of f if and only if $\lim_{z\to\infty}\frac{f(z)}{z} = 0$.
Prove that ∞ is a pole of f if and only if $\lim_{z\to\infty} f(z) = \infty$.
Let ∞ be a pole of f and let N be the largest $n \geq 1$ with $a_n \neq 0$. Then we say that ∞ is a pole of f of order N. Prove that ∞ is a pole of f of order N if and only if there is a g holomorphic in $D_0(R, +\infty) \cup \{\infty\}$ so that $g(\infty) \neq 0$ and $f(z) = z^N g(z)$ for every $z \in D_0(R, +\infty)$. Moreover, prove that ∞ is a pole of f of order N if and only if the $\lim_{z\to\infty}\frac{f(z)}{z^N}$ exists and it is finite and $\neq 0$.
Prove that ∞ is an essential singularity of f if and only if the $\lim_{z\to\infty} f(z)$ does not exist.
(ii) Let $r = \frac{p}{q}$ be a rational function and let n be the degree of the polynomial p and m be the degree of the polynomial q. Prove that ∞ is a removable singularity of r if $m \geq n$ and that it is a pole of r of order $n - m$ if $n > m$. In particular, a polynomial p of degree $n \geq 1$ has a pole of order n at ∞.
(iii) What kind of an isolated singularity is ∞ for the following functions?

$$e^z, \quad e^{\frac{1}{z}}, \quad z^2 e^{\frac{1}{z}}, \quad \sin z, \quad \sin\tfrac{1}{z}, \quad z^5 \sin\tfrac{1}{z}.$$

(iv) What kind of an isolated singularity is ∞ for any holomorphic branch of $(z^2-1)^{\frac{1}{2}}$ in the region $\mathbb{C} \setminus [-1, 1]$? (For the existence of such a branch look at exercise 4.4.9.)
(v) Is ∞ an isolated singularity of $\frac{1}{\sin z}$ or of $\tan z$?

5.9 Maximum principle.

Maximum principle. *Let f be holomorphic in the region $\Omega \subseteq \mathbb{C}$ and $M = \sup_{z\in\Omega}|f(z)|$. If there is $z_0 \in \Omega$ so that $|f(z_0)| = M$, then f is constant in Ω.*

Proof. We take any $z \in \Omega$ for which $|f(z)| = M$. We consider an open disc $D_z(R) \subseteq \Omega$ and any r with $0 < r < R$. The mean value property of f says that $f(z) = \frac{1}{2\pi}\int_0^{2\pi} f(z + re^{it})\,dt$. Since $|f(z + re^{it})| \leq M$ for every $t \in [0, 2\pi]$, we have

$$M = |f(z)| = \left|\tfrac{1}{2\pi}\textstyle\int_0^{2\pi} f(z + re^{it})\,dt\right| \leq \tfrac{1}{2\pi}\textstyle\int_0^{2\pi} |f(z + re^{it})|\,dt \leq M.$$

Hence, $\frac{1}{2\pi}\int_0^{2\pi}|f(z+re^{it})|\,dt = M$ and, since $|f(z+re^{it})|$ is a continuous function of t, we get $|f(z+re^{it})| = M$ for every $t\in[0,2\pi]$. Now, r is arbitrary in the interval $(0,R)$ and we find that $|f(z+re^{it})| = M$ for every $t\in[0,2\pi]$ and every $r\in(0,R)$. So we get $|f(w)| = M$ for every $w\in D_z(R)$. We proved that, if $|f(z)| = M$ for a $z\in\Omega$, then this equality holds in a neighborhood of z. Now we define

$$B = \{z\in\Omega\,|\,|f(z)| = M\}, \qquad C = \{z\in\Omega\,|\,|f(z)| < M\}$$

and it is clear that $B\cup C = \Omega$.
If $z\in B$, then $|f(z)| = M$ and hence the same is true at every point in a neighborhood of z. Therefore z is not a limit point of C. Moreover, if $z\in C$ then $|f(z)| < M$ and, by the continuity of f, the same is true in a neighborhood of z. Hence z is not a limit point of B.
If both B and C are non-empty, then they form a decomposition of Ω. But Ω is connected and, since $z_0\in B$, we get that $C=\emptyset$. Therefore, $|f(z)| = M$ for every $z\in\Omega$.
Now we shall prove that f is constant in Ω. If $M=0$, then clearly $f=0$ in Ω. So let us assume that $M>0$. If u and v are the real and the imaginary part of f, then u^2+v^2 is constant M^2 in Ω and hence $u\frac{\partial u}{\partial x}+v\frac{\partial v}{\partial x}=0$ and $u\frac{\partial u}{\partial y}+v\frac{\partial v}{\partial y}=0$ in Ω. Using the C-R equations, we get

$$u\frac{\partial u}{\partial x}+v\frac{\partial v}{\partial x}=0, \quad v\frac{\partial u}{\partial x}-u\frac{\partial v}{\partial x}=0$$

in Ω. Viewing this as a system with unknowns $\frac{\partial u}{\partial x}$, $\frac{\partial v}{\partial x}$, we see that its determinant is $u^2+v^2 = M^2>0$, and we find that $\frac{\partial u}{\partial x}=0$ and $\frac{\partial v}{\partial x}=0$ in Ω. Therefore, $f'=\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}=0$ in Ω and so f is constant in the region Ω. □

That was a first version of the maximum principle. There is a second version. In this second version the region Ω is a subset of $\mathbb{C}$, but when we consider $\overline{\Omega}$ or $\partial\Omega$ we shall think of them as subsets of $\widehat{\mathbb{C}}$. In other words, if Ω is unbounded, then we assume that $\overline{\Omega}$ and $\partial\Omega$ contain the point ∞. This assumption holds until the end of this subsection, including the exercises.

Maximum principle. *Let f be holomorphic in the region Ω and continuous in $\overline{\Omega}$. Then either f is constant in $\overline{\Omega}$ or $|f|$ has a maximum value in $\overline{\Omega}$, say M, attained at a point of $\partial\Omega$ and $|f(z)|<M$ for every $z\in\Omega$. In every case, $|f|$ has a maximum value in $\overline{\Omega}$ which is attained at a point of $\partial\Omega$. In other words we have*

$$\max_{z\in\overline{\Omega}}|f(z)| = \max_{\zeta\in\partial\Omega}|f(\zeta)|.$$

Proof. If f is constant in $\overline{\Omega}$, then $|f|$ is also constant, say M, in $\overline{\Omega}$. Then, obviously, M is the maximum value of $|f|$ and it is attained (everywhere and hence) at every point of $\partial\Omega$.
Now we assume that f is not constant in $\overline{\Omega}$. This implies easily that f is not constant in Ω either. Now, $|f|$ is continuous in the compact set $\overline{\Omega}$ and hence attains its maximum value, say M, at some point $z_0\in\overline{\Omega}$. I.e. we have $|f(z_0)| = M$ and $|f(z)|\le M$ for every $z\in\overline{\Omega}$.
If any such z_0 belongs to Ω, then the previous maximum principle implies that f is constant in Ω and we arrive at a contradiction. We conclude that $z_0\in\partial\Omega$ and $|f(z)|<M$ for every $z\in\Omega$. □

The second version of the maximum principle is usually applied in the simplified form:

Let f be holomorphic in the region Ω and continuous in $\overline{\Omega}$. If $|f(\zeta)|\le M$ for every $\zeta\in\partial\Omega$, then $|f(z)|\le M$ for every $z\in\Omega$.

Besides the maximum principle, we have the **minimum principle**. It is stated in two versions which can be found in exercise 5.9.1. Here we state a usefull simplified form:

Let f be holomorphic in the region Ω and continuous in $\overline{\Omega}$ so that $f(z)\ne 0$ for every $z\in\overline{\Omega}$. If $|f(\zeta)|\ge m$ for every $\zeta\in\partial\Omega$, then $|f(z)|\ge m$ for every $z\in\Omega$.

The proof is a trivial application of the previous simplified form of the maximum principle to the function $\frac{1}{f}$ which is holomorphic in Ω and continuous in $\overline{\Omega}$.

Exercises.

5.9.1. (i) Let f be holomorphic in the region $\Omega \subseteq \mathbb{C}$ so that $f(z) \neq 0$ for every $z \in \Omega$ and $m = \inf_{z\in\Omega} |f(z)|$. If there is $z_0 \in \Omega$ so that $|f(z_0)| = m$, then f is constant in Ω.
(ii) Let f be holomorphic in the region Ω and continuous in $\overline{\Omega}$ so that $f(z) \neq 0$ for every $z \in \overline{\Omega}$. Then either f is constant in $\overline{\Omega}$ or $|f|$ has a minimum value in $\overline{\Omega}$, say m, attained at a point of $\partial\Omega$ and $|f(z)| > m$ for every $z \in \Omega$. In every case, $|f|$ has a minimum value in $\overline{\Omega}$ which is attained at a point of $\partial\Omega$. In other words we have $\min_{z\in\overline{\Omega}} |f(z)| = \min_{\zeta\in\partial\Omega} |f(\zeta)|$.
Both (i) and (ii) are called **minimum principle**.

5.9.2. Let f be holomorphic in $\mathbb{D}$ and continuous in $\overline{\mathbb{D}}$ so that $|f(z)| > 1$ for every $z \in \mathbb{T}$ and $f(0) = 1$. Does f have a root in $\mathbb{D}$?

5.9.3. Let f be holomorphic in the region Ω and $\lim_{z\to\zeta} f(z) = 0$ for every $\zeta \in \partial\Omega$. Prove that f is constant 0 in Ω.

5.9.4. Let f be holomorphic in the region $\Omega \subseteq \mathbb{C}$ and $K = \sup_{z\in\Omega} \operatorname{Re} f(z)$. If there is $z_0 \in \Omega$ so that $\operatorname{Re} f(z_0) = K$, prove that f is constant in Ω.

5.9.5. Prove the fundamental theorem of algebra using the maximum principle.

5.9.6. Let f_n, f be holomorphic in the region Ω and continuous in $\overline{\Omega}$. If $f_n \to f$ uniformly in $\partial\Omega$, prove that $f_n \to f$ uniformly in $\overline{\Omega}$.

5.9.7. Let R be a square region with center z_0. Let f be holomorphic in R and continuous in $\overline{R}$. If $|f(z)| \leq m$ for every z in one of the four sides of R and $|f(z)| \leq M$ for every z in the other three sides of R, prove that $|f(z_0)| \leq \sqrt[4]{mM^3}$.

5.9.8. Let $\Omega = \{x + iy \,|\, -\frac{\pi}{2} < y < \frac{\pi}{2}\}$ and $f(z) = e^{e^z}$. Then f is holomorphic in Ω and continuous in $\overline{\Omega} = \{x + iy \,|\, -\frac{\pi}{2} \leq y \leq \frac{\pi}{2}\}$. Prove that $|f(x - i\frac{\pi}{2})| = |f(x + i\frac{\pi}{2})| = 1$ for every $x \in \mathbb{R}$ and that $\lim_{x\to+\infty} f(x) = +\infty$. Does this contradict the maximum principle?

5.9.9. Let f be holomorphic in the region Ω and continuous in $\overline{\Omega}$.
(i) If $|f|$ is constant in $\partial\Omega$, prove that either f has at least one root in Ω or f is constant in Ω.
(ii) If $\operatorname{Re} f$ or $\operatorname{Im} f$ is constant in $\partial\Omega$, prove that f is constant in Ω.

5.9.10. (i) Let $z_0 \in \mathbb{D}$, $|\lambda| = 1$ and

$$T(z) = \frac{z-z_0}{1-\overline{z_0}z}$$

for $z \in \overline{\mathbb{D}}$. This l.f.t. appears in exercise 4.1.8 and we know that T is holomorphic in $\mathbb{D}$ and continuous in $\overline{\mathbb{D}}$, and that $T(z) \in \mathbb{D}$ for every $z \in \mathbb{D}$, and $T(z) \in \mathbb{T}$ for every $z \in \mathbb{T}$.
Now let $z_1, \ldots, z_n \in \mathbb{D}$ and $|\lambda| = 1$ and

$$B(z) = \lambda \prod_{k=1}^{n} \frac{z-z_k}{1-\overline{z_k}z}$$

for $z \in \overline{\mathbb{D}}$. Then B is holomorphic in $\mathbb{D}$ and continuous in $\overline{\mathbb{D}}$. Prove that $B(z) \in \mathbb{D}$ for every $z \in \mathbb{D}$, and that $B(z) \in \mathbb{T}$ for every $z \in \mathbb{T}$.
Every function B of this form is called **(finite) Blaschke product**.
(ii) Prove the converse of (i). I.e. let f be holomorphic in $\mathbb{D}$ and continuous in $\overline{\mathbb{D}}$ and let $f(z) \in \mathbb{D}$ for every $z \in \mathbb{D}$ and $f(z) \in \mathbb{T}$ for every $z \in \mathbb{T}$. If f is non-constant, prove that there is $n \in \mathbb{N}$ and $z_1, \ldots, z_n \in \mathbb{D}$ and λ with $|\lambda| = 1$ so that $f(z) = \lambda \prod_{k=1}^{n} \frac{z-z_k}{1-\overline{z_k}z}$ for every $z \in \overline{\mathbb{D}}$.

5.9.11. Let f be holomorphic in the region Ω so that $\limsup_{\Omega\ni z\to\zeta} |f(z)| \leq M$ for every $\zeta \in \partial\Omega$. Prove that $|f(z)| \leq M$ for every $z \in \Omega$. Moreover, if $|f(z)| = M$ for at least one $z \in \Omega$, prove that f is constant in Ω.

5.9.12. Let the complex function f be holomorphic in the region Ω and continuous in $\overline{\Omega}$. If U is an open set so that $\overline{U} \subseteq \Omega$, prove that $\max_{z\in\partial U} |f(z)| \leq \max_{z\in\partial\Omega} |f(z)|$. If equality holds, prove that f is constant in $\overline{\Omega}$.

5.9.13. Let f be holomorphic in $D_0(R_1, R_2)$ and $a \in \mathbb{R}$. Prove that $|z|^a|f(z)|$ has no maximum value in $D_0(R_1, R_2)$, except if $a \in \mathbb{Z}$ and there is c so that $f(z) = cz^{-a}$ for every $z \in D_0(R_1, R_2)$.

5.9.14. Let f, g be holomorphic in $\mathbb{C}$ and $|f(z)| \leq |g(z)|$ for every z. Prove that there is μ so that $f(z) = \mu g(z)$ for every z.

5.9.15. Let f be holomorphic in $\mathbb{D}$. Prove that there is a sequence (z_n) in $\mathbb{D}$ so that $|z_n| \to 1$ and $(f(z_n))$ is bounded.

5.9.16. (i) Let f be holomorphic and non-constant in the region $\Omega \subseteq \mathbb{C}$. For every $\mu > 0$ prove that

$$\overline{\{z \in \Omega \,|\, |f(z)| < \mu\}} \cap \Omega = \{z \in \Omega \,|\, |f(z)| \leq \mu\}.$$

(ii) Let p be a polynomial of degree $n \geq 1$. Prove that for every $\mu > 0$ the set $\{z \,|\, |p(z)| < \mu\}$ has at most n connected components and each of them contains at least one root of p. How do these connected components behave when $\mu \to 0+$ and when $\mu \to +\infty$?

5.9.17. The three circles theorem of Hadamard. Let f be holomorphic in $D_{z_0}(R_1, R_2)$ and let

$$M(r) = \max_{z\in C_{z_0}(r)} |f(z)|$$

for $R_1 < r < R_2$. Prove that $\ln M(r)$ is a convex function of $\ln r$ in (R_1, R_2). I.e. prove that, if $R_1 < r_1 < r_2 < R_2$ and $\ln r = (1-t)\ln r_1 + t \ln r_2$ for $0 < t < 1$, then $\ln M(r) \leq (1-t)\ln M(r_1) + t \ln M(r_2)$. Another way to express this is:

$$M(r) \leq M(r_1)^{\frac{\ln r_2 - \ln r}{\ln r_2 - \ln r_1}} M(r_2)^{\frac{\ln r - \ln r_1}{\ln r_2 - \ln r_1}}$$

when $R_1 < r_1 < r < r_2 < R_2$.

5.9.18. The three lines theorem. Let f be holomorphic and bounded in the vertical zone $K = \{x + iy \,|\, X_1 < x < X_2\}$ and let

$$M(x) = \sup_{y\in\mathbb{R}} |f(x + iy)|$$

for $X_1 < x < X_2$. Prove that $\ln M(x)$ is a convex function of x in (X_1, X_2). I.e. prove that, if $X_1 < x_1 < x_2 < X_2$ and $x = (1-t)x_1 + tx_2$ for $0 < t < 1$, then $\ln M(x) \leq (1-t)\ln M(x_1) + t \ln M(x_2)$. Another way to express this is:

$$M(x) \leq M(x_1)^{\frac{x_2 - x}{x_2 - x_1}} M(r_2)^{\frac{x - x_1}{x_2 - x_1}}$$

when $X_1 < x_1 < x < x_2 < X_2$.

5.9.19. The Phragmén-Lindelöf theorem. Let f, ϕ be holomorphic in the region Ω and let ϕ be bounded in Ω and have no root in Ω. Let also $A \cap B = \emptyset$ and $A \cup B = \partial\Omega$. If
(i) $\overline{\lim}_{\Omega\ni z\to\zeta} |f(z)| \leq M$ for every $\zeta \in A$ and
(ii) $\overline{\lim}_{\Omega\ni z\to\zeta} |f(z)||\phi(z)|^\epsilon \leq M$ for every $\zeta \in B$ and every $\epsilon > 0$,
then prove that $|f(z)| \leq M$ for every $z \in \Omega$. If, moreover, f is non-constant in Ω, prove that $|f(z)| < M$ for every $z \in \Omega$.

5.10 The open mapping theorem.

Open mapping theorem. *If f is holomorphic and not constant in the region Ω, then $f(U)$ is open for every open $U \subseteq \Omega$.*

Proof. Let $U \subseteq \Omega$ be open. We shall prove that $f(U)$ is also open, i.e. that every $w_0 \in f(U)$ is an interior point of $f(U)$.
Since $w_0 \in f(U)$ there is some $z_0 \in U$ so that $f(z_0) = w_0$. Since U is open, there is $r > 0$ so that $\overline{D}_{z_0}(r) \subseteq U$. Since f is non-constant in Ω, the solution z_0 of the equation $f(z) = w_0$ is isolated. Therefore, we may take r small enough so that $f(z) = w_0$ has no solution in $\overline{D}_{z_0}(r)$ except z_0. Thus, $f(\zeta) \neq w_0$ for every $\zeta \in C_{z_0}(r)$ and by the continuity of $|f - w_0|$ we get that there is some $\epsilon > 0$ so that $|f(\zeta) - w_0| \geq \epsilon$ for every $\zeta \in C_{z_0}(r)$. Now, we consider any $w \notin f(\overline{D}_{z_0}(r))$ and we have that

$$|f(\zeta) - w| \geq |f(\zeta) - w_0| - |w_0 - w| \geq \epsilon - |w_0 - w|$$

for every $\zeta \in C_{z_0}(r)$. But the function $f - w$ is holomorphic in $D_{z_0}(r)$ and continuous in $\overline{D}_{z_0}(r)$ and also $f(z) - w \neq 0$ for every $z \in \overline{D}_{z_0}(r)$. Therefore, by (the simplified form of) the minimum principle at the end of section 5.9, we get

$$|w_0 - w| = |f(z_0) - w| \geq \epsilon - |w_0 - w|.$$

Thus $|w_0 - w| \geq \frac{\epsilon}{2}$ and we have proved that any $w \notin f(\overline{D}_{z_0}(r))$ satisfies $|w_0 - w| \geq \frac{\epsilon}{2}$. This implies that every $w \in D_{w_0}(\frac{\epsilon}{2})$ belongs to $f(\overline{D}_{z_0}(r))$. Hence $D_{w_0}(\frac{\epsilon}{2}) \subseteq f(\overline{D}_{z_0}(r)) \subseteq f(U)$ and so w_0 is an interior point of $f(U)$. $\square$

Exercises.

5.10.1. Prove the first maximum principle using the open mapping theorem.

5.11 Local mapping properties.

Proposition 5.8. *Let f be holomorphic in the open set Ω and let $z_0 \in \Omega$ with $f'(z_0) \neq 0$. Then there is an open set $U \subseteq \Omega$ containing z_0 so that $W = f(U)$ is an open set containing $w_0 = f(z_0)$ and so that the function $f : U \to W$ is one-to-one. Moreover, the function $f^{-1} : W \to U$ is holomorphic in W.*

Proof. Since f' is continuous, there is $r > 0$ so that $|f'(z) - f'(z_0)| \leq \frac{1}{2}|f'(z_0)|$ for every $z \in D_{z_0}(r)$. This implies $|f'(z)| \geq |f'(z_0)| - |f'(z) - f'(z_0)| \geq \frac{1}{2}|f'(z_0)| > 0$ and hence $f'(z) \neq 0$ for every $z \in D_{z_0}(r)$. Furthermore,

$$|f(z_2) - f(z_1) - f'(z_0)(z_2 - z_1)| = \Big| \int_{[z_1,z_2]} (f'(z) - f'(z_0))\, dz \Big| \leq \tfrac{1}{2}|z_2 - z_1||f'(z_0)|$$

for every $z_1, z_2 \in D_{z_0}(r)$. This implies

$$|f(z_2) - f(z_1)| \geq |f'(z_0)(z_2 - z_1)| - |f(z_2) - f(z_1) - f'(z_0)(z_2 - z_1)| \geq \tfrac{1}{2}|z_2 - z_1||f'(z_0)| > 0$$

for every $z_1, z_2 \in D_{z_0}(r)$ with $z_1 \neq z_2$.
Now we take $U = D_{z_0}(r)$. From the open mapping theorem we have that the set $W = f(U)$ is open. We have proved that $f' \neq 0$ in U and that $f : U \to W$ is one-to-one and onto and so the inverse mapping $f^{-1} : W \to U$ is defined. Now it is easy to see that this inverse mapping is continuous in W. Indeed let $w \in W$. Then there is (a unique) $z \in U$ so that $f(z) = w$. We take any $\epsilon > 0$ small enough so that $D_z(\epsilon) \subseteq U$. Then the set $f(D_z(\epsilon))$ is open and contains w. Hence there is $\delta > 0$ so that $D_w(\delta) \subseteq f(D_z(\epsilon))$. Then for every $w' \in D_w(\delta)$ the (unique) $z' \in U$ which satisfies $f(z') = w'$ is contained in $D_z(\epsilon)$. This says that for every $w' \in W$ with $|w' - w| < \delta$ we have $|f^{-1}(w') - f^{-1}(w)| = |z' - z| < \epsilon$ and the function $f^{-1} : W \to U$ is continuous at every $w \in W$. Now, proposition 3.4 implies that $f^{-1} : W \to U$ is holomorphic in W. $\square$

Theorem 5.2. *Let f be holomorphic in the region Ω and let $z_0 \in \Omega$ and $w_0 = f(z_0)$. Let z_0 be a solution of $f(z) = w_0$ of multiplicity N. Then there is an open set $U \subseteq \Omega$ containing z_0 so that $W = f(U)$ is an open set containing $w_0 = f(z_0)$ and so that the function $f : U \to W$ is N-to-one.*

Proof. We know that there is a disc $D_{z_0}(R)$ and a function g holomorphic in $D_{z_0}(R)$ so that

$$f(z) - w_0 = (z - z_0)^N g(z)$$

for every $z \in D_{z_0}(R)$ and $g(z_0) \neq 0$. By the continuity of g we have that there is $r \leq R$ so that $g(z) \neq 0$ for every $z \in D_{z_0}(r)$. Then the function $\frac{g'}{g}$ is holomorphic in $D_{z_0}(r)$ and the theorem of Cauchy in convex regions implies that $\oint_\gamma \frac{g'(z)}{g(z)}\, dz = 0$ for every closed curve γ in $D_{z_0}(r)$. Now, theorem 4.1 implies that there is a holomorphic branch of $\log g$ in $D_{z_0}(r)$ and then example 4.4.4 says that there is a holomorphic branch of $g^{1/N}$ in $D_{z_0}(r)$. I.e. there is a function ϕ holomorphic in $D_{z_0}(r)$ so that $\phi(z)^N = g(z)$ for every $z \in D_{z_0}(r)$. Now we consider the function $h(z) = (z - z_0)\phi(z)$. This is holomorphic in $D_{z_0}(r)$ and we have that

$$f(z) - w_0 = h(z)^N$$

for every $z \in D_{z_0}(r)$. Moreover, $h'(z_0) = \phi(z_0) \neq 0$. Proposition 5.8, applied to h, implies that there is an open set $U_0 \subseteq D_{z_0}(r)$ containing z_0 so that $W_0 = h(U_0)$ is an open set containing $h(z_0) = 0$ and so that the function $h : U_0 \to W_0$ is one-to-one. Now, we consider a disc $D_0(r_0) \subseteq W_0$ and the open set $U = h^{-1}(D_0(r_0)) \subseteq U_0$. Then $h : U \to D_0(r_0)$ is holomorphic in U, onto $D_0(r_0)$ and one-to-one in U. Moreover, we have that $f(z) - w_0 = h(z)^N$ for every $z \in U$. Since the N-th power $w = \zeta^N$ maps the disc $D_0(r_0)$ onto the disc $D_0(r_0^N)$ and in an N-to-one manner, we conclude that $f : U \to W$ is N-to-one, where W is the disc $D_{w_0}(r_0^N)$. □

In the proof of theorem 5.2 if we take any linear segment $[w_0, w]$ in the disc $D_{w_0}(r_0^N)$, where w is a point of the circle $C_{w_0}(r_0^N)$, then, through the mapping $w = w_0 + \zeta^N$, this linear segment corresponds to N linear segments $[0, z_1], \dots, [0, z_N]$ in the disc $D_0(r_0)$, where $z_1, \dots, z_N$ are N points on the circle $C_0(r_0)$. These N linear segments form N successive angles at 0 all equal to $\frac{2\pi}{N}$. Now the one-to-one function $h^{-1} : D_0(r_0) \to U$ maps these linear segments onto N curves $\gamma_1, \dots, \gamma_N$ with common initial endpoint z_0 and N corresponding final endpoints on ∂U. Since $h'(z_0) \neq 0$, the conformality of h at z_0 implies that $\gamma_1, \dots, \gamma_N$ form N successive angles at z_0 all equal to $\frac{2\pi}{N}$. The N successive "angular" regions $U_1, \dots, U_N$ in U between the curves $\gamma_1, \dots, \gamma_N$ are mapped by h onto the corresponding succesive angular regions $A_1, \dots, A_N$ in $D_0(r_0)$ between the linear segments $[0, z_1], \dots, [0, z_N]$ and these are then mapped by the mapping $w = w_0 + \zeta^N$ onto the same region $B = D_{w_0}(r_0^N) \setminus [w_0, w]$. We conclude that f, which is the composition of the two mappings, maps each of $U_1, \dots, U_N$ in U onto B and in an one-to-one manner.

Exercises.

5.11.1. Let f be holomorphic in $D_0(R)$, $f'(0) \neq 0$ and $n \in \mathbb{N}$. Prove that there is $r > 0$ and there is g holomorphic in $D_0(r)$ so that $f(z^n) = f(0) + g(z)^n$ for every $z \in D_0(r)$.

5.11.2. Let Ω_1, Ω_2 be two regions, let $f : \Omega_1 \to \Omega_2$ and $g : \Omega_2 \to \mathbb{C}$ be non-constant functions and let $h = g \circ f$.
(i) If f, h are holomorphic in Ω_1, is g holomorphic in Ω_2?
(ii) If g, h are holomorphic in Ω_2, Ω_1, respectively, is f holomorphic in Ω_1?

5.11.3. If f is holomorphic and one-to-one in $\mathbb{C}$, prove that there are $a \neq 0$ and b so that $f(z) = az + b$ for every z.

5.12 Uniform convergence in compact sets and holomorphy.

The theorem of Weierstrass. *Let every f_n be holomorphic in the open set $\Omega \subseteq \mathbb{C}$. If $f_n \to f$ uniformly in every compact subset of Ω, then f is also holomorphic in Ω and for every $k \in \mathbb{N}$ we have that $f_n^{(k)} \to f^{(k)}$ uniformly in every compact subset of Ω.*

Proof. We take any $z_0 \in \Omega$. Then there is a closed disc $\overline{D}_{z_0}(R)$ contained in Ω and for every n we have

$$f_n(z) = \tfrac{1}{2\pi i} \oint_{C_{z_0}(R)} \tfrac{f_n(\zeta)}{\zeta - z}\, d\zeta \tag{5.10}$$

for every $z \in D_{z_0}(R)$. Since $C_{z_0}(R)$ is a compact subset of Ω, we have that $f_n \to f$ uniformly in $C_{z_0}(R)$. We also have that $f_n(z) \to f(z)$ for every $z \in D_{z_0}(R)$. Hence

$$f(z) = \tfrac{1}{2\pi i} \oint_{C_{z_0}(R)} \tfrac{f(\zeta)}{\zeta - z}\, d\zeta \tag{5.11}$$

for every $z \in D_{z_0}(R)$. The right side of this equality is a holomorphic function of z in $D_{z_0}(R)$ and so the left side, $f(z)$, is also holomorphic in $D_{z_0}(R)$. Thus, f is holomorphic at every $z_0 \in \Omega$. Now, from the variants of (5.10) and (5.11) for derivatives, we have for every $z \in D_{z_0}(\frac{R}{2})$ that

$$\begin{aligned} |f_n^{(k)}(z) - f^{(k)}(z)| &= \Big|\tfrac{k!}{2\pi i} \oint_{C_{z_0}(R)} \tfrac{f_n(\zeta) - f(\zeta)}{(\zeta - z)^{k+1}}\, d\zeta\Big| \le \tfrac{k!}{2\pi}\, \tfrac{\|f_n - f\|_{C_{z_0}(R)}}{(R/2)^{k+1}}\, 2\pi R \\ &= \tfrac{2^{k+1} k!}{R^k}\, \|f_n - f\|_{C_{z_0}(R)}. \end{aligned}$$

Hence,

$$\|f_n^{(k)} - f^{(k)}\|_{D_{z_0}(R/2)} \le \tfrac{2^{k+1} k!}{R^k}\, \|f_n - f\|_{C_{z_0}(R)}$$

and so $f_n^{(k)} \to f^{(k)}$ uniformly in $D_{z_0}(\frac{R}{2})$.
We proved that every $z \in \Omega$ has a neighborhood $D_z(r_z)$ in which $f_n^{(k)}$ converges uniformly to f. Now, if $K \subseteq \Omega$ is compact, there are $z_1, \ldots, z_n \in K$ so that $K \subseteq D_{z_1}(r_{z_1}) \cup \cdots \cup D_{z_n}(r_{z_n})$. Since $f_n^{(k)} \to f^{(k)}$ uniformly in each $D_{z_j}(r_{z_j})$, we conclude that $f_n^{(k)} \to f^{(k)}$ uniformly in K. □

The theorem of Hurwitz. *Let every f_n be holomorphic in the region $\Omega \subseteq \mathbb{C}$ and $f_n \to f$ uniformly in every compact subset of Ω. If $f_n(z) \neq 0$ for every n and every $z \in \Omega$, then either $f(z) \neq 0$ for every $z \in \Omega$ or $f(z) = 0$ for every $z \in \Omega$.*

First proof. The theorem of Weierstrass implies that f is holomorphic in Ω. We assume that f is not identically 0 in Ω and we shall prove that $f(z) \neq 0$ for every $z \in \Omega$.
We take any $z_0 \in \Omega$. Even if $f(z_0) = 0$, we know that z_0 is an isolated root of f and hence there is $r > 0$ so that $f(\zeta) \neq 0$ for every $\zeta \in C_{z_0}(r)$. By the continuity of f we get that there is some $\delta > 0$ so that $|f(\zeta)| \ge \delta$ for every $\zeta \in C_{z_0}(r)$. Now, we have that $f_n \to f$ uniformly in $\overline{D}_{z_0}(r)$ and so there is n so that

$$|f_n(z) - f(z)| \le \tfrac{\delta}{4} \tag{5.12}$$

for every $z \in \overline{D}_{z_0}(r)$. Therefore

$$|f_n(\zeta)| \ge |f(\zeta)| - |f_n(\zeta) - f(\zeta)| \ge \delta - \tfrac{\delta}{4} = \tfrac{3\delta}{4}$$

for every $\zeta \in C_{z_0}(r)$. Since f_n is holomorphic in $D_{z_0}(r)$ and continuous in $\overline{D}_{z_0}(r)$ and $f_n(z) \neq 0$ for every $z \in \overline{D}_{z_0}(r)$, by the minimum principle we have that $|f_n(z)| \ge \frac{3\delta}{4}$ for every $z \in D_{z_0}(r)$. This and (5.12) imply

$$|f(z)| \ge |f_n(z)| - |f_n(z) - f(z)| \ge \tfrac{3\delta}{4} - \tfrac{\delta}{4} = \tfrac{\delta}{2}$$

for every $z \in D_{z_0}(r)$. Thus there is no root of f in the disc $D_{z_0}(r)$. In particular, $f(z_0) \neq 0$.
Second proof. We follow the first proof up to the point that we get $|f(\zeta)| \ge \delta$ for every $\zeta \in C_{z_0}(r)$.

Now, we have that $f_n \to f$ uniformly in $C_{z_0}(r)$ and the theorem of Weierstrass implies that also $f_n' \to f'$ uniformly in $C_{z_0}(r)$. Therefore, $\frac{f_n'}{f_n} \to \frac{f'}{f}$ uniformly in $C_{z_0}(r)$ and hence

$$\tfrac{1}{2\pi i}\oint_{C_{z_0}(r)} \tfrac{f_n'(\zeta)}{f_n(\zeta)}\,d\zeta \to \tfrac{1}{2\pi i}\oint_{C_{z_0}(r)} \tfrac{f'(\zeta)}{f(\zeta)}\,d\zeta.$$

By the argument principle, the left side is equal to the number of roots of f_n in the disc $D_{z_0}(r)$ and hence it is equal to 0. Thus, the right side is also equal to 0 and, by the argument principle again, there is no root of f in the disc $D_{z_0}(r)$. In particular, $f(z_0) \neq 0$. $\square$

We now recall some definitions for collections of complex functions defined in a subset of a general metric space: here our metric space will be $\mathbb{C}$.

Let $A \subseteq \mathbb{C}$ and $\mathcal{F}$ be a collection of complex functions defined in the set A. We say that $\mathcal{F}$ is **bounded** at some $z \in A$ if there is M so that $|f(z)| \leq M$ for every $f \in \mathcal{F}$. We say that $\mathcal{F}$ is **equicontinuous** at some $z \in A$ if for every $\epsilon > 0$ there is $\delta > 0$ so that $|f(w) - f(z)| < \epsilon$ for every $w \in A$ with $|w - z| < \delta$ and for every $f \in \mathcal{F}$.

We observe that if $\mathcal{F}$ is equicontinuous at some $z \in A$, then every $f \in \mathcal{F}$ is continuous at z and that the δ which corresponds to ϵ in the definition of continuity at z does not depend on the particular f, i.e. δ is *uniform* over $f \in \mathcal{F}$.

Let $A \subseteq \mathbb{C}$ and $\mathcal{F}$ be a collection of functions defined in the set A. We say that $\mathcal{F}$ is **locally bounded** at some $z \in A$ if there are $\delta > 0$ and M so that $|f(w)| \leq M$ for every $w \in A$ with $|w - z| < \delta$ and for every $f \in \mathcal{F}$.

The theorem of Montel. *Let $\Omega \subseteq \mathbb{C}$ be open and $\mathcal{F}$ be a collection of holomorphic functions in Ω. Then the following are equivalent:*
(i) For every sequence (f_n) in $\mathcal{F}$ there is a subsequence (f_{n_k}) and a function f holomorphic in Ω so that $f_{n_k} \to f$ uniformly in every compact subset of Ω.
(ii) $\mathcal{F}$ is locally bounded at every $z \in \Omega$.

Proof. (i) ⇒ (ii) Assume that $\mathcal{F}$ is not locally bounded at some $z \in \Omega$. Then for every $n \in \mathbb{N}$ there is $z_n \in \Omega$ and $f_n \in \mathcal{F}$ with

$$|z_n - z| < \tfrac{1}{n}, \quad |f_n(z_n)| > n.$$

Now, there is a subsequence (f_{n_k}) of (f_n) and a function f holomorphic in Ω so that $f_{n_k} \to f$ uniformly in every compact subset of Ω. Since $z_n \to z$, the set $K = \{z_n \,|\, n \in \mathbb{N}\} \cup \{z\}$ is a compact subset of Ω and hence $f_{n_k} \to f$ uniformly in K. Moreover, the continuity of f implies that $f(z_{n_k}) \to f(z)$. But then

$$\|f_{n_k} - f\|_K \geq |f_{n_k}(z_{n_k}) - f(z_{n_k})| \geq |f_{n_k}(z_{n_k})| - |f(z_{n_k})| \to +\infty$$

and we arrive at a contradiction.
Another course goes as follows. By the Arzela-Ascoli theorem, (i) implies that $\mathcal{F}$ is bounded and equicontinuous at every $z \in \Omega$. This easily implies that $\mathcal{F}$ is locally bounded at every $z \in \Omega$. Indeed, there is M so that $|f(z)| \leq M$ for every $f \in \mathcal{F}$. Moreover, there is $\delta > 0$ so that $|f(w) - f(z)| < 1$ for every $w \in A$ with $|w - z| < \delta$ and for every $f \in \mathcal{F}$. Hence

$$|f(w)| \leq |f(z)| + |f(w) - f(z)| \leq M + 1$$

for every $w \in A$ with $|w - z| < \delta$ and for every $f \in \mathcal{F}$. So $\mathcal{F}$ is locally bounded at every $z \in \Omega$.
(ii) ⇒ (i) By the Arzela-Ascoli theorem and by the theorem of Weierstrass it is enough to prove that $\mathcal{F}$ is bounded and equicontinuous at every $z \in \Omega$.
It is clear that local boundedness of $\mathcal{F}$ at every $z \in \Omega$ implies that $\mathcal{F}$ is bounded at every $z \in \Omega$. Now we take any $z \in \Omega$ and then there is $r > 0$ and M so that $|f(z)| \leq M$ for every $z \in \overline{D}_z(r)$ and every $f \in \mathcal{F}$. Thus, for every $w \in D_z(\frac{r}{2})$ and every $f \in \mathcal{F}$ we have

$$|f'(w)| = \big|\tfrac{1}{2\pi i}\oint_{C_z(r)} \tfrac{f(\zeta)}{(\zeta - w)^2}\,d\zeta\big| \leq \tfrac{1}{2\pi}\tfrac{M}{(r/2)^2}\,2\pi r = \tfrac{4M}{r}.$$

This implies that for every $w \in D_z(\frac{r}{2})$ and every $f \in \mathcal{F}$ we have

$$|f(w) - f(z)| = \big| \int_{[z,w]} f'(\zeta)\, d\zeta \big| \le \tfrac{4M}{r}\, |w - z|.$$

Hence, for every $\epsilon > 0$ we may take $\delta = \min\{\frac{r\epsilon}{4M}, \frac{r}{2}\}$ and then for every $z \in D_z(\delta)$ and every $f \in \mathcal{F}$ we get

$$|f(w) - f(z)| \le \tfrac{4M}{r}\, |w - z| < \tfrac{4M}{r}\, \delta \le \epsilon.$$

Thus, $\mathcal{F}$ is equicontinuous at z. $\square$

Exercises.

5.12.1. Prove that $\sum_{n=-\infty}^{+\infty} \frac{1}{(z+n)^2}$ converges uniformly in every compact subset of $\mathbb{C} \setminus \mathbb{Z}$.

5.12.2. Prove that $\sum_{n=0}^{+\infty} \frac{z^n}{z^{2n}+1}$ converges uniformly in every compact subset of $\mathbb{C} \setminus \mathbb{T}$.

5.12.3. Prove that $\sum_{n=0}^{+\infty} (\frac{z}{z+1})^n$ converges uniformly in every compact subset of $\{z \mid \operatorname{Re} z > -\frac{1}{2}\}$.

5.12.4. We define $t^z = e^{z \ln t}$ for every $z \in \mathbb{C}$ and every $t > 0$.
(i) Prove that $\sum_{n=1}^{+\infty} \frac{1}{n^z}$ converges absolutely for every z in $\{z \mid \operatorname{Re} z > 1\}$ and diverges for every z in $\{z \mid \operatorname{Re} z \le 1\}$.
(ii) Let $\delta > 0$. Prove that $\sum_{n=1}^{+\infty} \frac{1}{n^z}$ converges uniformly in $\{z \mid \operatorname{Re} z \ge 1 + \delta\}$.
The function $\zeta : \{z \mid \operatorname{Re} z > 1\} \to \mathbb{C}$ with

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

for every z with $\operatorname{Re} z > 1$ is the famous **ζ function** of Riemann and it is connected with one of the most difficult unsolved problems of mathematics.
(iii) Prove that

$$\zeta'(z) = -\sum_{n=2}^{+\infty} \frac{\ln n}{n^z}$$

for every z with $\operatorname{Re} z > 1$.

5.12.5. Let (f_n) be a sequence of functions holomorphic in the region $\Omega \subseteq \mathbb{C}$ which is locally bounded at every $z \in \Omega$. If every f_n has no roots in Ω and $f_n(z_0) \to 0$ for some $z_0 \in \Omega$, prove that $f_n \to 0$ uniformly in every compact subset of Ω.

5.12.6. Let (f_n) be a sequence of functions holomorphic in the region $\Omega \subseteq \mathbb{C}$ which is locally bounded at every $z \in \Omega$ and let $E \subseteq \Omega$ have an accumulation point in Ω. If $\lim_{n\to+\infty} f_n(z)$ is a complex number for every $z \in E$, prove that (f_n) converges to some function uniformly in every compact subset of Ω.

5.12.7. Let (f_n) be a sequence of functions holomorphic in the open set $\Omega \subseteq \mathbb{C}$. If $\lim_{n\to+\infty} f_n(z)$ is a complex number for every $z \in \Omega$, use the theorem of Baire to prove that there is an open set $H \subseteq \Omega$ which is dense in Ω and so that (f_n) converges to some function uniformly in every compact subset of H.

5.12.8. Let $\Omega \subseteq \mathbb{C}$ be a region and (f_n) be a sequence of functions holomorphic in Ω with $\operatorname{Re} f_n > 0$ in Ω for every n.
(i) If $(f_n(z_0))$ is bounded for some $z_0 \in \Omega$, prove that there is a subsequence (f_{n_k}) which converges to some function uniformly in every compact subset of Ω.
(ii) If $(f_n(z_0))$ is unbounded for some $z_0 \in \Omega$, prove that there is a subsequence (f_{n_k}) so that $f_{n_k} \to \infty$ uniformly in every compact subset of Ω.

5.12.9. Let f_n, f be holomorphic in $D_{z_0}(R)$ and $f_n \to f$ uniformly in every compact subset of $D_{z_0}(R)$. If $f_n(z) = \sum_{k=0}^{+\infty} a_{k,n}(z - z_0)^k$ and $f(z) = \sum_{k=0}^{+\infty} a_k(z - z_0)^k$ are the corresponding Taylor series, prove that $a_{k,n} \to a_k$ for every k.

5.12.10. Let $\mathcal{F}$ be a collection of functions holomorphic in $D_{z_0}(R)$. We denote $a_k(f) = \frac{f^{(k)}(z_0)}{k!}$ the k-th Taylor coefficient of each $f \in \mathcal{F}$. Prove that the following are equivalent:
(i) For every sequence (f_n) in $\mathcal{F}$ there is a subsequence (f_{n_j}) which converges to some function uniformly in every compact subset of $D_{z_0}(R)$.
(ii) There are M_k so that $\overline{\lim}\sqrt[k]{M_k} \leq \frac{1}{R}$ and $|a_k(f)| \leq M_k$ for every k and every $f \in \mathcal{F}$.

5.12.11. A theorem of Montel. Let $-\infty < a < x_0 < b < +\infty$ and f be bounded and holomorphic in the vertical zone

$$\Omega = \{z = x + iy \,|\, a < x < b, y > 0\}.$$

If $\lim_{y\to+\infty} f(x_0 + iy) = A \in \mathbb{C}$, prove that for every $\epsilon > 0$ we have

$$\lim_{y\to+\infty} \sup_{x\in[a+\epsilon,b-\epsilon]} |f(x + iy) - A| = 0.$$

5.12.12. Let $\Omega \subseteq \mathbb{C}$ be open, $M \geq 0$, $1 \leq p < +\infty$, and $\mathcal{F}$ be the collection of all functions f holomorphic in Ω with

$$\iint_\Omega |f(z)|^p \, dxdy \leq M \qquad (z = x + iy).$$

Using exercise 5.3.9, prove that $\mathcal{F}$ is locally bounded at every $z \in \Omega$.

5.12.13. Let $\mathcal{F}$ be a collection of holomorphic functions in the open set $\Omega \subseteq \mathbb{C}$ with the property: for every sequence (f_n) in $\mathcal{F}$ there is a subsequence (f_{n_k}) which converges to some function uniformly in every compact subset of Ω. Prove that the collection $\mathcal{F}' = \{f' \,|\, f \in \mathcal{F}\}$ has the same property. Is the converse true?

5.12.14. Let $\Omega \subseteq \mathbb{C}$ be open, $\overline{D}_{z_0}(r) \subseteq \Omega$, f_n, f be holomorphic in Ω and $f_n \to f$ uniformly in $C_{z_0}(r)$. If f has no root in $C_{z_0}(r)$ and has exactly k roots in $D_{z_0}(r)$, prove that every f_n, after some value of the index n, has exactly k roots in $D_{z_0}(r)$.

5.12.15. Let (f_n) be a sequence of holomorphic functions in the region $\Omega \subseteq \mathbb{C}$ so that $f_n \to f$ uniformly in every compact subset of Ω. If every f_n has at most k roots in Ω, prove that either f has also at most k roots in Ω or that f is identically 0 in Ω.

5.12.16. Let f_n, f be holomorphic in the open set $\Omega \subseteq \mathbb{C}$ and $f_n \to f$ uniformly in every compact subset of Ω. Prove that

$$\{z \in \Omega \,|\, f(z) = 0\} = \Omega \cap \bigcap_{n=1}^{+\infty} \left(\overline{\bigcup_{k=n}^{+\infty} \{z \in \Omega \,|\, f_k(z) = 0\}}\right).$$

Chapter 6

Global behaviour of holomorphic functions.

6.1 Index of a closed curve with respect to a point.

6.1.1 The piecewise smooth case

Let $A \subseteq \mathbb{C}$, and $g : A \to \mathbb{C} \setminus \{0\}$ be continuous in A. We say that the function h is a **continuous branch of arg g** in A if h is continuous in A and for every $w \in A$ we have that $h(w)$ is an element of arg $g(w)$ or, equivalently,

$$g(w) = |g(w)|e^{ih(w)}$$

for every $w \in A$.

We recall the notion of a continuous branch of $\log g$. We say that f is a continuous branch of $\log g$ if f is continuous in A and $f(w)$ is an element of $\log g(w)$ or, equivalently,

$$e^{f(w)} = g(w)$$

for every $w \in A$.

Proposition 6.1. *Let $A \subseteq \mathbb{C}$ and $g : A \to \mathbb{C} \setminus \{0\}$ be continuous in A. Then there is a one-to-one correspondence between continuous branches of* $\log g$ *and continuous branches of* $\arg g$ *in A.*

Proof. If h is a continuous branch of $\arg g$ in A, then the function

$$f = \ln|g| + ih \tag{6.1}$$

is a continuous branch of $\log g$ in A. Indeed, $e^{f(w)} = e^{\ln|g(w)|}e^{ih(w)} = |g(w)|e^{ih(w)} = g(w)$ for every $w \in A$ and f is continuous in A.
Conversely, if f is a continuous branch of $\log g$ in A, then h, defined through (6.1), is a continuous branch of $\arg g$ in A. Indeed, $|g(w)|e^{ih(w)} = |g(w)|e^{f(w)}e^{-\ln|g(w)|} = g(w)$ for every $w \in A$ and h is continuous in A. □

In other words, relation (6.1) says that, if we have a continuous branch f of $\log g$ in A, then the imaginary part h of f is a continuous branch of $\arg g$ in A. Conversely, if we have a continuous branch h of $\arg g$ in A, then the function f with imaginary part h and real part $\ln|g|$ is a continuous branch of $\log g$ in A.

The next result is analogous to proposition 4.6 and their proofs are almost identical.

Proposition 6.2. *Let $g : A \to \mathbb{C} \setminus \{0\}$ be continuous in $A \subseteq \mathbb{C}$.*
(i) If h_1 is a continuous branch of $\arg g$ *in A and $h_2 - h_1 = k2\pi$ in A, where k is a fixed integer, then h_2 is also a continuous branch of* $\arg g$ *in A.*
(ii) If, moreover, A is connected and h_1, h_2 are continuous branches of $\arg g$ *in A, then $h_2 - h_1 = k2\pi$ in A, where k is a fixed integer. In particular, if $h_1(w_0) = h_2(w_0)$ for some $w_0 \in A$, then $h_1 = h_2$ in A.*

Now we consider a piecewise smooth curve γ (not necessarily closed). Then there is a succession of points $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$ so that γ is continuously differentiable in every $[t_{k-1}, t_k]$. We consider an arbitrary fixed $z \notin \gamma^*$ and we define

$$f(t) = \int_a^t \frac{\gamma'(s)}{\gamma(s)-z}\, ds$$

for $t \in [a, b]$. Then f is continuous in $[a, b]$ and differentiable at every point of continuity of $\frac{\gamma'}{\gamma - z}$. So in every (t_{k-1}, t_k) we have $\frac{d}{dt}\big((\gamma(t) - z)e^{-f(t)}\big) = \gamma'(t)e^{-f(t)} - (\gamma(t) - z)f'(t)e^{-f(t)} = 0$. Thus, $(\gamma(t) - z)e^{-f(t)}$ is constant in each (t_{k-1}, t_k) with a constant value which *a priori* depends on k, but since this function is continuous in $[a, b]$, it is constant in $[a, b]$. Hence there is $c \in \mathbb{C}$ so that $(\gamma(t) - z)e^{-f(t)} = c$ for every $t \in [a, b]$. Since $c \neq 0$, there is $d \in \mathbb{C}$ so that $e^d = c$, and thus $e^{f(t)+d} = \gamma(t) - z$ for every $t \in [a, b]$. Now we redefine f by adding to it the constant d, i.e. we write

$$f(t) = \int_a^t \frac{\gamma'(s)}{\gamma(s)-z}\, ds + d \tag{6.2}$$

for every $t \in [a, b]$ and we have

$$e^{f(t)} = \gamma(t) - z$$

for every $t \in [a, b]$. In other words, the function f is a *continuous branch* of $\log(\gamma - z)$ in $[a, b]$. Now, the real part of f is $\ln |\gamma - z|$ and, if we denote h the imaginary part of f, i.e.

$$h(t) = \operatorname{Im} \int_a^t \frac{\gamma'(s)}{\gamma(s)-z}\, ds + \operatorname{Im} d \tag{6.3}$$

for every $t \in [a, b]$, then h is a continuous branch of $\arg(\gamma - z)$ in $[a, b]$. Loosely speaking, $h(t)$ is a continuously varying angle of the continuously varying vector $\overrightarrow{z\,\gamma(t)}$, as this vector turns around its fixed base point z following its variable tip $\gamma(t)$ which moves over the trajectory of the curve γ from its initial point $\gamma(a)$ towards its final point $\gamma(b)$. This is the reason why the expression

$$h(b) - h(a) = \operatorname{Im} \int_a^b \frac{\gamma'(s)}{\gamma(s)-z}\, ds = \operatorname{Im} \int_\gamma \frac{1}{\zeta - z}\, d\zeta,$$

a consequence of (6.3), is called **total increment of argument** or **total increment of angle** over the curve γ with respect to z.

Let us consider the important special case of a *closed* curve γ, i.e. when $\gamma(b) = \gamma(a)$. This implies $\gamma(b) - z = \gamma(a) - z$, and hence

$$\operatorname{Re} f(b) = \ln |\gamma(b) - z| = \ln |\gamma(a) - z| = \operatorname{Re} f(a). \tag{6.4}$$

It also implies that $h(b) - h(a)$ is an integer multiple of 2π: indeed, both $h(b), h(a)$ are values of $\arg(\gamma(b) - z) = \arg(\gamma(a) - z)$. Then the integer

$$n(\gamma; z) = \frac{h(b)-h(a)}{2\pi} \tag{6.5}$$

is called **rotation number** or **index** of γ with respect to z_0. It represents the number of complete rotations of the continuously varying vector $\overrightarrow{z\,\gamma(t)}$ as $\gamma(t)$ moves over the trajectory of the curve from its initial point towards its final point.

If we recall that h is the imaginary part of f, then (6.2), (6.4) and (6.5) give

$$n(\gamma; z) = \frac{1}{2\pi i} \int_a^b \frac{\gamma'(s)}{\gamma(s)-z}\, ds = \frac{1}{2\pi i} \oint_\gamma \frac{1}{\zeta - z}\, d\zeta. \tag{6.6}$$

Let γ be a closed piecewise smooth curve and $z \notin \gamma^*$. We say that γ **surrounds** z if

$$n(\gamma; z) \neq 0.$$

Example 6.1.1. We take $n \in \mathbb{Z}$ and the closed curve γ with parametric equation $\gamma(t) = z_0 + re^{int}$, $t \in [0, 2\pi]$. It is visually clear that, if $n \neq 0$ and t increases in the interval $[0, 2\pi]$, then $\gamma(t)$ describes $|n|$ times the circle $C_{z_0}(r)$ in the positive direction, if $n > 0$, and in the negative direction, if $n < 0$. In the case $n = 0$, then $\gamma(t)$ is constant and describes $|n| = 0$ times the circle $C_{z_0}(r)$. All these agree with the result of the calculation:

$$n(\gamma; z_0) = \tfrac{1}{2\pi i} \oint_\gamma \tfrac{1}{\zeta - z_0}\, d\zeta = \tfrac{1}{2\pi i} \int_0^{2\pi} \tfrac{1}{re^{int}}\, rine^{int}\, dt = n.$$

The next three propositions are trivial consequences of (6.6) and of basic properties of curvilinear integrals.

Proposition 6.3. *Let γ_1, γ_2 be closed piecewise smooth curves with the same endpoints and $z \notin \gamma_1^*$, $z \notin \gamma_2^*$. Then $\gamma_1 \dot{+} \gamma_2$ is defined and it is also a closed piecewise smooth curve and $z \notin (\gamma_1 \dot{+} \gamma_2)^*$ and*

$$n(\gamma_1 \dot{+} \gamma_2; z) = n(\gamma_1; z) + n(\gamma_2; z).$$

Proposition 6.4. *Let γ_1, γ_2 be closed piecewise smooth curves with $z \notin \gamma_1^*$, $z \notin \gamma_2^*$, so that each is a reparametrization of the other. Then*

$$n(\gamma_2; z) = n(\gamma_1; z).$$

Proposition 6.5. *Let γ be a closed piecewise smooth curve and $z \notin \gamma^*$. Then*

$$n(\neg\, \gamma; z) = -n(\gamma; z).$$

Let B be a bounded set and $A = \mathbb{C} \setminus B$. We know that A is equal to the union of its distinct (and hence mutually disjoint) connected components. Since B is bounded, it is contained in some disc $\overline{D}_0(R)$. Then the connected ring $D_0(R, +\infty)$ is contained in A and hence it is contained in (exactly) one of the connected components, call it M, of A. All other connected components of A are disjoint from M and so they are contained in $\overline{D}_0(R)$. Therefore, M is the unbounded connected component of A and all other connected components of A are bounded.

In particular if γ is a curve, then the open set $\mathbb{C} \setminus \gamma^*$ has one unbounded connected component and all its other connected components are bounded.

Proposition 6.6. *Let γ be a closed piecewise smooth curve. Then the integer-valued function $n(\gamma; z)$ is constant in every connected component of the open set $\mathbb{C} \setminus \gamma^*$. We also have that $n(\gamma; z) = 0$ for every z in the unbounded connected component of $\mathbb{C} \setminus \gamma^*$.*

Proof. Proposition 4.12 implies that $n(\gamma; z)$, as given by (6.6), is a holomorphic function of z in $\mathbb{C} \setminus \gamma^*$. Now, let Ω be any connected component of $\mathbb{C} \setminus \gamma^*$. The function $n(\gamma; z)$ is continuous and integer valued in Ω and, since $n(\gamma; z)$ has the intermediate value property in Ω, it has to be constant in Ω.
Finally, let Ω be the unbounded connected component of $\mathbb{C} \setminus \gamma^*$. We shall prove that $n(\gamma; z) = 0$ for every $z \in \Omega$. If $\gamma^* \subseteq \overline{D}_0(R)$, then (6.6) for $|z| > R$ implies $|n(\gamma; z)| \leq \frac{1}{2\pi} \frac{l(\gamma)}{|z| - R}$. Thus, $\lim_{z \to \infty} n(\gamma; z) = 0$ and since $n(\gamma; z)$ is constant in Ω, it has to be equal to 0 in Ω. $\square$

Proposition 6.6 says that *if z_1, z_2 are in the same connected component of the complement of the trajectory of the closed piecewise smooth curve γ, then the number of complete rotations of γ around z_1 is equal to the number of complete rotations of γ around z_2.*

Example 6.1.2. We consider the same closed curve as in example 6.1.1.
We have seen that $n(\gamma; z_0) = n$. Since $\gamma^* = C_{z_0}(r)$, the complement of γ^* has two connected components: the disc $D_{z_0}(r)$ and the unbounded ring $D_{z_0}(r, +\infty)$. Thus, $n(\gamma; z) = n(\gamma; z_0) = n$ when $z \in D_{z_0}(r)$. Also, $n(\gamma; z) = 0$ when $z \in D_{z_0}(r, +\infty)$.

Proposition 6.7. *Let Ω be a region and $z \notin \Omega$. A holomorphic branch of* $\log(\zeta - z)$ *(as a function of ζ) exists in Ω if and only if $n(\gamma; z) = 0$ for every closed piecewise smooth curve γ in Ω.*

Proof. A direct consequence of theorem 4.1 applied to $g(\zeta) = \zeta - z$. □

Example 6.1.3. We consider the region $\Omega = \mathbb{C} \setminus l$, where l is any halfline with vertex z. We know that a holomorphic branch of $\log(\zeta - z)$ exists in Ω and hence $n(\gamma; z) = 0$ for every closed curve γ in Ω. This is geometrically obvious: *since γ is in Ω, it does not intersect the halfline l with vertex z, and hence it cannot make any complete rotation around z.*

Cauchy's formula for derivatives and closed curves in convex regions. *If f is holomorphic in the convex region Ω and γ is a closed piecewise smooth curve in Ω, then for all $n \in \mathbb{N}_0$ we have*

$$n(\gamma; z) f^{(n)}(z) = \frac{n!}{2\pi i} \oint_\gamma \frac{f(\zeta)}{(\zeta - z)^{n+1}} \, d\zeta$$

for every $z \in \Omega \setminus \gamma^*$.

Proof. The function $F(\zeta) = \frac{f(\zeta) - f(z)}{\zeta - z}$ is holomorphic in $\Omega \setminus \{z\}$. Since z is a root of $f(\zeta) - f(z)$, the singularity z of F is removable. So we define F at z as $F(z) = \lim_{\zeta \to z} \frac{f(\zeta) - f(z)}{\zeta - z} = f'(z)$ and then F becomes holomorphic in Ω. Now we apply the theorem of Cauchy in convex regions and get

$$\oint_\gamma \frac{f(\zeta) - f(z)}{\zeta - z} \, d\zeta = \oint_\gamma F(\zeta) \, d\zeta = 0$$

for every $z \in \Omega \setminus \gamma^*$. This implies

$$\frac{1}{2\pi i} \oint_\gamma \frac{f(\zeta)}{\zeta - z} \, d\zeta = f(z) \frac{1}{2\pi i} \oint_\gamma \frac{1}{\zeta - z} \, d\zeta = f(z) n(\gamma; z) \tag{6.7}$$

for every $z \in \Omega \setminus \gamma^*$. This is the result of the statement in the case $n = 0$.
Now, if $z \in \Omega \setminus \gamma^*$, then z is contained in one connected component of $\mathbb{C} \setminus \gamma^*$ and, since all connected components of $\mathbb{C} \setminus \gamma^*$ are open, there is a small disc $D_z(r)$ which is contained in one connected component of $\mathbb{C} \setminus \gamma^*$. Therefore, the index $n(\gamma; w)$ is a constant function of w in $D_z(r)$, i.e. $n(\gamma; w) = n(\gamma; z)$ for every $w \in D_z(r)$. This implies that all derivatives of $n(\gamma; w)$ vanish at z and so when we differentiate (6.7) we get $\frac{n!}{2\pi i} \oint_\gamma \frac{f(\zeta)}{(\zeta - z)^{n+1}} \, d\zeta = f^{(n)}(z) n(\gamma; z)$ for $n \geq 1$. □

A particular instance of the last result is Cauchy's formula for derivatives and circles. Indeed, when the curve γ describes the circle $C_{z_0}(R)$ once in the positive direction we have $n(\gamma; z) = 1$ for all $z \in D_{z_0}(R)$. We originally proved the result in the case of a circle, using corollary 5.1. We now have a "new" proof using that z is a removable singularity of $\frac{f(\zeta) - f(z)}{\zeta - z}$. We have also introduced the notion of the index of a closed curve. This new proof together with the introduction of the notion of index allows us to generalize the case of a circle to the case of a more general closed piecewise smooth curve. There is still a restriction in the sense that the curve has to be contained in a *convex* region in which the function is holomorphic. This is because our proof is based on Cauchy's theorem in *convex* regions. In this chapter we shall replace this restriction on the region with a restriction on the curve.

6.1.2 The general case

In the general case of a curve γ, which is not necessarily piecewise smooth, the notions of *total increment of argument* over γ and of *index* of γ cannot be based on integrals of the form $\int_\gamma \frac{1}{\zeta - z} \, d\zeta$ any more.

Proposition 6.8. *Let $g_1, g_2 : A \to \mathbb{C} \setminus \{0\}$ be continuous in $A \subseteq \mathbb{C}$.*
(i) If f_1, f_2 are continuous branches of $\log g_1$, $\log g_2$ *in A, then $f_1 + f_2$ is a continuous branch of* $\log(g_1 g_2)$ *in A.*
(ii) If h_1, h_2 are continuous branches of $\arg g_1$, $\arg g_2$ *in A, then $h_1 + h_2$ is a continuous branch of* $\arg(g_1 g_2)$ *in A.*

Proof. (i) $f_1 + f_2$ is continuous in A and also $e^{f_1(w)+f_2(w)} = e^{f_1(w)}e^{f_2(w)} = g_1(w)g_2(w)$ for every $w \in A$.
(ii) Just as in (i). $\square$

Proposition 6.9 is the first existence result of this section.

Proposition 6.9. *Let $g : [a,b] \to \mathbb{C} \setminus \{0\}$ be continuous in the interval $[a,b]$. Then there is a continuous branch of $\log g$ and a continuous branch of $\arg g$ in $[a,b]$.*

Proof. It is enough to prove the existence of a continuous branch of $\log g$.
Since g is continuous in $[a,b]$, there is $\epsilon > 0$ so that $|g(t)| \geq \epsilon$ for every $t \in [a,b]$. Now, g is also uniformly continuous in $[a,b]$ and hence there is $\delta > 0$ so that $|g(t') - g(t'')| < \epsilon$ for every $t', t'' \in [a,b]$ with $|t' - t''| < \delta$. We take successive points $a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b$ so that $t_k - t_{k-1} < \delta$ for every $k = 1, \ldots, n$. Then, for every $k = 1, \ldots, n$ we have

$$\{g(t) \,|\, t \in [t_{k-1}, t_k]\} \subseteq D_{g(t_k)}(\epsilon).$$

Since $|g(t_k)| \geq \epsilon$, the disc $D_{g(t_k)}(\epsilon)$ does not contain 0 and hence a continuous branch of log is defined in this disc. Then example 4.3.3 implies that there is a continuous branch, say f_k, of $\log g$ in $[t_{k-1}, t_k]$.
Now, f_1 is a continuous branch of $\log g$ in $[t_0, t_1]$ and f_2 is a continuous branch of $\log g$ in $[t_1, t_2]$. Then $f_2(t_1) - f_1(t_1) = m2\pi i$ for some $m \in \mathbb{Z}$. We replace the function f_2 with the function $f_2 - m2\pi i$ and the new function f_2 is also a continuous branch of $\log g$ in $[t_1, t_2]$ with $f_2(t_1) = f_1(t_1)$. Working with the (new) function f_2 and the function f_3 which is a continuous branch of $\log g$ in $[t_2, t_3]$, we see as before that $f_3(t_2) - f_2(t_2) = l2\pi i$ for some $l \in \mathbb{Z}$. We replace the function f_3 with the function $f_3 - l2\pi i$ and the new function f_3 is also a continuous branch of $\log g$ in $[t_2, t_3]$ with $f_3(t_2) = f_2(t_2)$. We continue inductively and finally we end up with continuous branches f_k of $\log g$ in $[t_{k-1}, t_k]$ for every $k = 1, \ldots, n$ so that $f_k(t_k) = f_{k+1}(t_k)$ for every $k = 1, \ldots, n-1$. Therefore, the function $f : [a,b] \to \mathbb{C}$, which is defined to be equal to f_k in the corresponding interval $[t_{k-1}, t_k]$, is continuous in $[a,b]$. Moreover, f is a continuous branch of $\log g$ in every $[t_{k-1}, t_k]$ and hence in $[a,b]$. $\square$

We consider any curve $\gamma : [a,b] \to \mathbb{C}$ and $z \notin \gamma^*$. Then the function $\gamma - z : [a,b] \to \mathbb{C} \setminus \{0\}$ is continuous in $[a,b]$ and, according to proposition 6.9, there is a continuous branch f of $\log(\gamma - z)$ and a continuous branch h of $\arg(\gamma - z)$ in $[a,b]$ related by

$$f = \ln|\gamma - z| + ih. \tag{6.8}$$

Then the functions $f + k2\pi i$ and $h + k2\pi$, where k is an arbitrary, but constant, integer, are also continuous branches of $\log(\gamma - z)$ and $\arg(\gamma - z)$ in $[a,b]$. Moreover, since $[a,b]$ is connected, these are all the continuous branches of $\log(\gamma - z)$ and $\arg(\gamma - z)$ in $[a,b]$.

Now, let h be any continuous branch of $\arg(\gamma - z)$ in $[a,b]$. We observe that the expression $h(b) - h(a)$ is independent of the particular choice of h. Indeed, if h_1 is another continuous branch of $\arg(\gamma - z)$ in $[a,b]$, then there is a constant integer k so that $h_1 = h + k2\pi$ in $[a,b]$ and then we have $h_1(b) - h_1(a) = (h(b) + k2\pi) - (h(a) + k2\pi) = h(b) - h(a)$. The expression

$$\Delta \arg(\gamma - z) = h(b) - h(a)$$

is called **total increment of argument** or **total increment of angle** over the curve γ with respect to z.

Observe that in the previous subsection, i.e. when γ is piecewise smooth, we had a specific construction of a continuous branch h of $\arg(\gamma - z)$ in $[a,b]$ and $\Delta \arg(\gamma - z)$ was given by means of a curvilinear integral: $\mathrm{Im} \int_\gamma \frac{1}{\zeta - z}\, d\zeta$.

Now, assume that γ is closed, i.e. $\gamma(b) = \gamma(a)$. This implies $\gamma(b) - z = \gamma(a) - z$, and hence $\ln|\gamma(b) - z| = \ln|\gamma(a) - z|$. It also implies that $h(b)$ and $h(a)$ differ by some integer multiple of

2π, since both $h(b), h(a)$ are values of $\arg(\gamma(b) - z) = \arg(\gamma(a) - z)$. Therefore the expresion $\Delta \arg(\gamma - z) = h(b) - h(a)$ is an integer multiple of 2π. Then the integer

$$n(\gamma; z) = \frac{\Delta \arg(\gamma - z)}{2\pi}$$

is called **rotation number** or **index** of γ with respect to z.

Again, we remark that when the closed curve γ is piecewise smooth we have from the previous subsection an expression of $n(\gamma, z)$, namely (6.6), by means of a curvilinear integral.

Proposition 6.10. *Let γ_1, γ_2 be closed curves with the same endpoints and $z \notin \gamma_1^*$, $z \notin \gamma_2^*$. Then $\gamma_1 \dot{+} \gamma_2$ is defined and it is also a closed curve and $z \notin (\gamma_1 \dot{+} \gamma_2)^*$ and*

$$n(\gamma_1 \dot{+} \gamma_2; z) = n(\gamma_1; z) + n(\gamma_2; z).$$

Proof. Let $\gamma_1 : [a, b] \to \mathbb{C}$ and $\gamma_2 : [b, c] \to \mathbb{C}$ be the two curves and $h_1 : [a, b] \to \mathbb{R}$ and $h_2 : [b, c] \to \mathbb{R}$ be continuous branches of $\arg(\gamma_1 - z)$ and $\arg(\gamma_2 - z)$. We may redefine h_2 by adding to it an appropriate integer multiple of 2π so that $h_2(b) = h_1(b)$. Then the function $h : [a, c] \to \mathbb{R}$ which equals h_1 in $[a, b]$ and h_2 in $[b, c]$ is a continuous branch of $\log((\gamma_1 \dot{+} \gamma_2) - z)$ in $[a, c]$. Therefore, $h(c) - h(a) = h(c) - h(b) + h(b) - h(a) = h_2(c) - h_2(b) + h_1(b) - h_1(a)$ and hence $n(\gamma_1 \dot{+} \gamma_2; z) = n(\gamma_1; z) + n(\gamma_2; z)$. □

Proposition 6.11. *Let γ_1, γ_2 be closed curves with $z \notin \gamma_1^*$, $z \notin \gamma_2^*$, so that each is a reparametrization of the other. Then*

$$n(\gamma_2; z) = n(\gamma_1; z).$$

Proof. Let $\gamma_1 : [a, b] \to \mathbb{C}$ and $\gamma_2 : [c, d] \to \mathbb{C}$ be the two curves and $\sigma : [c, d] \to [a, b]$ be the change of parameter so that $\gamma_2 = \gamma_1 \circ \sigma$. If h is a continuous branch of $\arg(\gamma_1 - z)$ in $[a, b]$, then $h \circ \sigma$ is a continuous branch of $\arg(\gamma_2 - z)$ in $[c, d]$. Indeed, from $e^{ih(t)} = \frac{\gamma_1(t) - z}{|\gamma_1(t) - z|}$ for every $t \in [a, b]$ we get $e^{ih(\sigma(s))} = \frac{\gamma_1(\sigma(s)) - z}{|\gamma_1(\sigma(s)) - z|} = \frac{\gamma_2(s) - z}{|\gamma_2(s) - z|}$ for every $s \in [c, d]$.
So from $h(\sigma(d)) - h(\sigma(c)) = h(b) - h(a)$ we get $n(\gamma_2; z) = n(\gamma_1; z)$. □

Proposition 6.12. *Let γ be a closed curve and $z \notin \gamma^*$. Then*

$$n(\neg\, \gamma; z) = -n(\gamma; z).$$

Proof. Let $\gamma, \neg\, \gamma : [a, b] \to \mathbb{C}$ be the two curves. Then $\neg\, \gamma(t) = \gamma(a + b - t)$ for every $t \in [a, b]$. If h is a continuous branch of $\arg(\gamma - z)$ in $[a, b]$, then the function $k(t) = h(a + b - t)$ is a continuous branch of $\arg(\neg\, \gamma - z)$ in $[a, b]$. Indeed, from $e^{ih(t)} = \frac{\gamma(t) - z}{|\gamma(t) - z|}$ for every $t \in [a, b]$ we get $e^{ik(t)} = e^{ih(a+b-t)} = \frac{\gamma(a+b-t) - z}{|\gamma(a+b-t) - z|} = \frac{\neg\, \gamma(t) - z}{|\neg\, \gamma(t) - z|}$ for every $t \in [a, b]$.
So from $k(b) - k(a) = h(a) - h(b)$ we get $n(\neg\, \gamma; z) = -n(\gamma; z)$. □

Proposition 6.13. *Let $A \subseteq \mathbb{C}$ and $z \notin A$. If a continuous branch of* $\log(\zeta - z)$ *(as a function of ζ) exists in A then $n(\gamma; z) = 0$ for every closed curve γ in A.*

Proof. Let $\phi(\zeta)$ be a continuous branch of $\log(\zeta - z)$ in A and $\gamma : [a, b] \to A$ be a closed curve with $z \notin \gamma^*$. Then $h = \phi \circ \gamma$ is a continuous branch of $\log(\gamma - z)$ in $[a, b]$. Indeed, for every $t \in [a, b]$ we have $e^{ih(t)} = e^{i\phi(\gamma(t))} = \frac{\gamma(t) - z}{|\gamma(t) - z|}$.
Now, $\gamma(b) = \gamma(a)$ implies $h(b) = \phi(\gamma(b)) = \phi(\gamma(a)) = h(a)$ and $n(\gamma; z) = \frac{h(b) - h(a)}{2\pi} = 0$. □

Example 6.1.4. We consider the set $A = \mathbb{C} \setminus l$, where l is any halfline with vertex z. We know that a continuous branch of $\log(\zeta - z)$ exists in A and hence $n(\gamma; z) = 0$ for every closed curve γ in A.

Proposition 6.14. *Let $\gamma_1, \gamma_2 : [a, b] \to \mathbb{C}$ be closed curves such that $|\gamma_1(t) - \gamma_2(t)| < |\gamma_2(t) - z|$ for every $t \in [a, b]$. Then*

$$n(\gamma_1; z) = n(\gamma_2; z).$$

Proof. From $|\gamma_1(t) - \gamma_2(t)| < |\gamma_2(t) - z|$ for every $t \in [a, b]$ we easily get that $z \notin \gamma_1^*$ and $z \notin \gamma_2^*$. We also have

$$\left|\tfrac{\gamma_1(t)-z}{\gamma_2(t)-z} - 1\right| < 1$$

for every $t \in [a, b]$. Now, we apply again the argument of example 4.3.3. We consider the function $g : [a, b] \to D_1(1)$ with $g(t) = \frac{\gamma_1(t)-z}{\gamma_2(t)-z}$ for every $t \in [a, b]$. Let q be a continuous branch of log in $D_1(1)$. Then $f = q \circ g$ is a continuous branch of $\log g$ in $[a, b]$. Since the curves γ_1, γ_2 are closed, we have that $g(b) = g(a)$ and hence $f(b) = q(g(b)) = q(g(a)) = f(a)$. According to (6.8), the imaginary part h of f is a continuous branch of $\arg g$ in $[a, b]$ and from $f(b) = f(a)$ we get $h(b) = h(a)$.
Now let h_2 be a continuous branch of $\arg(\gamma_2 - z)$ in $[a, b]$. Since, $\gamma_1 - z = (\gamma_2 - z)g$ in $[a, b]$, proposition 6.8 implies that $h_1 = h_2 + h$ is a continuous branch of $\arg(\gamma_1 - z)$ in $[a, b]$. Therefore, $h_1(b) - h_1(a) = h_2(b) - h_2(a) + h(b) - h(a) = h_2(b) - h_2(a)$ and hence $n(\gamma_1; z) = n(\gamma_2; z)$ ☐

For every closed curve γ with $z \notin \gamma^*$ we may consider the translated closed curve $\gamma_z = \gamma - z$ with $0 \notin \gamma_z^*$. It is obvious that $n(\gamma; z) = n(\gamma_z; 0)$.

Proposition 6.15. *Let γ be a closed curve. Then the integer-valued function $n(\gamma; z)$ is constant in every connected component of the open set $\mathbb{C} \setminus \gamma^*$. We also have that $n(\gamma; z) = 0$ for every z in the unbounded connected component of $\mathbb{C} \setminus \gamma^*$.*

Proof. Let $\gamma : [a, b] \to \mathbb{C}$ be the curve and let $z \notin \gamma^*$. Then there is some disc $D_z(r)$ contained in $\mathbb{C} \setminus \gamma^*$ and hence $|w - z| < r \leq |\gamma(t) - z|$ for every $t \in [a, b]$ and every $w \in D_z(r)$.
We take any $w \in D_z(r)$ and we consider the translated curves $\gamma_z = \gamma - z$ and $\gamma_w = \gamma - w$. Then we have that $|\gamma_w(t) - \gamma_z(t)| = |w - z| < |\gamma_z(t)|$ for every $t \in [a, b]$ and proposition 6.14 implies that $n(\gamma; w) = n(\gamma_w; 0) = n(\gamma_z; 0) = n(\gamma; z)$.
We just proved that the function $n(\gamma; z)$ is locally constant in $\mathbb{C} \setminus \gamma^*$. Of course, this implies that $n(\gamma; z)$ is continuous in $\mathbb{C} \setminus \gamma^*$. Now, let Ω be a connected component of $\mathbb{C} \setminus \gamma^*$. Since $n(\gamma; z)$ is continuous and integer-valued in the connected set Ω, it is constant in Ω.
Now, let Ω be the unbounded connected component of $\mathbb{C} \setminus \gamma^*$. We take a disc $\overline{D}_0(R)$ which contains γ^*. As we saw in the previous subsection, the connected ring $D_0(R, +\infty)$ is contained in Ω. We take any $z \in D_0(R, +\infty)$ (and hence $z \in \Omega$) and then obviously there is a halfline l with vertex z which does not intersect the disc $\overline{D}_0(R)$ and hence it does not intersect γ^* either. From example 6.1.4 we have that $n(\gamma; z) = 0$. Therefore $n(\gamma, z) = 0$ for every $z \in \Omega$. ☐

Exercises.

6.1.1. (i) Consider closed curves γ_1, γ_2 and z not on their trajectories. Assume that there are successive points $w_1^{(1)}, \ldots, w_n^{(1)}, w_{n+1}^{(1)} = w_1^{(1)}$ of γ_1^* and successive points $w_1^{(2)}, \ldots, w_n^{(2)}, w_{n+1}^{(2)} = w_1^{(2)}$ of γ_2^* and curves $\sigma_1, \ldots, \sigma_n, \sigma_{n+1} = \sigma_1$ so that every σ_j goes from $w_j^{(1)}$ to $w_j^{(2)}$ and so that, for each $j = 1, \ldots, n$, the part of γ_1 between $w_j^{(1)}, w_{j+1}^{(1)}$, the part of γ_2 between $w_j^{(2)}, w_{j+1}^{(2)}$, σ_j and σ_{j+1} are all in a convex subregion D_j of $\mathbb{C} \setminus \{z\}$. Prove that $n(\gamma_1; z) = n(\gamma_2; z)$.
(ii) Take a point z and two halflines l, m with vertex z. Let $A \in l$, $A \neq z$ and $B \in m$, $B \neq z$. Consider any curve γ_1 from A to B in one of the two angular regions defined by l, m and any curve γ_2 from B to A in the second angular region defined by l, m. Consider the closed curve $\gamma = \gamma_1 \dot{+} \gamma_2$. Prove that $n(\gamma; z) = \pm 1$.

6.1.2. If γ_1, γ_2 are closed curves in $\mathbb{C} \setminus \{0\}$ then $\gamma_1\gamma_2$ is a closed curve in $\mathbb{C} \setminus \{0\}$. Prove that $\Delta \arg(\gamma_1\gamma_2) = \Delta \arg \gamma_1 + \Delta \arg \gamma_2$.

6.1.3. Let $F \subseteq \mathbb{C}$ be closed and connected, $\pm 1 \in F$ and $\Omega = \mathbb{C} \setminus F$. Prove that there is a holomorphic branch of $\log \frac{z-1}{z+1}$ in Ω. Prove also that there is a holomorphic branch of $(z^2-1)^{1/2}$ in Ω.

6.2 Homotopy.

Let $\gamma_0, \gamma_1 : [a,b] \to \mathbb{C}$ be two curves. We say that γ_1 is **homotopic** to γ_0 if there is a continuous function $F : [a,b] \times [0,1] \to \mathbb{C}$ so that $F(t,0) = \gamma_0(t)$ and $F(t,1) = \gamma_1(t)$ for every $t \in [a,b]$. The function F is called a **homotopy** from γ_0 to γ_1.

For each $s \in [0,1]$ the function $\gamma_s : [a,b] \to \mathbb{C}$, given by

$$\gamma_s(t) = F(t,s)$$

for $t \in [a,b]$, is continuous and hence it is a curve. We shall call it *intermediate curve* between γ_0 and γ_1.

Since $[a,b] \times [0,1]$ is compact, the homotopy F is uniformly continuous. Thus for every $\epsilon > 0$ there is $\delta > 0$ so that $|F(t',s') - F(t'',s'')| < \epsilon$ when $\sqrt{(t'-t'')^2 + (s'-s'')^2} < \delta$. Therefore, if $|s'-s''| < \delta$ then we have $|\gamma_{s'}(t) - \gamma_{s''}(t)| < \epsilon$ for every $t \in [a,b]$, i.e. the curves $\gamma_{s'}$ and $\gamma_{s''}$ are uniformly close. We see that *when s increases in $[0,1]$ the curves γ_s form a continuously varying family of curves, starting with γ_0 and ending with γ_1.* To be more precise, we have a mapping

$$[0,1] \ni s \mapsto \gamma_s \in C([a,b]),$$

which is continuous from $[0,1]$ with the euclidean distance to $C([a,b])$ with the uniform distance:

$$|s'-s''| < \delta \quad \Rightarrow \quad \|\gamma_{s'} - \gamma_{s''}\|_{[a,b]} = \max_{t \in [a,b]} |\gamma_{s'}(t) - \gamma_{s''}(t)| < \epsilon.$$

If all curves γ_s are closed, i.e. if $F(a,s) = F(b,s)$ for every $s \in [0,1]$, then we say that F is a **homotopy with closed intermediate curves**. If all curves γ_s have the same initial endpoint and the same final endpoint, i.e. if $F(a,s)$ is constant and $F(b,s)$ is constant for $s \in [0,1]$, then we say that F is a **homotopy with fixed endpoints**.

If all curves γ_s are in the same set A, then we say that F is a **homotopy in** A.

We may define a relation between curves in a set A: we write $\gamma_0 \equiv \gamma_1$ if there is a homotopy in A from γ_0 to γ_1. It is easy to see that this is an equivalence relation:
(i) Every curve $\gamma : [a,b] \to A$ is homotopic to itself through the homotopy $F : [a,b] \times [0,1] \to A$ given by $F(t,s) = \gamma(t)$.
(ii) If $F : [a,b] \times [0,1] \to A$ is a homotopy from γ_0 to γ_1, i.e. if $F(t,0) = \gamma_0(t)$ and $F(t,1) = \gamma_1(t)$ for $t \in [a,b]$, then the function $G : [a,b] \times [0,1] \to A$ given by $G(t,s) = F(t,1-s)$ is a homotopy from γ_1 to γ_0. In fact G is continuous and $G(t,0) = \gamma_1(t)$ and $G(t,1) = \gamma_0(t)$ for $t \in [a,b]$.
(iii) If $F : [a,b] \times [0,1] \to A$ is a homotopy from γ_0 to γ_1, i.e. if $F(t,0) = \gamma_0(t)$ and $F(t,1) = \gamma_1(t)$ for $t \in [a,b]$, and if $G : [a,b] \times [0,1] \to A$ is a homotopy from γ_1 to γ_2, i.e. if $G(t,0) = \gamma_1(t)$ and $G(t,1) = \gamma_2(t)$ for $t \in [a,b]$, then $H : [a,b] \times [0,1] \to A$, given by

$$H(t,s) = \begin{cases} F(t,2s), & t \in [a,b],\, s \in [0,\frac{1}{2}] \\ G(t,2s-1), & t \in [a,b],\, s \in [\frac{1}{2},1] \end{cases}$$

is a homotopy from γ_0 to γ_2. Indeed, H is continuous and $H(t,0) = \gamma_0(t)$ and $H(t,1) = \gamma_2(t)$ for $t \in [a,b]$.

Furthermore, the previous argument shows that the relation of homotopy with closed intermediate curves and the relation of homotopy with fixed endpoints are both equivalence relations.

Example 6.2.1. If the set A is convex, every two curves in A are homotopic in A. Indeed, let $\gamma_0, \gamma_1 : [a,b] \to A$ be two curves in A. Since $\gamma_0(t), \gamma_1(t) \in A$ and A is convex, the linear segment $[\gamma_0(t), \gamma_1(t)]$ is contained in A. Now, if we define $F : [a,b] \times [0,1] \to \mathbb{C}$ by

$$F(t,s) = (1-s)\gamma_0(t) + s\gamma_1(t),$$

then F is continuous and all its values are in A. Moreover, $F(t,0)=\gamma_0(t)$ and $F(t,1)=\gamma_1(t)$ for $t\in[a,b]$. Therefore, F is a homotopy in A from γ_0 to γ_1. It is easy to see that, if γ_0 and γ_1 are closed, then all intermediate curves are closed. Also, if γ_0 and γ_1 have the same initial endpoint and the same final endpoint, then all intermediate curves have the same initial endpoint and the same final endpoint.

Proposition 6.16. *Let f be holomorphic in the open set Ω.*
(i) If γ_0,γ_1 are piecewise smooth curves in Ω with the same initial endpoint and the same final endpoint and if there is a homotopy in Ω, with fixed endpoints, between γ_0 and γ_1, then

$$\int_{\gamma_0} f(z)\,dz=\int_{\gamma_1} f(z)\,dz.$$

(ii) If γ_0,γ_1 are closed piecewise smooth curves in Ω and if there is a homotopy in Ω, with closed intermediate curves, between γ_0 and γ_1, then

$$\oint_{\gamma_0} f(z)\,dz=\oint_{\gamma_1} f(z)\,dz.$$

Proof. Let $F:[a,b]\times[0,1]\to\Omega$ be the homotopy in Ω from γ_0 to γ_1.
Then the subset $F([a,b]\times[0,1])$ of Ω is compact and hence there is $\epsilon>0$ so that $|z-w|\geq\epsilon$ for every $z\in F([a,b]\times[0,1])$ and every $w\in\Omega^c$.
Moreover, since F is uniformly continuous, there is $\delta>0$ so that $|F(t',s')-F(t'',s'')|<\epsilon$ if $|t'-t''|<\delta$ and $|s'-s''|<\delta$.
Now, we take intermediate points $a=t_0<t_1<\ldots<t_{n-1}<t_n=b$ and $0=s_0<s_1<\ldots<s_{m-1}<s_m=1$ so that $t_k-t_{k-1}<\delta$ and $s_l-s_{l-1}<\delta$ for all k and l. Then every rectangle $[t_{k-1},t_k]\times[s_{l-1},s_l]$ is mapped by F in the disc $D_{F(t_{k-1},s_{l-1})}(\epsilon)$ which is contained in Ω. Since f is holomorphic in this disc, its curvilinear integral over any closed curve in this disc is equal to 0.
Now we denote $\gamma_{0,k}$ and $\gamma_{1,k}$ the restrictions of γ_0 and γ_1 in $[t_{k-1},t_k]$. We also denote $\sigma_{k,l}$ the linear segment $[F(t_{k-1},s_l),F(t_k,s_l)]$ for $k=1,\ldots,n$ and $l=1,\ldots,m-1$. Finally, we denote $\rho_{k,l}$ the linear segment $[F(t_k,s_{l-1}),F(t_k,s_l)]$ for $k=0,\ldots,n$ and $l=1,\ldots,m$. Then for every $k=1,\ldots,n$ we have

$$\begin{aligned}
\int_{\gamma_{0,k}} f(z)\,dz-\int_{\sigma_{k,1}} f(z)\,dz&=\int_{\rho_{k-1,1}} f(z)\,dz-\int_{\rho_{k,1}} f(z)\,dz\\
\int_{\sigma_{k,l-1}} f(z)\,dz-\int_{\sigma_{k,l}} f(z)\,dz&=\int_{\rho_{k-1,l}} f(z)\,dz-\int_{\rho_{k,l}} f(z)\,dz\qquad\text{for } l=2,\ldots,m-1\\
\int_{\sigma_{k,m-1}} f(z)\,dz-\int_{\gamma_{1,k}} f(z)\,dz&=\int_{\rho_{k-1,m}} f(z)\,dz-\int_{\rho_{k,m}} f(z)\,dz.
\end{aligned}$$

Adding these m equalities and then adding for $k=1,\ldots,n$ and considering cancellations, we find

$$\int_{\gamma_0} f(z)\,dz-\int_{\gamma_1} f(z)\,dz=\sum_{l=1}^m\int_{\rho_{0,l}} f(z)\,dz-\sum_{l=1}^m\int_{\rho_{n,l}} f(z)\,dz. \tag{6.9}$$

(i) Since all intermediate curves have the same initial endpoint and the same final endpoint, we see that all linear segments $\rho_{0,l}$ and $\rho_{n,l}$ are single point sets and hence all integrals in the right side of (6.9) are equal to 0. Thus, $\int_{\gamma_0} f(z)\,dz=\int_{\gamma_1} f(z)\,dz$.
(ii) Since all intermediate curves are closed, we have $F(a,s)=F(b,s)$ for every $s\in[0,1]$. Therefore, for each l the linear segments $\rho_{0,l}$ and $\rho_{n,l}$ coincide and again the right side of (6.9) is equal to 0. Thus, $\oint_{\gamma_0} f(z)\,dz=\oint_{\gamma_1} f(z)\,dz$. ☐

Proposition 6.17. *Let γ_0,γ_1 be two closed curves in $\mathbb{C}\setminus\{z\}$. If there is a homotopy in $\mathbb{C}\setminus\{z\}$, with closed intermediate curves, between γ_0 and γ_1, then*

$$n(\gamma_0;z)=n(\gamma_1;z).$$

Proof. First case: the two curves are piecewise continuous.
Then we just apply proposition 6.16(ii) to $f(\zeta)=\frac{1}{\zeta-z}$ and $\Omega=\mathbb{C}\setminus\{z\}$.
Second case: the two curves are not necessarily piecewise continuous.

Let $F : [a, b] \times [0, 1] \to \mathbb{C} \setminus \{z\}$ be a homotopy with closed intermediate curves, between γ_0 and γ_1. Since F is continuous and $[a, b] \times [0, 1]$ is compact and F does not take the value z, there is $\epsilon > 0$ so that $|F(t, s) - z| \geq \epsilon$ for every $t \in [a, b]$ and $s \in [0, 1]$. Also, since F is uniformly continuous, there is $\delta > 0$ so that $|s' - s''| < \delta$ implies $|\gamma_{s'}(t) - \gamma_{s''}(t)| < \epsilon$ for every $t \in [a, b]$, where γ_s is the intermediate curve corresponding to $s \in [0, 1]$. Then $|\gamma_{s'}(t) - \gamma_{s''}(t)| < |\gamma_{s''}(t) - z|$ for every $t \in [a, b]$ and proposition 6.14 implies that $n(\gamma_{s'}; z) = n(\gamma_{s''}; z)$. Now we take successive points $0 = s_0 < s_1 < \ldots < s_{n-1} < s_n = 1$ so that $s_k - s_{k-1} < \delta$ for every $k = 1, \ldots, n$. Then we have $n(\gamma_{s_{k-1}}; z) = n(\gamma_{s_k}; z)$ for every $k = 1, \ldots, n$ and hence $n(\gamma_0; z) = n(\gamma_1; z)$. $\square$

Exercises.

6.2.1. Let A be arcwise connected and $\gamma_1(t) = z_1$ and $\gamma_2(t) = z_2$ be two constant curves in A. If a curve γ is homotopic in A to γ_1, prove that γ is homotopic in A to γ_2.

6.2.2. If γ is a closed curve in $\mathbb{C} \setminus \{0\}$, prove that γ is homotopic in $\mathbb{C} \setminus \{0\}$ to a closed curve whose trajectory is contained in the circle $\mathbb{T}$.

6.2.3. (i) Let f be continuous in $\overline{D}_0(R)$. We define $\gamma(t) = f(Re^{it})$ for every $t \in [0, 2\pi]$. Prove that, if $n(\gamma; w) \neq 0$, then $w \in f(D_0(R))$. I.e. $\{w \,|\, w \text{ is surrounded by } \gamma\} \subseteq f(D_0(R))$.
(ii) Using the result of (i), prove the fundamental theorem of algebra.

6.2.4. Let $p \in A$ and let $\mathcal{M}_p(A)$ be the set of all closed curves in A with both of their endpoints at p. If $\gamma_1, \gamma_2 \in \mathcal{M}_p(A)$, then clearly $\gamma_1 \dot{+} \gamma_2 \in \mathcal{M}_p(A)$. Also, if $\gamma \in \mathcal{M}_p(A)$, then $\neg\,\gamma \in \mathcal{M}_p(A)$.
(i) Prove that the relation of homotopy in A with closed intermediate curves and fixed endpoints ($= p$) is an equivalence relation in $\mathcal{M}_p(A)$. The set of all equivalence classes is denoted $\mathcal{H}_p(A) = \{[\gamma] \,|\, \gamma \in \mathcal{M}_p(A)\}$.
(ii) If $\gamma, \gamma_1, \gamma_2 \in \mathcal{M}_p(A)$, we define $[\gamma_1] + [\gamma_2] = [\gamma_1 \dot{+} \gamma_2]$ and $-[\gamma] = [\neg\gamma]$. Prove that these are well-defined and that $\mathcal{H}_p(A)$ with these operations is a group, whose neutral element is $[\gamma_p]$, where γ_p is the constant curve p.
(iii) If A is arcwise connected, prove that for every $p, q \in A$ the groups $\mathcal{H}_p(A)$ and $\mathcal{H}_q(A)$ are isomorphic. In this case we write $\mathcal{H}(A)$. (See exercise 6.2.1.)
(iv) Prove that $\mathcal{H}(\mathbb{C}) \cong \{0\}$, $\mathcal{H}(\mathbb{C} \setminus \{0\}) \cong \mathbb{Z}$, $\mathcal{H}(\mathbb{T}) \cong \mathbb{Z}$.

6.2.5. Let $z_1, z_2, z_3, w_1, w_2, w_3$ be distinct points. Is it possible to join every z_k with every w_j with simple curves γ_{kj} whose trajectories are mutually disjoint?

6.3 Combinatorial results for curves and square nets.

Lemma 6.1. *Let $\Sigma = \{\sigma_1, \ldots, \sigma_n\}$ be a set of curves (not necessarily closed) and let $A = \{a_1, \ldots, a_m\}$ be the set of their endpoints ($m \leq 2n$). We assume that for every point of A the number of the curves in Σ that arrive at this point is the same as the number of the curves in Σ that leave from this point. Then we can partition Σ into subsets $\Sigma_1, \ldots, \Sigma_k$ so that each Σ_j consists of successive curves and the sum γ_j of the curves in Σ_j is a closed curve.*

Proof. We describe an algorithm for the partitioning of Σ.
We start with σ_1. The final endpoint of σ_1 is the initial endpoint of at least one curve in Σ. If the final endpoint of σ_1 coincides with its initial endpoint, then σ_1 is closed and we stop the process. If this is not the case, then, renumbering if necessary the curves $\sigma_2, \ldots, \sigma_n$, we may assume that the final endpoint of σ_1 coincides with the initial endpoint of σ_2. If the final endpoint of σ_2 coincides with the initial endpoint of σ_1, then the sum of σ_1, σ_2 is a closed curve and we stop the process. If the final endpoint of σ_2 coincides with its initial endpoint, then σ_2 is a closed curve and we stop the process. If the final endpoint of σ_2 is not the initial point of either σ_1 or σ_2, then renumbering if necessary the curves $\sigma_3, \ldots, \sigma_n$, we may assume that the final endpoint of σ_2 coincides with

the initial endpoint of σ_3. Then, exactly as before, we examine whether the final endpoint of σ_3 coincides with the initial endpoint of σ_1 or of σ_2 or of σ_3. Then, respectively, the sum of $\sigma_1, \sigma_2, \sigma_3$ or the sum of σ_2, σ_3 or σ_3 by itself is a closed curve and we stop the process. If the final endpoint of σ_3 is not the initial endpoint of either σ_1 or σ_2 or σ_3, then renumbering if necessary the curves $\sigma_4, \dots, \sigma_n$, we may assume that the final endpoint of σ_3 coincides with the initial endpoint of σ_4. Now, it is clear that this process will eventually stop, because we have only finitely many curves. Therefore, we shall eventually find successive curves $\sigma_1, \sigma_2, \dots, \sigma_{k-1}, \sigma_k$ ($1 \leq k \leq n$) so that the final endpoint of σ_k coincides with the initial endpoint of one of the same curves $\sigma_1, \sigma_2, \dots, \sigma_{k-1}, \sigma_k$. Let the final endpoint of σ_k coincide with the initial endpoint of σ_l for some l with $1 \leq l \leq k$. Then the sum of $\sigma_l, \sigma_{l+1}, \dots, \sigma_{k-1}, \sigma_k$ is a closed curve and we stop the process. Now we set

$$\Sigma_1 = \{\sigma_l, \sigma_{l+1}, \dots, \sigma_{k-1}, \sigma_k\}$$

and call γ_1 the closed curve which is the sum of $\sigma_l, \sigma_{l+1}, \dots, \sigma_{k-1}, \sigma_k$. Then we drop the curves of Σ_1 from Σ, i.e. we consider the set

$$\Sigma' = \Sigma \setminus \Sigma_1 = \{\sigma_1, \dots, \sigma_{l-1}, \sigma_{k+1}, \dots, \sigma_n\}.$$

Each endpoint of the curves in Σ' is one of the points of $A = \{a_1, \dots, a_m\}$, say it is a_j. Then the number of the curves in Σ that arrive at a_j is the same as the number of the curves in Σ that leave from a_j. But the curves $\sigma_l, \sigma_{l+1}, \dots, \sigma_{k-1}, \sigma_k$ are successive and hence if one of them arrives at a_j then the next one leaves from a_j. Therefore, the remaining curves, i.e. those in Σ', have the same property: the number of the curves in Σ' that arrive at a_j is the same as the number of the curves in Σ' that leave from a_j. Thus Σ' has the same property as the original Σ.
Now we continue our algorithm with Σ'. We find a subset Σ_2 of Σ' which consists of successive curves and we call γ_2 the closed curve which is the sum of the curves in Σ_2. Then we drop the curves of Σ_2 from Σ', i.e. we consider the set

$$\Sigma'' = \Sigma' \setminus \Sigma_2 = \Sigma \setminus (\Sigma_1 \cup \Sigma_2).$$

We go on until we exhaust the original Σ. □

Lemma 6.2. *We take any $\delta > 0$ and two perpendicular lines. For each of them we consider all its parallel lines at distances equal to integer multiples of δ. The result is a net of closed square regions of sidelength δ which cover the plane and have disjoint interiors. We choose any of those closed square regions, say $Q_1, \dots, Q_l$. We consider the closed boundary curves $\partial Q_1, \dots, \partial Q_l$ with their positive direction. Each of them is the sum of four corresponding linear segments, considered as curves with the same direction. We drop the linear segments (with necessarily opposite directions) which are common to any two neighboring square regions from among the $Q_1, \dots, Q_l$ and we consider the set $\Sigma = \{\sigma_1, \dots, \sigma_n\}$ of all the remaining linear segments, i.e. those which belong to only one of $Q_1, \dots, Q_l$. Then we can partition Σ into subsets $\Sigma_1, \dots, \Sigma_k$ so that each Σ_j consists of successive linear segments and the sum γ_j of the linear segments in Σ_j is a closed curve.*

Proof. It is enough to prove that Σ has the property described in lemma 6.1, i.e. that for every point of intersection a of our lines the number of the curves in Σ that arrive at a is the same as the number of the curves in Σ that leave from a. This can be done easily, considering cases for the number, 0 or 1 or 2 or 3 or 4, of the squares among $Q_1, \dots, Q_l$ which have a as one of their corners. □

6.4 The theorem of Cauchy in general open sets.

Let $\sigma_1, \dots, \sigma_n$ be any curves (not necessarily closed) and $k_1, \dots, k_n$ be any integers (not necessarily non-negative). Then we say that the curves $\sigma_1, \dots, \sigma_n$ considered $k_1, \dots, k_n$ times, respectively, form a **chain** Σ. The integer k_j is called **multiplicity** of the corresponding σ_j in the chain Σ. If every σ_j is closed, then Σ is called **closed chain** or **cycle**. If every σ_j is in a set A, then we say that Σ is in A.

If a curve σ is not among the curves which constitute a chain Σ, we may include it among those curves by assigning multiplicity 0 to σ. And now we may introduce the algebraic structure of a *module* in the set of all chains in the following manner. If Σ' and Σ'' are two chains, we may assume that they are formed by the same collection $\sigma_1, \ldots, \sigma_n$ of curves. If $k_1', \ldots, k_n'$ and $k_1'', \ldots, k_n''$ are the corresponding multiplicities in the chains Σ' and Σ'', then we define $\Sigma' + \Sigma''$ to be the chain which consists of $\sigma_1, \ldots, \sigma_n$ with multiplicities $k_1' + k_1'', \ldots, k_n' + k_n''$. Moreover, if k is an integer and Σ is a chain formed by the curves $\sigma_1, \ldots, \sigma_n$ with multiplicities $k_1, \ldots, k_n$, then we define $k\Sigma$ to be the chain formed by $\sigma_1, \ldots, \sigma_n$ with multiplicities $kk_1, \ldots, kk_n$. It is very easy to show that, under this addition of chains and this multiplication of chains and integers, the set of chains is a $\mathbb{Z}$-module. The opposite $-\Sigma$ of a chain Σ is $(-1)\Sigma$ and the neutral element of addition is the chain which contains no curve (or any curves all with multiplicities 0).

If Σ is a chain formed by the curves $\sigma_1, \ldots, \sigma_n$ with multiplicities $k_1, \ldots, k_n$, we immediately see that, under the above definitions of addition and multiplication, we have $\Sigma = k_1\sigma_1 + \cdots + k_n\sigma_n$. Here we consider each σ_j as a chain consisting of only one curve with multiplicity 1.

It is obvious that if Σ', Σ'' are cycles and k', k'' are integers then $k'\Sigma' + k''\Sigma''$ is a cycle. Therefore the set of cycles is a $\mathbb{Z}$-submodule of the $\mathbb{Z}$-module of all chains.

Now we consider a chain Σ formed by the piecewise smooth curves $\sigma_1, \ldots, \sigma_n$ with multiplicities $k_1, \ldots, k_n$ and a continuous $\phi : \sigma_1^* \cup \cdots \cup \sigma_n^* \to \mathbb{C}$. We define the **curvilinear integral** of ϕ over Σ by

$$\int_\Sigma \phi(z)\, dz = \sum_{j=1}^n k_j \int_{\sigma_j} \phi(z)\, dz.$$

If Σ is a cycle, we may use the notation

$$\oint_\Sigma \phi(z)\, dz.$$

It is easy to show that

$$\int_{k'\Sigma' + k''\Sigma''} \phi(z)\, dz = k' \int_{\Sigma'} \phi(z)\, dz + k'' \int_{\Sigma''} \phi(z)\, dz.$$

This says that integration "respects" the linear structure of the $\mathbb{Z}$-module of chains.

If Σ is a cycle formed by the closed curves $\sigma_1, \ldots, \sigma_n$ with multiplicities $k_1, \ldots, k_n$ and z does not belong to $\sigma_1^* \cup \cdots \cup \sigma_n^*$ we define the **rotation number** or **index** of Σ with respect to z by

$$n(\Sigma; z) = \sum_{j=1}^n k_j n(\sigma_j; z).$$

We may say that $n(\Sigma; z)$ is the total number of rotations around z of the closed curves forming Σ, taking into account their multiplicities.

Again, it is easy to show that

$$n(k'\Sigma' + k''\Sigma''; z) = k'n(\Sigma'; z) + k''n(\Sigma''; z)$$

for every z which does not belong to the trajectories of the curves forming the cycles Σ' and Σ'', and this says that the index "respects" the linear structure of the $\mathbb{Z}$-module of cycles.

Combining the last two definitions, we easily see that the index of a cycle consisting of closed piecewise smooth curves is given by the same integral form which gives the index of a closed piecewise smooth curve:

$$n(\Sigma; z) = \frac{1}{2\pi i} \oint_\Sigma \frac{1}{\zeta - z}\, d\zeta.$$

Indeed, $n(\Sigma; z) = \sum_{j=1}^n k_j n(\sigma_j; z) = \sum_{j=1}^n k_j \frac{1}{2\pi i} \oint_{\sigma_j} \frac{1}{\zeta - z}\, d\zeta = \frac{1}{2\pi i} \oint_\Sigma \frac{1}{\zeta - z}\, d\zeta$.

Now we state a basic definition.

Let Σ be a cycle in the open set Ω. We say that Σ is **null-homologous** in Ω if $n(\Sigma; z) = 0$ for every $z \in \Omega^c$.

In other words, a cycle Σ in Ω is null-homologous in Ω if the total number of rotations of the curves forming Σ, taking into account their multiplicities, around every point of the complement of Ω is zero.

It is easy to see that if the cycles Σ', Σ'' are null-homologous in Ω and k', k'' are integers then the cycle $k'\Sigma' + k'\Sigma''$ is null-homologous in Ω. Thus, the set $\mathcal{C}_0(\Omega)$ of all cycles which are null-homologous in Ω is a $\mathbb{Z}$-submodule of the $\mathbb{Z}$-module $\mathcal{C}(\Omega)$ of all cycles in Ω. Hence we may form the *quotient* $\mathbb{Z}$-module

$$\mathcal{H}(\Omega) = \mathcal{C}(\Omega)/\mathcal{C}_0(\Omega).$$

The elements of $\mathcal{H}(\Omega)$ are the classes $[\Sigma]$ of all cycles Σ in Ω, i.e. $\Sigma \in \mathcal{C}(\Omega)$, described as

$$[\Sigma] = \{\Sigma + \Sigma' \,|\, \Sigma' \in \mathcal{C}_0(\Omega)\}.$$

Now we introduce an equivalence relation among the cycles in Ω. We say that the cycles Σ_1, Σ_2 are **homologous** in Ω and we write $\Sigma_1 \sim \Sigma_2$ if $\Sigma_1 - \Sigma_2$ is null-homologous in Ω i.e. if $\Sigma_1 - \Sigma_2 \in \mathcal{C}_0(\Omega)$. Of course this means that $n(\Sigma_1 - \Sigma_2; z) = 0$ or equivalently $n(\Sigma_1; z) = n(\Sigma_2; z)$ for every $z \in \Omega^c$. If O is the zero-cycle, then clearly Σ is null-homologous in Ω if and only if $\Sigma \sim$ O. Another way to describe the elements $[\Sigma]$ of $\mathcal{H}(\Omega)$ is

$$[\Sigma] = \{\Sigma' \in \mathcal{C}(\Omega) \,|\, \Sigma' - \Sigma \in \mathcal{C}_0(\Omega)\} = \{\Sigma' \in \mathcal{C}(\Omega) \,|\, \Sigma' \sim \Sigma\}.$$

The algebraic operations in the *quotient* $\mathbb{Z}$-module $\mathcal{H}(\Omega)$ are as follows:

$$[\Sigma'] + [\Sigma''] = [\Sigma' + \Sigma''], \qquad k[\Sigma] = [k\Sigma].$$

We shall not go further into this algebraic point of view, since it does not have much to offer in our study of complex analysis. We shall keep in mind, though, the definition and notation of $\Sigma' + \Sigma''$ and $k\Sigma$ and from time to time we shall feel free to make certain mild algebraic comments.

Proposition 6.18. *Let Ω be an open set and $K \subseteq \Omega$ be compact. Then there are closed piecewise smooth curves $\gamma_1, \ldots, \gamma_k$ in $\Omega \setminus K$ so that for every f holomorphic in Ω we have*

$$f(z) = \textstyle\sum_{j=1}^{k} \frac{1}{2\pi i} \oint_{\gamma_j} \frac{f(\zeta)}{\zeta - z}\, d\zeta. \tag{6.10}$$

for every $z \in K$, and

$$0 = \textstyle\sum_{j=1}^{k} \frac{1}{2\pi i} \oint_{\gamma_j} \frac{f(\zeta)}{\zeta - z}\, d\zeta. \tag{6.11}$$

for every $z \in \Omega^c$.

Proof. There is $\delta > 0$ so that $|z - w| \geq 2\delta$ for every $z \in K$ and every $w \in \Omega^c$. For this $\delta > 0$ we consider the net of closed square regions of lemma 6.2 and we take all closed square regions $Q_1, \ldots, Q_l$ of the net which intersect K. Each Q_m intersects K and its diameter is equal to $\sqrt{2}\delta$. Therefore, the distance of every point of Q_m from K is $\leq \sqrt{2}\delta$. Since $\sqrt{2}\delta < 2\delta$, we see that Q_m is contained in Ω. Thus, all $Q_1, \ldots, Q_l$ are contained in Ω. As in lemma 6.2, we consider the set $\Sigma = \{\sigma_1, \ldots, \sigma_n\}$ of all boundary linear segments of $Q_1, \ldots, Q_l$ which belong to *only one* of $Q_1, \ldots, Q_l$ and we partition Σ into subsets $\Sigma_1, \ldots, \Sigma_k$ so that each Σ_j consists of successive linear segments and the sum γ_j of the linear segments in Σ_j is a closed curve.
Now consider any of the linear segments $\sigma_1, \ldots, \sigma_n$, say σ_j. Then σ_j belongs to one of $Q_1, \ldots, Q_l$, say Q_m. Since Q_m is contained in Ω, we have that σ_j is also contained in Ω. If σ_j intersects K, then both closed square regions of our net which lie on the two sides of σ_j intersect K and hence both are among $Q_1, \ldots, Q_l$. This is impossible because σ_j belongs to only one of $Q_1, \ldots, Q_l$. Therefore, σ_j does not intersect K and hence it is contained in $\Omega \setminus K$. Finally, since each of $\gamma_1, \ldots, \gamma_k$ is the sum of certain of the $\sigma_1, \ldots, \sigma_n$, we get that all $\gamma_1, \ldots, \gamma_k$ are in $\Omega \setminus K$.
Now we take any $z \in K$. Then z belongs to one of $Q_1, \ldots, Q_l$, say Q_m. Let us assume that z is an interior point of Q_m. Since the closed square region Q_m is contained in Ω, there is a slightly

larger open square region Q' which is also contained in Ω. Now f is holomorphic in the convex region Q' and Cauchy's formula in section 6.1 says that

$$f(z) = \tfrac{1}{2\pi i} \oint_{\partial Q_m} \tfrac{f(\zeta)}{\zeta - z}\, d\zeta, \tag{6.12}$$

because the index of ∂Q_m with respect to z is equal to 1. Now we take any closed square region Q_p with $p \neq m$. Then z is not contained in Q_p and again we may find an open square region Q' slightly larger than Q_p which is contained in Ω and which does not contain z. Then $\frac{f(\zeta)}{\zeta - z}$ is a holomorphic function of ζ in the convex region Q' and hence

$$0 = \tfrac{1}{2\pi i} \oint_{\partial Q_p} \tfrac{f(\zeta)}{\zeta - z}\, d\zeta \tag{6.13}$$

for $p \neq m$. We add (6.12) and (6.13) for all values of p and we get

$$f(z) = \textstyle\sum_{p=1}^{l} \tfrac{1}{2\pi i} \oint_{\partial Q_p} \tfrac{f(\zeta)}{\zeta - z}\, d\zeta. \tag{6.14}$$

Now we split the integral over each ∂Q_p in four integrals over the boundary linear segments of ∂Q_p and we get $4l$ integrals. If a linear segment belongs to two neighboring closed square regions, then it appears twice among the integrals, with opposite directions, and hence the two integrals cancel. Therefore, the remaining integrals will be only over the boundary linear segments which belong to exactly one of $Q_1, \dots, Q_l$, i.e. the linear segments of the set $\Sigma = \{\sigma_1, \dots, \sigma_n\}$. Thus (6.14) becomes

$$f(z) = \textstyle\sum_{\sigma \in \Sigma} \tfrac{1}{2\pi i} \int_{\sigma} \tfrac{f(\zeta)}{\zeta - z}\, d\zeta.$$

The subsets $\Sigma_1, \dots, \Sigma_k$ form a partition of Σ and hence

$$f(z) = \textstyle\sum_{j=1}^{k} \sum_{\sigma \in \Sigma_j} \tfrac{1}{2\pi i} \int_{\sigma} \tfrac{f(\zeta)}{\zeta - z}\, d\zeta.$$

Finally, since γ_j is the sum of the successive linear segments $\sigma \in \Sigma_j$, we end up with (6.10).
Now let z be a boundary point of Q_m. Then we may consider a variable point z' in the interior of Q_m so that $z' \to z$. We have proved (6.10) for z', i.e.

$$f(z') = \textstyle\sum_{j=1}^{k} \tfrac{1}{2\pi i} \oint_{\gamma_j} \tfrac{f(\zeta)}{\zeta - z'}\, d\zeta.$$

Proposition 4.12 implies the continuity of the right side as a function of z'. Therefore, taking the limit as $z' \to z$, we end up again with (6.10).
Now we consider any $z \in \Omega^c$. Then z does not belong to any of $Q_1, \dots, Q_l$ and so we get (6.13) for all values of p. Adding we find (6.14) with $f(z)$ replaced by 0. Now, following the same steps as before (splitting each ∂Q_p in four linear segments etc.), we end up with (6.11). □

Lemma 6.3. *Let γ be a piecewise smooth curve, K be a compact set so that $K \cap \gamma^* = \emptyset$ and f be a complex function continuous in γ^*. Then for every ϵ there are points $\zeta_0, \zeta_1, \dots, \zeta_{m-1}, \zeta_m$ of γ^* so that*

$$\big| \textstyle\int_{\gamma} \tfrac{f(\zeta)}{\zeta - z}\, d\zeta - \sum_{l=1}^{m} \tfrac{f(\zeta_l)}{\zeta_l - z}(\zeta_l - \zeta_{l-1}) \big| \leq \epsilon$$

for every $z \in K$.

Proof. Since $K \cap \gamma^* = \emptyset$, there is some $\rho > 0$ so that

$$|\zeta - z| \geq \rho \tag{6.15}$$

for every $\zeta \in \gamma^*$ and every $z \in K$.
We have that $\gamma : [a, b] \to \gamma^*$ and $f : \gamma^* \to \mathbb{C}$ are continuous and hence $f \circ \gamma : [a, b] \to \mathbb{C}$ is also continuous. Therefore, there is $M \geq 0$ so that

$$|f(\gamma(t))| \leq M \tag{6.16}$$

for every $t\in[a,b]$ and also there is $\delta>0$ so that

$$|\gamma(t')-\gamma(t'')|\le\tfrac{\rho^2\epsilon}{2Ml(\gamma)},\quad |f(\gamma(t'))-f(\gamma(t''))|\le\tfrac{\rho\epsilon}{2l(\gamma)} \tag{6.17}$$

for every $t',t''\in[a,b]$ with $|t'-t''|<\delta$.
Now we take succesive points $a=t_0<t_1<\ldots<t_{m-1}<t_m=b$ so that $t_l-t_{l-1}<\delta$ for every $l=1,\ldots,m$. Then (6.15), (6.16) and (6.17) imply that for every $t\in[t_{l-1},t_l]$ and every $z\in K$ we have

$$\begin{aligned}\big|\tfrac{f(\gamma(t))}{\gamma(t)-z}-\tfrac{f(\gamma(t_l))}{\gamma(t_l)-z}\big| &\le \big|\tfrac{f(\gamma(t))}{\gamma(t)-z}-\tfrac{f(\gamma(t_l))}{\gamma(t)-z}\big|+\big|\tfrac{f(\gamma(t_l))}{\gamma(t)-z}-\tfrac{f(\gamma(t_l))}{\gamma(t_l)-z}\big|\\ &=\tfrac{|f(\gamma(t))-f(\gamma(t_l))|}{|\gamma(t)-z|}+\tfrac{|f(\gamma(t_l))||\gamma(t)-\gamma(t_l)|}{|\gamma(t)-z||\gamma(t_l)-z|}\\ &\le\tfrac{\rho\epsilon}{2l(\gamma)\rho}+\tfrac{M\rho^2\epsilon}{2Ml(\gamma)\rho^2}=\tfrac{\epsilon}{l(\gamma)}.\end{aligned} \tag{6.18}$$

The points $\zeta_l=\gamma(t_l)$ are in γ^* and by (6.18) we finally get

$$\begin{aligned}\big|\int_\gamma\tfrac{f(\zeta)}{\zeta-z}\,d\zeta-\sum_{l=1}^m\tfrac{f(\zeta_l)}{\zeta_l-z}(\zeta_l-\zeta_{l-1})\big| &=\big|\sum_{l=1}^m\int_{t_{l-1}}^{t_l}\big(\tfrac{f(\gamma(t))}{\gamma(t)-z}-\tfrac{f(\gamma(t_l))}{\gamma(t_l)-z}\big)\gamma'(t)\,dt\big|\\ &\le\sum_{l=1}^m\int_{t_{l-1}}^{t_l}\big|\tfrac{f(\gamma(t))}{\gamma(t)-z}-\tfrac{f(\gamma(t_l))}{\gamma(t_l)-z}\big||\gamma'(t)|\,dt\\ &\le\sum_{l=1}^m\int_{t_{l-1}}^{t_l}\tfrac{\epsilon}{l(\gamma)}|\gamma'(t)|\,dt\\ &=\tfrac{\epsilon}{l(\gamma)}\int_a^b|\gamma'(t)|\,dt=\epsilon\end{aligned}$$

for every $z\in K$. $\square$

The actual points $\zeta_0,\zeta_1,\ldots,\zeta_{m-1},\zeta_m$ of γ^*, which were constructed in the proof of lemma 6.3, are obviously successive in the direction of γ from $\gamma(a)$ towards $\gamma(b)$. The actual content of lemma 6.3 is the approximation of curvilinear integrals by Riemann sums in a concrete situation. For the more general picture (but with no parameter z) look at exercise 2.2.8.

Proposition 6.19. *Let Ω be an open set, $K\subseteq\Omega$ be compact and f be holomorphic in Ω. Then for every $\epsilon>0$ there is a function g which is a linear combination of functions (of z) of the form $\frac{1}{z-\zeta}$ with $\zeta\in\Omega\setminus K$ so that $\|f-g\|_K\le\epsilon$.*

Proof. We consider the closed piecewise smooth curves $\gamma_1,\ldots,\gamma_k$ in $\Omega\setminus K$ which are provided by proposition 6.18. If f is holomorphic in Ω, then (6.10) holds for every $z\in K$. Lemma 6.3 for $\epsilon=\frac{1}{nk}$ implies that in each γ_j^* there are points $\zeta_{j,0},\zeta_{j,1},\ldots,\zeta_{j,m_j-1},\zeta_{j,m_j}$ so that

$$\big|\int_{\gamma_j}\tfrac{f(\zeta)}{\zeta-z}\,d\zeta-\sum_{l=1}^{m_j}\tfrac{f(\zeta_{j,l})}{\zeta_{j,l}-z}(\zeta_{j,l}-\zeta_{j,l-1})\big|\le\tfrac{2\pi\epsilon}{k}$$

for every $z\in K$. Now, the points $\zeta_{j,l}$ ($1\le j\le k$, $1\le l\le m_j$) are in $\Omega\setminus K$ and we have

$$\begin{aligned}&\big|2\pi i f(z)-\sum_{j=1}^k\sum_{l=1}^{m_j}\tfrac{f(\zeta_{j,l})}{\zeta_{j,l}-z}(\zeta_{j,l}-\zeta_{j,l-1})\big|\\ &\qquad\le\big|\sum_{j=1}^k\big(\oint_{\gamma_j}\tfrac{f(\zeta)}{\zeta-z}\,d\zeta-\sum_{l=1}^{m_j}\tfrac{f(\zeta_{j,l})}{\zeta_{j,l}-z}(\zeta_{j,l}-\zeta_{j,l-1})\big)\big|\\ &\qquad\le\sum_{j=1}^k\big|\oint_{\gamma_j}\tfrac{f(\zeta)}{\zeta-z}\,d\zeta-\sum_{l=1}^{m_j}\tfrac{f(\zeta_{j,l})}{\zeta_{j,l}-z}(\zeta_{j,l}-\zeta_{j,l-1})\big|\\ &\qquad\le\sum_{j=1}^k\tfrac{2\pi\epsilon}{k}=2\pi\epsilon\end{aligned}$$

for every $z\in K$. So if we denote $a_{j,l}=-\frac{f(\zeta_{j,l})}{2\pi i}(\zeta_{j,l}-\zeta_{j,l-1})$, we have that

$$\big|f(z)-\sum_{j=1}^k\sum_{l=1}^{m_j}\tfrac{a_{j,l}}{z-\zeta_{j,l}}\big|\le\epsilon$$

for every $z\in K$. Now the function $g(z)=\sum_{j=1}^k\sum_{l=1}^{m_j}\frac{a_{j,l}}{z-\zeta_{j,l}}$ is a linear combination of functions (of z) of the form $\frac{1}{z-\zeta}$ with $\zeta\in\Omega\setminus K$ and $\|f-g\|_K\le\epsilon$. $\square$

Lemma 6.4. *Let $\Omega \subseteq \mathbb{C}$ be open, $R > 0$, $\delta > 0$. Then*

$$K = \{z \in \Omega \,|\, |z| \leq R, |z - w| \geq \delta \text{ for every } w \in \Omega^c\}$$

is a compact subset of Ω.

Proof. It is clear that $K \subseteq \Omega$. Also K is bounded since $K \subseteq \overline{D}_0(R)$.
Now, let $z_n \in K$ for all n and $z_n \to z$. From $|z_n| \leq R$ for all n we get $|z| \leq R$. Also, for every $w \in \Omega^c$, from $|z_n - w| \geq \delta$ for all n we get $|z - w| \geq \delta$. Therefore, $z \in K$ and hence K is closed. ☐

The theorem of Cauchy in general open sets. *If f is holomorphic in the open set Ω and if the cycle Σ, consisting of closed piecewise smooth curves, is null-homologous in Ω, then*

$$\oint_\Sigma f(z)\,dz = 0.$$

First proof. Let the cycle Σ consist of the closed piecewise smooth curves $\sigma_1, \ldots, \sigma_n$ with multiplicities $k_1, \ldots, k_n$. Since $\sigma_1^* \cup \cdots \cup \sigma_n^*$ is a compact subset of Ω, there is $\delta > 0$ so that every point of $\sigma_1^* \cup \cdots \cup \sigma_n^*$ has a distance $\geq 2\delta$ from Ω^c and there is $R > 0$ so that $\sigma_1^* \cup \cdots \cup \sigma_n^*$ is contained in the closed disc $\overline{D}_0(R)$. We consider the set

$$K = \{z \in \Omega \,|\, |z| \leq R, |z - w| \geq 2\delta \text{ for every } w \in \Omega^c\}.$$

Lemma 6.4 says that K is a compact subset of Ω. Moreover, $\sigma_1^* \cup \cdots \cup \sigma_n^* \subseteq K$.
Now, take any ζ in $\Omega \setminus K$. Then either $\zeta \notin \overline{D}_0(R)$ or the distance of ζ from Ω^c is $< 2\delta$. If $\zeta \notin \overline{D}_0(R)$, then, since Σ is in $\overline{D}_0(R)$, we have that $n(\Sigma; \zeta) = 0$. If the distance of ζ from Ω^c is $< 2\delta$, then there is $w \in \Omega^c$ so that $|\zeta - w| < 2\delta$. Then every point of the linear segment $[\zeta, w]$ has distance $< 2\delta$ from w and hence from Ω^c. Thus $[\zeta, w]$ is not contained in K which implies that $[\zeta, w]$ is in the complement of $\sigma_1^*, \ldots, \sigma_n^*$. Since $[\zeta, w]$ is connected and it is contained in the complement of every σ_j^* we have that $n(\sigma_j; \zeta) = n(\sigma_j; w)$ for every $j = 1, \ldots, n$. Therefore,

$$n(\Sigma; \zeta) = \textstyle\sum_{j=1}^n k_j\, n(\sigma_j; \zeta) = \sum_{j=1}^n k_j\, n(\sigma_j; w) = n(\Sigma; w) = 0$$

because $w \in \Omega^c$ and Σ is null-homologous in Ω. With this compact set K we form the closed curves $\gamma_1, \ldots, \gamma_k$ in $\Omega \setminus K$, which are described in proposition 6.18. According to proposition 6.18 we have

$$f(z) = \textstyle\sum_{l=1}^k \frac{1}{2\pi i} \oint_{\gamma_l} \frac{f(\zeta)}{\zeta - z}\,d\zeta$$

for every $z \in \sigma_1^* \cup \cdots \cup \sigma_n^*$. Hence

$$\begin{aligned}
\oint_\Sigma f(z)\,dz &= \textstyle\sum_{j=1}^n k_j \oint_{\sigma_j} f(z)\,dz = \sum_{j=1}^n k_j \oint_{\sigma_j} \big(\sum_{l=1}^k \frac{1}{2\pi i} \oint_{\gamma_l} \frac{f(\zeta)}{\zeta - z}\,d\zeta\big)\,dz \\
&= -\textstyle\sum_{l=1}^k \oint_{\gamma_l} \big(\sum_{j=1}^n k_j \frac{1}{2\pi i} \oint_{\sigma_j} \frac{1}{z - \zeta}\,dz\big) f(\zeta)\,d\zeta \\
&= -\textstyle\sum_{l=1}^k \oint_{\gamma_l} \big(\sum_{j=1}^n k_j\, n(\sigma_j; \zeta)\big) f(\zeta)\,d\zeta \\
&= -\textstyle\sum_{l=1}^k \oint_{\gamma_l} n(\Sigma; \zeta) f(\zeta)\,d\zeta.
\end{aligned} \tag{6.19}$$

Finally, when ζ belongs to any of $\gamma_1^*, \ldots, \gamma_k^*$ then ζ belongs to $\Omega \setminus K$ and so $n(\Sigma; \zeta) = 0$. Now (6.19) implies $\oint_\Sigma f(z)\,dz = 0$.
Second proof. We start with the same compact $K \subseteq \Omega$ as in the first proof and we observe, exactly as before, that $n(\Sigma; \zeta) = 0$ for every $\zeta \in \Omega \setminus K$. Now, proposition 6.19 implies that, for every $\epsilon > 0$, there is a function g which is a linear combination of functions (of z) of the form $\frac{1}{z - \zeta}$ with $\zeta \in \Omega \setminus K$ and so that $\|f - g\|_K \leq \epsilon$. Let

$$g(z) = \textstyle\sum_{l=1}^m \frac{a_l}{z - \zeta_l}$$

with $\zeta_1, \dots, \zeta_m \in \Omega \setminus K$. Then we get

$$\oint_\Sigma g(z)\,dz = \sum_{l=1}^m a_l \oint_\Sigma \frac{1}{z-\zeta_l}\,dz = \sum_{l=1}^m a_l 2\pi i n(\Sigma;\zeta_l) = 0.$$

Now we apply proposition 6.19 with $\epsilon = \frac{1}{n}$ and we get a sequence (g_n) of functions, of the same type as the g we just considered, so that $g_n \to f$ uniformly in K. Since Σ is in K, we have that $\oint_\Sigma g_n(z)\,dz \to \oint_\Sigma f(z)\,dz$ and, finally, since $\oint_\Sigma g_n(z)\,dz = 0$ for every n, we conclude that $\oint_\Sigma f(z)\,dz = 0$. □

It is interesting to see that the assumption of our last result is at the same time a special case of it. Indeed, if we take any $w \in \Omega^c$, then the function $f(z) = \frac{1}{z-w}$ is holomorphic in Ω and the theorem of Cauchy implies that $\oint_\Sigma \frac{1}{z-w}\,dz = 0$. But this says that $n(\Sigma;w) = 0$. In other words, we have the following situation. The assumption that Σ is null-homologous in Ω is equivalent to the validity of the theorem of Cauchy for the very particular holomorphic functions of the form $f(z) = \frac{1}{z-w}$ for every $w \in \Omega^c$. Therefore the real content of the theorem of Cauchy is that *the validity of $\oint_\Sigma f(z)\,dz = 0$ for the special holomorphic functions in Ω of the form $f(z) = \frac{1}{z-w}$ for every $w \in \Omega^c$ implies its validity for every function f which is holomorphic in Ω.*

Example 6.4.1. Let γ be any closed piecewise smooth curve in the *convex* region Ω and let $w \in \Omega^c$. Then w is contained in the unbounded connected component of $\mathbb{C} \setminus \gamma^*$ and proposition 6.6 implies that $n(\gamma;w) = 0$. Hence γ is null-homologous in Ω. Now the theorem of Cauchy for general open sets says that $\oint_\gamma f(z)\,dz = 0$ for every f holomorphic in Ω. We conclude that *the theorem of Cauchy for convex regions is a corrolary of the theorem of Cauchy for general open sets*.

Example 6.4.2. We consider the open set $D_{z_0}(R_1,R_2)$ with $0 \le R_1 < R_2 \le +\infty$. We consider the closed curve γ which describes the circle $C_{z_0}(r)$, with $R_1 < r < R_2$, once and in the positive direction. This curve is *not* null-homologous in $D_{z_0}(R_1,R_2)$. Indeed, z_0 is in the complement of $D_{z_0}(R_1,R_2)$ and $n(\gamma;z_0) = \frac{1}{2\pi i}\oint_{C_{z_0}(r)} \frac{1}{z-z_0}\,dz = 1$. So we do not expect that $\oint_\gamma f(z)\,dz = 0$ is true for *every* f which is holomorphic in $D_{z_0}(R_1,R_2)$. In fact, this is certainly not true for $f(z) = \frac{1}{z-z_0}$ which *is* holomorphic in $D_{z_0}(R_1,R_2)$.

Example 6.4.3. We consider the same open set $D_{z_0}(R_1,R_2)$ as in the previous example and an arbitrary closed piecewise smooth curve γ in $D_{z_0}(R_1,R_2)$. We shall see how we can evaluate $\oint_\gamma f(z)\,dz$ with a minimum of effort for any f holomorphic in $D_{z_0}(R_1,R_2)$. It is clear that, depending on the specific curve γ, it may be difficult to evaluate the integral using a parametric equation of γ.
Let us assume that the shape of the trajectory and the direction of γ allow us to count the number of rotations of γ around z_0, i.e. we assume that we know the integer $k = n(\gamma;z_0)$.
Since $\overline{D}_{z_0}(R_1)$ is one of the two connected components of the complement of $D_{z_0}(R_1,R_2)$, we have that $n(\gamma;z) = k$ for every $z \in \overline{D}_{z_0}(R_1)$. On the other hand, we have that $n(\gamma;z) = 0$ for every z in the unbounded connected component of the complement of $D_{z_0}(R_1,R_2)$, which is $\overline{D}_{z_0}(R_2,+\infty)$. Now we take a closed piecewise smooth curve γ_1 in $D_{z_0}(R_1,R_2)$ such that the $\oint_{\gamma_1} f(z)\,dz$ may be much easier to evaluate than the original $\oint_\gamma f(z)\,dz$. For instance, we may consider γ_1 to describe the circle $C_{z_0}(r)$ with $R_1 < r < R_2$ once and in the positive direction. In this case we have that $n(\gamma_1;z) = 1$ for every $z \in \overline{D}_{z_0}(R_1)$ and $n(\gamma_1;z) = 0$ for every $z \in \overline{D}_{z_0}(R_2,+\infty)$. Now we form the cycle $\Sigma = 1\,\gamma + (-k)\,\gamma_1$ and we have

$$n(\Sigma;z) = 1\,n(\gamma;z) + (-k)\,n(\gamma_1;z) = k + (-k) = 0$$

for every $z \in \overline{D}_{z_0}(R_1)$ and also

$$n(\Sigma;z) = 1\,n(\gamma;z) + (-k)\,n(\gamma_1;z) = 0 + 0 = 0$$

for every $z \in \overline{D}_{z_0}(R_2, +\infty)$. Therefore, Σ is null-homologous in $D_{z_0}(R_1, R_2)$ and the theorem of Cauchy implies

$$0 = \oint_\Sigma f(z)\,dz = 1\,\oint_\gamma f(z)\,dz + (-k)\,\oint_{\gamma_1} f(z)\,dz$$

and hence

$$\oint_\gamma f(z)\,dz = k\oint_{\gamma_1} f(z)\,dz = k\oint_{C_{z_0}(r)} f(z)\,dz.$$

We see that the evaluation of $\oint_\gamma f(z)\,dz$ has been reduced to the evaluation of the possibly much simpler integral $\oint_{C_{z_0}(r)} f(z)\,dz$ and the evaluation of the index $n(\gamma; z_0)$.
We shall generalize this technique in the following sections and chapters.

Now we generalize Cauchy's formulas for derivatives.

Cauchy's formula for derivatives and closed curves in general open sets. *If f is holomorphic in the open set Ω and if the cycle Σ, consisting of closed piecewise smooth curves, is null-homologous in Ω, then for all $n \in \mathbb{N}_0$ we have*

$$n(\Sigma; z)f^{(n)}(z) = \tfrac{n!}{2\pi i}\oint_\Sigma \tfrac{f(\zeta)}{(\zeta-z)^{n+1}}\,d\zeta$$

for every $z \in \Omega$ which does not belong to the trajectory of any closed curve forming Σ.

Proof. The function $F(\zeta) = \frac{f(\zeta)-f(z)}{\zeta-z}$ is holomorphic in $\Omega\setminus\{z\}$. Since z is a root of $f(\zeta)-f(z)$, the singularity z of F is removable. So we may define F at z as $F(z) = \lim_{\zeta\to z}\frac{f(\zeta)-f(z)}{\zeta-z} = f'(z)$ and then F becomes holomorphic in Ω. Now we apply the theorem of Cauchy in general open sets and get

$$\oint_\Sigma \tfrac{f(\zeta)-f(z)}{\zeta-z}\,d\zeta = \oint_\Sigma F(\zeta)\,d\zeta = 0,$$

which implies

$$\tfrac{1}{2\pi i}\oint_\Sigma \tfrac{f(\zeta)}{\zeta-z}\,d\zeta = f(z)\tfrac{1}{2\pi i}\oint_\Sigma \tfrac{1}{\zeta-z}\,d\zeta = f(z)n(\Sigma; z)$$

for every $z \in \Omega$ which does not belong to the trajectory of any closed curve forming Σ. This is the result of the statement in the case $n = 0$. For derivatives of order $n \geq 1$ we differentiate both sides of the last formula, just as in the proof of the same theorem in convex sets, using the fact that the index of Σ is constant in a neighborhood of z. $\square$

Exercises.

6.4.1. Let f be holomorphic in $\mathbb{D}\setminus\{0\}$. If the closed piecewise smooth curve γ is in $\mathbb{D}\setminus\{0\}$ and $n(\gamma; 0) = 0$, evaluate $\oint_\gamma f(z)\,dz$.

6.4.2. Let f be holomorphic in $\mathbb{C}$ and $f(1) = 6$, $f(-1) = 10$. Prove that, if γ is any closed piecewise smooth curve in $\mathbb{C}\setminus\{-1, 1\}$, then $\frac{1}{2\pi i}\oint_\gamma \frac{f(z)}{z^2-1}\,dz$ can take every integral value.

6.4.3. Let $f(z) = (\frac{1}{z} + \frac{a}{z^3})e^z$ for $z \neq 0$. Find all the values of a so that $\oint_\gamma f(z)\,dz = 0$ for every closed piecewise smooth curve γ in $\mathbb{C}\setminus\{0\}$.

6.4.4. (i) Find all possible values of $\oint_\gamma \frac{2z-1}{z^2-z}\,dz$, where γ is an arbitrary closed piecewise smooth curve in $\mathbb{C}\setminus\{0, 1\}$.
(ii) Find all possible values of $\int_\gamma \frac{2z-1}{z^2-z}\,dz$, where γ is an arbitrary piecewise smooth curve in $\mathbb{C}\setminus\{0, 1\}$ with initial endpoint $-i$ and final endpoint i.

6.4.5. Find all possible values of $\oint_\gamma \frac{\cos z}{z^2-\pi z}\,dz$, where γ is an arbitrary closed piecewise smooth curve in $\mathbb{C}\setminus\{0, \pi\}$.

6.4.6. Let f be holomorphic in the open set Ω and γ be a closed piecewise smooth curve null-homologous in Ω. Let also $n(\gamma; z_0) \neq 0$.
(i) If A is the connected component of $\mathbb{C}\setminus\gamma^*$ which contains z_0, prove that $A \subseteq \Omega$ and $\partial A \subseteq \gamma^*$.
(ii) If $|f(\zeta)| \leq 1$ for every $\zeta \in \gamma^*$, prove that $|f(z_0)| \leq 1$.

6.5 The residue theorem.

Let z_0 be an isolated singularity of f and let

$$\sum_{-\infty}^{+\infty} a_n(z-z_0)^n$$

be the Laurent series of f in the ring $D_{z_0}(R)\setminus\{z_0\}$. Then the coefficient a_{-1} is called **residue** of f at z_0 and we denote

$$\text{Res}(f;z_0)=a_{-1}=\tfrac{1}{2\pi i}\oint_{C_{z_0}(r)} f(\zeta)\,d\zeta$$

for $0<r<R$.

Example 6.5.1. If z_0 is a removable singularity of f, then $a_n=0$ for every $n<0$ and in particular $\text{Res}(f;z_0)=0$.

Example 6.5.2. Every function of the form $f(z)=\frac{1}{(z-z_0)^N}$ with $N\geq 2$ has residue 0 at z_0.

Example 6.5.3. If z_0 is a pole of f of order $N\geq 1$, then we can find "easily" the residue of f at z_0. Indeed, there is a function g holomorphic in a disc $D_{z_0}(R)$ so that $g(z_0)\neq 0$ and $f(z)=\frac{g(z)}{(z-z_0)^N}$ for every $z\in D_{z_0}(R)\setminus\{z_0\}$. From the Taylor series $\sum_{n=0}^{+\infty} b_n(z-z_0)^n$ of g we see that

$$\text{Res}(f;z_0)=b_{N-1}=\tfrac{g^{(N-1)}(z_0)}{(N-1)!}.$$

For instance, if $N=1$, then $\text{Res}(f;z_0)=g(z_0)$ and, if $N=2$, then $\text{Res}(f;z_0)=g'(z_0)$.

Example 6.5.4. We consider a power series of the form

$$\sum_{-\infty}^{n=-1} a_n(z-z_0)^n$$

and we assume that its radius of convergence is 0, i.e. that it converges in the ring $D_{z_0}(0,+\infty)$. If f is the holomorphic function defined by the power series in $D_{z_0}(0,+\infty)$, then

$$\tfrac{1}{2\pi i}\oint_\gamma f(\zeta)\,d\zeta=n(\gamma;z_0)a_{-1}=n(\gamma;z_0)\,\text{Res}(f;z_0)$$

for every closed piecewise smooth curve γ in $\mathbb{C}\setminus\{z_0\}$. Indeed, since the power series converges uniformly in the compact set γ^* which is contained in its ring of convergence, we have

$$\tfrac{1}{2\pi i}\oint_\gamma f(\zeta)\,d\zeta=\sum_{-\infty}^{n=-1}\tfrac{a_n}{2\pi i}\oint_\gamma(\zeta-z_0)^n\,d\zeta=\tfrac{a_{-1}}{2\pi i}\oint_\gamma\tfrac{1}{\zeta-z_0}\,d\zeta=n(\gamma;z_0)\,\text{Res}(f;z_0),$$

where, for $n\leq -2$ we used the result of example 4.5.3. Of course, this result holds for a general cycle Σ which consists of closed piecewise smooth curves γ in $\mathbb{C}\setminus\{z_0\}$.

The residue theorem is a generalization of the last example.

The residue theorem. *Let f be holomorphic, except for isolated singularities, in the open set Ω and Σ be a cycle which is null-homologous in Ω and so that no isolated singularity of f is in the trajectory of any of the closed curves forming Σ. Then $n(\Sigma;z)\neq 0$ for at most finitely many isolated singularities z of f. Moreover, if Σ consists of closed piecewise smooth curves, then*

$$\tfrac{1}{2\pi i}\oint_\Sigma f(\zeta)\,d\zeta=\sum_{z\text{ sing. of }f} n(\Sigma;z)\ \text{Res}(f;z),$$

where the sum, extended over all isolated singularities of f in Ω, is finite.

First proof. Exactly as in the proof of the theorem of Cauchy in general open sets, we see that there is a compact set $K \subseteq \Omega$ so that $n(\Sigma; z) = 0$ for every $z \in \Omega \setminus K$. Now, since all singularities of f are isolated, there can be only finitely many of them in K. Let $z_1, \ldots, z_n$ be the singularities of f in K. Then every other singularity z of f is in $\Omega \setminus K$ and hence $n(\Sigma; z) = 0$.
We define the integers

$$p_1 = n(\Sigma; z_1), \ldots, p_n = n(\Sigma; z_n)$$

and then

$$\textstyle\sum_{z \text{ sing. of } f} n(\Sigma; z) \operatorname{Res}(f; z) = \sum_{k=1}^{n} p_k \operatorname{Res}(f; z_k).$$

Therefore, it is enough to prove

$$\tfrac{1}{2\pi i} \oint_{\Sigma} f(\zeta)\, d\zeta = \textstyle\sum_{k=1}^{n} p_k \operatorname{Res}(f; z_k). \tag{6.20}$$

Since every $z_1, \ldots, z_n$ is an isolated singularity, there are disjoint closed discs $\overline{D}_{z_k}(r_k)$ for $k = 1, \ldots, n$ so that each of them contains no singularity of f except its center. We denote γ_k the closed curve which describes the circle $C_{z_k}(r_k)$ once and in the positive direction. We consider the cycle

$$\Sigma' = \Sigma + (-p_1)\, \gamma_1 + \cdots + (-p_n)\, \gamma_n$$

and the open set

$$\Omega' = \Omega \setminus \{z \in \Omega \,|\, z \text{ singularity of } f\}.$$

Clearly, f is holomorphic in Ω' and we shall prove that the cycle Σ' is null-homologous in Ω', i.e. $n(\Sigma'; z) = 0$ for every $z \notin \Omega'$. If $z \notin \Omega'$, then either $z \notin \Omega$ or $z = z_1, \ldots, z_n$ or z is any other isolated singularity of f in Ω.
If $z \notin \Omega$ or if z is any isolated singularity of f in Ω different from $z_1, \ldots, z_n$, then $n(\Sigma; z) = 0$ and $n(\gamma_k; z) = 0$ for every k. Therefore

$$n(\Sigma'; z) = n(\Sigma; z) - p_1 n(\gamma_1; z) - \cdots - p_n n(\gamma_n; z) = 0.$$

If $z = z_{k_0}$ for some k_0, then $n(\Sigma; z) = n(\Sigma; z_{k_0}) = p_{k_0}$ and $n(\gamma_{k_0}; z) = n(\gamma_{k_0}; z_{k_0}) = 1$ and $n(\gamma_k; z) = n(\gamma_k; z_{k_0}) = 0$ for every $k \neq k_0$. Therefore

$$n(\Sigma'; z) = n(\Sigma; z) - p_1 n(\gamma_1; z) - \cdots - p_n n(\gamma_n; z) = p_{k_0} - p_{k_0} = 0.$$

Thus, Σ' is null-homologous in Ω'. Since f is holomorphic in Ω', the theorem of Cauchy implies $\oint_{\Sigma'} f(\zeta)\, d\zeta = 0$. Hence

$$\oint_{\Sigma} f(\zeta)\, d\zeta = \textstyle\sum_{k=1}^{n} p_k \oint_{\gamma_k} f(\zeta)\, d\zeta = 2\pi i \sum_{k=1}^{n} p_k \operatorname{Res}(f; z_k)$$

and we proved (6.20).
Second proof. We follow the first proof up to the point where we considered the isolated singularities $z_1, \ldots, z_n$ of f. I.e. $n(\Sigma; z) = 0$ for every isolated singularity of f different from $z_1, \ldots, z_n$. Now, we consider the corresponding singular parts $s_1, \ldots, s_n$ of f at $z_1, \ldots, z_n$. Then we know from section 5.8 that $f - s_k$ is holomorphic at z_k and also that s_k is holomorphic in $\mathbb{C} \setminus \{z_k\}$. Hence the function

$$g = f - s_1 - \ldots - s_n$$

is holomorphic in Ω except at the isolated singularities of f which are different from $z_1, \ldots, z_n$. We consider the open set

$$\Omega'' = \Omega \setminus \{z \in \Omega \,|\, z \text{ is a singularity of } f, z \neq z_1, \ldots, z_n\}.$$

and then g is holomorphic in Ω''. Also, Σ is null-homologous in Ω''. Therefore, the theorem of Cauchy implies that $\frac{1}{2\pi i} \oint_{\Sigma} g(\zeta)\, d\zeta = 0$ and hence

$$\begin{aligned} \tfrac{1}{2\pi i} \oint_{\Sigma} f(\zeta)\, d\zeta &= \textstyle\sum_{k=1}^{n} \tfrac{1}{2\pi i} \oint_{\Sigma} s_k(\zeta)\, d\zeta = \sum_{k=1}^{n} n(\Sigma; z_k) \operatorname{Res}(s_k; z_k) \\ &= \textstyle\sum_{k=1}^{n} n(\Sigma; z_k) \operatorname{Res}(f; z_k), \end{aligned}$$

where for the second equality we used the result of example 6.5.4. □

Exercises.

6.5.1. Find the singular parts as well as the residues of

$$\tfrac{1}{z^2+5z+6}, \quad \tfrac{1}{(z^2-1)^2}, \quad e^z + e^{1/z}, \quad \tfrac{\cos z - 1}{z^4}, \quad \tfrac{1}{\sin z}, \quad \tan z, \quad \tfrac{1}{\sin^2 z}, \quad \tfrac{1}{e^z - 1}$$

at their isolated singularities.

6.5.2. If $f = gh$, where g is holomorphic at z_0 and h has a pole of order 1 at z_0, prove that $\mathrm{Res}(f; z_0) = g(z_0)\,\mathrm{Res}(h; z_0)$.

6.5.3. Let $f = \frac{g}{h}$, where g, h are holomorphic in a neighborhood of z_0. Assume that z_0 is a root of h of multiplicity N and not a root of g. Then z_0 is a pole of f of order N.
(i) If $N = 1$, prove that $\mathrm{Res}(f; z_0) = \frac{g(z_0)}{h'(z_0)}$.
(ii) If $N = 2$, prove that $\mathrm{Res}(f; z_0) = \frac{6g'(z_0)h''(z_0) - 2g(z_0)h'''(z_0)}{3h''(z_0)^2}$.

6.5.4. If $z_1, \ldots, z_n \in D_0(R)$ are distinct and f is holomorphic in an open set containing $\overline{D}_0(R)$ and $p(z) = (z - z_1)\cdots(z - z_n)$, prove that

$$\oint_{C_0(R)} \tfrac{f(z)}{(z-z_1)\cdots(z-z_n)}\, dz = 2\pi i\big(\tfrac{f(z_1)}{p'(z_1)} + \cdots + \tfrac{f(z_n)}{p'(z_n)}\big).$$

6.5.5. If $n \in \mathbb{N}$, evaluate $\oint_{C_0(n)} \tan(\pi z)\, dz$.

6.5.6. Let $r = \frac{p}{q}$ be a rational function with $\deg q \geq \deg p + 2$. If $z_1, \ldots, z_n$ are the distinct roots of q, prove that $\sum_{k=1}^{n} \mathrm{Res}(r; z_k) = 0$. What is the value of $\sum_{k=1}^{n} \mathrm{Res}(r; z_k)$ if $\deg q = \deg p + 1$?

6.5.7. If $f(z) = e^{z + (1/z)}$, prove that $\mathrm{Res}(f; 0) = \sum_{n=0}^{+\infty} \frac{1}{n!(n+1)!}$.

6.5.8. (i) Prove that there is $m > 0$ so that $|\sin(\pi z)| \geq m$ and $|\tan(\pi z)| \geq m$ for every $z \in \partial R_n$, where R_n is the square region with corners at the points $\pm(n + \frac{1}{2}) \pm i(n + \frac{1}{2})$, $n \in \mathbb{N}$.
(ii) Let f be holomorphic in $D_0(R, +\infty)$ for some $R > 0$ and let $\lim_{z\to\infty} zf(z)$ be a complex number. Prove that

$$\lim_{n\to+\infty} \oint_{\partial R_n} \tfrac{f(z)}{\sin(\pi z)}\, dz = 0, \quad \lim_{n\to+\infty} \oint_{\partial R_n} \tfrac{f(z)}{\tan(\pi z)}\, dz = 0.$$

(iii) Let f be holomorphic in $\mathbb{C}$ except for poles $z_1, \ldots, z_N \notin \mathbb{Z}$ and let $\lim_{z\to\infty} zf(z)$ be a complex number. Prove that

$$\lim_{n\to+\infty} \sum_{k=-n}^{n} f(k) = -\pi \sum_{j=1}^{N} \mathrm{Res}\,\big(\tfrac{f(z)}{\tan(\pi z)}; z_j\big),$$

$$\lim_{n\to+\infty} \sum_{k=-n}^{n} (-1)^k f(k) = -\pi \sum_{j=1}^{N} \mathrm{Res}\,\big(\tfrac{f(z)}{\sin(\pi z)}; z_j\big).$$

(iv) If $w \notin \mathbb{Z}$, prove that

$$-\tfrac{1}{w} + \sum_{k=-\infty}^{+\infty}\big(\tfrac{1}{k-w} - \tfrac{1}{k}\big) = \lim_{n\to+\infty} \sum_{k=-n}^{n} \tfrac{1}{k-w} = -\tfrac{\pi}{\tan(\pi w)}.$$

(v) If $w \notin \mathbb{Z}$, prove that

$$\sum_{k=-\infty}^{+\infty} \tfrac{1}{(k-w)^2} = \tfrac{\pi^2}{\sin^2(\pi w)}$$

and then that

$$\sum_{n=1}^{+\infty} \tfrac{1}{k^2} = \tfrac{\pi^2}{6}.$$

(vi) If $a > 0$, prove that

$$\sum_{k=1}^{+\infty} \tfrac{1}{k^2+a^2} = -\tfrac{1}{2a^2} + \tfrac{\pi}{2a}\,\tfrac{e^{\pi a} + e^{-\pi a}}{e^{\pi a} - e^{-\pi a}}, \quad \sum_{k=1}^{+\infty} \tfrac{(-1)^k}{k^2+a^2} = -\tfrac{1}{2a^2} - \tfrac{\pi}{a}\,\tfrac{1}{e^{\pi a} - e^{-\pi a}}.$$

6.6 Evaluation of integrals.

The residue theorem is a powerful tool for the evaluation of integrals, because it reduces this evaluation to the location of the isolated sinularities of the function to be integrated and to the evaluation of the corresponding residues. Let us see some characteristic examples.

Example 6.6.1. *Evaluation of* $\int_{-\infty}^{+\infty} r(x)\,dx$*, where* $r=\frac{p}{q}$ *is a rational function,* $\deg q\geq\deg p+2$*,* q *has no real roots and the coefficients of* p,q *are real numbers.*
Let $p(x)=a_nx^n+\cdots+a_1x+a_0$, with $a_n\neq 0$, and $q(x)=b_mx^m+\cdots+b_1x+b_0$, with $b_m\neq 0$, and $m\geq n+2$. Then r is continuous in $\mathbb{R}$ and the generalized integral $\int_{-\infty}^{+\infty} r(x)\,dx$ converges. To see this, we observe that $\lim_{z\to\infty} z^{m-n}r(z)=\frac{a_n}{b_m}$. Hence, if $c=\frac{|a_n|}{|b_m|}>0$, there is $R_0>0$ so that

$$\tfrac{c}{2}\leq |z|^{m-n}|r(z)|\leq 2c \tag{6.21}$$

when $|z|\geq R_0$. Now, since $m-n\geq 2$, we get

$$\int_{-\infty}^{-R_0}|r(x)|\,dx\leq 2c\int_{-\infty}^{-R_0}\tfrac{1}{|x|^{m-n}}\,dx<+\infty,\quad \int_{R_0}^{+\infty}|r(x)|\,dx\leq 2c\int_{R_0}^{+\infty}\tfrac{1}{x^{m-n}}\,dx<+\infty.$$

Thus, the integrals $\int_{-\infty}^{-R_0} r(x)\,dx$, $\int_{R_0}^{+\infty} r(x)\,dx$ converge absolutely and so they converge. Moreover, r is continuous in $[-R_0,R_0]$ and so the integral $\int_{-\infty}^{+\infty} r(x)\,dx$ also converges.
We consider the roots of q in the *upper halfplane* and let them be $z_1,\dots,z_M$, where $M\leq m$. We take any $R>R_0$ so that $z_1,\dots,z_M$ are contained in the disc $D_0(R)$. We apply the residue theorem with $r=\frac{p}{q}$ which is holomorphic in $\mathbb{C}$ except for the roots of q and with the closed curve γ_R which is the sum of the linear segment $[-R,R]$, with parametric equation $z=x$, $x\in[-R,R]$, and of the curve σ_R, with parametric equation $z=Re^{it}$, $t\in[0,\pi]$, which describes the upper semicircle of $C_0(R)$ from R to $-R$. The trajectory of γ_R contains no isolated singularity of r. Since γ_R rotates around each of $z_1,\dots,z_M$ once and in the positive direction, the residue theorem implies

$$\tfrac{1}{2\pi i}\oint_{\gamma_R} r(z)\,dz=\mathrm{Res}(r;z_1)+\cdots+\mathrm{Res}(r;z_M).$$

We have that $\oint_{\gamma_R} r(z)\,dz=\int_{[-R,R]} r(z)\,dz+\int_{\sigma_R} r(z)\,dz$ and hence

$$\int_{-R}^{R} r(x)\,dx=\int_{[-R,R]} r(z)\,dz=2\pi i(\mathrm{Res}(r;z_1)+\cdots+\mathrm{Res}(r;z_M))-\int_{\sigma_R} r(z)\,dz.$$

Since $R>R_0$, (6.21) and $m\geq n+2$ imply

$$\Big|\int_{\sigma_R} r(z)\,dz\Big|\leq \tfrac{2c}{R^{m-n}}\,\pi R\to 0$$

when $R\to+\infty$, and we conclude that

$$\int_{-\infty}^{+\infty} r(x)\,dx=2\pi i(\mathrm{Res}(r;z_1)+\cdots+\mathrm{Res}(r;z_M)).$$

Thus, to evaluate $\int_{-\infty}^{+\infty} r(x)\,dx$ we only need to find the residues of r at the poles $z_1,\dots,z_M$ of r in the upper halfplane.

Example 6.6.2. *Evaluation of* pv $\int_{-\infty}^{+\infty} r(x)\,dx$*, where* $r=\frac{p}{q}$ *is a rational function,* $\deg q=\deg p+1$*,* q *has no real root and the coefficients of* p,q *are real numbers.*
Let $p(x)=a_nx^n+\cdots+a_1x+a_0$, with $a_n\neq 0$, and $q(x)=b_{n+1}x^{n+1}+\cdots+b_1x+b_0$, with $b_{n+1}\neq 0$. It easy to see that the generalized integral $\int_{-\infty}^{+\infty} r(x)\,dx$ does not converge. Indeed, we recall the estimate (6.21), i.e. $|r(z)|\geq\frac{c}{2|z|}$ when $|z|\geq R_0$. Therefore, for real $z=x$ we have that $|r(x)|\geq\frac{c}{2x}$ when $x\geq R_0$. Now, r has constant sign in $[R_0,+\infty)$ and hence

$$\Big|\int_{R_0}^{+\infty} r(x)\,dx\Big|=\int_{R_0}^{+\infty}|r(x)|\,dx\geq\tfrac{c}{2}\int_{R_0}^{+\infty}\tfrac{1}{x}\,dx=+\infty.$$

Thus, $\int_{R_0}^{+\infty} r(x)\,dx = +\infty$ or $-\infty$ and, similarly, $\int_{-\infty}^{-R_0} r(x)\,dx = +\infty$ or $-\infty$.
Since the generalized integral diverges, we examine its **principal value**, i.e.

$$\text{pv}\int_{-\infty}^{+\infty} r(x)\,dx = \lim_{R\to+\infty}\int_{-R}^{R} r(x)\,dx.$$

It is easy to see that $r(z) - \frac{a_n}{b_{n+1}}\frac{1}{z}$ is a rational function whose denominator has degree *two units* larger than the degree of its numerator. According to the previous example, there is $R_0 > 0$ so that

$$\big|r(z) - \tfrac{a_n}{b_{n+1}}\tfrac{1}{z}\big| \le \tfrac{C}{|z|^2} \tag{6.22}$$

when $|z| \ge R_0$. As in the previous example, we consider the roots $z_1, \dots, z_M$ of q in the upper halfplane and we take $R > R_0$ so that $z_1, \dots, z_M$ are contained in $D_0(R)$. We apply the residue theorem with $r = \frac{p}{q}$ and the same closed curve γ_R and we get

$$\tfrac{1}{2\pi i}\oint_{\gamma_R} r(z)\,dz = \text{Res}(r; z_1) + \cdots + \text{Res}(r; z_M).$$

Now, $\oint_{\gamma_R} r(z)\,dz = \int_{[-R,R]} r(z)\,dz + \int_{\sigma_R} r(z)\,dz$ and hence

$$\int_{-R}^{R} r(x)\,dx = 2\pi i(\text{Res}(r; z_1) + \cdots + \text{Res}(r; z_M)) - \int_{\sigma_R}\big(r(z) - \tfrac{a_n}{b_{n+1}}\tfrac{1}{z}\big)\,dz - \tfrac{a_n}{b_{n+1}}\int_{\sigma_R}\tfrac{1}{z}\,dz.$$

The last term is

$$\tfrac{a_n}{b_{n+1}}\int_{\sigma_R}\tfrac{1}{z}\,dz = \tfrac{a_n}{b_{n+1}}\int_0^{\pi}\tfrac{1}{Re^{it}}\,iRe^{it}\,dt = i\pi\,\tfrac{a_n}{b_{n+1}}.$$

Since $R > R_0$, we have from (6.22) that

$$\big|\int_{\sigma_R}(r(z) - \tfrac{a_n}{b_{n+1}}\tfrac{1}{z})\,dz\big| \le \tfrac{C}{R^2}\,\pi R \to 0$$

when $R \to +\infty$ and we finally get

$$\text{pv}\int_{-\infty}^{+\infty} r(x)\,dx = 2\pi i(\text{Res}(r; z_1) + \cdots + \text{Res}(r; z_M)) - i\pi\,\tfrac{a_n}{b_{n+1}}.$$

Example 6.6.3. *Evaluation of* pv $\int_{-\infty}^{+\infty} r(x)\,dx$, *where* $r = \frac{p}{q}$ *is a rational function,* $\deg q \ge \deg p + 1$, *the real roots of* q *have multiplicity* 1 *and the coefficients of* p, q *are real numbers.*
Let $p(x) = a_n x^n + \cdots + a_1 x + a_0$, with $a_n \ne 0$, and $q(x) = b_m x^m + \cdots + b_1 x + b_0$, with $b_m \ne 0$, and $m \ge n+1$. We assume that the real roots of q are $x_1, \dots, x_N$ with $x_1 < \ldots < x_N$ and that these are not roots of p. We take $\epsilon_0 > 0$ so that the intervals $[x_1-\epsilon_0, x_1+\epsilon_0], \dots, [x_N-\epsilon_0, x_N+\epsilon_0]$ around the real roots of q are disjoint. In order for $\int_{-\infty}^{+\infty} r(x)\,dx$ to converge, the generalized integrals $\int_{x_k-\epsilon_0}^{x_k} r(x)\,dx$ and $\int_{x_k}^{x_k+\epsilon_0} r(x)\,dx$ must converge for every x_k. This is *not* correct. Indeed, we write $r(z) = \frac{p(z)}{(z-x_k)q_k(z)} = \frac{g_k(z)}{z-x_k}$, where q_k is a polynomial with $q_k(x_k) \ne 0$ and where $g_k = \frac{p}{q_k}$ is a rational function holomorphic at x_k. Since $\lim_{z\to x_k} g_k(z) = g_k(x_k) \ne 0$, there is ϵ_k with $0 < \epsilon_k \le \epsilon_0$ so that $|g_k(z)| \ge \frac{1}{2}|g_k(x_k)|$ for every z with $|z - x_k| \le \epsilon_k$. Hence, $|r(z)| \ge \frac{1}{2}\frac{|g_k(x_k)|}{|z-x_k|}$ for every z with $0 < |z - x_k| \le \epsilon_k$. The function r has constant sign in $(x_k, x_k + \epsilon_k]$. Therefore,

$$\big|\int_{x_k}^{x_k+\epsilon_k} r(x)\,dx\big| = \int_{x_k}^{x_k+\epsilon_k}|r(x)|\,dx \ge \tfrac{|g_k(x_k)|}{2}\int_{x_k}^{x_k+\epsilon_k}\tfrac{1}{x-x_k}\,dx = +\infty$$

and the generalized integral $\int_{x_k}^{x_k+\epsilon_k} r(x)\,dx$ does not converge. Similarly, $\int_{x_k-\epsilon_k}^{x_k} r(x)\,dx$ does not converge either. This is why we examine the **principal value** of $\int_{-\infty}^{+\infty} r(x)\,dx$, i.e.

$$\begin{aligned}\text{pv}\int_{-\infty}^{+\infty} r(x)\,dx = \lim_{R\to+\infty,\epsilon\to0+}\big(&\int_{-R}^{x_1-\epsilon} r(x)\,dx + \int_{x_1+\epsilon}^{x_2-\epsilon} r(x)\,dx + \cdots\\ &\cdots + \int_{x_{N-1}+\epsilon}^{x_N-\epsilon} r(x)\,dx + \int_{x_N+\epsilon}^{R} r(x)\,dx\big) = \lim_{R\to+\infty,\epsilon\to0+} I(R,\epsilon).\end{aligned} \tag{6.23}$$

We evaluate $I(R,\epsilon)$ using a variant of the curve γ_R of the previous examples: the curve $\gamma_{R,\epsilon}$, which is the sum of the linear segments $[-R,x_1-\epsilon],[x_1+\epsilon,x_2-\epsilon],\ldots,[x_{N-1}+\epsilon,x_N-\epsilon],[x_N+\epsilon,R]$, of the curve σ_R, which describes the upper semicircle of $C_0(R)$ from R to $-R$, and of the curves $\sigma_{1,\epsilon},\ldots,\sigma_{N,\epsilon}$, where each $\sigma_{k,\epsilon}$ describes the upper semicircle of the corresponding $C_{x_k}(\epsilon)$ from $x_k-\epsilon$ to $x_k+\epsilon$. We just take R large enough and ϵ small enough so that the curve $\gamma_{R,\epsilon}$ rotates once and in the positive direction around each of the roots $z_1,\ldots,z_M$ of q in the upper halfplane. Then $\gamma_{R,\epsilon}$ rotates no times around each of the remaining roots of q. The residue theorem implies that

$$\oint_{\gamma_{R,\epsilon}} r(z)\,dz = 2\pi i(\mathrm{Res}(r;z_1)+\cdots+\mathrm{Res}(r;z_M))$$

and hence

$$\begin{aligned} I(R,\epsilon) = 2\pi i(\mathrm{Res}(r;z_1)+\cdots+\mathrm{Res}(r;z_M)) &- \textstyle\int_{\sigma_R} r(z)\,dz \\ - \textstyle\int_{\sigma_{1,\epsilon}} r(z)\,dz - \cdots &- \textstyle\int_{\sigma_{N,\epsilon}} r(z)\,dz. \end{aligned} \tag{6.24}$$

Now, x_k is a pole of r of order 1 and r can be written $r(z)=\frac{c_k}{z-x_k}+f_k(z)$ for $z\neq x_k$ in a disc with center x_k, where f_k is holomorphic at x_k and $c_k=\mathrm{Res}(r;x_k)$. Since f_k is bounded in a disc with center x_k, there is $M_k\geq 0$ and $\epsilon'_k>0$ so that $|f_k(z)|\leq M_k$ for $|z-x_k|\leq\epsilon'_k$. Thus, $0<\epsilon\leq\epsilon'_k$ implies $|\int_{\sigma_{k,\epsilon}} f_k(z)\,dz|\leq M_k\pi\epsilon$ and hence $\lim_{\epsilon\to 0+}\int_{\sigma_{k,\epsilon}} f_k(z)\,dz=0$. Therefore,

$$\begin{aligned} \textstyle\int_{\sigma_{k,\epsilon}} r(z)\,dz &= c_k \textstyle\int_{\sigma_{k,\epsilon}} \frac{1}{z-x_k}\,dz + \int_{\sigma_{k,\epsilon}} f_k(z)\,dz \\ &= -\pi i c_k + \textstyle\int_{\sigma_{k,\epsilon}} f_k(z)\,dz \to -\pi i c_k \end{aligned} \tag{6.25}$$

when $\epsilon\to 0+$. The limit of $\int_{\sigma_R} r(z)\,dz$ when $R\to+\infty$ has been evaluated in the previous two examples:

$$\lim_{R\to+\infty}\textstyle\int_{\sigma_R} r(z)\,dz = \begin{cases} 0, & \text{if } m\geq n+2 \\ i\pi\,\frac{a_n}{b_{n+1}}, & \text{if } m=n+1 \end{cases} \tag{6.26}$$

Now, (6.23), (6.24), (6.25) and (6.26) imply

$$\begin{aligned} \mathrm{pv}\textstyle\int_{-\infty}^{+\infty} r(x)\,dx &= 2\pi i(\mathrm{Res}(r;z_1)+\cdots+\mathrm{Res}(r;z_M)) \\ &\quad + \pi i(\mathrm{Res}(r;x_1)+\cdots+\mathrm{Res}(r;x_N)) - \begin{cases} 0, & \text{if } m\geq n+2 \\ i\pi\,\frac{a_n}{b_{n+1}}, & \text{if } m=n+1 \end{cases} \end{aligned}$$

Example 6.6.4. *Evaluation of $\int_{-\infty}^{+\infty} r(x)\cos x\,dx$, $\int_{-\infty}^{+\infty} r(x)\sin x\,dx$ (or of their principal values), where $r=\frac{p}{q}$ is a rational function,* $\deg q\geq\deg p+1$, *the real roots of q (if they exist) have multiplicity 1 and the coefficients of p,q are real numbers.*

Since the coefficients of p,q are real, we have that $r(x)\in\mathbb{R}$ for every $x\in\mathbb{R}$ which is not a root of q. Hence,

$$\textstyle\int_{-\infty}^{+\infty} r(x)\cos x\,dx = \mathrm{Re}\int_{-\infty}^{+\infty} r(x)e^{ix}\,dx, \qquad \int_{-\infty}^{+\infty} r(x)\sin x\,dx = \mathrm{Im}\int_{-\infty}^{+\infty} r(x)e^{ix}\,dx$$

and we evaluate $\int_{-\infty}^{+\infty} r(x)e^{ix}\,dx$ (or its principal value).

The method of evaluation has been described already in the previous three examples. We use either the curve γ_R or the curve $\gamma_{R,\epsilon}$ and we evaluate the residues of $r(z)e^{iz}$ at the roots of q.

We shall concentrate on the important specific generalized integral

$$\textstyle\int_0^{+\infty} \frac{\sin x}{x}\,dx = \frac{1}{2}\int_{-\infty}^{+\infty}\frac{\sin x}{x}\,dx.$$

(Equality holds because $\frac{\sin x}{x}$ is even.) We shall evaluate $\mathrm{pv}\int_{-\infty}^{+\infty}\frac{e^{ix}}{x}\,dx$ instead of $\int_{-\infty}^{+\infty}\frac{\sin x}{x}\,dx$. Observe that $\frac{e^{ix}}{x}=\frac{\cos x}{x}+i\frac{\sin x}{x}$ diverges at 0 because its real part $\frac{\cos x}{x}$ diverges at 0. The imaginary part $\frac{\sin x}{x}$ converges at 0 and, in fact, if we define $\frac{\sin x}{x}$ at 0 to have value $\lim_{x\to 0}\frac{\sin x}{x}=1$, then it

becomes continuous at 0.
The function $\frac{e^{iz}}{z}$ is holomorphic in $\mathbb{C}$ except for a pole at 0 of order 1. We consider the closed curve $\gamma_{R,\epsilon}$ which is the sum of the linear segments $[-R,-\epsilon]$ and $[\epsilon,R]$, of the curve σ_R, which describes the upper semicircle of $C_0(R)$ from R to $-R$, and of the curve σ_ϵ, which describes the upper semicircle of $C_0(\epsilon)$ from $-\epsilon$ to ϵ. Then $\gamma_{R,\epsilon}$ does not rotate around the pole 0 of $\frac{e^{iz}}{z}$. The residue theorem implies $\oint_{\gamma_{R,\epsilon}} \frac{e^{iz}}{z}\,dz = 0$ and hence

$$\int_{-R}^{-\epsilon} \frac{e^{ix}}{x}\,dx + \int_{\epsilon}^{R} \frac{e^{ix}}{x}\,dx = -\int_{\sigma_R} \frac{e^{iz}}{z}\,dz - \int_{\sigma_\epsilon} \frac{e^{iz}}{z}\,dz. \tag{6.27}$$

Now,

$$\int_{\sigma_R} \frac{e^{iz}}{z}\,dz = \int_0^\pi \frac{e^{iRe^{it}}}{Re^{it}}\, iRe^{it}\,dt = i\int_0^\pi e^{-R\sin t + iR\cos t}\,dt$$

and

$$\begin{aligned} \Big|\int_{\sigma_R} \frac{e^{iz}}{z}\,dz\Big| &\le \int_0^\pi e^{-R\sin t}\,dt = 2\int_0^{\pi/2} e^{-R\sin t}\,dt \le 2\int_0^{\pi/2} e^{-\frac{2R}{\pi}t}\,dt \\ &= \tfrac{\pi}{R}(1-e^{-R}) \to 0 \end{aligned} \tag{6.28}$$

when $R \to +\infty$. For the second inequality we used the well known inequality $\sin t \ge \frac{2t}{\pi}$ for $0 \le t \le \frac{\pi}{2}$. From the Laurent series of $\frac{e^{iz}}{z}$ at 0 we see that $\frac{e^{iz}}{z} = \frac{1}{z} + h(z)$ for $z \ne 0$, where h is holomorphic in $\mathbb{C}$. Now, h is bounded in a neighborhood of 0, i.e. there is $M \ge 0$ so that $|h(z)| \le 1$ when $|z| \le 1$. Hence, for $\epsilon \le 1$ we have $|\int_{\sigma_\epsilon} h(z)\,dz| \le M\pi\epsilon \to 0$ when $\epsilon \to 0+$. Therefore

$$\begin{aligned} \int_{\sigma_\epsilon} \frac{e^{iz}}{z}\,dz &= \int_{\sigma_\epsilon} \frac{1}{z}\,dz + \int_{\sigma_\epsilon} h(z)\,dz \\ &= -\pi i + \int_{\sigma_\epsilon} h(z)\,dz \to -\pi i \end{aligned} \tag{6.29}$$

when $\epsilon \to 0+$. From (6.27), (6.28) and (6.29):

$$\mathrm{pv}\int_{-\infty}^{+\infty} \frac{e^{ix}}{x}\,dx = \lim_{\epsilon\to 0+, R\to+\infty} \Big(\int_{-R}^{-\epsilon} \frac{e^{ix}}{x}\,dx + \int_{\epsilon}^{R} \frac{e^{ix}}{x}\,dx\Big) = \pi i.$$

Since $\frac{\cos x}{x}$ is odd and $\frac{\sin x}{x}$ is even, we get $\int_{-R}^{-\epsilon} \frac{e^{ix}}{x}\,dx + \int_{\epsilon}^{R} \frac{e^{ix}}{x}\,dx = 2i\int_{\epsilon}^{R} \frac{\sin x}{x}\,dx$ and hence

$$\int_0^{+\infty} \frac{\sin x}{x}\,dx = \lim_{\epsilon\to 0+, R\to+\infty} \int_{\epsilon}^{R} \frac{\sin x}{x}\,dx = \frac{\pi}{2}.$$

Example 6.6.5. We shall evaluate $\int_0^{+\infty} \frac{\ln x}{x^2+4}\,dx$.
We consider the holomorphic branch of the logarithm, which we shall denote $\log z$, in the open region $\Omega = \mathbb{C} \setminus \{iy \,|\, y \le 0\}$ and which takes the value 0 at 1. This branch is given by

$$\log z = \ln r + i\theta$$

for $z = re^{i\theta}$ with $r > 0$ and $-\frac{\pi}{2} < \theta < \frac{3\pi}{2}$. The function $\frac{\log z}{z^2+4}$ is holomorphic in Ω except for the point $2i$ which is a pole of order 1. Indeed, we write $\frac{\log z}{z^2+4} = \frac{(\log z)/(z+2i)}{z-2i} = \frac{g(z)}{z-2i}$ and we have that $g(z) = \frac{\log z}{z+2i}$ is holomorphic in Ω with $g(2i) = \frac{\pi}{8} - \frac{\ln 2}{4}\,i$. Moreover, $\mathrm{Res}(\frac{\log z}{z^2+4}; 2i) = g(2i) = \frac{\pi}{8} - \frac{\ln 2}{4}\,i$. Now we consider the closed curve $\gamma_{R,\epsilon}$ of the previous example. We take R large enough and ϵ small enough so that $\gamma_{R,\epsilon}$ rotates once and in the positive direction around the pole $2i$. From the residue theorem we have that

$$\oint_{\gamma_{R,\epsilon}} \frac{\log z}{z^2+4}\,dz = 2\pi i\,\mathrm{Res}(\tfrac{\log z}{z^2+4}; 2i) = \frac{\pi\ln 2}{2} + \frac{\pi^2}{4}\,i.$$

Taking real parts of both sides, we find

$$2\int_{\epsilon}^{R} \frac{\ln x}{x^2+4}\,dx = \frac{\pi\ln 2}{2} - \mathrm{Re}\int_{\sigma_R} \frac{\log z}{z^2+4}\,dz - \mathrm{Re}\int_{\sigma_\epsilon} \frac{\log z}{z^2+4}\,dz.$$

Now,

$$\Big|\int_{\sigma_R} \frac{\log z}{z^2+4}\,dz\Big| \le \frac{\ln R+\pi}{R^2-4}\,\pi R \to 0, \quad \Big|\int_{\sigma_\epsilon} \frac{\log z}{z^2+4}\,dz\Big| \le \frac{\ln\epsilon+\pi}{4-\epsilon^2}\,\pi\epsilon \to 0$$

when $R \to +\infty$ and $\epsilon \to 0+$. Hence

$$\int_0^{+\infty} \frac{\ln x}{x^2+4}\,dx = \lim_{\epsilon\to 0+, R\to+\infty} \int_{\epsilon}^{R} \frac{\ln x}{x^2+4}\,dx = \frac{\pi\ln 2}{4}.$$

Example 6.6.6. We shall evaluate $\int_0^{+\infty} \frac{x^{a-1}}{x+1}\,dx$ when $0 < a < 1$.
We write x^2 instead of x:

$$\int_0^{+\infty} \frac{x^{a-1}}{x+1}\,dx = 2\int_0^{+\infty} \frac{x^{2a-1}}{x^2+1}\,dx = 2\int_0^{+\infty} \frac{x^b}{x^2+1}\,dx$$

with $b = 2a-1$ and $-1 < b < 1$.
We consider the holomorphic branch $\log z$ of the previous example in the same region Ω. The function $h(z) = e^{b\log z}$ is holomorphic in Ω and, if $z = x > 0$, we have $h(x) = e^{b\ln x} = x^b$. The function $\frac{h(z)}{z^2+1}$ is holomorphic in Ω except for a pole at i of order 1. Indeed, we write $\frac{h(z)}{z^2+1} = \frac{h(z)/(z+i)}{z-i} = \frac{g(z)}{z-i}$ and we have that $g(z) = \frac{h(z)}{z+i}$ is holomorphic in Ω with $g(i) = \frac{h(i)}{2i} = \frac{e^{\frac{b\pi}{2}i}}{2i}$. Moreover, $\mathrm{Res}(\frac{h(z)}{z^2+1}; i) = g(i) = \frac{e^{\frac{b\pi}{2}i}}{2i}$. Now we consider the same closed curve $\gamma_{R,\epsilon}$ of the previous example. The residue theorem implies

$$\oint_{\gamma_{R,\epsilon}} \frac{h(z)}{z^2+1}\,dz = 2\pi i\,\mathrm{Res}(\tfrac{h(z)}{z^2+1}; i) = \pi e^{\frac{b\pi}{2}i},$$

and hence

$$(e^{b\pi i}+1)\int_\epsilon^R \frac{x^b}{x^2+1}\,dx = \pi e^{\frac{b\pi}{2}i} - \int_{\sigma_R} \frac{h(z)}{z^2+1}\,dz - \int_{\sigma_\epsilon} \frac{h(z)}{z^2+1}\,dz.$$

Now

$$\Big|\int_{\sigma_R} \frac{h(z)}{z^2+1}\,dz\Big| \le \frac{R^b}{R^2-1}\,\pi R \to 0, \quad \Big|\int_{\sigma_\epsilon} \frac{h(z)}{z^2+1}\,dz\Big| \le \frac{\epsilon^b}{1-\epsilon^2}\,\pi\epsilon \to 0$$

when $R \to +\infty$ and $\epsilon \to 0+$. Hence

$$\int_0^{+\infty} \frac{x^{a-1}}{x+1}\,dx = 2\int_0^{+\infty} \frac{x^b}{x^2+1}\,dx = \lim_{\epsilon\to 0+, R\to+\infty} 2\int_\epsilon^R \frac{x^b}{x^2+1}\,dx = \frac{2\pi e^{\frac{b\pi}{2}i}}{e^{b\pi i}+1} = \frac{\pi}{\sin a\pi}.$$

We shall evaluate $\int_0^{+\infty} \frac{x^{a-1}}{x+1}\,dx$ in a different way.
We consider the holomorphic branch of the logarithm, which we shall denote $\log z$ again, in the (different) region $\Omega = \mathbb{C} \setminus \{x \,|\, x \ge 0\}$ and which takes the value $i\pi$ at -1. This branch is given by

$$\log z = \ln r + i\theta$$

for $z = re^{i\theta}$ with $r > 0$ and $0 < \theta < 2\pi$. The function $h(z) = e^{(a-1)\log z}$ is holomorphic in Ω, and hence $\frac{h(z)}{z+1}$ is holomorphic in Ω except at the point -1 which is a pole of order 1. Indeed, we have $\mathrm{Res}(\frac{h(z)}{z+1}; -1) = h(-1) = e^{(a-1)\pi i}$. We also consider the closed curve $\gamma_{R,\epsilon,\delta}$ which is the sum of the curve $\sigma_{R,\delta}$, which describes the arc of $C_0(R)$ from $Re^{i\delta}$ to $Re^{i(2\pi-\delta)}$ in the positive direction, of the curve $\sigma_{\epsilon,\delta}$, which describes the arc of $C_0(\epsilon)$ from $\epsilon e^{i(2\pi-\delta)}$ to $\epsilon e^{i\delta}$ in the negative direction, of the linear segment $[\epsilon e^{i\delta}, Re^{i\delta}]$ and of the linear segment $[Re^{i(2\pi-\delta)}, \epsilon e^{i(2\pi-\delta)}]$. The residue theorem implies that

$$\oint_{\gamma_{R,\epsilon,\delta}} \frac{h(z)}{z+1}\,dz = 2\pi i\,\mathrm{Res}(\tfrac{h(z)}{z+1}; -1) = 2\pi i e^{(a-1)\pi i}$$

and hence

$$\int_{[\epsilon e^{i\delta}, Re^{i\delta}]} \frac{h(z)}{z+1}\,dz + \int_{[Re^{i(2\pi-\delta)}, \epsilon e^{i(2\pi-\delta)}]} \frac{h(z)}{z+1}\,dz$$
$$= 2\pi i e^{(a-1)\pi i} - \int_{\sigma_{R,\delta}} \frac{h(z)}{z+1}\,dz - \int_{\sigma_{\epsilon,\delta}} \frac{h(z)}{z+1}\,dz.$$

Now, $\big|\int_{\sigma_{R,\delta}} \frac{h(z)}{z+1}\,dz\big| \le \frac{2\pi R^a}{R-1}$ and $\big|\int_{\sigma_{\epsilon,\delta}} \frac{h(z)}{z+1}\,dz\big| \le \frac{2\pi\epsilon^a}{1-\epsilon}$. Therefore

$$\Big|\int_{[\epsilon e^{i\delta}, Re^{i\delta}]} \frac{h(z)}{z+1}\,dz + \int_{[Re^{i(2\pi-\delta)}, \epsilon e^{i(2\pi-\delta)}]} \frac{h(z)}{z+1}\,dz - 2\pi i e^{(a-1)\pi i}\Big| \le \frac{2\pi R^a}{R-1} + \frac{2\pi\epsilon^a}{1-\epsilon}. \tag{6.30}$$

We have

$$\int_{[\epsilon e^{i\delta}, Re^{i\delta}]} \frac{h(z)}{z+1}\,dz = e^{ia\delta}\int_\epsilon^R \frac{r^{a-1}}{re^{i\delta}+1}\,dr.$$

Keeping ϵ and R fixed, we take the limit when $\delta \to 0+$. Clearly, $e^{ia\delta} \to 1$. Also, $\frac{1}{re^{i\delta}+1} \to \frac{1}{r+1}$ uniformly in $[\epsilon, R]$ and hence

$$\int_{[\epsilon e^{i\delta}, Re^{i\delta}]} \frac{h(z)}{z+1}\, dz \to \int_\epsilon^R \frac{r^{a-1}}{r+1}\, dr \tag{6.31}$$

when $\delta \to 0+$. We also have

$$\int_{[Re^{i(2\pi-\delta)}, \epsilon e^{i(2\pi-\delta)}]} \frac{h(z)}{z+1}\, dz = -e^{ia(2\pi-\delta)} \int_\epsilon^R \frac{r^{a-1}}{re^{-i\delta}+1}\, dr.$$

Keeping ϵ and R fixed, we take the limit when $\delta \to 0+$. Exactly as with (6.31), we get

$$\int_{[Re^{i(2\pi-\delta)}, \epsilon e^{i(2\pi-\delta)}]} \frac{h(z)}{z+1}\, dz \to -e^{i2a\pi} \int_\epsilon^R \frac{r^{a-1}}{r+1}\, dr \tag{6.32}$$

when $\delta \to 0+$. From (6.30), (6.31) and (6.32) we get

$$\left|(1 - e^{i2a\pi}) \int_\epsilon^R \frac{r^{a-1}}{r+1}\, dr - 2\pi i e^{(a-1)\pi i}\right| \leq \frac{2\pi R^a}{R-1} + \frac{2\pi\epsilon^a}{1-\epsilon}.$$

Finally, we let $\epsilon \to 0+$ and $R \to +\infty$ and we conclude that

$$\int_0^{+\infty} \frac{x^{a-1}}{x+1}\, dx = \lim_{\epsilon\to 0+, R\to+\infty} \int_\epsilon^R \frac{r^{a-1}}{r+1}\, dr = \frac{2\pi i e^{(a-1)\pi i}}{1-e^{i2a\pi}} = \frac{\pi}{\sin a\pi}.$$

Example 6.6.7. *Evaluation of $\int_0^{2\pi} r(\cos\theta, \sin\theta)\, d\theta$, where $r(s,t)$ is a rational function of two variables.*
We parametrize $C_0(1)$ with $z = e^{i\theta}$, $\theta \in [0, 2\pi]$, and we have $\cos\theta = \frac{1}{2}(z + \frac{1}{z})$, $\sin\theta = \frac{1}{2i}(z - \frac{1}{z})$ and $\frac{dz}{d\theta} = ie^{i\theta} = iz$. Hence

$$\int_0^{2\pi} r(\cos\theta, \sin\theta)\, d\theta = \frac{1}{i} \oint_{C_0(1)} r\left(\frac{z^2+1}{2z}, \frac{z^2-1}{2iz}\right) \frac{1}{z}\, dz.$$

The function $s(z) = r(\frac{z^2+1}{2z}, \frac{z^2-1}{2iz})\frac{1}{z}$ is a rational function of z. We apply the residue theorem after we evaluate the residues of s at its poles in the disc $D_0(1)$.

Exercises.

6.6.1. Evaluate

$$\int_{-\infty}^{+\infty} \frac{1}{x^2+1}\, dx, \quad \int_{-\infty}^{+\infty} \frac{1}{(x^2+1)(x^2+4)}\, dx, \quad \int_{-\infty}^{+\infty} \frac{1}{(x^2+1)^2}\, dx, \quad \int_{-\infty}^{+\infty} \frac{x^4}{1+x^8}\, dx,$$

$$\text{pv}\int_{-\infty}^{+\infty} \frac{x+1}{x^2+1}\, dx, \quad \text{pv}\int_{-\infty}^{+\infty} \frac{x^3}{x^4-4x^2+5}\, dx, \quad \text{pv}\int_{-\infty}^{+\infty} \frac{x^2+3}{x(x^2+1)}\, dx,$$

$$\int_{-\infty}^{+\infty} \frac{\cos x}{(x^2+1)(x^2+4)}\, dx, \quad \int_{-\infty}^{+\infty} \frac{x^3 \sin x}{x^4+1}\, dx, \quad \text{pv}\int_{-\infty}^{+\infty} \frac{\cos x}{x(x^2+1)}\, dx,$$

$$\int_0^{2\pi} \frac{1}{(1-a\cos\theta)^2}\, d\theta \ (0 < a < 1), \quad \int_0^{2\pi} \frac{\cos 2\theta}{1-2a\cos\theta+a^2}\, d\theta \ (0 < a < 1),$$

$$\int_0^{\pi/2} \frac{1}{a+\sin^2\theta}\, d\theta \ (|a| > 1), \quad \int_0^{+\infty} \frac{x^a}{x^2+3x+2}\, dx \ (|a| < 1),$$

$$\int_0^{+\infty} \frac{\ln x}{(x^2+1)(x^2+4)}\, dx, \quad \int_0^{+\infty} \frac{\ln^2 x}{x^2+1}\, dx, \quad \int_0^{+\infty} \frac{\ln(1+x^2)}{x^{1+a}}\, dx \ (0 < a < 2),$$

$$\int_{-\infty}^{+\infty} \frac{\cos x}{e^x+e^{-x}}\, dx, \quad \int_0^{+\infty} \frac{1}{x^3+8}\, dx, \quad \int_0^{+\infty} \frac{x}{x^4+16}\, dx, \quad \int_0^{2\pi} \frac{1}{2+\cos\theta}\, d\theta.$$

6.7 The argument principle. The theorem of Rouché.

A function f is called **meromorphic** in the open set Ω if it is holomorphic in Ω except at certain points in Ω which are poles of f.

Let f be meromorphic in the open set Ω. If $w \in \mathbb{C}$, we shall denote A_w the set of solutions of $f(z) = w$, i.e.

$$A_w = \{z \in \Omega \mid f(z) = w\}.$$

If f is not constant in any connected component of Ω, then the solutions of $f(z) = w$ are isolated points.

Also, letting f have the value ∞ at each of its poles in Ω, so that f becomes continuous at its poles considered as a function from Ω to $\widehat{\mathbb{C}}$, we denote A_∞ the set of solutions of $f(z) = \infty$, i.e.

$$A_\infty = \{z \in \Omega \mid f(z) = \infty\} = \{z \in \Omega \mid z \text{ is a pole of } f\}.$$

The argument principle. *Let $w \in \mathbb{C}$. We assume that f is meromorphic in the open set Ω and that it is not constant in any connected component of Ω. Also let Σ be a cycle, which consists of closed piecewise smooth curves and which is null-homologous in Ω, so that no element of $A_w \cup A_\infty$ is in the trajectory of any of the closed curves forming Σ. Then $n(\Sigma; z) \neq 0$ for at most finitely many elements of $A_w \cup A_\infty$ and so the sums*

$$\textstyle\sum_{z \in A_w} n(\Sigma; z)\, m(z), \quad \sum_{z \in A_\infty} n(\Sigma; z)\, m(z),$$

where $m(z)$ is the corresponding multiplicity of $z \in A_w \cup A_\infty$, are finite. Moreover,

$$n(f(\Sigma); w) = \tfrac{1}{2\pi i} \oint_\Sigma \tfrac{f'(\zeta)}{f(\zeta) - w}\, d\zeta = \textstyle\sum_{z \in A_w} n(\Sigma; z)\, m(z) - \sum_{z \in A_\infty} n(\Sigma; z)\, m(z). \tag{6.33}$$

Furthermore, even if the closed curves which form Σ are not necessarily piecewise smooth, then the left and the right side of (6.33) are still equal.

Proof. At first we assume that the closed curves forming Σ are all piecewise continuous.
We apply the residue theorem to the function $\frac{f'}{f-w}$. The isolated singularities of this function are the elements of $A_w \cup A_\infty$.
If $m(z)$ is the multiplicity of $z \in A_w$, then there is a g holomorphic in some neighborhood $D_z(r)$ of z so that $f(\zeta) - w = (\zeta - z)^{m(z)} g(\zeta)$ when $\zeta \in D_z(r)$ and also $g(z) \neq 0$. Since $g(z) \neq 0$, we may assume that r is small enough so that $g(\zeta) \neq 0$ when $\zeta \in D_z(r)$. Therefore

$$\tfrac{f'(\zeta)}{f(\zeta) - w} = \tfrac{m(z)}{\zeta - z} + \tfrac{g'(\zeta)}{g(\zeta)}$$

when $\zeta \in D_z(r) \setminus \{z\}$. Since $\frac{g'}{g}$ is holomorphic in $D_z(r)$, we have that z is a pole of $\frac{f'}{f-w}$ of order 1 with residue $m(z)$.
If $m(z)$ is the order of $z \in A_\infty$, there is a g holomorphic in some neighborhood $D_z(r)$ of z so that $f(\zeta) - w = \frac{g(\zeta)}{(\zeta - z)^{m(z)}}$ when $\zeta \in D_z(r)$ and also $g(z) \neq 0$. Since $g(z) \neq 0$, we may assume that r is small enough so that $g(\zeta) \neq 0$ when $\zeta \in D_z(r)$. Hence

$$\tfrac{f'(\zeta)}{f(\zeta) - w} = \tfrac{-m(z)}{\zeta - z} + \tfrac{g'(\zeta)}{g(\zeta)}$$

when $\zeta \in D_z(r) \setminus \{z\}$. Since $\frac{g'}{g}$ is holomorphic in $D_z(r)$, we have that z is a pole of $\frac{f'}{f-w}$ of order 1 with residue $-m(z)$.
Now, the residue theorem implies the second equality in (6.33). The first equality is a matter of a simple change of variable. If $\zeta = \gamma(t)$, $t \in [a, b]$, is the parametric equation of any curve γ forming Σ, then the parametric equation of $f(\gamma)$ is $\eta = f(\gamma(t))$, $t \in [a, b]$, and hence:

$$n(f(\gamma); w) = \tfrac{1}{2\pi i} \oint_{f(\gamma)} \tfrac{1}{\eta - w}\, d\eta = \tfrac{1}{2\pi i} \int_a^b \tfrac{f'(\gamma(t))\gamma'(t)}{f(\gamma(t)) - w}\, dt = \tfrac{1}{2\pi i} \oint_\gamma \tfrac{f'(\zeta)}{f(\zeta) - w}\, d\zeta.$$

The rest is simple if we recall that $\Sigma = n_1\gamma_1 + \cdots + n_k\gamma_k$ and $f(\Sigma) = n_1 f(\gamma_1) + \cdots + n_k f(\gamma_k)$.
Now we assume that the curves γ which form Σ are not necessarily piecewise smooth.
We consider any of the closed curves which form Σ with parametric equation $\zeta = \gamma(t)$, $t \in [a,b]$, and the corresponding $f(\gamma)$ with parametric equation $\eta = f(\gamma(t))$, $t \in [a,b]$. The set $A_w \cup A_\infty$ has no accumulation point in Ω. Thus, the set $A_w \cup A_\infty \cup \Omega^c$ is closed and we also have that it is disjoint from γ^*. Therefore, there is $\epsilon_1 > 0$ so that

$$|\gamma(t) - z| \geq 2\epsilon_1 \tag{6.34}$$

for every $t \in [a,b]$ and every $z \in A_w \cup A_\infty \cup \Omega^c$. We consider the set

$$K = \{z \,|\, |z - \gamma(t)| \leq \epsilon_1 \text{ for at least one } t \in [a,b]\}$$

and we easily see that K is a compact subset of $\Omega \setminus (A_w \cup A_\infty)$ and hence f is continuous in K. Also, we have $f(z) \neq w$ for every $z \in K$ and γ^* is a subset of K and hence there is $\epsilon_2 > 0$ so that

$$|f(\gamma(t)) - w| \geq \epsilon_2 \tag{6.35}$$

for every $t \in [a,b]$. Since f is continuous in K, there is δ_1 with $0 < \delta_1 \leq \epsilon_1$ so that

$$|f(z') - f(z'')| < \epsilon_2 \tag{6.36}$$

for every $z', z'' \in K$ with $|z' - z''| < \delta_1$. Finally, there is $\delta > 0$ so that

$$|\gamma(t') - \gamma(t'')| < \delta_1 \tag{6.37}$$

for every $t', t'' \in [a,b]$ with $|t' - t''| < \delta$.
Now we take successive points $a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b$ so that $t_k - t_{k-1} < \delta$ for every k and we consider the polygonal curve $\sigma : [a,b] \to \mathbb{C}$ consisting of the successive linear segments $[\gamma(t_{k-1}), \gamma(t_k)]$. It is easy to see that we have

$$|\sigma(t) - \gamma(t)| < \delta_1 \leq \epsilon_1 \tag{6.38}$$

for every $t \in [a,b]$. Indeed, if $t \in [t_{k-1}, t_k]$, then, because of (6.37), we have

$$\begin{aligned}
|\sigma(t) - \gamma(t)| &= \big|\big(\tfrac{t_k - t}{t_k - t_{k-1}}\gamma(t_{k-1}) + \tfrac{t - t_{k-1}}{t_k - t_{k-1}}\gamma(t_k)\big) - \gamma(t)\big| \\
&\leq \tfrac{t_k - t}{t_k - t_{k-1}}|\gamma(t_{k-1}) - \gamma(t)| + \tfrac{t - t_{k-1}}{t_k - t_{k-1}}|\gamma(t_k) - \gamma(t)| \\
&< \tfrac{t_k - t}{t_k - t_{k-1}}\delta_1 + \tfrac{t - t_{k-1}}{t_k - t_{k-1}}\delta_1 = \delta_1 \leq \epsilon_1.
\end{aligned}$$

Now, (6.34), (6.38) imply

$$|\sigma(t) - \gamma(t)| < |\gamma(t) - z|$$

for every $t \in [a,b]$ and every $z \in A_w \cup A_\infty$. Proposition 6.14 implies $n(\gamma; z) = n(\sigma; z)$ for every $z \in A_w \cup A_\infty$ and hence

$$\begin{aligned}
\textstyle\sum_{z \in A_w} n(\gamma; z)\, m(z) - &\textstyle\sum_{z \in A_\infty} n(\gamma; z)\, m(z) \\
&= \textstyle\sum_{z \in A_w} n(\sigma; z)\, m(z) - \sum_{z \in A_\infty} n(\sigma; z)\, m(z).
\end{aligned} \tag{6.39}$$

Also, (6.38) implies $\sigma(t) \in K$ for every $t \in [a,b]$ and, because of (6.36),

$$|f(\sigma(t)) - f(\gamma(t))| < \epsilon_2$$

for every $t \in [a,b]$. But then (6.35) implies

$$|f(\sigma(t)) - f(\gamma(t))| < |f(\gamma(t)) - w|$$

for every $t \in [a, b]$. Proposition 6.14 again implies

$$n(f(\gamma); w) = n(f(\sigma); w). \tag{6.40}$$

Since the curve σ is piecewise smooth, we have from the first part of the proof that

$$n(f(\sigma); w) = \sum_{z \in A_w} n(\sigma; z)\, m(z) - \sum_{z \in A_\infty} n(\sigma; z)\, m(z). \tag{6.41}$$

Now, (6.39), (6.40) and (6.41) imply the equality of the left and the right side of (6.33) for each γ forming Σ and the proof is finished by addition over all such γ. $\square$

The geometric content of the argument principle is described as follows. *The number of rotations of $f(\Sigma)$ around w is equal to the total number of rotations of Σ around the solutions of $f(z) = w$ minus the total number of rotations of Σ around the poles of f.* When we count the solutions of $f(z) = w$ and the poles of f we take into account their multiplicities. We count $m(z)$ points at every point $z \in A_w \cup A_\infty$ which has multiplicity $m(z)$.

If f has no poles in Ω, i.e. if f is holomorphic in Ω, then the argument principle says that *the number of rotations of $f(\Sigma)$ around w is equal to the total number of rotations of Σ around the solutions of $f(z) = w$.* In fact, if Σ is such that for every z not in the trajectories of the curves forming Σ we have either $n(\Sigma; z) = 1$ or $n(\Sigma; z) = 0$, then *the number of rotations of $f(\Sigma)$ around w is equal to the number of solutions of $f(z) = w$ which are surrounded by Σ.*

The theorem of Rouché. *Let $w \in \mathbb{C}$. We assume that f, g are holomorphic in the open set Ω and that they are not constant in any connected component of Ω. We also consider Σ to be a cycle which is null-homologous in Ω. If $|f(\zeta) - g(\zeta)| < |g(\zeta) - w|$ for every ζ in the trajectories of the closed curves forming Σ, then*

$$\sum_{z \in A_{w,f}} n(\Sigma; z)\, m_f(z) = \sum_{z \in A_{w,g}} n(\Sigma; z)\, m_g(z),$$

where $m_f(z)$ and $m_g(z)$ are the corresponding multiplicities and $A_{w,f} = \{z \in \Omega \,|\, f(z) = w\}$, $A_{w,g} = \{z \in \Omega \,|\, g(z) = w\}$.

Proof. We observe that the condition $|f(\zeta) - g(\zeta)| < |g(\zeta) - w|$ for every ζ in the trajectories of the closed curves forming Σ implies that no element of $A_{w,f} \cup A_{w,g}$ is in these trajectories. The function $h = \frac{f-w}{g-w}$ is holomorphic in Ω except for the elements of $A_{w,g}$, which are either poles or removable singularities of h. From (6.33) we have

$$n(h(\Sigma); 0) = \sum_{z \in A_{0,h}} n(\Sigma; z)\, m_h(z) - \sum_{z \in A_{\infty,h}} n(\Sigma; z)\, m_h(z). \tag{6.42}$$

If $z \in A_{w,f} \setminus A_{w,g}$, then $z \in A_{0,h}$ and $m_h(z) = m_f(z)$. Similarly, if $z \in A_{w,g} \setminus A_{w,f}$, then $z \in A_{\infty,h}$ and $m_h(z) = m_g(z)$. Finally, if $z \in A_{w,f} \cap A_{w,g}$, then we have three cases. If $m_f(z) > m_g(z)$, then $z \in A_{0,h}$ and $m_h(z) = m_f(z) - m_g(z)$. If $m_f(z) < m_g(z)$, then $z \in A_{\infty,h}$ and $m_h(z) = m_g(z) - m_f(z)$. If $m_f(z) = m_g(z)$, then $z \notin A_{0,h} \cup A_{\infty,h}$ and $m_h(z) = 0$. All these imply

$$\begin{aligned}\sum_{z \in A_{0,h}} n(\Sigma; z)\, m_h(z) - &\sum_{z \in A_{\infty,h}} n(\Sigma; z)\, m_h(z) \\ &= \sum_{z \in A_{w,f}} n(\Sigma; z)\, m_f(z) - \sum_{z \in A_{w,g}} n(\Sigma; z)\, m_g(z)\end{aligned}$$

and from (6.42) we get

$$\sum_{z \in A_{w,f}} n(\Sigma; z)\, m_f(z) - \sum_{z \in A_{w,g}} n(\Sigma; z)\, m_g(z) = n(h(\Sigma); 0).$$

Now, our hypothesis says that $|h(z) - 1| < 1$ for every z in the trajectories of the curves forming Σ. Therefore, the cycle $h(\Sigma)$ is in the disc $D_1(1)$ and hence $n(h(\Sigma); 0) = 0$. $\square$

Example 6.7.1. We shall find the number of roots of $f(z) = z^7 - 2z^5 + 6z^3 - z + 1$ in $\mathbb{D}$. We consider $g(z) = 6z^3$ and we have

$$|f(z) - g(z)| = |z^7 - 2z^5 - z + 1| \leq |z|^7 + 2|z|^5 + |z| + 1 = 5 < 6|z|^3 = |g(z)|$$

for every $z \in \mathbb{T}$. Now we apply the theorem of Rouché with $w = 0$ and Σ consisting of only the curve γ which describes $\mathbb{T}$ once and in the positive direction. We have $n(\gamma; z) = 1$ for every $z \in \mathbb{D}$ and $n(\gamma; z) = 0$ for every $z \notin \overline{\mathbb{D}}$. The only solution of $g(z) = 0$ in $\mathbb{D}$ is $z = 0$ with multiplicity $m_g(0) = 3$. Therefore

$$\textstyle\sum_{z \in A_{0,g}} n(\gamma; z)\, m_g(z) = \sum_{z \in A_{0,g} \cap \mathbb{D}} m_g(z) = 3.$$

Moreover,

$$\textstyle\sum_{z \in A_{0,f}} n(\gamma; z)\, m_f(z) = \sum_{z \in A_{0,f} \cap \mathbb{D}} m_f(z).$$

Now the theorem of Rouché implies that $\sum_{z \in A_{0,f} \cap \mathbb{D}} m_f(z) = 3$ and hence f has three roots in $\mathbb{D}$.

Exercises.

6.7.1. Let f be holomorphic in $D_{z_0}(R)$, let $0 < r < R$ and assume that there is no solution of $f(z) = w$ in $C_{z_0}(r)$. If $k \in \mathbb{N}$, describe the content of

$$\tfrac{1}{2\pi i} \oint_{C_{z_0}(r)} \tfrac{f'(z)}{f(z) - w}\, z^k\, dz.$$

6.7.2. Let f be holomorphic in $\mathbb{D}$ and continuous in $\overline{\mathbb{D}}$ and let $|f(z)| < 1$ for every $z \in \mathbb{T}$. Prove that the equation $f(z) = z^n$ has exactly n solutions in $\mathbb{D}$.

6.7.3. Find the number of roots of
(i) $z^4 - 6z + 3$ in $D_0(1, 2)$.
(ii) $z^4 + 8z^3 + 3z^2 + 8z + 3$ in $\{z \mid \mathrm{Re}\, z > 0\}$.

6.7.4. Let $z_1, \dots, z_n \in \mathbb{D}$ and $|\lambda| = 1$. In $\mathbb{C} \setminus \{\frac{1}{\overline{z_1}}, \dots, \frac{1}{\overline{z_k}}\}$ we consider the function

$$B(z) = \lambda \textstyle\prod_{k=1}^{n} \frac{z - z_k}{1 - \overline{z_k}\, z}.$$

We know from exercise 5.9.10 that $B(z) \in \mathbb{D}$ for every $z \in \mathbb{D}$ and that $B(z) \in \mathbb{T}$ for every $z \in \mathbb{T}$.
(i) Find the index with respect to 0 of the curve with parametric equation $w = B(e^{it})$, $t \in [0, 2\pi]$.
(ii) Prove that for every $w \in \mathbb{D}$ the equation $B(z) = w$ has exactly n solutions and all of them are in $\mathbb{D}$.

6.7.5. Prove that the set of all meromorphic functions in the region Ω is an algebraic field.

6.7.6. Let f be holomorphic in the open set Ω. We assume that γ is a closed piecewise smooth curve in Ω, that $\mathbb{C} \setminus \gamma^*$ has only one bounded connected component U and that $n(\gamma; z) = 1$ for every $z \in U$. We also assume that $\mathbb{C} \setminus f(\gamma)^*$ has only one bounded connected component V and that $n(f(\gamma); w) = N$ for every $w \in V$.
(i) If $f(z) \notin f(\gamma)^*$ for every $z \in U$, prove that f is N-to-one from U onto V.
(ii) If moreover $N = 1$, we may consider the inverse function $f^{-1} : V \to U$. Prove that

$$f^{-1}(w) = \tfrac{1}{2\pi i} \oint_{\gamma} \tfrac{\zeta f'(\zeta)}{f(\zeta) - w}\, d\zeta$$

for every $w \in V$.

6.7.7. Let $f(z) = \sum_{n=0}^{+\infty} a_n z^n$ for $z \in \mathbb{D}$ and let $F \subseteq \mathbb{D}$ be compact with $0 \in F$. If m is the number of roots of f in F, prove that $\min_{z \in \partial F} |f(z)| \leq |a_0| + |a_1| + \cdots + |a_m|$.

6.7.8. Let f be holomorphic in $\mathbb{D}$ and continuous in $\overline{\mathbb{D}}$. Assume that the restriction of f in $\mathbb{T}$ is one-to-one and hence the curve $\gamma(t) = f(e^{it})$, $t \in [0, 2\pi]$, is closed and simple. Using Jordan's theorem, prove that f is one-to-one in $\mathbb{D}$ and that it maps $\mathbb{D}$ onto the interior region of γ. Also prove that γ has the positive direction.

Chapter 7

Simply connected regions and the theorem of Riemann.

7.1 Conformal equivalence.

If $\Omega \subseteq \mathbb{C}$ is a region and f is holomorphic and not constant in Ω, then by the open mapping theorem $f(\Omega)$ is also a region.

Proposition 7.1. *Let f be holomorphic and one-to-one in the region $\Omega \subseteq \mathbb{C}$. Then $f(\Omega)$ is also a region, $f'(z) \neq 0$ for every $z \in \Omega$ and f^{-1} is holomorphic in $f(\Omega)$.*

Proof. If $f'(z_0) = 0$ for some $z \in \Omega$, then theorem 5.2 implies that there is $N \geq 2$ so that f is N-to-one in some open set $U \subseteq \Omega$ containing z_0. Hence $f'(z) \neq 0$ for every $z \in \Omega$.
Now let $w_0 \in f(\Omega)$ and consider the unique $z_0 \in \Omega$ so that $f(z_0) = w_0$. Then proposition 5.8 implies that there are two open sets, $U \subseteq \Omega$ and $W \subseteq f(\Omega)$ with $z_0 \in U$ and $w_0 \in W$ so that $f^{-1} : W \to U$ is holomorphic. Thus f^{-1} is holomorphic at every $w_0 \in f(\Omega)$. □

Let f be holomorphic and one-to-one in the region $\Omega \subseteq \mathbb{C}$. Since $f'(z) \neq 0$ for every $z \in \Omega$ and due to the discussion in section 3.3, we say that f is a **conformal mapping** of Ω.

Two regions $\Omega_1, \Omega_2 \subseteq \mathbb{C}$ are called **conformally equivalent** if there is $f : \Omega_1 \to \Omega_2$ holomorphic and one-to-one from Ω_1 onto Ω_2.

If $f : \Omega_1 \to \Omega_2$ is holomorphic and one-to-one from Ω_1 onto Ω_2, then $f^{-1} : \Omega_2 \to \Omega_1$ is also holomorphic and one-to-one from Ω_2 onto Ω_1. It is easy to see that conformal equivalence between regions in $\mathbb{C}$ is an equivalence relation.

The Schwarz lemma. *Let $f : \mathbb{D} \to \mathbb{D}$ be holomorphic in $\mathbb{D}$ and $f(0) = 0$. Then*
(i) $|f(z)| \leq |z|$ for every $z \in \mathbb{D}$,
(ii) $|f'(0)| \leq 1$.
If equality holds in (i) for at least one $z \in \mathbb{D} \setminus \{0\}$ or in (ii), then there is a constant c with $|c| = 1$ so that $f(z) = cz$ for every $z \in \mathbb{D}$.

Proof. Since $f(0) = 0$, the function $\frac{f(z)}{z}$ has a removable singularity at 0 and we may define the function g by

$$g(z) = \begin{cases} \frac{f(z)}{z}, & \text{if } z \in \mathbb{D}, z \neq 0 \\ f'(0), & \text{if } z = 0 \end{cases}$$

Then g is holomorphic in $\mathbb{D}$.
We take any $z \in \mathbb{D}$ and we take any r so that $|z| < r < 1$. By the maximum principle we have

$$|g(z)| \leq \max_{\zeta \in C_0(r)} |g(\zeta)| = \max_{\zeta \in C_0(r)} \frac{|f(\zeta)|}{|\zeta|} \leq \frac{1}{r}.$$

Hence, $|g(z)| \le \frac{1}{r}$ and since this is true for every r with $|z| < r < 1$, we conclude that $|g(z)| \le 1$. Of course this implies (i) and (ii).
Now, assume that equality holds in (i) for at least one $z \in \mathbb{D} \setminus \{0\}$ or in (ii). Then $|g(z)| = 1$ for at least one $z \in \mathbb{D}$ and the maximum principle implies that g is a constant c in $\mathbb{D}$ with $|c| = 1$. Hence $f(z) = cz$ for every $z \in \mathbb{D}$. $\square$

Example 7.1.1. Let $z_0 \in \mathbb{D}$ and $|\lambda| = 1$. We consider the function $T : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ given by

$$T(z) = \begin{cases} \lambda \, \frac{z - z_0}{1 - \overline{z_0}\, z}, & \text{if } z \in \mathbb{C},\, z \neq \frac{1}{\overline{z_0}} \\ \infty, & \text{if } z = \frac{1}{\overline{z_0}} \\ -\frac{\lambda}{\overline{z_0}}, & \text{if } z = \infty \end{cases}$$

Then T is a linear fractional transformation and hence it is one-to-one from $\widehat{\mathbb{C}}$ onto $\widehat{\mathbb{C}}$ and holomorphic in $\widehat{\mathbb{C}} \setminus \{\frac{1}{\overline{z_0}}\}$. The inverse function $T^{-1} : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ is given by

$$T^{-1}(w) = \begin{cases} \mu \, \frac{w - w_0}{1 - \overline{w_0}\, w}, & \text{if } w \in \mathbb{C},\, w \neq \frac{1}{\overline{w_0}} \\ \infty, & \text{if } w = \frac{1}{\overline{w_0}} \\ -\frac{\mu}{\overline{w_0}}, & \text{if } w = \infty \end{cases}$$

where $\mu = \frac{1}{\lambda}$ and $w_0 = -\lambda z_0$. Since $|\mu| = 1$ and $w_0 \in \mathbb{D}$, the inverse function T^{-1} is of the same form as T.
For simplicity, we shall follow the same practice as with all l.f.t. and we shall only write

$$T(z) = \lambda \, \frac{z - z_0}{1 - \overline{z_0}\, z},$$

understanding that $T(\frac{1}{\overline{z_0}}) = \infty$ and $T(\infty) = -\frac{\lambda}{\overline{z_0}}$ whenever this is needed.
We easily see that

$$T(\mathbb{D}) = \mathbb{D}, \quad T(\mathbb{T}) = \mathbb{T}.$$

Indeed,

$$1 - |T(z)|^2 = 1 - \frac{|z - z_0|^2}{|1 - \overline{z_0}\, z|^2} = \frac{1 + |z|^2|z_0|^2 - |z|^2 - |z_0|^2}{|1 - \overline{z_0}\, z|^2} = \frac{(1 - |z|^2)(1 - |z_0|^2)}{|1 - \overline{z_0}\, z|^2},$$

which implies that $|T(z)| < 1$ if $|z| < 1$, that $|T(z)| = 1$ if $|z| = 1$ and that $|T(z)| > 1$ if $|z| > 1$. Thus, $T(\mathbb{D}) \subseteq \mathbb{D}$, $T(\mathbb{T}) \subseteq \mathbb{T}$ and $T(\widehat{\mathbb{C}} \setminus \overline{\mathbb{D}}) \subseteq \widehat{\mathbb{C}} \setminus \overline{\mathbb{D}}$. But, since T is onto $\widehat{\mathbb{C}}$, all these inclusions are equalities.
Another simple property of T is

$$T(z_0) = 0.$$

We also have

$$T'(z) = \lambda \, \frac{1 - |z_0|^2}{(1 - \overline{z_0}\, z)^2}$$

for every $z \neq \frac{1}{\overline{z_0}}$. Thus, $T'(z_0) = \frac{\lambda}{1 - |z_0|^2}$ and hence

$$\operatorname{Arg} T'(z_0) = \operatorname{Arg} \lambda.$$

If we restrict T in $\mathbb{D}$ we see that T is a conformal mapping of $\mathbb{D}$ onto $\mathbb{D}$.
All functions T are called **Möbius transformations**.

The next proposition describes all conformal mappings of $\mathbb{D}$ onto $\mathbb{D}$: they are just the Möbius transformations.

Proposition 7.2. *Let $z_0 \in \mathbb{D}$ and $\theta_0 \in (-\pi, \pi]$. Then the function $T : \mathbb{D} \to \mathbb{D}$ given by*

$$T(z) = e^{i\theta_0} \, \frac{z - z_0}{1 - \overline{z_0}\, z}$$

for every $z \in \mathbb{D}$ is a conformal mapping of $\mathbb{D}$ onto $\mathbb{D}$. Moreover, T is the unique conformal mapping of $\mathbb{D}$ onto $\mathbb{D}$ satisfying

$$T(z_0) = 0, \quad \operatorname{Arg} T'(z_0) = \theta_0.$$

Proof. From the discussion in example 7.1.1 we have all properties of the function T. So we only need to prove the uniqueness of T.
Let S be another conformal mapping of $\mathbb{D}$ onto $\mathbb{D}$ satisfying $S(z_0) = 0$ and $\operatorname{Arg} S'(z_0) = \theta_0$.
Then the function $f = S \circ T^{-1} : \mathbb{D} \to \mathbb{D}$ is holomorphic in $\mathbb{D}$ and satisfies $f(0) = 0$ and $f'(0) = \frac{S'(z_0)}{T'(z_0)} > 0$. By the Schwarz lemma we get $|f'(0)| \leq 1$.
But also the function $g = T \circ S^{-1} : \mathbb{D} \to \mathbb{D}$ is holomorphic in $\mathbb{D}$ and satisfies $g(0) = 0$ and $g'(0) > 0$. Again, by the Schwarz lemma we get $|g'(0)| \leq 1$.
Now, the functions f and g are mutually inverse and hence $g'(0) = \frac{1}{f'(0)}$. Therefore, $|f'(0)| = |g'(0)| = 1$ and the Schwarz lemma implies that there is some c with $|c| = 1$ so that $f(w) = cw$ for every $w \in \mathbb{D}$. Now, $c = f'(0) > 0$ implies $c = 1$. Hence, $f(w) = w$ for every $w \in \mathbb{D}$ and finally $S(z) = T(z)$ for every $z \in \mathbb{D}$. □

Exercises.

7.1.1. Let T, S be two Möbius transformations. Prove that $S \circ T$ is a Möbius transformation.

7.1.2. Let f be a conformal mapping of the region $\Omega \subseteq \mathbb{C}$ onto $\mathbb{D}$ with $f(z_0) = 0$ for some $z_0 \in \Omega$ and let $g : \Omega \to \mathbb{D}$ be holomorphic in Ω with $g(z_0) = 0$. Prove that $|g'(z_0)| \leq |f'(z_0)|$. What can you conclude if $|g'(z_0)| = |f'(z_0)|$?

7.1.3. Let $f : \mathbb{D} \to \mathbb{D}$ be holomorphic in $\mathbb{D}$. Prove that:
(i) $\big|\frac{f(z_1)-f(z_2)}{1-\overline{f(z_2)}f(z_1)}\big| \leq \big|\frac{z_1-z_2}{1-\overline{z_2}z_1}\big|$ for every $z_1, z_2 \in \mathbb{D}$.
(ii) $\frac{|f'(z)|}{1-|f(z)|^2} \leq \frac{1}{1-|z|^2}$ for every $z \in \mathbb{D}$.
Prove that, if equality holds in (i) for at least one pair of $z_1, z_2 \in \mathbb{D}$ with $z_1 \neq z_2$ or in (ii) for at least one $z \in \mathbb{D}$, then f is a Möbius transformation and then equalities in (i) and (ii) hold identically.

7.1.4. (See exercise 7.1.3.) For every piecewise smooth curve $\gamma : [a, b] \to \mathbb{D}$ we define the **hyperbolic length** of γ by

$$l_h(\gamma) = \int_a^b \frac{|\gamma'(t)|}{1-|\gamma(t)|^2}\,dt.$$

(i) If $f : \mathbb{D} \to \mathbb{D}$ is holomorphic in $\mathbb{D}$, and γ is a piecewise smooth curve in $\mathbb{D}$, prove that $l_h(f(\gamma)) \leq l_h(\gamma)$. If, moreover, f is a Möbius transformation, prove that $l_h(f(\gamma)) = l_h(\gamma)$.
(ii) If $z_1, z_2 \in \mathbb{D}$ and $z_1 \neq z_2$, prove that among all piecewise smooth curves in $\mathbb{D}$ with endpoints z_1 and z_2 the one with the smallest hyperbolic length is the arc of the circle which contains z_1, z_2 and which is orthogonal to $\mathbb{T}$. This smallest hyperbolic length is called **hyperbolic distance** of z_1, z_2 and it is equal to

$$d_h(z_1, z_2) = \frac{1}{2}\ln\frac{1+\big|\frac{z_1-z_2}{1-\overline{z_2}z_1}\big|}{1-\big|\frac{z_1-z_2}{1-\overline{z_2}z_1}\big|}.$$

Prove that d_h is a metric in $\mathbb{D}$, the so-called **hyperbolic metric**, which is equivalent to the euclidean metric in $\mathbb{D}$.
(iii) Consider sequences (z'_n) and (z''_n) in $\mathbb{D}$ so that $z'_n \to \zeta$ for some $\zeta \in \mathbb{T}$ and so that $d_h(z'_n, z''_n) \leq M$ for every n. Prove that $z''_n \to \zeta$.

7.1.5. (See exercise 7.1.4.) Let $f : \mathbb{D} \to \mathbb{D}$ be holomorphic in $\mathbb{D}$. Prove that $d_h(f(z_1), f(z_2)) \leq d_h(z_1, z_2)$ for all $z_1, z_2 \in \mathbb{D}$. If, moreover, equality $d_h(f(z_1), f(z_2)) = d_h(z_1, z_2)$ for at least one pair of $z_1, z_2 \in \mathbb{D}$ with $z_1 \neq z_2$, prove that f is a Möbius transformation and then equality $d_h(f(z_1), f(z_2)) = d_h(z_1, z_2)$ holds for all $z_1, z_2 \in \mathbb{D}$.

7.1.6. Find all $f : \mathbb{D} \to \mathbb{D}$ holomorphic in $\mathbb{D}$ with $f(0) = \frac{1}{2}$ and $f'(0) = \frac{3}{4}$.

7.1.7. Prove that for every M, N with $0 < M < N$ there is $P = P(M, N) < N$ with this property: if f is holomorphic in $D_{z_0}(R)$ with $|f(z_0)| < M$ and $|f(z)| < N$ for every $z \in D_{z_0}(R)$, then $|f(z)| < P$ for every $z \in D_{z_0}(\frac{R}{2})$.

7.1.8. Let $f : \mathbb{D} \to \mathbb{D}$ be holomorphic in $\mathbb{D}$ with $f(0) = 0$ and $|f'(0)| < 1$. For every n we define $f_n = \underbrace{f \circ f \circ \cdots \circ f \circ f}_{n \text{ times}}$. Prove that $f_n(z) \to 0$ for every $z \in \mathbb{D}$.

7.2 Simply connected regions and the theorem of Riemann.

Let Ω be a region in $\mathbb{C}$. We say that Ω is **topologically simply connected** if $\widehat{\mathbb{C}}\setminus\Omega$ is connected. We say that Ω is **homologically simply connected** if $n(\gamma; z) = 0$ for every closed curve γ in Ω and every $z \in \Omega^c$. We say that Ω is **homotopically simply connected** if for every closed curve γ in Ω there is a homotopy with closed intermediate curves in Ω from γ to a constant curve.

Example 7.2.1. The region $\Omega = D_{z_0}(R_1, R_2)$ with $0 \leq R_1 < R_2 \leq +\infty$ is not simply connected in any of the three senses.
The set $\widehat{\mathbb{C}} \setminus \Omega = \overline{D}_{z_0}(R_1) \cup \overline{D}_{z_0}(R_2, +\infty) \cup \{\infty\}$ has two connected components.
If $R_1 < r < R_2$, the closed curve γ in Ω which describes the circle $C_{z_0}(r)$ once in the positive direction has $n(\gamma; z_0) = 1$.
For the same closed curve γ in Ω there is no homotopy with closed intermediate curves in Ω from γ to a constant curve. Indeed, if there was such a homotopy from γ to a constant curve γ_1, then proposition 6.17 would imply that $n(\gamma; z_0) = n(\gamma_1; z_0) = 0$, which is wrong.

If Ω is a region, its complement $\widehat{\mathbb{C}} \setminus \Omega$ in the sphere of Riemann is compact with respect to the chordal metric. Therefore, the connected components of $\widehat{\mathbb{C}} \setminus \Omega$ are all compact sets with respect to the chordal metric. One of them contains the point ∞. Every other connected component is a compact subset of $\mathbb{C}$ either with respect to the chordal metric or with respect to the euclidean metric, since the two metrics are equivalent in $\mathbb{C}$. Hence, all connected components of $\widehat{\mathbb{C}} \setminus \Omega$, besides the one which contains ∞, are closed and bounded subsets of $\mathbb{C}$. We continue with some mathematically imprecise thoughts, which may help the understanding of the three notions of simple connectedness. In visually simple cases of regions Ω, like the one in example 7.2.1, the closed and bounded components of the complement of Ω appear as "holes" of Ω. Thus, naively speaking, a region Ω is *topologically simply connected* if it has no "holes". On the other hand, the region Ω is *homologically simply connected* if no closed curve in Ω surrounds any point in the complement of Ω. If Ω has a "hole" then, naively speaking again, one can find a closed curve surrounding the "hole", exactly as in example 7.2.1, and then Ω is not *homologically simply connected*. In the same case and naively speaking again, a closed curve surrounding a "hole" of Ω cannot be shrunk continuously to a point (i.e. to a constant curve) so that all intermediate closed curves are in Ω: it seems that some intermediate curves must intersect the "hole". Thus, Ω is not *homotopically simply connected*.

Example 7.2.2. A set $A \subseteq \mathbb{C}$ is called **star-shaped** if there is a specific $z_0 \in A$ so that $[z_0, z] \subseteq A$ for every $z \in A$. The point z_0 is called center of A.
Now, let Ω be any open star-shaped set and let z_0 be a center of Ω. We consider any $z \in \Omega^c$ and the halfline l_z with vertex z which is opposite to the halfine with vertex z going through z_0. Then $l_z \subseteq \Omega^c$ and hence $\widehat{l_z} = l_z \cup \{\infty\} \subseteq \widehat{\mathbb{C}} \setminus \Omega$. Therefore, $\widehat{\mathbb{C}} \setminus \Omega$ is the union of the connected subsets $\widehat{l_z}$, $z \in \Omega^c$, of $\widehat{\mathbb{C}}$ all of which have ∞ as a common point. Thus $\widehat{\mathbb{C}} \setminus \Omega$ is connected. We conclude that every open star-shaped set is topologically simply connected.
If γ is a closed curve in Ω and $z \in \Omega^c$, then γ is in $\mathbb{C} \setminus l_z$, where l_z is the halfline of the previous paragraph. Hence $n(\gamma; z) = 0$. Therefore every open star-shaped set is homologically simply connected.
Finally, if $\gamma : [a, b] \to \Omega$ is a closed curve in Ω, then the function $F : [a, b] \times [0, 1] \to \Omega$ defined by

$$F(t, s) = (1 - s)\gamma(t) + sz_0$$

is a homotopy with closed intermediate curves in Ω from γ to the constant curve z_0. Therefore every open star-shaped set is homotopically simply connected.

Example 7.2.3. The region $\Omega = \mathbb{C} \setminus \big(\overline{D}_0(1) \cup (-\infty, -1]\big)$ is not star-shaped but it is topologically simply connected. Indeed, $\widehat{\mathbb{C}} \setminus \Omega = \overline{D}_0(1) \cup (-\infty, -1] \cup \{\infty\}$ is connected.
Moreover, if γ is a closed curve in Ω, then $\Omega^c = \overline{D}_0(1) \cup (-\infty, -1]$ is connected and it is contained

in the connected component of $\mathbb{C} \setminus \gamma^*$. Hence $n(\gamma; z) = 0$ for every $z \in \Omega^c$. Therefore, Ω is homologically simply connected.
Finally, let $\gamma : [a, b] \to \Omega$ be a closed curve in Ω. Then the function $F : [a, b] \times [0, 1] \to \Omega$ defined by

$$F(t, s) = (1 - s)\gamma(t) + 2s\frac{\gamma(t)}{|\gamma(t)|}$$

is a homotopy with closed intermediate curves in Ω from γ to the closed curve $\gamma_1 : [a, b] \to \Omega \cap C_0(2)$ given by $\gamma_1(t) = 2\frac{\gamma(t)}{|\gamma(t)|}$. Now we consider the function $G : [a, b] \times [0, 1] \to \Omega \cap C_0(2)$ defined by

$$G(t, s) = 2e^{(1-s)\,\mathrm{Arg}(\gamma(t))}.$$

Then G is a homotopy with closed intermediate curves in Ω from γ_1 to the constant curve 2. Hence there is a homotopy with closed intermediate curves in Ω from γ to the constant curve 2 and we conclude that Ω is homotopically simply connected.

Proposition 7.3. *Let Ω be a region in $\mathbb{C}$.*
(i) Ω is topologically simply connected if and only if it is homologically simply connected.
(ii) If Ω is homotopically simply connected then it is topologically and homologically simply connected.

Proof. (i) Assume that Ω is topologically simply connected. Let γ be any closed curve in Ω and let U be the unbounded connected component of $\mathbb{C} \setminus \gamma^*$. Then it is easy to see that $U \cup \{\infty\}$ is a connected component of $\widehat{\mathbb{C}} \setminus \gamma^*$. Since $\gamma^* \subseteq \Omega$, we have $\widehat{\mathbb{C}} \setminus \Omega \subseteq \widehat{\mathbb{C}} \setminus \gamma^*$ and hence the connected set $\widehat{\mathbb{C}} \setminus \Omega$ is contained in only one connected component of $\widehat{\mathbb{C}} \setminus \gamma^*$. Since $\widehat{\mathbb{C}} \setminus \Omega$ contains ∞, we conclude that $\widehat{\mathbb{C}} \setminus \Omega \subseteq U \cup \{\infty\}$. Therefore $\Omega^c \subseteq U$. Since $n(\gamma; z) = 0$ for every $z \in U$, we have $n(\gamma; z) = 0$ for every $z \in \Omega^c$.
Now, assume that Ω is homologically simply connected and, to arrive at a contradiction, assume that Ω is not topologically simply connected. Then $\widehat{\mathbb{C}} \setminus \Omega$ is not connected and so there is a decomposition B, C of $\widehat{\mathbb{C}} \setminus \Omega$. Let $\infty \in C$ (the case $\infty \in B$ is the same). Then ∞ is not a limit point of B and hence B is a bounded subset of $\mathbb{C}$. Since $\widehat{\mathbb{C}} \setminus \Omega$ is closed, both B, C are closed and hence B is a compact subset of $\mathbb{C}$. The complement of $\Omega' = \Omega \cup B$ is the closed set C and hence Ω' is open. Now we apply proposition 6.18 to the open set Ω', to the compact subset B of Ω' and to the constant function $f(z) = 1$, and we get that there are closed curves $\gamma_1, \dots, \gamma_k$ in $\Omega' \setminus B = \Omega$ so that

$$1 = n(\gamma_1; z) + \cdots + n(\gamma_k; z)$$

for every $z \in B$. We fix any $z_0 \in B$ and then for at least one of the closed curves $\gamma_1, \dots, \gamma_k$, say γ_j, in Ω we have that $n(\gamma_j; z_0) \neq 0$. Therefore Ω is not homologically simply connected and we arrived at a contradiction.
(ii) Assume that Ω is homotopically simply connected and let γ be any closed curve in Ω and $z \in \Omega^c$. Then there is a homotopy with closed intermediate curves in Ω and hence in $\mathbb{C} \setminus \{z\}$ from γ to a constant curve γ_1 in Ω. Proposition 6.17 implies that $n(\gamma; z) = n(\gamma_1; z) = 0$. Thus, Ω is homologically simply connected. $\square$

Later on, at the end of this section, we shall prove that topological and homological simple connectedness imply homotopical simple connectedness and thus all senses of simple connectedness are equivalent.

The theorem of Cauchy in simply connected regions. *If f is holomorphic in the region $\Omega \subseteq \mathbb{C}$ which is simply connected in any of the three senses, then for every cycle Σ in Ω which consists of closed piecewise smooth curves we have*

$$\oint_\Sigma f(z)\,dz = 0.$$

Proof. Immediate from the theorem of Cauchy in general open sets and proposition 7.3. $\square$

In the same manner we have versions of Cauchy's formulas for derivatives of any order, of the residue theorem, of the argument principle and of the theorem of Rouché for regions Ω which are simply connected in any of the three senses. In all these cases we do not have to assume that the cycles Σ in Ω are null-homologous in Ω: every cycle in a simply connected Ω is automatically null-homologous in Ω. In this respect it might be desirable to recall the algebraic facts mentioned at the beginning of section 6.4. If the region Ω is simply connected in any of the three senses, then the $\mathbb{Z}$-module $\mathcal{C}(\Omega)$ of all cycles in Ω is identical to its $\mathbb{Z}$-submodule $\mathcal{C}_0(\Omega)$ of all cycles which are null-homologous in Ω. Therefore the quotient $\mathbb{Z}$-module $\mathcal{H}(\Omega)$ is trivial:

$$\mathcal{H}(\Omega) = \mathcal{C}(\Omega)/\mathcal{C}_0(\Omega) = \{[\mathrm{O}]\},$$

where O is the zero cycle in Ω. I.e. $\mathcal{H}(\Omega)$ consists only of its zero element.

Proposition 7.4. *Let the region $\Omega \subseteq \mathbb{C}$ be simply connected in any of the three senses. Then*
(i) every f holomorphic in Ω has a primitive in Ω.
(ii) for every $g : \Omega \to \mathbb{C} \setminus \{0\}$ there is a holomorphic branch of $\log g$ *in Ω.*

Proof. (i) An application of proposition 4.10 and the theorem of Cauchy in simply connected regions.
(ii) An application of theorem 4.1 and the theorem of Cauchy in simply connected regions. □

Proposition 7.5. *Let the regions $\Omega_1, \Omega_2 \subseteq \mathbb{C}$ be conformally equivalent.*
(i) If Ω_1 is topologically or homologically simply connected, then Ω_2 is also topologically or homologically simply connected.
(ii) If Ω_1 is homotopically simply connected, then Ω_2 is also homotopically simply connected.

Proof. Let $f : \Omega_1 \to \Omega_2$ be holomorphic and one-to-one from Ω_1 onto Ω_2.
(i) Consider any closed piecewise smooth curve γ in Ω_2 and any $w_0 \in \Omega_2^c$. Consider also the closed piecewise smooth curve $f^{-1}(\gamma)$ in Ω_1. Then, after a simple change of variables, we have

$$n(\gamma; w_0) = \tfrac{1}{2\pi i} \oint_\gamma \tfrac{1}{w-w_0}\, dw = \tfrac{1}{2\pi i} \oint_{f^{-1}(\gamma)} \tfrac{f'(z)}{f(z)-w_0}\, dz = 0$$

by the theorem of Cauchy in the topologically or homologically simply connected region Ω. Therefore, Ω_2 is homologically (and hence also topologically) simply connected.
(ii) Let $\gamma : [a,b] \to \Omega_2$ be any closed curve in Ω_2. Then $f^{-1}(\gamma) : [a,b] \to \Omega_1$ is a closed curve in Ω_1. Since Ω_1 is homotopically simply connected, there is a homotopy $F : [a,b] \times [0,1] \to \Omega_1$ with closed intermediate curves so that

$$F(t,0) = f^{-1}(\gamma)(t) = f^{-1}(\gamma(t)), \quad F(t,1) = z_0$$

for every $t \in [a,b]$. Then $f \circ F : [a,b] \times [0,1] \to \Omega_2$ is a homotopy with closed intermediate curves so that

$$(f \circ F)(t,0) = \gamma(t), \quad (f \circ F)(t,1) = f(z_0)$$

for every $t \in [a,b]$. Therefore, γ is homotopic with closed intermediate curves to a constant curve in Ω_2. □

The theorem of Riemann. *Let $\Omega \subsetneq \mathbb{C}$ be a region which is simply connected in any of the three senses, $z_0 \in \Omega$ and $\theta_0 \in (-\pi, \pi]$. Then there is a unique conformal mapping f of Ω onto $\mathbb{D}$ with*

$$f(z_0) = 0, \quad \mathrm{Arg}\, f'(z_0) = \theta_0.$$

Proof. Step 1. We take any $a \in \Omega^c$. Since the function $z - a$ is holomorphic in Ω and has no root in Ω, proposition 7.4 implies that there is a holomorphic branch g of $\log(z-a)$ in Ω. I.e. $g : \Omega \to \mathbb{C}$ is holomorphic in Ω and

$$e^{g(z)} = z - a$$

for every $z \in \Omega$.
Now, g is one-to-one in Ω. Indeed, if $g(z_1) = g(z_2)$, then $e^{g(z_1)} = e^{g(z_2)}$ and hence $z_1 = z_2$.
We consider $w_0' = g(z_0) + 2\pi i$ and then we have $w_0' \notin \overline{g(\Omega)}$. Indeed, if $w_0' \in \overline{g(\Omega)}$, then there are $z_n \in \Omega$ so that $g(z_n) \to w_0'$. Hence $z_n - a = e^{g(z_n)} \to e^{w_0'} = e^{g(z_0)} = z_0 - a$ and thus $z_n \to z_0$. But then $g(z_n) \to g(z_0)$ which implies $w_0' = g(z_0)$ and we arrive at a contradiction.
Since $w_0' \notin \overline{g(\Omega)}$, there is $r_0 > 0$ so that $|g(z) - w_0'| > r_0$ for every $z \in \Omega$.
We consider the function $\chi : \Omega \to \mathbb{D}$ given by

$$\chi(z) = \tfrac{r_0}{g(z) - w_0'}$$

for every $z \in \Omega$. Then χ is holomorphic and one-to-one in Ω. In particular, $\chi'(z_0) \neq 0$.
Now we consider the Möbius transformation $R : \mathbb{D} \to \mathbb{D}$ given by

$$R(w) = \tfrac{|\chi'(z_0)|}{\chi'(z_0)}\, e^{i\theta_0}\, \tfrac{w - \chi(z_0)}{1 - \overline{\chi(z_0)}\, w}$$

for every $w \in \mathbb{D}$. (Look again at example 7.1.1 and at proposition 7.2 for the properties of Möbius transformations. They appear many times in this proof.) Then the function

$$h = R \circ \chi : \Omega \to \mathbb{D}$$

is holomorphic and one-to-one in Ω and satisfies $h(z_0) = R(\chi(z_0)) = 0$ and

$$h'(z_0) = R'(\chi(z_0))\chi'(z_0) = \tfrac{|\chi'(z_0)| e^{i\theta_0}}{1 - |\chi(z_0)|^2}$$

and hence $\operatorname{Arg} h'(z_0) = \theta_0$.
Step 2. We consider the set

$$\mathcal{F} = \{h \,|\, h : \Omega \to \mathbb{D}, h \text{ is holomorphic and one-to-one in } \Omega, h(z_0) = 0, \operatorname{Arg} h'(z_0) = \theta_0\}.$$

The result of step 1 implies that $\mathcal{F}$ is a non-empty subset of $H(\Omega)$. We also define

$$\alpha = \sup_{h \in \mathcal{F}} |h'(z_0)|.$$

Since, $h'(z_0) \neq 0$ for every $h \in \mathcal{F}$, we have that $\alpha > 0$ (but $\alpha = +\infty$ is not excluded *a priori*).
There is a sequence (h_n) in $\mathcal{F}$ so that $|h_n'(z_0)| \to \alpha$. For every $h \in \mathcal{F}$ we have that $|h(z)| < 1$ for every $z \in \Omega$ and hence $\mathcal{F}$ is obviously locally bounded in Ω. Montel's theorem implies that there is a subsequence (h_{n_k}) such that $h_{n_k} \to f$ uniformly in every compact subset of Ω for some f holomorphic in Ω. Obviously, we have $0 = h_{n_k}(z_0) \to f(z_0)$ and so $f(z_0) = 0$. The theorem of Weierstrass implies that $h_{n_k}' \to f'$ uniformly in every compact subset of Ω. Hence, $h_{n_k}'(z_0) \to f'(z_0)$ and thus $|f'(z_0)| = \alpha$ (hence $\alpha < +\infty$) and $\operatorname{Arg} f'(z_0) = \theta_0$. Since $f'(z_0) \neq 0$, we have that f is not constant in Ω. Now, for every $z \in \Omega$ we have $|h_{n_k}(z)| < 1$ for every n_k and hence $|f(z)| \leq 1$. If $|f(z)| = 1$ for some $z \in \Omega$, the maximum principle implies that f is constant in Ω and we just saw that this is wrong. Therefore, $f : \Omega \to \mathbb{D}$. Next, we take any $z_1, z_2 \in \Omega$ with $z_1 \neq z_2$. Since $h_{n_k}(z_2) \to f(z_2)$, we get that $h_{n_k} - h_{n_k}(z_2) \to f - f(z_2)$ uniformly in every compact subset of Ω and hence in every compact subset of $\Omega \setminus \{z_2\}$. Each h_{n_k} is one-to-one in Ω and so $h_{n_k} - h_{n_k}(z_2)$ has no root in $\Omega \setminus \{z_2\}$. Since $f - f(z_2)$ is not identically 0 in $\Omega \setminus \{z_2\}$, the theorem of Hurwitz implies that $f - f(z_0)$ has no root in $\Omega \setminus \{z_2\}$. Thus $f(z_1) - f(z_2) \neq 0$ and we conclude that f is one-to-one in Ω.
We proved that there is $f \in \mathcal{F}$ with $|f'(z_0)| = \alpha$.
Step 3. Assume that there is some $b \in \mathbb{D} \setminus f(\Omega)$.
We consider the Möbius transformation $T : \mathbb{D} \to \mathbb{D}$ given by

$$T(w) = \tfrac{w - b}{1 - \bar{b}\, w}$$

for every $w \in \mathbb{D}$, and then the function

$$\phi = T \circ f : \Omega \to \mathbb{D}.$$

Then ϕ is holomorphic and one-to-one in Ω. Since $f(z) \neq b$ for every $z \in \Omega$, we have that $\phi(z) \neq 0$ for every $z \in \Omega$. But Ω is simply connected in any of the three senses, and so proposition 7.4 implies that there is a holomorphic branch of $\log \phi$ and hence a holomorphic branch ψ of $\phi^{1/2}$ in Ω. I.e. there is $\psi : \Omega \to \mathbb{D}$ which is holomorphic in Ω and satisfies

$$\psi(z)^2 = \phi(z)$$

for every $z \in \Omega$. It is easy to see that ψ is one-to-one in Ω, because ϕ is one-to-one in Ω.
Now we consider the Möbius transformation $S : \mathbb{D} \to \mathbb{D}$ given by

$$S(w) = \tfrac{|\psi'(z_0)|}{\psi'(z_0)} \, e^{i\theta_0} \, \tfrac{w - \psi(z_0)}{1 - \overline{\psi(z_0)}\, w} \qquad \text{for every } w \in \mathbb{D}$$

and then the function

$$h = S \circ \psi : \Omega \to \mathbb{D}.$$

Then h is holomorphic and one-to-one in Ω. We also see easily that $h(z_0) = S(\psi(z_0)) = 0$ and

$$h'(z_0) = S'(\psi(z_0))\psi'(z_0) = \tfrac{|\psi'(z_0)| e^{i\theta_0}}{1 - |\psi(z_0)|^2}$$

and hence $\operatorname{Arg} h'(z_0) = \theta_0$. Thus, $h \in \mathcal{F}$.
Now we have altogether that $f, \phi, \psi, h : \Omega \to \mathbb{D}$, that $T, S : \mathbb{D} \to \mathbb{D}$ and that

$$\phi = T \circ f, \qquad h = S \circ \psi, \qquad \phi = F \circ \psi,$$

where $F : \mathbb{D} \to \mathbb{D}$ is given by $F(w) = w^2$ for every w. All these functions, except F, are one-to-one. We consider now the holomorphic function $\Phi : \mathbb{D} \to \mathbb{D}$, given by

$$\Phi = T^{-1} \circ F \circ S^{-1},$$

and then we have

$$f = \Phi \circ h.$$

Now, $\Phi(0) = (T^{-1} \circ F \circ S^{-1})(0) = (T^{-1} \circ F)(\psi(z_0)) = T^{-1}(\phi(z_0)) = f(z_0) = 0$ and

$$|f'(z_0)| = |\Phi'(h(z_0))||h'(z_0)| = |\Phi'(0)||h'(z_0)|. \tag{7.1}$$

Then the Schwartz lemma implies that $|\Phi'(0)| \leq 1$.
If $|\Phi'(0)| = 1$, then there is c with $|c| = 1$ so that $\Phi(z) = cz$ for every $z \in \mathbb{D}$. This implies that $F(w) = T(cS(w))$ for every $w \in \mathbb{D}$. This is wrong because the right side is one-to-one in $\mathbb{D}$. We conclude that $|\Phi'(0)| < 1$ and (7.1) implies that

$$|h'(z_0)| > |f'(z_0)| = \alpha.$$

This contradicts the definition of α and the fact that $h \in \mathcal{F}$. Therefore, there is no $b \in \mathbb{D} \setminus f(\Omega)$ and hence f is onto $\mathbb{D}$.
We proved the existence of a function $f : \Omega \to \mathbb{D}$ which is conformal from Ω onto $\mathbb{D}$ and which satisfies $f(z_0) = 0$ and $\operatorname{Arg} f'(z_0) = \theta_0$.
Step 4. To prove the uniqueness of f, we repeat the argument in the proof of proposition 7.2. Let $f_1, f_2 : \Omega \to \mathbb{D}$ be conformal from Ω onto $\mathbb{D}$ with $f_1(z_0) = f_2(z_0) = 0$ and $\operatorname{Arg} f_1'(z_0) = \operatorname{Arg} f_2'(z_0) = \theta_0$.
Then the function $f = f_2 \circ f_1^{-1} : \mathbb{D} \to \mathbb{D}$ is holomorphic in $\mathbb{D}$ and satisfies $f(0) = 0$ and $f'(0) = \frac{f_2'(z_0)}{f_1'(z_0)} > 0$. By the Schwarz lemma we get $|f'(0)| \leq 1$.

The function $g = f_1 \circ f_2^{-1} : \mathbb{D} \to \mathbb{D}$ is also holomorphic in $\mathbb{D}$ and satisfies $g(0) = 0$ and $g'(0) = \frac{f_1'(z_0)}{f_2'(z_0)} > 0$. Again, by the Schwarz lemma we get $|g'(0)| \leq 1$.
But the functions f and g are mutually inverse and hence $g'(0) = \frac{1}{f'(0)}$. Therefore, $|f'(0)| = |g'(0)| = 1$ and the Schwarz lemma implies that there is some c with $|c| = 1$ so that $f(w) = cw$ for every $w \in \mathbb{D}$. Now, $c = f'(0) > 0$ implies $c = 1$. Hence, $f(w) = w$ for every $w \in \mathbb{D}$ and finally $f_2(z) = f_1(z)$ for every $z \in \mathbb{D}$. ☐

Proposition 7.6. *Let $\Omega \subseteq \mathbb{C}$ be a region. If Ω is topologically or homologically simply connected, then it is also homotopically simply connected.*

Proof. If $\Omega = \mathbb{C}$, then Ω is obviously homotopically simply connected. If $\Omega \subsetneqq \mathbb{C}$, then, by the theorem of Riemann, Ω is conformally equivalent to $\mathbb{D}$. Since $\mathbb{D}$ is homotopically simply connected, proposition 7.5 implies that Ω is also homotopically simply connected. ☐

We just proved that all three senses of simple connectedness are equivalent. From now on, we shall use the term **simply connected** for a region without having to distinguish between the three senses.

Proposition 7.7. *Every simply connected region $\Omega \subsetneqq \mathbb{C}$ is conformally equivalent with $\mathbb{D}$. The simply connected region $\mathbb{C}$ is conformally equivalent only with itself.*

Proof. The first part is a simple application of the theorem of Riemann.
If $\mathbb{C}$ is conformally equivalent with some simply connected region $\Omega \subsetneqq \mathbb{C}$, then, by the first part, $\mathbb{C}$ is conformally equivalent with $\mathbb{D}$. Thus, there is a holomorphic $f : \mathbb{C} \to \mathbb{D}$ which is one-to-one in $\mathbb{C}$. But Liouville's theorem implies that f is constant and we arrive at a contradiction. ☐

Exercises.

7.2.1. We know that there is no holomorphic $f : \mathbb{C} \to \mathbb{D}$ which is one-to-one in $\mathbb{C}$. Find some $f : \mathbb{C} \to \mathbb{D}$ which is one-to-one and onto so that f and f^{-1} are both continuous.

7.2.2. Are the regions $D_0(1,3) \setminus [1,3]$ and $\mathbb{C} \setminus \big((-\infty, -2] \cup [-\frac{1}{2}, \frac{1}{2}] \cup [2, +\infty)\big)$ simply connected? Which are the possible values of $\oint_\gamma (z + \frac{1}{z})\, dz$, where γ is a closed piecewise smooth curve (i) in the first set? (ii) in the second set?

7.2.3. Let f be holomorphic in the simply connected region Ω except for isolated singularities in Ω. Prove that (i) and (ii) are equivalent:
(i) $e^{\oint_\gamma f(z)\, dz} = 1$ for every closed piecewise smooth curve γ in Ω whose trajectory contains no isolated singularity of f.
(ii) $\mathrm{Res}(f; z) \in \mathbb{Z}$ for every isolated singularity z of f in Ω.
If f satisfies (i), (ii) and it is holomorphic at $z_0 \in \Omega$, define $F(z) = e^{\int_\gamma f(\zeta)\, d\zeta}$ for every $z \in \Omega$, where γ is any piecewise smooth curve in Ω from z_0 to z and whose trajectory contains no isolated singularity of f.
Prove that F is well-defined and holomorphic in Ω except for the isolated singularities of f.
Prove that every point in Ω is either a point of holomorphy or a pole of F if and only if all isolated singularities of f in Ω are simple poles of f.

7.2.4. Let $\mathbb{H}_+ = \{z \mid \mathrm{Im}\, z > 0\}$, $z_0 \in \mathbb{H}_+$, $\theta_0 \in (-\pi, \pi]$. Find the unique conformal mapping f of $\mathbb{H}_+$ onto $\mathbb{D}$ with $f(z_0) = 0$ and $\mathrm{Arg}\, f'(z_0) = \theta_0$.

7.2.5. Find a conformal mapping of $\{z \mid \mathrm{Re}\, z > 0, \mathrm{Im}\, z > 0\}$ onto $\mathbb{D}$.

7.2.6. (i) Find a conformal mapping between two angular regions.
(ii) Find a conformal mapping between an angular region and an open zone.
(iii) Find a conformal mapping between an angular region and the intersection of two open discs or the intersection of an open disc and an open halfplane.

7.2.7. (i) Find a conformal mapping f of $\widehat{\mathbb{C}} \setminus [-1,1]$ onto $\mathbb{D}$, with $f(\infty)=0$.
(ii) Let $0<a<\pi$ and τ be the arc of $\mathbb{T}$ with endpoints e^{-ia} and e^{ia}. Find a conformal mapping f of $\widehat{\mathbb{C}} \setminus \tau$ onto $\mathbb{D}$, with $f(\infty)=0$.

7.2.8. Find a conformal mapping of $[-1,1]\times[-1,1]$ onto $D_0(1)$.

7.2.9. Prove that there is no conformal mapping of $\mathbb{D}$ onto $\mathbb{D}\setminus\{0\}$.

7.2.10. Let $\mathbb{H}_+=\{z \,|\, \operatorname{Im} z>0\}$ and let $f:\mathbb{H}_+\to\mathbb{H}_+$ be holomorphic in $\mathbb{H}_+$. Prove that:
(i) $\big|\frac{f(z_1)-f(z_2)}{f(z_1)-\overline{f(z_2)}}\big| \leq \big|\frac{z_1-z_2}{z_1-\overline{z_2}}\big|$ for every $z_1,z_2\in\mathbb{H}_+$.
(ii) $\frac{|f'(z)|}{\operatorname{Im} f(z)} \leq \frac{1}{\operatorname{Im} z}$ for every $z\in\mathbb{H}_+$.
Prove that, if equality holds in (i) for at least one pair of $z_1,z_2\in\mathbb{H}_+$ with $z_1\neq z_2$ or in (ii) for at least one $z\in\mathbb{H}_+$, then there is $z_0\in\mathbb{H}_+$ and λ with $|\lambda|=1$ so that $\frac{f(z)-i}{f(z)+i}=\lambda\frac{z-z_0}{z-\overline{z_0}}$ for every $z\in\mathbb{H}_+$ and then equalities in (i) and (ii) hold identically.

7.2.11. Let $\mathbb{H}_+=\{z \,|\, \operatorname{Im} z>0\}$ and let $f:\mathbb{H}_+\to\mathbb{D}$ be holomorphic in $\mathbb{H}_+$ with $f(i)=0$. Prove that $|f(z)|\leq\big|\frac{i-z}{i+z}\big|$ for every $z\in\mathbb{H}_+$ and $|f'(i)|\leq\frac{1}{2}$.

7.2.12. Let $\Omega\subseteq\mathbb{C}$ be a simply connected region, $z_0\in\Omega$ and f,g be conformal mappings of Ω onto $\mathbb{D}$ with $f(z_0)=g(z_0)$ for some $z_0\in\Omega$. Find a relation between f,g.

7.2.13. Let $\Omega_1,\Omega_2\subseteq\mathbb{C}$ be two regions and f be a conformal mapping of Ω_1 onto Ω_2. If (z_n) is in Ω_1 and $z_n\to z\in\partial\Omega_1$, prove that every limit point of $(f(z_n))$ belongs to $\partial\Omega_2$. Is it necessary for $(f(z_n))$ to converge?

7.2.14. (i) Let $f,g:\mathbb{D}\to\Omega$ be holomorphic in $\mathbb{D}$ so that f is one-to-one in $\mathbb{D}$ and onto Ω. If $f(0)=g(0)$, prove that $g(D_0(r))\subseteq f(D_0(r))$ for every r with $0<r<1$.
(ii) Let $\Omega=\{w=u+iv \,|\, -1<u<1\}$. Find the conformal mapping f of $\mathbb{D}$ onto Ω with $f(0)=0$ and $f'(0)>0$. If $g:\mathbb{D}\to\Omega$ is holomorphic in $\mathbb{D}$ with $g(0)=0$, prove that $|g(z)|\leq\frac{2}{\pi}\ln\frac{1+|z|}{1-|z|}$ for every $z\in\mathbb{D}$.

7.2.15. Let $\Omega\subsetneqq\mathbb{C}$ be a simply connected region, $z_0\in\Omega$ and let $\mathcal{F}$ be the collection of all holomorphic $f:\Omega\to\mathbb{D}$ with $f(z_0)=0$ and which are one-to-one in Ω. We fix $a\in\Omega$, $a\neq z_0$ and we define $m=\sup_{f\in\mathcal{F}}|f(a)|$. Prove that there is $f_0\in\mathcal{F}$ so that $|f_0(a)|=m$ and that such a f_0 is a conformal mapping of Ω onto $\mathbb{D}$.

7.2.16. Let $\Omega\subsetneqq\mathbb{C}$ be a simply connected region so that $\overline{z}\in\Omega$ for every $z\in\Omega$. Let $z_0\in\Omega\cap\mathbb{R}$ and let f be the conformal mapping of Ω onto $\mathbb{D}$ with $f(z_0)=0$ and $f'(z_0)>0$. Let $\Omega_+=\{z\in\Omega \,|\, \operatorname{Im} z>0\}$, $\Omega_-=\{z\in\Omega \,|\, \operatorname{Im} z<0\}$, $\mathbb{D}_+=\{z\in\mathbb{D} \,|\, \operatorname{Im} z>0\}$ and $\mathbb{D}_-=\{z\in\mathbb{D} \,|\, \operatorname{Im} z<0\}$. Prove that $f(\Omega_+)=\mathbb{D}_+$, $f(\Omega_-)=\mathbb{D}_-$ and $f(\Omega\cap\mathbb{R})=(-1,1)$.

7.3 Multiply connected regions.

The region $\Omega\subseteq\mathbb{C}$ is called **m-tuply connected** if $\widehat{\mathbb{C}}\setminus\Omega$ has exactly m connected components.

If the region $\Omega\subseteq\mathbb{C}$ is m-tuply connected and $A_1,\dots,A_m$ are the connected components of $\widehat{\mathbb{C}}\setminus\Omega$, then one of these components, say A_m, contains ∞. Since $A_1,\dots,A_m$ are closed subsets of $\widehat{\mathbb{C}}\setminus\Omega$ (with respect to the chordal metric), they are compact. Thus, $A_1,\dots,A_{m-1}$ are compact subsets of $\mathbb{C}$ (either with respect to the chordal metric or with respect to the euclidean metric).

Example 7.3.1. Let the closed discs $\overline{D}_{z_1}(r_1),\dots,\overline{D}_{z_{m-1}}(r_{m-1})$ be disjoint and contained in the open disc $D_{z_0}(R_0)$. Then the region $D_{z_0}(R_0)\setminus\big(\overline{D}_{z_1}(r_1)\cup\cdots\cup\overline{D}_{z_{m-1}}(r_{m-1})\big)$ is m-tuply connected. Some (or all) of the inner discs may reduce to single points.

Example 7.3.2. The region $\mathbb{C}\setminus\{z_0\}$ is doubly connected. The region $\mathbb{C}\setminus[a,b]$ is doubly connected.

Proposition 7.8. *Let $m \geq 2$ and $\Omega \subseteq \mathbb{C}$ be an m-tuply connected region and let $A_1, \ldots, A_{m-1}$ be the connected components of $\widehat{\mathbb{C}} \setminus \Omega$ which do not contain ∞. Then:*
(i) There are cycles $\Sigma_1, \ldots, \Sigma_{m-1}$ in Ω so that for every $k = 1, \ldots, m-1$ we have: $n(\Sigma_k, z) = 1$ for every $z \in A_k$ and $n(\Sigma_k; z) = 0$ for every $z \in \Omega^c \setminus A_k$.
(ii) Let $\Sigma_1, \ldots, \Sigma_{m-1}$ be any cycles in Ω as in (i) and take any $z_k \in A_k$ for $k = 1, \ldots, m-1$. Then for every f holomorphic in Ω and every cycle Σ in Ω we have

$$\oint_{\Sigma} f(z)\, dz = n(\Sigma; z_1) \oint_{\Sigma_1} f(z)\, dz + \cdots + n(\Sigma; z_{m-1}) \oint_{\Sigma_{m-1}} f(z)\, dz. \tag{7.2}$$

Proof. (i) We take any of the connected components $A_1, \ldots, A_{m-1}$, say A_k, and the set $\Omega' = \Omega \cup A_k$. Then $\Omega' \subseteq \mathbb{C}$ and $\widehat{\mathbb{C}} \setminus \Omega' = A_1 \cup \cdots \cup A_{k-1} \cup A_{k+1} \cup \cdots A_m$ is a closed subset of $\widehat{\mathbb{C}}$ (with respect to the chordal metric) and hence Ω' is an open subset of $\mathbb{C}$ (either with respect to the chordal metric or with respect to the euclidean metric). Now we apply proposition 6.18 to the open set Ω', to the compact $A_k \subseteq \Omega'$ and to the constant function $f(z) = 1$, and we get a cycle Σ_k in $\Omega' \setminus A_k = \Omega$ such that $n(\Sigma_k; z) = 1$ for every $z \in A_k$ and $n(\Sigma_k; z) = 0$ for every $z \in (\Omega')^c = \Omega^c \setminus A_k$.
(ii) We concider any cycle Σ in Ω and the integers

$$p_k = n(\Sigma; z_k), \qquad k = 1, \ldots, m-1.$$

Now we define the cycle $\Sigma' = p_1\Sigma_1 + \cdots + p_{m-1}\Sigma_{m-1}$ and we get

$$n(\Sigma; z) = n(\Sigma'; z)$$

for every $z \in \Omega^c$. Indeed, if $z \in A_k$ for any $k = 1, \ldots, m-1$ then

$$n(\Sigma'; z) = p_1 n(\Sigma_1; z) + \cdots + p_{m-1} n(\Sigma_{m-1}; z) = p_k = n(\Sigma; z_k) = n(\Sigma; z)$$

since z, z_k belong to the connected set A_k which is in the complement of all the trajectories of the closed curves forming Σ in Ω. Also, if $z \in A_m$ then

$$n(\Sigma'; z) = p_1 n(\Sigma_1; z) + \cdots + p_{m-1} n(\Sigma_{m-1}; z) = 0 = n(\Sigma; z).$$

Thus, $n(\Sigma - \Sigma'; z) = 0$ for every $z \in \Omega^c$, i.e. $\Sigma - \Sigma'$ is null-homologous in Ω. So, if f is holomorphic in Ω then $\oint_{\Sigma - \Sigma'} f(z)\, dz = 0$ and hence

$$\oint_{\Sigma} f(z)\, dz = \oint_{\Sigma'} f(z)\, dz = p_1 \oint_{\Sigma_1} f(z)\, dz + \cdots + p_{m-1} \oint_{\Sigma_{m-1}} f(z)\, dz$$

and this is (7.2). $\square$

In the course of the proof of proposition 7.8 we saw that for every cycle Σ in Ω there are integers $p_1, \ldots, p_{m-1}$ so that $\Sigma - (p_1\Sigma_1 + \cdots + p_{m-1}\Sigma_{m-1})$ is null-homologous in Ω, i.e. $\Sigma - (p_1\Sigma_1 + \cdots + p_{m-1}\Sigma_{m-1}) \in \mathcal{C}_0(\Omega)$. This says that in $\mathcal{H}(\Omega) = \mathcal{C}(\Omega)/\mathcal{C}_0(\Omega)$ we have

$$[\Sigma] = [p_1\Sigma_1 + \cdots + p_{m-1}\Sigma_{m-1}] = p_1[\Sigma_1] + \cdots + p_{m-1}[\Sigma_{m-1}].$$

In other words, the elements $[\Sigma_1], \ldots, [\Sigma_{m-1}]$ of $\mathcal{H}(\Omega)$ produce the $\mathbb{Z}$-module $\mathcal{H}(\Omega)$. On the other hand, if for some integers $p_1, \ldots, p_{m-1}$ we have

$$p_1[\Sigma_1] + \cdots + p_{m-1}[\Sigma_{m-1}] = [\mathrm{O}],$$

the zero element of $\mathcal{H}(\Omega)$, then $[p_1\Sigma_1 + \cdots + p_{m-1}\Sigma_{m-1}] = [\mathrm{O}]$ and hence $p_1\Sigma_1 + \cdots + p_{m-1}\Sigma_{m-1}$ is null-homologous in Ω. This implies that

$$p_1 n(\Sigma_1; z) + \cdots + p_{m-1} n(\Sigma_{m-1}; z) = 0$$

for every $z \in \Omega^c$. If $z \in A_k$ for any $k = 1, \dots, m-1$, then $n(\Sigma_j; z) = 0$ for $j \neq k$ and $n(\Sigma_k; z) = 1$. Therefore, the last formula reduces to $p_k = 0$. This means that the elements $[\Sigma_1], \dots, [\Sigma_{m-1}]$ of $\mathcal{H}(\Omega)$ are linearly independent and we conclude that they form a basis of $\mathcal{H}(\Omega)$. Hence

$$\dim \mathcal{H}(\Omega) = m - 1.$$

We say that the elements $[\Sigma_1], \dots, [\Sigma_{m-1}]$ form a **homology basis** of $\mathcal{H}(\Omega)$ and that the cycles $\Sigma_1, \dots, \Sigma_{m-1}$ form a homology basis in Ω.

The above complement the case of a simply connected region Ω, where $m = 1$ and $\mathcal{H}(\Omega) = \{[\mathrm{O}]\}$ and hence $\dim \mathcal{H}(\Omega) = 0$.

Exercises.

7.3.1. Let $\Omega \subseteq \mathbb{C}$ be an m-tuply connected region and let $A_1, \dots, A_{m-1}$ be the connected components of $\widehat{\mathbb{C}} \setminus \Omega$ which do not contain ∞. We take any $z_k \in A_k$ for $k = 1, \dots, m-1$. Prove that for every f holomorphic in Ω which has no roots in Ω there are $n_1, \dots, n_{m-1} \in \mathbb{N}_0$ so that there is a holomorphic branch in Ω of $\log g$, where $g(z) = \frac{f(z)}{(z-z_1)^{n_1} \cdots (z-z_{m-1})^{n_{m-1}}}$ for every $z \in \Omega$.

7.3.2. Let γ be a closed curve. Prove that every bounded connected component of $\mathbb{C} \setminus \gamma^*$ is a simply connected region and that the unbounded connected component of $\mathbb{C} \setminus \gamma^*$ is a doubly connected region.

7.3.3. Let $\Omega \subseteq \mathbb{C}$ be a simply connected region and $z_1, \dots, z_{m-1} \in \Omega$. Prove that the region $\Omega \setminus \{z_1, \dots, z_{m-1}\}$ is m-tuply connected and find a homology basis of cycles in this region.

7.3.4. Let $\Omega \subseteq \mathbb{C}$ be a doubly connected region and let A_1, A_2 be the connected components of $\widehat{\mathbb{C}} \setminus \Omega$. If f is holomorphic in Ω, prove that there are f_1, f_2 so that $f = f_1 + f_2$ in Ω, and f_1 is holomorphic in $\Omega \cup A_1$ and f_2 is holomorphic in $\Omega \cup A_2$.

Chapter 8

Isolated singularities and roots.

8.1 Isolated singularities in the complex plane.

Let us generalize slightly the argument at the end of section 5.8. We consider a function f in $\mathbb{C}$ with a finite number of isolated singularities $z_1, \ldots, z_n$ and holomorphic in the rest of $\mathbb{C}$. The singular part of f at z_j has the form

$$s_j(z) = \sum_{-\infty}^{k=-1} a_{j,k}(z - z_j)^k = \sum_{k=1}^{+\infty} \frac{a_{j,-k}}{(z-z_j)^k}$$

and converges in $\widehat{\mathbb{C}} \setminus \{z_j\}$. We consider the function

$$h(z) = f(z) - (s_1(z) + \cdots + s_n(z)).$$

Then h is holomorphic in the set $\mathbb{C} \setminus \{z_1, \ldots, z_n\}$ and its only possible singularities are the points $z_1, \ldots, z_n$. We observe that every z_j is a removable singularity of $f(z) - s_j(z)$ and that all terms $s_1(z), \ldots, s_n(z)$, besides $s_j(z)$, are holomorphic at z_j. Therefore, every z_j is a removable singularity of the function h. So the function h has *no isolated singularities* and hence it is holomorphic in $\mathbb{C}$. Now, we have the identity

$$f(z) = s_1(z) + \cdots + s_n(z) + h(z),$$

which gives *the general form of a holomorphic function in* $\mathbb{C}$ *with the exception of finitely many isolated singularities*.

We shall generalize this to the case of a holomorphic function f in $\mathbb{C}$ with the exception of infinitely many isolated singularities. In this case, i.e. if the terms of the sequence (z_n) are the distinct isolated singularities of f in $\mathbb{C}$, it is necessary that $z_n \to \infty$. In the opposite case there would be a subsequence of (z_n) converging to some $z \in \mathbb{C}$ and then this z would be a non-isolated singularity of f.

We may obviously try to form the infinite sum $\sum_{n=1}^{+\infty} s_n(z)$, but this is doomed to failure in the general case since there is no guarantee that this series converges. The next theorem shows that we may subtract a suitable "correction term" from each $s_n(z)$ so as to make the series convergent.

The theorem of Mittag-Leffler. *Let the terms of the sequence* (z_n) *be distinct with* $z_n \to \infty$*. For each* z_n *we consider a power series of the form* $s_n(z) = \sum_{-\infty}^{k=-1} a_{n,k}(z - z_n)^k$*, which converges in* $\widehat{\mathbb{C}} \setminus \{z_n\}$*.*
(i) Then there are polynomials q_n *so that the series of functions*

$$\sum_{n=1}^{+\infty}(s_n - q_n)$$

has the property: for every compact set K *there is* n_0 *so that* $\sum_{n=n_0+1}^{+\infty}(s_n - q_n)$ *converges uniformly in* K*.*

(ii) If the polynomials q_n satisfy (i), then the function $F = \sum_{n=1}^{+\infty}(s_n - q_n)$ is holomorphic in $\mathbb{C}$, with the exception of the terms of (z_n), and its singular part at each z_n is s_n. Moreover, the most general holomorphic function in $\mathbb{C}$, with the exception of the terms of (z_n), and whose singular part at each z_n is s_n, is of the form

$$f = F + h = \sum_{n=1}^{+\infty}(s_n - q_n) + h,$$

where h is an arbitrary function holomorphic in $\mathbb{C}$. We also have that

$$f' = \sum_{n=1}^{+\infty}(s_n' - q_n') + h'.$$

Proof. (i) If $z_n = 0$, we just take $q_n = 0$. If $z_n \neq 0$, then the function s_n is holomorphic in the disc $D_0(|z_n|)$ and so its Taylor series at 0 converges to it uniformly in the smaller disc $D_0(|z_n|/2)$. Hence there is a partial sum q_n of this Taylor series, i.e. a polynomial, so that

$$\|s_n - q_n\|_{D_0(|z_n|/2)} \leq \tfrac{1}{2^n}.$$

Now let K be any compact set. Since K is bounded, there is $R > 0$ so that $K \subseteq D_0(R)$. Since $z_n \to \infty$, there is n_0 so that $|z_n| \geq 2R$ and hence $K \subseteq D_0(|z_n|/2)$ for every $n \geq n_0 + 1$. Thus

$$\|s_n - q_n\|_K \leq \|s_n - q_n\|_{D_0(|z_n|/2)} \leq \tfrac{1}{2^n}$$

for every $n \geq n_0+1$. The test of Weierstrass implies that $\sum_{n=n_0+1}^{+\infty}(s_n - q_n)$ converges uniformly in K.
(ii) We assume that the polynomials q_n satisfy (i) and we take any $z \in \mathbb{C}$. Since $\{z\}$ is compact, there is n_0 so that $\sum_{n=n_0+1}^{+\infty}(s_n(z) - q_n(z))$ converges. So if z is not equal to any of $z_1, \ldots, z_{n_0}$, then the sum $\sum_{n=1}^{+\infty}(s_n(z) - q_n(z))$ is finite and we define the function $F : \mathbb{C} \setminus \{z_n \,|\, n \in \mathbb{N}\} \to \mathbb{C}$ by

$$F = \sum_{n=1}^{+\infty}(s_n - q_n).$$

If z is not equal to any of the terms of (z_n), then, because of $z_n \to \infty$, there is a closed disc $\overline{D}_z(r)$ which contains no term of (z_n). Then there is n_0 so that $\sum_{n=n_0+1}^{+\infty}(s_n - q_n)$ converges uniformly in $\overline{D}_z(r)$ and so it defines a function holomorphic in $D_z(r)$. But the finite sum $\sum_{n=1}^{n_0}(s_n - q_n)$ is also holomorphic in $D_z(r)$ and hence F is holomorphic in $D_z(r)$. Moreover, by the uniform convergence of $\sum_{n=n_0+1}^{+\infty}(s_n - q_n)$ in $D_z(r)$, we have that the series of the derivatives also converges uniformly in $D_z(r)$ and hence

$$F'(z) = \sum_{n=1}^{+\infty}(s_n'(z) - q_n'(z)).$$

This equality holds at every z which is not equal to any of the terms of $\{z_n\}$.
If $z = z_k$ for some k, then there is a closed disc $\overline{D}_{z_k}(r)$ which contains only the term z_k of (z_n). Then there is n_0 so that $\sum_{n=n_0+1}^{+\infty}(p_n - q_n)$ converges uniformly in $\overline{D}_{z_k}(r)$ and so it defines a function holomorphic in $D_{z_k}(r)$. But the finite sum $\sum_{n=1}^{n_0}(s_n - q_n)$ is holomorphic in $D_{z_k}(r) \setminus \{z_k\}$ with singular part s_k at z_k. So F has the singular part s_k at z_k.
We conclude that F is holomorphic in $\mathbb{C}$ with the exception of the terms of (z_n) and that its singular part at each z_n is s_n.
Now let us consider an arbitrary holomorphic function f in $\mathbb{C}$ with the exception of the terms of (z_n) and whose singular part at each z_n is s_n. Then the function $h = f - F$ is holomorphic in $\mathbb{C}$ and hence $f = F + h$. □

The theorem of Mittag-Leffler describes the most general holomorphic function in $\mathbb{C}$ with the exception of preassigned isolated singularities and corresponding preassigned singular parts. In fact, the actual theorem of Mittag-Leffler is restricted to the case of meromorphic functions, i.e. to the case that all isolated singularities are poles.

Example 8.1.1. We consider the function $\frac{1}{\sin(\pi z)}$, which is meromorphic in $\mathbb{C}$. Its poles are the integers $n \in \mathbb{Z}$. Since $\sin(\pi z) = \pi z - \frac{\pi^3 z^3}{3!} + \cdots$ is the Taylor series of $\sin(\pi z)$ at 0 we have, for each $n \in \mathbb{Z}$, that

$$\sin(\pi z) = (-1)^n \sin(\pi z - n\pi) = (-1)^n \pi(z-n) - \frac{(-1)^n \pi^3 (z-n)^3}{3!} + \cdots = (z-n)g(z)$$

where g is holomorphic in $\mathbb{C}$ with $g(n) = (-1)^n \pi$. So the function $h = \frac{1}{g}$ is holomorphic at n with $h(n) = \frac{(-1)^n}{\pi}$ and

$$\frac{1}{\sin(\pi z)} = \frac{h(z)}{z-n} = \frac{(-1)^n}{\pi(z-n)} + h'(n) + \frac{h''(n)}{2!}(z-n) + \cdots$$

in a neighborhood of n. This says that n is a pole of order 1 of $\frac{1}{\sin(\pi z)}$ and the singular part at n is $s_n(z) = \frac{(-1)^n}{\pi(z-n)}$.
Now we take $n \neq 0$ and we write the Taylor series of $\frac{(-1)^n}{\pi(z-n)}$ at 0:

$$\frac{(-1)^n}{\pi(z-n)} = \frac{(-1)^{n+1}}{n\pi} \frac{1}{1-z/n} = \frac{(-1)^{n+1}}{n\pi} + \frac{(-1)^{n+1}}{n^2\pi} z + \frac{(-1)^{n+1}}{n^3\pi} z^2 + \cdots .$$

We consider the polynomial q_n to be the constant term of this Taylor series, i.e. $q_n(z) = \frac{(-1)^{n+1}}{n\pi}$. If $n = 0$, we just take $q_n = 0$.
Now we examine the uniform convergence of the series

$$\textstyle\sum_{n\in\mathbb{Z}}(s_n(z) - q_n(z)) = \frac{1}{\pi z} + \frac{1}{\pi}\sum_{n\in\mathbb{Z}\setminus\{0\}}(-1)^n\big(\frac{1}{z-n} + \frac{1}{n}\big).$$

If K is a compact set, then there is R so that $K \subseteq D_0(R)$. Now, if $n_0 + 1 \geq 2R$ and $z \in K$, then for every n with $|n| \geq n_0 + 1$ we have that $|z-n| \geq |n| - |z| \geq |n| - R \geq \frac{|n|}{2}$ and hence

$$|(-1)^n(\tfrac{1}{z-n} + \tfrac{1}{n})| = \frac{|z|}{|n||z-n|} \leq \frac{2R}{n^2}.$$

By the test of Weierstrass, $\sum_{n\in\mathbb{Z},|n|\geq n_0+1}(-1)^n\big(\frac{1}{z-n} + \frac{1}{n}\big)$ converges uniformly in K.
Now the theorem of Mittag-Leffler implies that

$$\frac{1}{\sin(\pi z)} = \frac{1}{\pi z} + \frac{1}{\pi}\textstyle\sum_{n\in\mathbb{Z}\setminus\{0\}}(-1)^n\big(\frac{1}{z-n} + \frac{1}{n}\big) + h(z), \tag{8.1}$$

where h is holomorphic in $\mathbb{C}$. We shall determine the function h.
Again, based on the theorem of Mittag-Leffler, we differentiate the last series to get

$$\frac{\pi\cos(\pi z)}{\sin^2(\pi z)} = \frac{1}{\pi z^2} + \frac{1}{\pi}\textstyle\sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{(-1)^n}{(z-n)^2} + h'(z) = \frac{1}{\pi}\sum_{n\in\mathbb{Z}} \frac{(-1)^n}{(z-n)^2} + h'(z). \tag{8.2}$$

The function $\frac{\pi\cos(\pi z)}{\sin^2(\pi z)}$ is 2-periodic and it is easy to prove that $\frac{1}{\pi}\sum_{n\in\mathbb{Z}} \frac{(-1)^n}{(z-n)^2}$ is also 2-periodic. Indeed,

$$\frac{1}{\pi}\textstyle\sum_{n\in\mathbb{Z}} \frac{(-1)^n}{(z+2-n)^2} = \frac{1}{\pi}\sum_{n\in\mathbb{Z}} \frac{(-1)^n}{(z-(n-2))^2} = \frac{1}{\pi}\sum_{n\in\mathbb{Z}} \frac{(-1)^{n+2}}{(z-n)^2} = \frac{1}{\pi}\sum_{n\in\mathbb{Z}} \frac{(-1)^n}{(z-n)^2}.$$

Therefore, h' is 2-periodic.
We restrict now our investigation in a *period-zone* $A = \{z + iy \mid -1 \leq x \leq 1\}$. Again, it is easy to prove that $\frac{\cos(\pi z)}{\sin^2(\pi z)} \to 0$ when $z \to \infty$ in A. Indeed, if $z = x + iy$ and $|x| \leq 1$, then we have

$$\Big|\frac{\cos(\pi z)}{\sin^2(\pi z)}\Big| = \frac{\sinh^2(\pi y) + \cos^2(\pi x)}{(\sinh^2(\pi y) + \sin^2(\pi x))^2} \leq \frac{\sinh^2(\pi y) + 1}{\sinh^4(\pi y)} \to 0 \tag{8.3}$$

when $|y| \to +\infty$. The same is true for $\sum_{n\in\mathbb{Z}} \frac{(-1)^n}{(z-n)^2}$. To see this we take $\epsilon > 0$ and then there is n_0 so that

$$\textstyle\sum_{n\in\mathbb{Z},|n|\geq n_0+1} \frac{1}{(|n|-1)^2} < \frac{\epsilon}{2}.$$

If $z \in A$, i.e. if $z = x + iy$ and $|x| \leq 1$, then $|z - n| \geq |x - n| \geq |n| - 1$ and hence

$$\big|\sum_{n\in\mathbb{Z},|n|\geq n_0+1} \tfrac{(-1)^n}{(z-n)^2}\big| \leq \sum_{n\in\mathbb{Z},|n|\geq n_0+1} \tfrac{1}{|z-n|^2} \leq \sum_{n\in\mathbb{Z},|n|\geq n_0+1} \tfrac{1}{(|n|-1)^2} < \tfrac{\epsilon}{2}. \tag{8.4}$$

Since $\frac{1}{(z-n\pi)^2} \to 0$ when $z \to \infty$, there is $y_0 > 0$ so that

$$\big|\sum_{n\in\mathbb{Z},|n|\leq n_0} \tfrac{(-1)^n}{(z-n)^2}\big| < \tfrac{\epsilon}{2} \tag{8.5}$$

when $z = x + iy$ and $|x| \leq 1$, $|y| > y_0$. From (8.4) and (8.5) we get

$$\big|\sum_{n\in\mathbb{Z}} \tfrac{(-1)^n}{(z-n)^2}\big| \leq \big|\sum_{n\in\mathbb{Z},|n|\leq n_0} \tfrac{(-1)^n}{(z-n)^2}\big| + \big|\sum_{n\in\mathbb{Z},|n|\geq n_0+1} \tfrac{(-1)^n}{(z-n)^2}\big| < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon$$

when $z = x + iy$ and $|x| \leq 1$, $|y| > y_0$. Therefore

$$\sum_{n\in\mathbb{Z}} \tfrac{(-1)^n}{(z-n)^2} \to 0 \tag{8.6}$$

when $z \to \infty$ in A. From (8.2), (8.3) and (8.6) we conclude that $h'(z) \to 0$ when $z \to \infty$ in A. This implies that h' is bounded in the period-zone A and since h' is 2-periodic we have that h' is bounded in $\mathbb{C}$. By the theorem of Liouville, h' is constant in $\mathbb{C}$. But since $h'(z) \to 0$ when $z \to \infty$ in A, we get that $h' = 0$ in $\mathbb{C}$. This implies that h is constant in $\mathbb{C}$.
Now we go back to (8.1) and we observe that the terms $\frac{1}{\sin(\pi z)}$ and $\frac{1}{\pi z}$ are odd functions. The same is true for $\sum_{n\in\mathbb{Z}\setminus\{0\}}(-1)^n\big(\frac{1}{z-n} + \frac{1}{n}\big)$. Indeed,

$$\begin{aligned}\sum_{n\in\mathbb{Z}\setminus\{0\}}(-1)^n\big(\tfrac{1}{-z-n} + \tfrac{1}{n}\big) &= -\sum_{n\in\mathbb{Z}\setminus\{0\}}(-1)^n\big(\tfrac{1}{z+n} - \tfrac{1}{n}\big)\\ &= -\sum_{n\in\mathbb{Z}\setminus\{0\}}(-1)^{-n}\big(\tfrac{1}{z-n} + \tfrac{1}{n}\big)\\ &= -\sum_{n\in\mathbb{Z}\setminus\{0\}}(-1)^n\big(\tfrac{1}{z-n} + \tfrac{1}{n}\big).\end{aligned}$$

Hence h is an odd constant function and this implies that $h = 0$ in $\mathbb{C}$. So we end up with the identity

$$\tfrac{1}{\sin(\pi z)} = \tfrac{1}{\pi z} + \tfrac{1}{\pi}\sum_{n\in\mathbb{Z}\setminus\{0\}}(-1)^n\big(\tfrac{1}{z-n} + \tfrac{1}{n}\big) = \tfrac{1}{\pi z} + \tfrac{2z}{\pi}\sum_{n=1}^{+\infty} \tfrac{(-1)^n}{z^2-n^2}$$

from which, by differentiation, we get (8.2) (with $h' = 0$), i.e.

$$\tfrac{\cos(\pi z)}{\sin^2(\pi z)} = \tfrac{1}{\pi^2}\sum_{n\in\mathbb{Z}} \tfrac{(-1)^n}{(z-n)^2}.$$

In exactly the same manner we can prove the identity

$$\cot(\pi z) = \tfrac{\cos(\pi z)}{\sin(\pi z)} = \tfrac{1}{\pi z} + \tfrac{1}{\pi}\sum_{n\in\mathbb{Z}\setminus\{0\}}\big(\tfrac{1}{z-n} + \tfrac{1}{n}\big) = \tfrac{1}{\pi z} + \tfrac{2z}{\pi}\sum_{n=1}^{+\infty} \tfrac{1}{z^2-n^2} \tag{8.7}$$

from which, by differentiation, we get

$$\tfrac{1}{\sin^2(\pi z)} = \tfrac{1}{\pi^2}\sum_{n\in\mathbb{Z}} \tfrac{1}{(z-n)^2}.$$

Exercises.

8.1.1. Using (8.7) and the Laurent series of $\cot(\pi z)$ at 0, find the values of

$$\sum_{n=1}^{+\infty} \tfrac{1}{n^2}, \quad \sum_{n=1}^{+\infty} \tfrac{1}{n^4}, \quad \sum_{n=1}^{+\infty} \tfrac{1}{n^6}.$$

8.1.2. Express $\sum_{n\in\mathbb{Z}} \frac{1}{(z-n)^2+a^2}$ in closed form.

8.2 Infinite products.

Let (z_n) be a sequence in $\mathbb{C}$. The expression

$$\prod_{n=1}^{+\infty} z_n$$

is called **infinite product** of the $z_1, z_2, \ldots$. We consider three cases.
First case. $z_n \neq 0$ for every n.
We denote $p_n = z_1 \cdots z_n$ the n-th *partial product* of the $z_1, z_2, \ldots$. If $p_n \to p$ for some $p \in \widehat{\mathbb{C}}$, we write

$$\prod_{n=1}^{+\infty} z_n = p$$

and we say that p is the product of the $z_1, z_2, \ldots$. If $p \neq 0$ and $p \neq \infty$, we say that the infinite product **converges** to p. If $p = 0$ or $p = \infty$, we say that the infinite product **diverges** to 0 or to ∞, respectively. If the sequence (p_n) does not have a limit, we say that the infinite product **diverges**.

Example 8.2.1. Let $z_n = 1 + \frac{1}{n}$ for every n. Then

$$p_n = (1+\tfrac{1}{1}) \cdots (1+\tfrac{1}{n}) = \tfrac{2}{1}\,\tfrac{3}{2} \cdots \tfrac{n}{n-1}\,\tfrac{n+1}{n} = n+1.$$

Hence $p_n \to \infty$ and so $\prod_{n=1}^{+\infty}(1+\frac{1}{n}) = \infty$. In this case the infinite product diverges to ∞.
Since all p_n are real, we may also say that $p_n \to +\infty$ and that the infinite product diverges to $+\infty$.

Example 8.2.2. Let $z_n = 1 - \frac{1}{n+1}$ for every n. Then

$$p_n = (1-\tfrac{1}{2}) \cdots (1-\tfrac{1}{n+1}) = \tfrac{1}{2}\,\tfrac{2}{3} \cdots \tfrac{n-1}{n}\,\tfrac{n}{n+1} = \tfrac{1}{n+1}.$$

So $p_n \to 0$ and $\prod_{n=1}^{+\infty}(1-\frac{1}{n+1}) = 0$. In this case the infinite product diverges to 0.

Example 8.2.3. Let $z_n = 1 - \frac{1}{(n+1)^2}$ for every n. Then

$$p_n = (1-\tfrac{1}{2^2}) \cdots (1-\tfrac{1}{(n+1)^2}) = \tfrac{1\cdot 3}{2^2}\,\tfrac{2\cdot 4}{3^2} \cdots \tfrac{(n-1)(n+1)}{n^2}\,\tfrac{n(n+2)}{(n+1)^2} = \tfrac{1}{2}\,\tfrac{n+2}{n+1}.$$

Therefore $p_n \to \frac{1}{2}$ and so $\prod_{n=1}^{+\infty}(1-\frac{1}{(n+1)^2}) = \frac{1}{2}$. In this case the infinite product converges to $\frac{1}{2}$.

Second case. There is m so that $z_n \neq 0$ for every $n \geq m+1$ and $z_n = 0$ for at least one $n \leq m$.
If the infinite product $\prod_{n=m+1}^{+\infty} z_n$ diverges, we say that the infinite product $\prod_{n=1}^{+\infty} z_n$ **diverges**. Now let $\prod_{n=m+1}^{+\infty} z_n = p'$ for some $p' \in \widehat{\mathbb{C}}$. If $p' \neq 0$ and $p' \neq \infty$, then $\prod_{n=m+1}^{+\infty} z_n$ converges to p', and we say that $\prod_{n=1}^{+\infty} z_n$ **converges** to $p = \big(\prod_{n=1}^{m} z_n\big)p' = 0p' = 0$. If $p' = 0$, then $\prod_{n=m+1}^{+\infty} z_n$ diverges to 0, and we say that $\prod_{n=1}^{+\infty} z_n$ **diverges** to $p = \big(\prod_{n=1}^{m} z_n\big)0 = 00 = 0$. If $p' = \infty$, then $\prod_{n=m+1}^{+\infty} z_n$ diverges to ∞, and we say that $\prod_{n=1}^{+\infty} z_n$ **diverges**.
Third case. There are infinitely many n so that $z_n = 0$.
Then we say that $\prod_{n=1}^{+\infty} z_n$ **diverges**.

Therefore, in any case, *the infinite product $\prod_{n=1}^{+\infty} z_n$ converges if and only if there is m so that $z_n \neq 0$ for every $n \geq m+1$ and the partial products $z_{m+1} \cdots z_n$ converge (as $n \to +\infty$) to some complex number* $\neq 0$. Moreover, if $\prod_{n=1}^{+\infty} z_n$ converges, its value is equal to 0 if and only if $z_n = 0$ for at least one n.

Proposition 8.1. *If $\prod_{n=1}^{+\infty} z_n$ converges, then $z_n \to 1$.*

Proof. There is m so that $z_n \neq 0$ for every $n \geq m+1$ and $\prod_{n=m+1}^{+\infty} z_n = p'$ where $p' \neq 0, \infty$. Then $z_{m+1} \cdots z_n \to p'$, as $n \to +\infty$, and thus $z_n = \frac{z_{m+1} \cdots z_{n-1} z_n}{z_{m+1} \cdots z_{n-1}} \to \frac{p'}{p'} = 1$. □

From now on we shall use the symbol

$$\prod_{n=1}^{+\infty}(1+a_n)$$

for the infinite product. According to the previous discussion, convergence of the infinite product implies that $a_n \to 0$.

There are three simple inequalities which play some role in the study of infinite products. The first two are:

$$1+a_1+\cdots+a_n \le (1+a_1)\cdots(1+a_n) \le e^{a_1+\cdots+a_n} \tag{8.8}$$

when $0 \le a_1,\dots,a_n$. The left inequality is easily proved by induction and the right inequality is based on the well known $1+x\le e^x$. The third inequality is:

$$1-a_1-\cdots-a_n \le (1-a_1)\cdots(1-a_n) \tag{8.9}$$

when $0\le a_1,\dots,a_n\le 1$. This is proved also by induction.

Lemma 8.1. *Let $a_n \ge 0$ for every n. Then $\prod_{n=1}^{+\infty}(1+a_n)$ converges if and only if $\sum_{n=1}^{+\infty} a_n$ converges.*

Proof. We set $p_n=(1+a_1)\cdots(1+a_n)$ for every n. Then the sequence (p_n) is increasing and we have $p_n\ge 1$ for every n. Then $p=\lim_{n\to+\infty}p_n$ exists and $1\le p\le+\infty$. We also denote $s=\sum_{n=1}^{+\infty}a_n$ and we have $0\le s\le+\infty$. Taking the limit in (8.8) we find $1+s\le p\le e^s$. Thus, $p<+\infty$ if and only if $s<+\infty$. □

Example 8.2.4. $\prod_{n=1}^{+\infty}(1+\frac{1}{n})=+\infty$, because $\sum_{n=1}^{+\infty}\frac{1}{n}=+\infty$.
$\prod_{n=1}^{+\infty}(1+\frac{1}{n^2})$ converges, because $\sum_{n=1}^{+\infty}\frac{1}{n^2}<+\infty$.

We say that the infinite product $\prod_{n=1}^{+\infty}(1+a_n)$ **converges absolutely** if the infinite product $\prod_{n=1}^{+\infty}(1+|a_n|)$ converges or, equivalently, if the series $\sum_{n=1}^{+\infty}|a_n|$ converges.

Criterion of absolute convergence. *If $\prod_{n=1}^{+\infty}(1+a_n)$ converges absolutely, then it converges.*

Proof. Since $\sum_{n=1}^{+\infty}|a_n|<+\infty$, we have that $a_n\to 0$ and so at most finitely many a_n are equal to -1.
We denote

$$p_n=(1+a_1)\cdots(1+a_n),\qquad P_n=(1+|a_1|)\cdots(1+|a_n|).$$

Then, if $n<m$, we have

$$\begin{aligned}|p_m-p_n| &= \big|\prod_{k=1}^{m}(1+a_k)-\prod_{k=1}^{n}(1+a_k)\big|\\ &= \big|\prod_{k=1}^{n}(1+a_k)\big(\prod_{k=n+1}^{m}(1+a_k)-1\big)\big|\\ &= \prod_{k=1}^{n}|1+a_k|\big|\prod_{k=n+1}^{m}(1+a_k)-1\big|\\ &\le \prod_{k=1}^{n}(1+|a_k|)\big(\prod_{k=n+1}^{m}(1+|a_k|)-1\big)\\ &= \prod_{k=1}^{m}(1+|a_k|)-\prod_{k=1}^{n}(1+|a_k|)=P_m-P_n.\end{aligned} \tag{8.10}$$

Since $\prod_{n=1}^{+\infty}(1+|a_n|)$ converges, we have that (P_n) is a Cauchy sequence. So the last inequality implies that (p_n) is also a Cauchy sequence and hence converges.
Now we have two cases.
Let $\sum_{n=1}^{+\infty}|a_n|<1$. Then

$$|p_n|=\prod_{k=1}^{n}|1+a_k|\ge\prod_{k=1}^{n}(1-|a_k|)\ge 1-\sum_{k=1}^{n}|a_k|\ge 1-\sum_{k=1}^{+\infty}|a_k|>0,$$

where for the second inequality we use (8.9). This implies $|\lim_{n\to+\infty}p_n|\ge 1-\sum_{k=1}^{+\infty}|a_k|>0$. Therefore $\lim_{n\to+\infty}p_n\ne 0$ and hence $\prod_{n=1}^{+\infty}(1+a_n)$ converges.
Now let $\sum_{n=1}^{+\infty}|a_n|\ge 1$. Then there is m so that $\sum_{n=m+1}^{+\infty}|a_n|<1$ and from the first case we have that $\prod_{n=m+1}^{+\infty}(1+a_n)$ converges. Hence $\prod_{n=1}^{+\infty}(1+a_n)$ also converges. □

From now on we consider *infinite products of functions*.

Proposition 8.2. *Let $a_n : A \to \mathbb{C}$ be bounded functions in A and let $\sum_{n=1}^{+\infty} |a_n|$ converge uniformly in A. Then $\prod_{n=1}^{+\infty}(1 + a_n)$ converges uniformly in A.*

Proof. Since $\sum_{n=1}^{+\infty} |a_n(z)|$ converges for every $z \in A$, we have that $\prod_{n=1}^{+\infty}(1+a_n(z))$ converges absolutely and so it converges for every $z \in A$. We define $p : A \to \mathbb{C}$ by

$$p(z) = \prod_{n=1}^{+\infty}(1 + a_n(z))$$

for every $z \in A$. The uniform convergence of $\sum_{n=1}^{+\infty} |a_n|$ in A implies that there is M so that $\sum_{n=1}^{+\infty} |a_n(z)| \leq M$ for every $z \in A$.
We set

$$p_n(z) = \prod_{k=1}^{n}(1 + a_k(z)), \quad S_n(z) = \sum_{k=1}^{n} |a_k(z)|, \quad S(z) = \sum_{k=1}^{+\infty} |a_k(z)|.$$

We apply (8.10) and we get

$$|p_m(z) - p_n(z)| \leq \prod_{k=1}^{n}(1 + |a_k(z)|)\big(\prod_{k=n+1}^{m}(1 + |a_k(z)|) - 1\big)$$

for $n < m$ and $z \in A$. We apply the right side of (8.8) and then we let $m \to +\infty$ to find

$$|p(z) - p_n(z)| \leq e^{S_n(z)}\big(e^{S(z)-S_n(z)} - 1\big) \leq e^M\big(e^{S(z)-S_n(z)} - 1\big) = e^M\big(e^{|S(z)-S_n(z)|} - 1\big)$$

for every n and $z \in A$. Therefore, $\|p - p_n\|_A \leq e^M\big(e^{\|S-S_n\|_A} - 1\big)$ for every n. Since $S_n \to S$ uniformly in A, we have that $p_n \to p$ uniformly in A. □

Now we state the analogue of the theorem of Weierstrass for the uniform convergence of series of holomorphic functions in compact sets.

Theorem 8.1. *Let (a_n) be a sequence of holomorphic functions in the region Ω, so that $\sum_{n=1}^{+\infty} |a_n|$ converges uniformly in every compact subset of Ω. Then $\prod_{n=1}^{+\infty}(1 + a_n)$ converges uniformly in every compact subset of Ω and it defines a function*

$$p = \prod_{n=1}^{+\infty}(1 + a_n),$$

which is holomorphic in Ω. Moreover, $p(z) = 0$ if and only if $a_n(z) = -1$ for at least one n. Finally, if none of the a_n is identically -1 in Ω, we have that

$$\frac{p'}{p} = \sum_{n=1}^{+\infty} \frac{a_n'}{1+a_n} \tag{8.11}$$

at every point in Ω which is not a root of p. The series in (8.11) has the property: for every compact $K \subseteq \Omega$ there is n_0 so that $\sum_{n=n_0+1}^{+\infty} \frac{a_n'}{1+a_n}$ converges uniformly in K.

Proof. Of course, proposition 8.2 implies that $\prod_{n=1}^{+\infty}(1 + a_n)$ converges uniformly in every compact subset of Ω. Every $p_n = \prod_{k=1}^{n}(1+a_k)$ is holomorphic in Ω. Since $p_n \to p$ uniformly in every compact subset of Ω, the theorem of Weierstrass implies that p is holomorphic in Ω. Moreover, for every $z \in \Omega$ we have $p(z) = \prod_{n=1}^{+\infty}(1 + a_n(z))$ and, since the product converges, we have that $p(z) = 0$ if and only if $a_n(z) = -1$ for at least one n.
Now, let us assume that none of the a_n is identically -1 in Ω. Then every root of the function $1 + a_n$ is isolated and hence the set of the roots of $1 + a_n$ is countable. From the first part of the theorem we have that the set of the roots of p is also countable and hence p is not identically 0 in Ω. In particular, the roots of p are isolated and if we take any compact $K \subseteq \Omega$ then there are only finitely many roots, say $z_1, \ldots, z_m$, of p in K. Now, by the convergence of the infinite product, for each $j = 1, \ldots, m$, there is n_j so that $a_n(z_j) \neq -1$ for every $n \geq n_j + 1$. If we set $n_0 = \max\{n_1, \ldots, n_m\}$, then we have that $a_n(z_j) \neq -1$ for every $n \geq n_0 + 1$ and for every

$j = 1, \dots, m$. Moreover, since p has no root in K other than $z_1, \dots, z_m$, we have that $a_n(z) \neq -1$ for every $n \geq n_0 + 1$ and for every $z \in K$.
Now we consider the infinite product $q = \prod_{n=n_0+1}^{+\infty}(1 + a_n)$ and the partial products $q_n = \prod_{k=n_0+1}^{n}(1 + a_k)$. Of course, we have that $q_n \to q$ uniformly in K and also $q_n' \to q'$ uniformly in K. We also have that q has no root in K and so there is $\delta > 0$ so that $|q(z)| \geq \delta$ for every $z \in K$. These imply that $\frac{q_n'}{q_n} \to \frac{q'}{q}$ uniformly in K. On the other hand, it is trivial to show that

$$\frac{q_n'}{q_n} = \sum_{k=n_0+1}^{n} \frac{a_k'}{1+a_k}$$

and so

$$\frac{q'}{q} = \sum_{n=n_0+1}^{+\infty} \frac{a_n'}{1+a_n} \tag{8.12}$$

uniformly in K. At last, from $p_n = \prod_{k=1}^{n_0}(1 + a_k)\, q_n$ and from $p = \prod_{k=1}^{n_0}(1 + a_k)\, q$, we also get

$$\frac{p_n'}{p_n} = \sum_{k=1}^{n_0} \frac{a_k'}{1+a_k} + \frac{q_n'}{q_n}, \qquad \frac{p'}{p} = \sum_{k=1}^{n_0} \frac{a_k'}{1+a_k} + \frac{q'}{q} \tag{8.13}$$

at every point in K which is not a root of p. From (8.12) and (8.13) we get (8.11) at every $z \in K$ which is not a root of p. Since K is an arbitrary compact subset of Ω, we conclude that (8.11) holds at every point in Ω which is not a root of p. □

Exercises.

8.2.1. Recall from the proof of the argument principle, that the roots of p are simple poles of $\frac{p'}{p}$. What are the corresponding residues? Now prove that (8.11) holds also at the roots of p.

8.2.2. Prove theorem 8.1 under the assumption of the uniform convergence of $\prod_{n=1}^{+\infty}(1 + a_n)$ in every compact subset of Ω. *Do not* assume that $\sum_{n=1}^{+\infty} |a_n|$ converges uniformly in every compact subset of Ω.

8.3 Holomorphic functions in the complex plane.

We know that every non-zero polynomial of degree n can be written as

$$p(z) = c(z - z_1)^{m_1} \cdots (z - z_k)^{m_k}$$

where $z_1, \dots, z_k$ are the distinct roots of p and $m_1, \dots, m_k$ are the corresponding multiplicities. In particular, $m_1 + \cdots + m_k = n$.

Let f be a non-zero function holomorphic in the region Ω and let $z_1, \dots, z_k$ be *all* the roots of f in Ω with corresponding multiplicities $m_1, \dots, m_k$. We know that we can factorize $(z - z_1)^{m_1}$ from f, i.e. that $f(z) = (z - z_1)^{m_1} g(z)$ for every $z \in \Omega$, where g is holomorphic in Ω with $g(z_1) \neq 0$. Now g has roots $z_2, \dots, z_n$ with corresponding multiplicities $m_2, \dots, m_n$. Similarly, $g(z) = (z - z_2)^{m_2} h(z)$ for every $z \in \Omega$, where h is holomorphic in Ω with $h(z_1) \neq 0$, $h(z_2) \neq 0$. Now h has roots $z_3, \dots, z_n$ with corresponding multiplicities $m_3, \dots, m_n$. Continuing inductively, we get that

$$f(z) = (z - z_1)^{m_1} \cdots (z - z_k)^{m_k} F(z)$$

for every $z \in \Omega$, where F is holomorphic and has no roots in Ω. If we do not want to show the multiplicities of the roots except for the (possible) root at 0 we may simply write

$$f(z) = z^m (z - z_1) \cdots (z - z_n) F(z),$$

where $m \geq 0$ is the multiplicity of the root 0 and $z_1, \dots, z_n$ are the remaining (not necessarily distinct) non-zero roots of f in Ω.

The question now is to generalize this situation in case f has infinitely many roots $0, z_1, z_2, \dots$. In this case the corresponding infinite product $z^m(z - z_1)(z - z_2) \cdots$ may not converge.

To prepare for what will follow, we rewrite the last identity in the form

$$f(z)=z^m(1-\tfrac{z}{z_1})\cdots(1-\tfrac{z}{z_n})F(z),$$

where the new F is the previous F multiplied by the non-zero number $(-1)^n z_1\cdots z_n$. We also note that if the region is *simply connected*, e.g. if $\Omega=\mathbb{D}$ or $\Omega=\mathbb{C}$, then, since F has no roots in Ω, there is a holomorphic branch g of F in Ω. So the last identity becomes

$$f(z)=z^m(1-\tfrac{z}{z_1})\cdots(1-\tfrac{z}{z_n})e^{g(z)}$$

for every $z\in\Omega$, where g is holomorphic in Ω. This is the most general form of a holomorphic function in the simply connected region Ω with finitely many preassigned roots (and no other roots).

In the following discussion we shall concentrate only in the case $\Omega=\mathbb{C}$.

Lemma 8.2. *We have* $|e^z-1|\leq\frac{8}{7}\,|z|$ *for every* z *with* $|z|\leq\frac{1}{4}$.

Proof. Since $2^{k-1}\leq k!$ when $k\geq 1$, we get

$$|e^z-1|=\big|\sum_{k=1}^{+\infty}\tfrac{z^k}{k!}\big|\leq\sum_{k=1}^{+\infty}\tfrac{|z|^k}{k!}=|z|\sum_{k=1}^{+\infty}\tfrac{|z|^{k-1}}{k!}\leq|z|\sum_{k=1}^{+\infty}\tfrac{|z|^{k-1}}{2^{k-1}}=\tfrac{|z|}{1-\frac{|z|}{2}}\leq\tfrac{8}{7}\,|z|$$

when $|z|\leq\frac{1}{8}$. □

We set

$$p_0(z)=1-z,\qquad p_m(z)=(1-z)e^{z+\frac{z^2}{2}+\cdots+\frac{z^m}{m}}\quad\text{when } m\geq 1.$$

Lemma 8.3. *For every* $m\geq 0$ *we have*

$$|p_m(z)-1|\leq\tfrac{3|z|^{m+1}}{m+1}$$

when $|z|\leq\frac{1}{2}$.

Proof. For $m=0$ we have $|p_0(z)-1|=|z|\leq 3|z|$.
Now let $m\geq 1$. If $|z|\leq\frac{1}{2}$, then by the Taylor series of $\mathrm{Log}(1-z)$ in the disc $D_0(1)$ we get

$$\begin{aligned}\big|\,\mathrm{Log}(1-z)+z+\tfrac{z^2}{2}+\cdots+\tfrac{z^m}{m}\big|&=\big|\sum_{k=m+1}^{+\infty}\tfrac{z^k}{k}\big|\leq\sum_{k=m+1}^{+\infty}\tfrac{|z|^k}{k}\leq\tfrac{|z|^{m+1}}{m+1}\sum_{k=0}^{+\infty}|z|^k\\&=\tfrac{|z|^{m+1}}{(m+1)(1-|z|)}\leq\tfrac{2|z|^{m+1}}{m+1}.\end{aligned}$$

If $|z|\leq\frac{1}{2}$, then $\frac{2|z|^{m+1}}{m+1}\leq\frac{1}{(m+1)2^m}\leq\frac{1}{4}$. Thus, lemma 8.2 implies

$$|p_m(z)-1|=|e^{\mathrm{Log}(1-z)+z+\frac{z^2}{2}+\cdots+\frac{z^m}{m}}-1|\leq\tfrac{8}{7}\,\tfrac{2|z|^{m+1}}{m+1}\leq\tfrac{3|z|^{m+1}}{m+1}$$

for $|z|\leq\frac{1}{2}$. □

The following is a theorem of Weierstrass.

Theorem 8.2. *Let* (z_n) *be a sequence of non-zero numbers so that* $z_n\to\infty$.
(i) Then there are integers $m_n\geq 0$ *so that*

$$\sum_{n=1}^{+\infty}\tfrac{1}{m_n+1}\big(\tfrac{R}{|z_n|}\big)^{m_n+1}<+\infty$$

for every $R>0$.
(ii) If the integers m_n *satisfy (i) then the function*

$$F(z)=\prod_{n=1}^{+\infty}p_{m_n}\big(\tfrac{z}{z_n}\big)$$

is holomorphic in $\mathbb{C}$ and its only roots are the terms of (z_n). The multiplicity of each z_k as a root of F is the same as the number of its appearances as a term of (z_n). Moreover, the most general holomorphic function in $\mathbb{C}$, whose only roots, besides 0, are the terms of (z_n) and so that the multiplicity of each z_k as a root of f is the same as the number of its appearances as a term of (z_n), is of the form

$$f(z) = z^m e^{g(z)} \prod_{n=1}^{+\infty} p_{m_n}\big(\tfrac{z}{z_n}\big) = z^m e^{g(z)} \prod_{n=1}^{+\infty} \big(1 - \tfrac{z}{z_n}\big) e^{\frac{z}{z_n} + \frac{1}{2}(\frac{z}{z_n})^2 + \cdots + \frac{1}{m_n}(\frac{z}{z_n})^{m_n}},$$

where $m \geq 0$ and g is an arbitrary function holomorphic in $\mathbb{C}$. We also have that

$$\tfrac{f'(z)}{f(z)} = \tfrac{m}{z} + g'(z) + \sum_{n=1}^{+\infty} \big(\tfrac{1}{z - z_n} + \tfrac{1}{z_n} + \tfrac{z}{z_n^2} + \cdots + \tfrac{z^{m_n - 1}}{z_n^{m_n}}\big)$$

at every z which is not a root of f.

Proof. (i) We may consider $m_n = n$ and then, since $z_n \to \infty$, for every $R > 0$ there is n_0 so that $|z_n| \geq 2R$ for every $n \geq n_0 + 1$. This implies that

$$\sum_{n=n_0+1}^{+\infty} \tfrac{1}{n+1} \big(\tfrac{R}{|z_n|}\big)^{n+1} \leq \sum_{n=n_0+1}^{+\infty} \tfrac{1}{(n+1)2^{n+1}} < +\infty.$$

(ii) Let the integers m_n satisfy (i). We consider any compact $K \subseteq \mathbb{C}$, and then there is $R > 0$ so that $K \subseteq \overline{D}_0(R)$. As in (i), there is n_0 so that $|z_n| \geq 2R$ for every $n \geq n_0 + 1$. Now lemma 8.3 implies that for every $z \in K$ we have

$$\big|p_{m_n}\big(\tfrac{z}{z_n}\big) - 1\big| \leq \tfrac{3}{m_n+1} \big(\tfrac{R}{|z_n|}\big)^{m_n+1} < +\infty.$$

By the test of Weierstrass, the series $\sum_{n=1}^{+\infty} |p_{m_n}(\frac{z}{z_n}) - 1|$ converges uniformly in K. Since this is true for an arbitrary compact $K \subseteq \mathbb{C}$, theorem 8.1 implies that the infinite product defines a function

$$F(z) = \prod_{n=1}^{+\infty} p_{m_n}\big(\tfrac{z}{z_n}\big)$$

holomorphic in $\mathbb{C}$. The roots of F are the roots of p_{m_n}, i.e. the terms of (z_n). Also, the multiplicity of each z_k as a root of F is the same as the number of its appearances as a term of (z_n). Theorem 8.1 also implies that

$$\tfrac{F'(z)}{F(z)} = \sum_{n=1}^{+\infty} \big(\tfrac{1}{z - z_n} + \tfrac{1}{z_n} + \tfrac{z}{z_n^2} + \cdots + \tfrac{z^{m_n - 1}}{z_n^{m_n}}\big)$$

at every z which is not a root of F.
Now let f be any holomorphic function in $\mathbb{C}$, whose only roots, besides 0, are the terms of (z_n) and so that the multiplicity of each z_k as a root of f is the same as the number of its appearances as a term of (z_n). Let $m \geq 0$ be the multiplicity of 0 as a root of f. Then the function $\frac{f(z)}{z^m F(z)}$ is holomorphic in $\mathbb{C}$ and has no roots. So there is some function g holomorphic in $\mathbb{C}$ so that $\frac{f(z)}{z^m F(z)} = e^{g(z)}$ for every z. Finally, from $f(z) = z^m e^{g(z)} F(z)$ we easily get that

$$\tfrac{f'(z)}{f(z)} = \tfrac{m}{z} + g'(z) + \tfrac{F'(z)}{F(z)}$$

and the proof is complete. ☐

The functions $p_{m_n}(\frac{z}{z_n})$ appearing in the product expansion of f and of F in theorem 8.2 are called **primary factors of Weierstrass**.

There is an important special case of theorem 8.2. It is the case when all integers $m_n \geq 0$ can be taken to be equal to the same integer $h \geq 0$. This means that

$$\sum_{n=1}^{+\infty} \tfrac{1}{|z_n|^{h+1}} < +\infty.$$

If this is true for some integer $h \geq 0$ and we consider the *smallest* such h, called **genus** of the sequence of roots (z_n), then the most general holomorphic function f in $\mathbb{C}$, whose only roots, besides 0, are the terms of (z_n) and so that the multiplicity of each z_k as a root of f is the same as the number of its appearances as a term of (z_n), is of the form

$$f(z) = z^m e^{g(z)} \prod_{n=1}^{+\infty} p_h\big(\tfrac{z}{z_n}\big) = z^m e^{g(z)} \prod_{n=1}^{+\infty} \big(1 - \tfrac{z}{z_n}\big) e^{\frac{z}{z_n} + \frac{1}{2}(\frac{z}{z_n})^2 + \cdots + \frac{1}{h}(\frac{z}{z_n})^h},$$

where $m \geq 0$ and g is an arbitrary function holomorphic in $\mathbb{C}$.

Example 8.3.1. The function $\sin(\pi z)$ is holomorphic in $\mathbb{C}$ and its roots are the integers $n \in \mathbb{Z}$. Each root is of multiplicity 1. For the non-zero roots we have that

$$\sum_{n\in\mathbb{Z}\setminus\{0\}} \tfrac{1}{|n|} = +\infty, \quad \sum_{n\in\mathbb{Z}\setminus\{0\}} \tfrac{1}{|n|^2} < +\infty.$$

Thus, we may use $h = 1$ in order to apply theorem 8.2 and we get that

$$\sin(\pi z) = z e^{g(z)} \prod_{n\in\mathbb{Z}\setminus\{0\}} \big(1 - \tfrac{z}{n}\big) e^{\frac{z}{n}}$$

for some g holomorphic in $\mathbb{C}$. We also have that

$$\tfrac{\pi \cos(\pi z)}{\sin(\pi z)} = \tfrac{1}{z} + g'(z) + \sum_{n\in\mathbb{Z}\setminus\{0\}} \big(\tfrac{1}{z-n} + \tfrac{1}{n}\big).$$

Now, (8.7) implies that $g'(z) = 0$ for every z and so g is constant in $\mathbb{C}$. Then e^g is a constant, say c, and then we have that

$$\tfrac{\sin(\pi z)}{z} = c \prod_{n\in\mathbb{Z}\setminus\{0\}} \big(1 - \tfrac{z}{n}\big) e^{\frac{z}{n}}$$

for every z. Both sides of this equality are holomorphic in $\mathbb{C}$ and setting $z = 0$ to it we get $c = \pi$. Therefore,

$$\sin(\pi z) = \pi z \prod_{n\in\mathbb{Z}\setminus\{0\}} \big(1 - \tfrac{z}{n}\big) e^{\frac{z}{n}} = \pi z \prod_{n=1}^{+\infty} \big(1 - \tfrac{z^2}{n^2}\big). \tag{8.14}$$

This is the **formula of Wallis**.

8.4 Euler's gamma function.

Lemma 8.4. *The limit*

$$\gamma = \lim_{n\to+\infty} \big(1 + \tfrac{1}{2} + \cdots + \tfrac{1}{n} - \ln n\big)$$

exists and $0 < \gamma < 1$. *The constant* γ *is the so-called* **Euler's constant**.

Proof. We have

$$\tfrac{1}{k+1} = \int_k^{k+1} \tfrac{1}{k+1}\, dt < \int_k^{k+1} \tfrac{1}{t}\, dt < \int_k^{k+1} \tfrac{1}{k}\, dt = \tfrac{1}{k}$$

and hence

$$\tfrac{1}{k+1} < \ln \tfrac{k+1}{k} < \tfrac{1}{k}. \tag{8.15}$$

We observe that

$$1 + \tfrac{1}{2} + \cdots + \tfrac{1}{n} - \ln(n+1) = \sum_{k=1}^{n} \big(\tfrac{1}{k} - \ln \tfrac{k+1}{k}\big)$$

is the n-th partial sum of the series

$$\sum_{k=1}^{+\infty} \big(\tfrac{1}{k} - \ln \tfrac{k+1}{k}\big).$$

Now (8.15) says that this series has positive terms and that it is dominated by $\sum_{k=1}^{+\infty}(\frac{1}{k} - \frac{1}{k+1}) = 1$. So the limit γ of the partial sums $1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln(n+1)$ exists and $0 < \gamma < 1$. □

As in example 8.3.1, we use the sequence $z_n = -n$, $n \in \mathbb{N}$, to apply theorem 8.2, and we form the function

$$f(z) = ze^{\gamma z} \prod_{n=1}^{+\infty} \left(1 + \tfrac{z}{n}\right) e^{-\frac{z}{n}}.$$

This function is holomorphic in $\mathbb{C}$ and has simple roots at the points $0, -1, -2, -3, \ldots$.

Definition. *We define the function*

$$\Gamma(z) = \tfrac{1}{f(z)} = z^{-1} e^{-\gamma z} \prod_{n=1}^{+\infty} \left(1 + \tfrac{z}{n}\right)^{-1} e^{\frac{z}{n}}.$$

This is Euler's **gamma function**.

It is clear that the gamma function is meromorphic in $\mathbb{C}$. It has simple poles at the points $0, -1, -2, -3, \ldots$ and it is otherwise holomorphic in $\mathbb{C}$. It is also clear that the gamma function has no roots in $\mathbb{C}$ and that $\Gamma(1) = 1$.

If we restrict the gamma function in $\mathbb{C} \setminus (-\infty, 0]$, then we have a holomorphic function which does not vanish in a simply connected region, and so there is a holomorphic branch of the logarithm of the gamma function in this region. We shall denote

$$\log \Gamma$$

this branch of the logarithm and we may uniquely determine it by setting $\log \Gamma(1) = \operatorname{Log} 1 = 0$. Now, it is clear from the defining formula of the gamma function that $\Gamma(x) > 0$ for every $x > 0$. And then it is obvious, by the uniqueness of branches of logarithms in connected sets, that

$$\log \Gamma(x) = \ln(\Gamma(x))$$

for all $x \in (0, +\infty)$.

Proposition 8.3.

$$\Gamma(z) = \lim_{n \to +\infty} \tfrac{n^z n!}{z(z+1)\cdots(z+n)}.$$

This is the **formula of Gauss** *for the gamma function.*

Proof. By the definition of $\Gamma(z)$ we have

$$\Gamma(z) = \lim_{n \to +\infty} z^{-1} e^{-\left(\sum_{k=1}^{n} \frac{1}{k} - \ln n\right) z} \prod_{k=1}^{n} \left(1 + \tfrac{z}{k}\right)^{-1} e^{\sum_{k=1}^{n} \frac{z}{k}} = \lim_{n \to +\infty} \tfrac{n^z n!}{z(z+1)\cdots(z+n)}.$$

□

Proposition 8.4. *The gamma function satisfies the functional equation*

$$\Gamma(z+1) = z\Gamma(z).$$

Proof. By the formula of Gauss,

$$\Gamma(z+1) = \lim_{n \to +\infty} \tfrac{n^{z+1} n!}{(z+1)(z+2)\cdots(z+1+n)} = z \lim_{n \to +\infty} \tfrac{n^z n!}{z(z+1)\cdots(z+n)} \tfrac{n}{z+1+n} = z\Gamma(z).$$

□

Proposition 8.5.

$$\Gamma(n) = (n-1)!$$

for every integer $n \in \mathbb{N}$.

Proof. Since $\Gamma(1) = 1$, the result is implied by the functional equation and a trivial induction. □

Proposition 8.6.

$$\Gamma(z)\Gamma(1-z) = \tfrac{\pi}{\sin(\pi z)}.$$

Proof. We use the functional equation to write

$$\begin{aligned}\Gamma(z)\Gamma(1-z) &= \Gamma(z)(-z)\Gamma(-z)\\ &= z^{-1}e^{-\gamma z}\textstyle\prod_{n=1}^{+\infty}\big(1+\frac{z}{n}\big)^{-1}e^{\frac{z}{n}}\,(-z)(-z)^{-1}e^{\gamma z}\prod_{n=1}^{+\infty}\big(1-\frac{z}{n}\big)^{-1}e^{-\frac{z}{n}}\\ &= z^{-1}\textstyle\prod_{n\in\mathbb{Z}\setminus\{0\}}\big(1-\frac{z}{n}\big)^{-1}e^{-\frac{z}{n}} = \frac{\pi}{\sin(\pi z)}.\end{aligned}$$

For the last equality we used the formula (8.14) of Wallis. ☐

Now substituting $z=\frac{1}{2}$ in the last equality we get

$$\Gamma\big(\tfrac{1}{2}\big) = \sqrt{\pi}.$$

Proposition 8.7.

$$\Gamma(z)\Gamma(z+\tfrac{1}{2}) = \sqrt{\pi}\,2^{1-2z}\Gamma(2z).$$

This is the **duplication formula**.

Proof. By the formula of Gauss,

$$\begin{aligned}\Gamma(z)\Gamma\big(z+\tfrac{1}{2}\big) &= \lim_{n\to+\infty}\tfrac{n^z n!}{z(z+1)\cdots(z+n)}\,\tfrac{n^{z+\frac{1}{2}}n!}{(z+\frac{1}{2})(z+\frac{1}{2}+1)\cdots(z+\frac{1}{2}+n)}\\ &= \lim_{n\to+\infty}\tfrac{(2n)^{2z}(2n)!}{(2z)(2z+1)(2z+2)(2z+3)\cdots(2z+2n)}\,\tfrac{2^{2n+2-2z}n^{\frac{1}{2}}(n!)^2}{(2n)!(2z+2n+1)}\\ &= \Gamma(2z)\lim_{n\to+\infty}\tfrac{2^{2n+2-2z}n^{\frac{1}{2}}(n!)^2}{(2n)!(2z+2n+1)}.\end{aligned}$$

On the other hand,

$$\sqrt{\pi} = \Gamma\big(\tfrac{1}{2}\big) = \lim_{n\to+\infty}\tfrac{n^{\frac{1}{2}}n!}{\frac{1}{2}(\frac{1}{2}+1)\cdots(\frac{1}{2}+n)} = \lim_{n\to+\infty}\tfrac{2^{2n+1}n^{\frac{1}{2}}(n!)^2}{(2n)!(2n+1)}.$$

These two equalities imply the duplication formula. ☐

Proposition 8.8.

$$\tfrac{\Gamma'(z)}{\Gamma(z)} = -\gamma-\tfrac{1}{z}-\textstyle\sum_{n=1}^{+\infty}\big(\frac{1}{z+n}-\frac{1}{n}\big),\qquad \big(\frac{\Gamma'(z)}{\Gamma(z)}\big)' = \sum_{n=0}^{+\infty}\frac{1}{(z+n)^2}.$$

Proof. We use theorem 8.2 of Weierstrass to calculate $\frac{\Gamma'(z)}{\Gamma(z)} = -\frac{f'(z)}{f(z)}$ and then we differentiate the resulting series termwise, due to its uniform convergence in compact sets. ☐

Of course, if we restrict in $\mathbb{C}\setminus(-\infty,0]$ then $\frac{\Gamma'(z)}{\Gamma(z)}$ and $\big(\frac{\Gamma'(z)}{\Gamma(z)}\big)'$ are the logarithmic derivatives $(\log\Gamma)'$ and $(\log\Gamma)''$.

The following proposition states some results for the restriction of the gamma function in $(0,+\infty)$. Some of them have been already proved.

Proposition 8.9. *(i)* $\Gamma(x)>0$ *for every* $x\in(0,+\infty)$.
(ii) $\lim_{x\to0+}\Gamma(x)=+\infty$ *and* $\lim_{x\to+\infty}\Gamma(x)=+\infty$.
(iii) $\Gamma(x+1)=x\Gamma(x)$ *for every* $x\in(0,+\infty)$.
(iv) $\Gamma(n)=(n-1)!$ *for every* $n\in\mathbb{N}$.
(v) $\log\Gamma=\ln\Gamma$ *is convex in* $(0,+\infty)$.
(vi) $\Gamma(x)\Gamma''(x)\geq(\Gamma'(x))^2$ *for every* $x\in(0,+\infty)$.

Proof. Only (ii), (v) and (vi) need to be proved.
(ii) We know that 0 is a pole of Γ and that $\Gamma(x) > 0$ for $x > 0$. This implies that $\lim_{x\to 0+}\Gamma(x) = +\infty$. A second proof uses (iii):

$$\lim_{x\to 0+}\Gamma(x) = \lim_{x\to 0+}\Gamma(x+1)\lim_{x\to 0+}\tfrac{1}{x} = \Gamma(1)(+\infty) = +\infty.$$

Now, if $n = [x]$, then $1 \leq x - n + 1 \leq 2$ and since Γ is continuous and positive in $[0,1]$, there is $\delta > 0$ so that $\Gamma(x-n+1) \geq \delta$. Then (iii) implies

$$\Gamma(x) = (x-1)(x-2)\cdots(x-n+1)\Gamma(x-n+1) \geq (n-1)(n-2)\cdots 1\delta = (n-1)!\delta \to +\infty$$

when $x \to +\infty$.
(v) Proposition 8.8 implies

$$(\operatorname{Log}\Gamma)''(x) = \big(\tfrac{\Gamma'(x)}{\Gamma(x)}\big)' = \textstyle\sum_{n=0}^{+\infty}\tfrac{1}{(x+n)^2} > 0$$

for every $x > 0$.
(vi) We just observe that

$$(\operatorname{Log}\Gamma)''(x) = \big(\tfrac{\Gamma'(x)}{\Gamma(x)}\big)' = \tfrac{\Gamma(x)\Gamma''(x)-(\Gamma'(x))^2}{(\Gamma(x))^2}$$

and the inequality to be proved is equivalent to the convexity of $\operatorname{Log}\Gamma$. $\square$

Now we consider the so called **Euler's second integral**:

$$\textstyle\int_0^{+\infty} t^{z-1}e^{-t}\,dt.$$

When we write t^{z-1} with $t > 0$, we mean $t^{z-1} = e^{(z-1)\ln t}$.

At first we consider the case when $z = x$ is real and > 0.

If $x > 1$, then $t^{x-1}e^{-t}$ as a function of t is continuous in $[0,+\infty)$. So $\int_0^1 t^{x-1}e^{-t}\,dt$ is a common integral. On the other hand $\int_1^{+\infty} t^{x-1}e^{-t}\,dt$ converges. To see this we consider $n \in \mathbb{N}$ so that $n > x$. Then $e^t = \sum_{k=0}^{+\infty}\frac{t^k}{k!}$ implies $e^t \geq \frac{t^n}{n!}$ for $t \geq 0$ and hence

$$0 \leq \textstyle\int_1^{+\infty} t^{x-1}e^{-t}\,dt \leq n!\int_1^{+\infty} t^{x-n-1}\,dt < +\infty.$$

Therefore $\int_0^{+\infty} t^{x-1}e^{-t}\,dt$ converges.

If $0 < x \leq 1$, then $t^{x-1}e^{-t}$ as a function of t is continuous in $(0,+\infty)$. The integral $\int_1^{+\infty} t^{x-1}e^{-t}\,dt$ is still convergent. Regarding $\int_0^1 t^{x-1}e^{-t}\,dt$ we see that

$$0 \leq \textstyle\int_0^1 t^{x-1}e^{-t}\,dt \leq \int_0^1 t^{x-1}\,dt < +\infty.$$

Therefore $\int_0^{+\infty} t^{x-1}e^{-t}\,dt$ is convergent again.

In the case when z is complex with $\operatorname{Re} z > 0$, we see that

$$\textstyle\int_0^{+\infty} |t^{z-1}e^{-t}|\,dt = \int_0^{+\infty} t^{\operatorname{Re} z-1}e^{-t}\,dt < +\infty$$

and so $\int_0^{+\infty} t^{z-1}e^{-t}\,dt$ converges and

$$\big|\textstyle\int_0^{+\infty} t^{z-1}e^{-t}\,dt\big| \leq \int_0^{+\infty} t^{\operatorname{Re} z-1}e^{-t}\,dt.$$

Now we are going to see that the function

$$F(z) = \textstyle\int_0^{+\infty} t^{z-1}e^{-t}\,dt$$

is holomorphic in the right half-plane $\{z \,|\, 0 < \operatorname{Re} z\}$. At first we consider the integral

$$F_{\delta,R}(z) = \int_\delta^R t^{z-1} e^{-t}\, dt$$

when $0 < \delta < 1 < R < +\infty$, and we leave it as an exercise to the reader to prove that $F_{\delta,R}$ is holomorphic in $\{z \,|\, 0 < \operatorname{Re} z\}$ and, in fact, that it is holomorphic in all of $\mathbb{C}$, with derivatives

$$F_{\delta,R}^{(n)}(z) = \int_\delta^R t^{z-1} e^{-t} \ln^n t\, dt.$$

We now prove that

$$\lim_{\delta \to 0+, R \to +\infty} F_{\delta,R} = F$$

uniformly in every compact subset of $\{z \,|\, 0 < \operatorname{Re} z\}$. In fact, since every compact subset of $\{z \,|\, 0 < \operatorname{Re} z\}$ is contained in some vertical zone of the form $\{z \,|\, a \le \operatorname{Re} z \le b\}$, $0 < a < b$, we shall prove that the convergence is uniform in every such zone. Indeed, for every z with $a \le \operatorname{Re} z \le b$ we get

$$|F_{\delta,R}(z) - F(z)| \le \int_0^\delta t^{\operatorname{Re} z-1} e^{-t}\, dt + \int_R^{+\infty} t^{\operatorname{Re} z-1} e^{-t}\, dt \le \int_0^\delta t^{a-1} e^{-t}\, dt + \int_R^{+\infty} t^{b-1} e^{-t}\, dt.$$

Thus

$$\|F_{\delta,R} - F\|_{\{z \,|\, a \le \operatorname{Re} z \le b\}} \le \int_0^\delta t^{a-1} e^{-t}\, dt + \int_R^{+\infty} t^{b-1} e^{-t}\, dt \to 0$$

when $\delta \to 0+$ and $R \to +\infty$.

The uniform convergence of the holomorphic functions $F_{\delta,R}$ to F in every compact subset of $\{z \,|\, 0 < \operatorname{Re} z\}$ implies that F is also holomorphic in $\{z \,|\, 0 < \operatorname{Re} z\}$ and also that

$$F^{(n)}(z) = \int_0^{+\infty} t^{z-1} e^{-t} \ln^n t\, dt$$

for all z with $\operatorname{Re} z > 0$.

It easy to prove the functional equation

$$F(z+1) = zF(z)$$

i.e. the same functional equation that the gamma function satisfies. Indeed, by a simple integration by parts we get

$$\begin{aligned} F(z+1) &= \int_0^{+\infty} t^z e^{-t}\, dt = -\int_0^{+\infty} t^z (e^{-t})'\, dt = \int_0^{+\infty} (t^z)' e^{-t}\, dt = z \int_0^{+\infty} t^{z-1} e^{-t}\, dt \\ &= zF(z). \end{aligned}$$

Since $F(1) = \int_0^{+\infty} e^{-t}\, dt = 1$, induction shows that $F(n) = (n-1)!$ for every $n \in \mathbb{N}$.

With the help of the functional equation, F, which is holomorphic in $\{z \,|\, 0 < \operatorname{Re} z\}$, can be extended as a meromorphic function in all of $\mathbb{C}$, having simple poles only at the integers $0, -1, -2, -3, \ldots$. To do this we consider any $n \in \mathbb{N}$ and the function

$$F_{-n}(z) = \frac{F(z+n)}{z(z+1)\cdots(z+n-1)}$$

in the half-plane $\{z \,|\, -n < \operatorname{Re} z\}$. This function is holomorphic in $\{z \,|\, -n < \operatorname{Re} z\}$ except at the points $0, -1, \ldots, -(n-1)$. Since $F(n), F(n-1), \ldots, F(1) \ne 0$, the points $0, -1, \ldots, -(n-1)$ are simple poles of F_{-n}. If $0 < \operatorname{Re} z$, we get $F_{-n}(z) = F(z)$ because of the functional equation satisfied by F. Thus, F_{-n} is an extension of F in $\{z \,|\, -n < \operatorname{Re} z\}$ and we trivially see that F_{-n} satisfies the functional equation $F_{-n}(z+1) = zF_{-n}(z)$ for all $z \in \{z \,|\, -n < \operatorname{Re} z\}$.
Now, if we take any $m, n \in \mathbb{N}$ with $n < m$, then the functions F_{-n} and F_{-m} are the same in the intersection of their domains of definition $\{z \,|\, -n < \operatorname{Re} z\}$ and $\{z \,|\, -m < \operatorname{Re} z\}$. This intersection is the smallest of the two half-planes, i.e. $\{z \,|\, -n < \operatorname{Re} z\}$. To see that F_{-n} and F_{-m} are the same in $\{z \,|\, -n < \operatorname{Re} z\}$, we either use their defining relations $F_{-n}(z) = \frac{F(z+n)}{z(z+1)\cdots(z+n-1)}$

and $F_{-m}(z) = \frac{F(z+m)}{z(z+1)\cdots(z+m-1)}$ together with the functional equation $F(z+1) = zF(z)$, or we think that both functions coincide with F in $\{z \,|\, 0 < \operatorname{Re} z\}$ and then use the principle of identity. Finally, since all the half-planes $\{z \,|\, -n < \operatorname{Re} z\}$, $n \in \mathbb{N}$, cover $\mathbb{C}$, we conclude that all the functions F_{-n}, $n \in \mathbb{N}$, determine a single function $F_{-\infty}$ defined in $\mathbb{C} \setminus \{0, -1, -2, \ldots\}$. Indeed, for any $z \neq 0, -1, -2, \ldots$ we take any $n \in \mathbb{N}$ so that $-n < \operatorname{Re} z$ and we set

$$F_{-\infty}(z) = F_{-n}(z).$$

The value $F_{-n}(z)$ does not depend upon the specific $n \in \mathbb{N}$ satisfying $-n < \operatorname{Re} z$. In fact, if $n, m \in \mathbb{N}$ are such that $-n < \operatorname{Re} z$ and $-m < \operatorname{Re} z$, then we have shown that $F_{-n}(z) = F_{-m}(z)$. Thus $F_{-\infty}$ is well defined in $\mathbb{C} \setminus \{0, -1, -2, \ldots\}$.
Of course we have that $F_{-\infty} = F$ in $\{z \,|\, 0 < \operatorname{Re} z\}$ and $F_{-\infty} = F_{-n}$ in $\{z \,|\, -n < \operatorname{Re} z\}$ for every $n \in \mathbb{N}$. Therefore, $F_{-\infty}$ is holomorphic in $\mathbb{C}$ except at the points $0, -1, -2, \ldots$ which are simple poles of it. Since $F_{-\infty}$ extends F, we shall denote it by the same symbol F. We easily see that it satisfies the functional equation $F(z+1) = zF(z)$ in its domain of definition.

Now we shall prove some results for the restriction of F in $(0, +\infty)$ which are analogous to results for the restriction of the gamma function Γ in $(0, +\infty)$.

Proposition 8.10. *(i)* $F(x) > 0$ *for every* $x \in (0, +\infty)$.
(ii) $\lim_{x \to 0+} F(x) = +\infty$ *and* $\lim_{x \to +\infty} F(x) = +\infty$.
(iii) $F(x+1) = xF(x)$ *for every* $x \in (0, +\infty)$.
(iv) $F(n) = (n-1)!$ *for every* $n \in \mathbb{N}$.
(v) $\ln F$ *is convex in* $(0, +\infty)$.
(vi) $F(x)F''(x) \geq (F'(x))^2$ *for every* $x \in (0, +\infty)$.

Proof. (i) Trivial.
(ii) Let $x > 0$. Then

$$F(x) \geq \int_0^1 t^{x-1} e^{-t}\, dt \geq e^{-1} \int_0^1 t^{x-1}\, dt = \tfrac{1}{ex} \to +\infty$$

when $x \to 0+$.
Now let $x \geq 1$. Then

$$F(x) \geq \int_2^{+\infty} t^{x-1} e^{-t}\, dt \geq 2^{x-1} \int_2^{+\infty} e^{-t}\, dt = 2^{x-1} e^{-2} \to +\infty$$

when $x \to +\infty$.
(iii)-(iv) They have been proved.
(v)-(vi) For every $x \in (0, +\infty)$ and every $a \in \mathbb{R}$ we get

$$a^2 F''(x) + 2aF'(x) + F(x) = \int_0^{+\infty} t^{x-1} (a \ln t + 1)^2 e^{-t}\, dt \geq 0.$$

This implies that $F''(x)F(x) \geq (F'(x))^2$ and so $(\ln F)''(x) \geq 0$ for every $x > 0$.
The inequality $F''(x)F(x) \geq (F'(x))^2$ can also be proved by the Schwarz inequality for integrals:

$$\begin{aligned}(F'(x))^2 &= \big(\int_0^{+\infty} t^{x-1} e^{-t} \ln t\, dt\big)^2 = \big(\int_0^{+\infty} t^{\frac{x-1}{2}} e^{-\frac{t}{2}} t^{\frac{x-1}{2}} e^{-\frac{t}{2}} \ln t\, dt\big)^2 \\ &\leq \int_0^{+\infty} t^{x-1} e^{-t}\, dt \int_0^{+\infty} t^{x-1} e^{-t} \ln^2 t\, dt = F(x)F''(x).\end{aligned}$$

$\square$

Our next task is to prove that the functions Γ and F are the same function. Below we give three proofs. The first is the simplest.

Proposition 8.11.

$$\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t}\, dt$$

for every $z \in \mathbb{C}$, $z \neq 0, -1, -2, -3, \ldots$.

First proof. Due the principle of identity, it is enough to prove $\Gamma(x) = F(x)$ for $x \in (0, +\infty)$. The functional equation for F implies that

$$\ln F(x+n+1) = \ln F(x) + \ln x + \ln(x+1) + \cdots + \ln(x+n)$$

for every $x > 0$ and $n \in \mathbb{N}$. Differentiating this twice, we get

$$(\ln F)''(x+n+1) = (\ln F)''(x) - \tfrac{1}{x^2} - \tfrac{1}{(x+1)^2} - \cdots - \tfrac{1}{(x+n)^2}.$$

Since $(\ln F)''(x+n+1) \geq 0$, we have that

$$(\ln F)''(x) \geq \tfrac{1}{x^2} + \tfrac{1}{(x+1)^2} + \cdots + \tfrac{1}{(x+n)^2}$$

for every $x > 0$ and $n \in \mathbb{N}$. Letting $n \to +\infty$ we get

$$(\ln F)''(x) \geq (\ln \Gamma)''(x)$$

for every $x > 0$.
Therefore, the function $h = \ln F - \ln \Gamma$ is convex in $(0, +\infty)$. But we also have that $F(n) = \Gamma(n) = (n-1)!$, and hence $h(n) = 0$, for every $n \in \mathbb{N}$. Now it is trivial for the reader to prove that $h(x) = 0$ for every $x \geq 1$. So $F(x) = \Gamma(x)$ for $x \geq 1$ and then the functional equation implies the same for $0 < x < 1$.
Second proof. As in the first proof, we shall prove that $\Gamma(x) = F(x)$ for $x \in (0, +\infty)$.
Succesive integrations by parts give the formula

$$\int_0^n t^{x-1}\big(1 - \tfrac{t}{n}\big)^n \, dt = \tfrac{n^x n!}{x(x+1)\cdots(x+n)}$$

and so it is enough to prove that

$$\lim_{n \to +\infty} \int_0^n t^{x-1}\big(1 - \tfrac{t}{n}\big)^n \, dt = \int_0^{+\infty} t^{x-1} e^{-t} \, dt.$$

We have

$$t^{x-1}\big(1 - \tfrac{t}{n}\big)^n \leq t^{x-1} e^{-\frac{t}{n} n} = t^{x-1} e^{-t} \tag{8.16}$$

for every $t > 0$ and $n \in \mathbb{N}$.
Now take $\epsilon > 0$. Then there is $a > 0$ so that

$$\int_a^{+\infty} t^{x-1} e^{-t} \, dt < \tfrac{\epsilon}{4}. \tag{8.17}$$

Now $t^{x-1}\big(1 - \frac{t}{n}\big)^n \to t^{x-1} e^{-t}$ uniformly in $[0, a]$. Indeed, for $n \geq 2a$ and $t \in [0, a]$ we have that $\frac{t}{n} \leq \frac{1}{2}$ and so

$$\begin{aligned}\big|t^{x-1}\big(1 - \tfrac{t}{n}\big)^n - t^{x-1} e^{-t}\big| &= t^{x-1}\big|e^{n \ln(1-(t/n))} - e^{-t}\big| = t^{x-1} e^{-t}\big(1 - e^{t + n \ln(1-(t/n))}\big) \\ &\leq t^{x-1} e^{-t}\big(1 - e^{-t^2/n}\big) \leq t^{x-1} e^{-t} \tfrac{t^2}{n} \leq \tfrac{a^{x+1}}{n}.\end{aligned}$$

Therefore $\int_0^a t^{x-1}(1 - \frac{t}{n})^n \, dt \to \int_0^a t^{x-1} e^{-t} \, dt$. Thus there is n_0 so that

$$\big| \int_0^a t^{x-1}\big(1 - \tfrac{t}{n}\big)^n \, dt - \int_0^a t^{x-1} e^{-t} \, dt\big| < \tfrac{\epsilon}{2}$$

for every $n \geq n_0$. This last relation together with (8.16) and (8.17) imply

$$\begin{aligned}&\big| \int_0^{+\infty} t^{x-1}\big(1 - \tfrac{t}{n}\big)^n \, dt - \int_0^{+\infty} t^{x-1} e^{-t} \, dt\big| \\ &\quad \leq \big| \int_0^a t^{x-1}\big(1 - \tfrac{t}{n}\big)^n \, dt - \int_0^a t^{x-1} e^{-t} \, dt\big| + \int_0^a t^{x-1}\big(1 - \tfrac{t}{n}\big)^n \, dt + \int_0^a t^{x-1} e^{-t} \, dt \\ &\quad \leq \big| \int_0^a t^{x-1}\big(1 - \tfrac{t}{n}\big)^n \, dt - \int_0^a t^{x-1} e^{-t} \, dt\big| + \int_0^a t^{x-1} e^{-t} \, dt + \int_0^a t^{x-1} e^{-t} \, dt \\ &\quad < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{4} + \tfrac{\epsilon}{4} = \epsilon\end{aligned}$$

for every $n \geq n_0$.
Third proof. We shall prove a rather crude estimate of $\Gamma(z)$ when $z = x + iy$, $0 \leq x \leq 1$. At first we use (8.15) to get

$$\sum_{k=m+1}^{n} \frac{1}{k} \leq \sum_{k=m+1}^{n} \ln \frac{k}{k-1} = \ln \frac{n}{m}.$$

Also

$$\sum_{k=m+1}^{+\infty} \frac{1}{k^2} \leq \sum_{k=m+1}^{+\infty} \int_{k-1}^{k} \frac{1}{t^2}\, dt = \int_{m}^{+\infty} \frac{1}{t^2}\, dt = \frac{1}{m}.$$

Now we consider $0 \leq x \leq 1$ and $|y| \geq 1$, and then we take $m = [2|y|] \in \mathbb{N}$. Then we have

$$\begin{aligned}\sum_{k=m+1}^{n} \Big(\tfrac{(x+k)^2+y^2}{k^2} - 1\Big) &= \sum_{k=m+1}^{n} \tfrac{2x}{k} + \sum_{k=m+1}^{n} \tfrac{x^2+y^2}{k^2} \leq 2x \ln \tfrac{n}{m} + \tfrac{x^2+y^2}{m} \\ &\leq 2x \ln \tfrac{n}{m} + 1 + \tfrac{|y|}{2}.\end{aligned}$$

This implies

$$\tfrac{((x+m+1)^2+y^2)\cdots((x+n)^2+y^2)}{(m+1)^2\cdots n^2} \leq e^{\sum_{k=m+1}^{n}\left(\frac{(x+k)^2+y^2}{k^2}-1\right)} \leq \big(\tfrac{n}{m}\big)^{2x} e^{1+\frac{|y|}{2}} \leq n^{2x} e^{1+\frac{|y|}{2}}.$$

On the other hand,

$$\begin{aligned}\tfrac{(x^2+y^2)((x+1)^2+y^2)\cdots((x+m)^2+y^2)}{1^2\cdots m^2} &\leq (1+|y|)^2 \Big(\tfrac{(1+m+|y|)^m}{m!}\Big)^2 \leq (1+|y|)^2 e^{2(1+m+|y|)} \\ &\leq (1+|y|)^2 e^{4+6|y|}.\end{aligned}$$

Multiplying the last two inequalities we get

$$\Big|\tfrac{z(z+1)\cdots(z+n)}{n^z n!}\Big| \leq (1+|y|) e^{\frac{5}{2}+\frac{13}{4}|y|}$$

and letting $n \to +\infty$ we conclude that

$$\tfrac{1}{|\Gamma(z)|} \leq (1+|y|) e^{\frac{5}{2}+\frac{13}{4}|y|}.$$

We observe that, if $0 \leq x \leq 1$ and $|y| \geq 1$, then

$$|F(z)| = \tfrac{|F(z+1)|}{|z|} \leq \int_0^{+\infty} t^x e^{-t}\, dt \leq \int_0^1 t^x e^{-t}\, dt + \int_1^{+\infty} t^x e^{-t}\, dt \leq \int_0^1 dt + \int_1^{+\infty} t e^{-t}\, dt \leq 2.$$

Now we consider the function

$$h(z) = \tfrac{F(z)}{\Gamma(z)}.$$

Since the poles of F and Γ are simple and coincide, h is holomorphic in $\mathbb{C}$. Also, since both F and Γ satisfy the same functional equation, we get that $h(z+1) = h(z)$ for every z, i.e. that h is 1-periodic. Now the estimates we have got for $|F(z)|$ and $\frac{1}{|\Gamma(z)|}$ imply that

$$|h(z)| \leq c(1+|y|) e^{\frac{13}{4}|y|}$$

for every $z = x + iy$ with $0 \leq x \leq 1$, where c is a constant.
Since h is 1-periodic, we may define the function

$$g(w) = h(z), \qquad \text{where } w = e^{2\pi i z}.$$

Then g is holomorphic in $\mathbb{C} \setminus \{0\}$. From the last estimate of $|h(z)|$ we easily get that

$$|g(w)| \leq c\big(1 + \tfrac{1}{2\pi} \ln \tfrac{1}{|w|}\big)|w|^{-a}, \quad \text{if } |w| \leq 1,$$

$$|g(w)| \leq c\big(1 + \tfrac{1}{2\pi} \ln |w|\big)|w|^{a}, \quad \text{if } |w| \geq 1,$$

where $a = \frac{13}{8\pi}$ and hence $0 < a < 1$. From the first relation we get $\lim_{w \to 0} w g(w) = 0$ and, by the criterion of Riemann, g can be considered holomorphic at 0. Moreover, if we consider the function $p(z) = g(\frac{1}{z})$, then from the second relation we have that $|p(z)| \leq c\big(1 + \frac{1}{2\pi} \ln \frac{1}{|z|}\big)|z|^{-a}$ for $|z| \leq 1$. The same argument as before shows that p can also be considered holomorphic at 0. In particular, p is bounded near 0 and so g is bounded near ∞. By the theorem of Liouville, we conclude that g is constant and of course this implies that $h = \frac{F}{\Gamma}$ is constant. Since $h(1) = 1$ we finally get that $F = \Gamma$. $\square$

Lemma 8.5. *We consider the* 1*-periodic functions in* $\mathbb{R}$*:*

$$F(t) = t - [t] - \tfrac{1}{2}, \quad G(t) = \tfrac{1}{2}(t - [t])(t - [t] - 1).$$

Then for every twice continuously differentiable function f *in* $[0, n]$ *we have*

$$\textstyle\sum_{k=0}^{n} f(k) - \int_0^n f(t)\,dt = \frac{f(0)+f(n)}{2} + \int_0^n f'(t)F(t)\,dt = \frac{f(0)+f(n)}{2} - \int_0^n f''(t)G(t)\,dt.$$

Proof. In every interval $[k-1, k]$, $k \in \mathbb{Z}$, we have that $F'(t) = 1$ and $G'(t) = F(t)$. Integration by parts implies

$$\tfrac{f(k-1)+f(k)}{2} - \textstyle\int_{k-1}^{k} f(t)\,dt = \int_{k-1}^{k} f'(t)F(t)\,dt = \int_{k-1}^{k} f''(t)G(t)\,dt$$

from which we get

$$f(k) - \textstyle\int_{k-1}^{k} f(t)\,dt = -\frac{f(k-1)-f(k)}{2} + \int_{k-1}^{k} f'(t)F(t)\,dt = -\frac{f(k-1)-f(k)}{2} - \int_{k-1}^{k} f''(t)G(t)\,dt.$$

Now we add these relations for $k = 1, \ldots, n$, and then we add $f(0)$ to both sides of the resulting equality. $\square$

Now we shall prove a famous asymptotic formula.
When we write $f(z) \sim g(z)$ as $z \to \infty$ we mean that $\lim_{z\to\infty} \frac{f(z)}{g(z)} = 1$.

Theorem 8.3. *Let* $0 < \delta < \pi$ *and* $G_\delta = \{z \,|\, z \neq 0, -\pi + \delta < \operatorname{Arg} z < \pi - \delta\}$*. Then*

$$\Gamma(z) \sim \sqrt{2\pi} e^{-z} z^{z-\frac{1}{2}} \quad \textit{as } z \to \infty \textit{ in } G_\delta.$$

This is **Stirling's asymptotic formula***.*
When we write $z^{z-\frac{1}{2}}$ *we mean* $z^{z-\frac{1}{2}} = e^{(z-\frac{1}{2})\operatorname{Log} z}$*.*

Proof. We apply lemma 8.5 to the function $f(t) = \operatorname{Log}(z+t)$ and we get

$$\textstyle\sum_{k=0}^{n} \operatorname{Log}(z+k) - \int_0^n \operatorname{Log}(z+t)\,dt = \frac{\operatorname{Log} z + \operatorname{Log}(z+n)}{2} + \int_0^n \frac{G(t)}{(z+t)^2}\,dt.$$

Now, $\operatorname{Log}(z+t)$ is the derivative of $(z+t)\operatorname{Log}(z+t) - t$ and so we have

$$\textstyle\sum_{k=0}^{n} \operatorname{Log}(z+k) - (z+n)\operatorname{Log}(z+n) + n + z\operatorname{Log} z = \frac{\operatorname{Log} z + \operatorname{Log}(z+n)}{2} + \int_0^n \frac{G(t)}{(z+t)^2}\,dt.$$

Thus

$$\textstyle\sum_{k=0}^{n} \operatorname{Log}(z+k) = \left(z + n + \frac{1}{2}\right)\operatorname{Log}(z+n) - \left(z - \frac{1}{2}\right)\operatorname{Log} z - n + \int_0^n \frac{G(t)}{(z+t)^2}\,dt.$$

We write the same formula for $z = 1$:

$$\textstyle\sum_{k=0}^{n} \operatorname{Log}(1+k) = \left(n + \frac{3}{2}\right)\operatorname{Log}(1+n) - n + \int_0^n \frac{G(t)}{(1+t)^2}\,dt.$$

Taking exponentials of both equalities we find

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

$$(n+1)! = (n+1)^{n+\frac{3}{2}} e^{-n} e^{I_n(1)},$$

where $I_n(z) = \int_0^n \frac{G(t)}{(z+t)^2}\,dt$.
We divide the last two equalities and we easily find

$$\tfrac{n^z n!}{z(z+1)\cdots(z+n)} = \left(\tfrac{n}{n+z}\right)^z \left(\tfrac{n+1}{n+z}\right)^{n+\frac{1}{2}} z^{z-\frac{1}{2}} e^{I_n(1)-I_n(z)}.$$

The left side of this equality has limit $\Gamma(z)$ when $n \to +\infty$. Moreover, $(\frac{n}{n+z})^z \to 1$ and $(\frac{n+1}{n+z})^{n+\frac{1}{2}} \to e^{1-z}$ when $n \to +\infty$. Finally, we observe that $-\frac{1}{8} \leq G(t) \leq 0$ for every t. This implies that the integral $I(z) = \int_0^{+\infty} \frac{G(t)}{(z+t)^2}\,dt$ converges absolutely:

$$|I(z)| \leq \int_0^{+\infty} \left|\frac{G(t)}{(z+t)^2}\right| dt \leq \frac{1}{8}\int_0^{+\infty} \frac{1}{|z+t|^2}\,dt < +\infty$$

for every z in $\mathbb{C} \setminus (-\infty, 0]$. We conclude that

$$\Gamma(z) = e^{1-z} z^{z-\frac{1}{2}} e^{I(1)-I(z)} = c e^{-z} z^{z-\frac{1}{2}} e^{-I(z)}$$

for every z in $\mathbb{C} \setminus (-\infty, 0]$, where $c = e^{1+I(1)}$ is a constant.
Now we shall prove that $I(z) \to 0$ when $z \to \infty$ in G_δ. We write $z = |z|e^{i\theta}$ with $\theta = \operatorname{Arg} z$ and we assume that $z \in G_\delta$, i.e. $-\pi + \delta < \theta < \pi - \delta$. Then

$$\begin{aligned}
|I(z)| &\leq \tfrac{1}{8}\int_0^{+\infty} \tfrac{1}{|z+t|^2}\,dt = \tfrac{1}{8|z|}\int_0^{+\infty} \tfrac{1}{|e^{i\theta}+u|^2}\,du \\
&= \tfrac{1}{8|z|}\int_0^{+\infty} \tfrac{1}{u^2+2u\cos\theta+1}\,du \leq \tfrac{1}{8|z|}\int_0^{+\infty} \tfrac{1}{u^2-2u\cos\delta+1}\,du \\
&= \tfrac{1}{8|z|}\int_0^{+\infty} \tfrac{1}{(u-\cos\delta)^2+\sin^2\delta}\,du = \tfrac{1}{8|z|}\int_{-\cos\delta}^{+\infty} \tfrac{1}{u^2+\sin^2\delta}\,du \\
&\leq \tfrac{1}{8|z|}\int_{-\infty}^{+\infty} \tfrac{1}{u^2+\sin^2\delta}\,du = \tfrac{1}{8|z|\sin\delta}\int_{-\infty}^{+\infty} \tfrac{1}{u^2+1}\,du \\
&= \tfrac{\pi}{8|z|\sin\delta}.
\end{aligned}$$

So we get that

$$\sup_{z\in G_\delta, |z|\geq r} |I(z)| \leq \tfrac{\pi}{8r\sin\delta}$$

and we conclude that $I(z) \to 0$ *uniformly* as $z \to \infty$ in G_δ.
We have got that $\Gamma(z) \sim c e^{-z} z^{z-\frac{1}{2}}$ as $z \to \infty$ in G_δ and the only thing that remains to be proved is that $c = \sqrt{2\pi}$.
We shall use the duplication formula

$$\Gamma(z)\Gamma(z + \tfrac{1}{2}) = \sqrt{\pi}\, 2^{1-2z}\Gamma(2z)$$

with $z = x \to +\infty$. We then have that

$$c e^{-x} x^{x-\frac{1}{2}} c e^{-x-\frac{1}{2}} \left(x + \tfrac{1}{2}\right)^x \sim \sqrt{\pi}\, 2^{1-2x} c e^{-2x} (2x)^{2x-\frac{1}{2}}$$

and this easily implies that $c = \sqrt{2\pi}$. □

Since $n! = \Gamma(n + 1)$, Stirling's asymptotic formula implies that

$$n! \sim \sqrt{2\pi}\, e^{-n} n^{n+\frac{1}{2}}$$

when $n \to +\infty$.

8.5 Riemann's zeta function.

Definition. *We define the function*

$$\zeta(z) = \sum_{n=1}^{+\infty} \tfrac{1}{n^z} = 1 + \tfrac{1}{2^z} + \tfrac{1}{3^z} + \tfrac{1}{4^z} + \cdots .$$

When we write n^z we mean $n^z = e^{z \ln n}$.
The function ζ is called **Riemann's zeta function**.

If $\operatorname{Re} z > 1$, then the series defining the zeta function converges absolutely. Indeed, if $x = \operatorname{Re} z > 1$, then

$$\sum_{n=1}^{+\infty} \left|\frac{1}{n^z}\right| = \sum_{n=1}^{+\infty} \frac{1}{n^x} < +\infty.$$

Now we can easily see that the series converges uniformly in every half-plane of the form $\{z \,|\, a \leq \operatorname{Re} z\}$ with $a > 1$ and hence in every compact subset of the half-plane $\{z \,|\, 1 < \operatorname{Re} z\}$. Indeed, this is a corollary of the test of Weierstrass: if $\operatorname{Re} z \geq a > 1$, then $\left|\frac{1}{n^z}\right| \leq \frac{1}{n^a}$, and $\sum_{n=1}^{+\infty} \frac{1}{n^a} < +\infty$. Since every $\frac{1}{n^z}$ is holomorphic, we conclude that ζ is holomorphic in the half-plane $\{z \,|\, 1 < \operatorname{Re} z\}$.

The following result connects the zeta function to number theory.

Let $p_1 < p_2 < \ldots < p_n < \ldots$ be the increasing sequence of the *prime numbers*. We consider the infinite product

$$\prod_{n=1}^{+\infty} \left(1 - \frac{1}{p_n^z}\right).$$

Since the sequence of the prime numbers is only part of the sequence of all natural numbers, the previous arguments show that the series $\sum_{n=1}^{+\infty} \left|\frac{1}{p_n^z}\right|$ converges uniformly in every compact subset of $\{z \,|\, 1 < \operatorname{Re} z\}$. Now theorem 8.1 implies that the infinite product

$$\prod_{n=1}^{+\infty} \left(1 - \frac{1}{p_n^z}\right)$$

defines a function holomorphic in $\{z \,|\, 1 < \operatorname{Re} z\}$.

Proposition 8.12. *For every z with* $\operatorname{Re} z > 1$ *we have*

$$\zeta(z) = \prod_{n=1}^{+\infty} \left(1 - \frac{1}{p_n^z}\right)^{-1}.$$

Proof. Let $\operatorname{Re} z > 1$. We observe that

$$\zeta(z)\left(1 - \frac{1}{2^z}\right) = \sum_{n=1}^{+\infty} \frac{1}{n^z} - \sum_{n=1}^{+\infty} \frac{1}{(2n)^z} = \sum_{n\in\mathbb{N}, 2\not| n} \frac{1}{n^z}.$$

Next

$$\zeta(z)\left(1 - \frac{1}{2^z}\right)\left(1 - \frac{1}{3^z}\right) = \sum_{n\in\mathbb{N}, 2\not| n} \frac{1}{n^z} - \sum_{n\in\mathbb{N}, 2\not| n} \frac{1}{(3n)^z} = \sum_{n\in\mathbb{N}, 2\not| n, 3\not| n} \frac{1}{n^z}.$$

So we see that

$$\zeta(z)\left(1 - \frac{1}{p_1^z}\right) \cdots \left(1 - \frac{1}{p_N^z}\right) = \sum_{n\in\mathbb{N}, p_1\not| n, \ldots, p_N\not| n} \frac{1}{n^z}$$

for every N.

Now, if $1 < n < p_{N+1}$, then n is divisible by at least one of $p_1, \ldots, p_N$ and so the last series does not include the term $\frac{1}{n^z}$. Therefore the first term of this series is 1 and the next term is $\frac{1}{p_{N+1}^z}$. This implies

$$\left|\zeta(z)\left(1 - \frac{1}{p_1^z}\right) \cdots \left(1 - \frac{1}{p_N^z}\right) - 1\right| \leq \sum_{n=p_{N+1}}^{+\infty} \frac{1}{n^x}$$

for $x = \operatorname{Re} z > 1$. Since $p_{N+1} \to +\infty$, we get $\sum_{n=p_{N+1}}^{+\infty} \frac{1}{n^x} \to 0$ when $N \to +\infty$. This finishes the proof. $\square$

A corollary of proposition 8.12 is that ζ *has no root in the half-plane* $\{z \,|\, 1 < \operatorname{Re} z\}$.

The next result relates the gamma function and the zeta function.

We consider the generalized integral

$$\int_0^{+\infty} \frac{t^{z-1}}{e^t - 1}\, dt$$

for $\operatorname{Re} z > 1$. The integral converges absolutely: if $x = \operatorname{Re} z > 1$ we have

$$\int_0^{+\infty} \left|\frac{t^{z-1}}{e^t - 1}\right| dt = \int_0^1 \frac{t^{x-1}}{e^t - 1}\, dt + \int_1^{+\infty} \frac{t^{x-1}}{e^t - 1}\, dt = \int_0^1 t^{x-2}\, dt + n! \int_1^{+\infty} t^{x-n-1}\, dt < +\infty,$$

where we use any $n \in \mathbb{N}$, $n > x$, and hence $e^t - 1 \geq \frac{t^n}{n!}$.

Proposition 8.13. *For every z with* $\operatorname{Re} z > 1$ *we have*

$$\zeta(z)\Gamma(z) = \int_0^{+\infty} \tfrac{t^{z-1}}{e^t-1}\, dt.$$

Proof. Let $\operatorname{Re} z > 1$. For every $n \in \mathbb{N}$ we have

$$\Gamma(z) = \int_0^{+\infty} e^{-s} s^{z-1}\, ds = n^z \int_0^{+\infty} e^{-nt} t^{z-1}\, dt.$$

Therefore

$$\zeta(z)\Gamma(z) = \sum_{n=1}^{+\infty} \int_0^{+\infty} e^{-nt} t^{z-1}\, dt = \int_0^{+\infty} \big(\sum_{n=1}^{+\infty} e^{-nt}\big) t^{z-1}\, dt = \int_0^{+\infty} \tfrac{t^{z-1}}{e^t-1}\, dt.$$

To justify the interchange of the sum and the integral we take any $N \in \mathbb{N}$ and we set

$$A_N = \sum_{n=N+1}^{+\infty} \int_0^{+\infty} e^{-nt} t^{z-1}\, dt, \quad B_N = \int_0^{+\infty} \big(\sum_{n=N+1}^{+\infty} e^{-nt}\big) t^{z-1}\, dt.$$

Then we have

$$|A_N| = \big|\sum_{n=N+1}^{+\infty} \tfrac{\Gamma(z)}{n^z}\big| \le |\Gamma(z)| \sum_{n=N+1}^{+\infty} \tfrac{1}{n^x} \to 0$$

when $N \to +\infty$. Also

$$\begin{aligned}
|B_N| &= \big|\int_0^{+\infty} \tfrac{e^{-Nt} t^{z-1}}{e^t-1}\, dt\big| \le \int_0^{+\infty} \tfrac{e^{-Nt} t^{x-1}}{e^t-1}\, dt \\
&= \int_0^{1/\sqrt{N}} \tfrac{e^{-Nt} t^{x-1}}{e^t-1}\, dt + \int_{1/\sqrt{N}}^{+\infty} \tfrac{e^{-Nt} t^{x-1}}{e^t-1}\, dt \\
&\le \int_0^{1/\sqrt{N}} \tfrac{t^{x-1}}{e^t-1}\, dt + e^{-\sqrt{N}} \int_{1/\sqrt{N}}^{+\infty} \tfrac{t^{x-1}}{e^t-1}\, dt \\
&\le \int_0^{1/\sqrt{N}} \tfrac{t^{x-1}}{e^t-1}\, dt + e^{-\sqrt{N}} \int_0^{+\infty} \tfrac{t^{x-1}}{e^t-1}\, dt \to 0
\end{aligned}$$

when $N \to +\infty$. Combining these two limits with the interchange of the sum of the first N terms and the integral, we get

$$\sum_{n=1}^{+\infty} \int_0^{+\infty} e^{-nt} t^{z-1}\, dt - \int_0^{+\infty} \big(\sum_{n=1}^{+\infty} e^{-nt}\big) t^{z-1}\, dt = A_N - B_N \to 0$$

when $N \to +\infty$. Since the left side does not depend upon N, it is equal to 0. $\square$

We consider the principal branch Log of the logarithm defined in the region $\mathbb{C} \setminus (-\infty, 0]$ by

$$\operatorname{Log} \zeta = \ln r + i\theta \quad \text{for } \zeta = re^{i\theta} \text{ with } -\pi < \theta < \pi.$$

We also define

$$\operatorname{Log}_+ \zeta = \ln(-\zeta) + i\pi, \quad \operatorname{Log}_- \zeta = \ln(-\zeta) - i\pi \quad \text{for } \zeta \in (-\infty, 0)$$

It is clear that

$$\lim_{\theta \to \pi-} \operatorname{Log}(re^{i\theta}) = \operatorname{Log}_+(-r), \quad \lim_{\theta \to -\pi+} \operatorname{Log}(re^{i\theta}) = \operatorname{Log}_-(-r) \tag{8.18}$$

uniformly in $(0, +\infty)$.

Now we consider the parametric equation $\gamma_-(t) = -t$, $t \in (-\infty, 0]$, and the parametric equation $\gamma_+(t) = t$, $t \in [0, +\infty)$. Then we take the curvilinear integral

$$\int_{\gamma_-} \tfrac{e^{(z-1)\operatorname{Log}_-(-\zeta)}}{e^\zeta - 1}\, d\zeta + \int_{\gamma_+} \tfrac{e^{(z-1)\operatorname{Log}_+(-\zeta)}}{e^\zeta - 1}\, d\zeta$$

where $\operatorname{Re} z > 1$.

The curve γ_- describes the positive x-axis $[0, +\infty)$ in the direction from $+\infty$ to 0, and then the curve γ_+ describes the same positive x-axis $[0, +\infty)$ in the direction from 0 back to $+\infty$. Now we calculate

$$\begin{aligned}
\int_{\gamma_-} & \frac{e^{(z-1)\,\mathrm{Log}_-(-\zeta)}}{e^\zeta - 1}\, d\zeta + \int_{\gamma_+} \frac{e^{(z-1)\,\mathrm{Log}_+(-\zeta)}}{e^\zeta - 1}\, d\zeta \\
&= -\int_{-\infty}^{0} \frac{(-t)^{z-1} e^{-i\pi(z-1)}}{e^{-t}-1}\, dt + \int_0^{+\infty} \frac{t^{z-1} e^{i\pi(z-1)}}{e^t - 1}\, dt \\
&= -\int_0^{+\infty} \frac{t^{z-1} e^{-i\pi(z-1)}}{e^t - 1}\, dt + \int_0^{+\infty} \frac{t^{z-1} e^{i\pi(z-1)}}{e^t - 1}\, dt \\
&= \left(e^{i\pi(z-1)} - e^{-i\pi(z-1)}\right) \int_0^{+\infty} \frac{t^{z-1}}{e^t - 1}\, dt \\
&= -2i \sin(\pi z) \int_0^{+\infty} \frac{t^{z-1}}{e^t - 1}\, dt \\
&= -2i \sin(\pi z) \zeta(z) \Gamma(z)
\end{aligned} \tag{8.19}$$

for $\mathrm{Re}\, z > 1$.

We now take any r with $0 < r < 2\pi$ and we consider the parametric equation $\gamma_-^r(t) = -t$, $t \in (-\infty, -r]$, the parametric equation $\gamma_+^r(t) = t$, $t \in [r, +\infty)$, and the parametric equation $\sigma^r(\theta) = re^{i\theta}$, $0 \le \theta \le 2\pi$. The curve γ_-^r describes the half-line $[r, +\infty)$ in the direction from $+\infty$ to r, then the curve σ^r describes the circle $C_0(r)$ in the positive direction from r back to r, and then the curve γ_+^r describes the half-line $[r, +\infty)$ in the direction from r back to $+\infty$.

And now we consider the curvilinear integral

$$\begin{aligned}
I_r(z) &= \int_{\gamma_-^r} \frac{e^{(z-1)\,\mathrm{Log}_-(-\zeta)}}{e^\zeta - 1}\, d\zeta + \oint_{C_0(r)} \frac{e^{(z-1)\,\mathrm{Log}(-\zeta)}}{e^\zeta - 1}\, d\zeta + \int_{\gamma_+^r} \frac{e^{(z-1)\,\mathrm{Log}_+(-\zeta)}}{e^\zeta - 1}\, d\zeta \\
&= \oint_{C_0(r)} \frac{e^{(z-1)\,\mathrm{Log}(-\zeta)}}{e^\zeta - 1}\, d\zeta - 2i \sin(\pi z) \int_r^{+\infty} \frac{t^{z-1}}{e^t - 1}\, dt.
\end{aligned}$$

The second equality above is implied by the same calculation as the one in (8.19).

We observe that $I_r(z)$ is defined for every z. Now the restriction $\mathrm{Re}\, z > 1$ is not needed because the curve defining $I_r(z)$ does not approach the point $\zeta = 0$ which is a root of $e^\zeta - 1$.

Proposition 8.14. *Let $0 < r < 2\pi$. Then $I_r(z)$ is holomorphic as a function of z in $\mathbb{C}$. Moreover*

$$\zeta(z) = -\frac{1}{2\pi i}\, \Gamma(1-z) I_r(z) \tag{8.20}$$

for every z with $\mathrm{Re}\, z > 1$.

Proof. We leave it as an exercise to the reader to prove that, when r is fixed in $(0, 2\pi)$, the function $I_r(z)$ is holomorphic as a function of z in $\mathbb{C}$.
Now we shall first prove that $I_r(z)$ is constant as a function of r in $(0, 2\pi)$. If we take r_1, r_2 with $0 < r_1 < r_2 < 2\pi$ then it is clear that

$$\begin{aligned}
I_{r_2}(z) - I_{r_1}(z) &= \oint_{C_0(r_2)} \frac{e^{(z-1)\,\mathrm{Log}(-\zeta)}}{e^\zeta - 1}\, d\zeta - \int_{\gamma_-^{r_1,r_2}} \frac{e^{(z-1)\,\mathrm{Log}_-(-\zeta)}}{e^\zeta - 1}\, d\zeta \\
&\quad - \oint_{C_0(r_1)} \frac{e^{(z-1)\,\mathrm{Log}(-\zeta)}}{e^\zeta - 1}\, d\zeta - \int_{\gamma_+^{r_1,r_2}} \frac{e^{(z-1)\,\mathrm{Log}_+(-\zeta)}}{e^\zeta - 1}\, d\zeta \\
&= \lim_{\delta \to 0+} \oint_{c_{r_1,r_2,\delta}} \frac{e^{(z-1)\,\mathrm{Log}(-\zeta)}}{e^\zeta - 1}\, d\zeta,
\end{aligned}$$

where $\gamma_-^{r_1,r_2}(t) = -t$, $t \in [-r_2, -r_1]$ and $\gamma_+^{r_1,r_2}(t) = t$, $t \in [r_1, r_2]$ and where $c_{r_1,r_2,\delta}$ is the closed curve which describes the arc of the circle $C(0, r_2)$ from $r_2 e^{i\delta}$ to $r_2 e^{i(2\pi-\delta)}$ in the positive direction, and then the linear segment from $r_2 e^{i(2\pi-\delta)}$ to $r_1 e^{i(2\pi-\delta)}$, and then the arc of the circle $C(0, r_1)$ from $r_1 e^{i(2\pi-\delta)}$ to $r_1 e^{i\delta}$ in the negative direction, and then the linear segment from $r_1 e^{i\delta}$ to $r_2 e^{i\delta}$. The above limit as $\delta \to 0+$ holds because of (8.18). Now we have

$$\oint_{c_{r_1,r_2,\delta}} \frac{e^{(z-1)\,\mathrm{Log}(-\zeta)}}{e^\zeta - 1}\, d\zeta = 0$$

because the closed curve $c_{r_1,r_2,\delta}$ is in $\mathbb{C} \setminus [0,+\infty)$ and includes no singularity of the function $\frac{e^{(z-1)\,\mathrm{Log}(-\zeta)}}{e^{\zeta}-1}$ which is meromorphic in $\mathbb{C} \setminus [0,+\infty)$. Thus $I_{r_2}(z) - I_{r_1}(z) = 0$ and so $I_r(z)$ is constant as a function of r in $(0, 2\pi)$.
Now it is easy to see that, if $\mathrm{Re}\, z > 1$, then

$$\lim_{r\to 0+} \oint_{C_0(r)} \frac{e^{(z-1)\,\mathrm{Log}(-\zeta)}}{e^{\zeta}-1}\, d\zeta = 0$$

Indeed, from the limit $\lim_{\zeta\to 0} \frac{e^{\zeta}-1}{\zeta} = 1$ we get that $\frac{|e^{\zeta}-1|}{|\zeta|} \geq \frac{1}{2}$ when $r = |\zeta|$ is close to 0. Now if $z = x + iy$ with $x > 1$, then

$$\Big| \oint_{C_0(r)} \frac{e^{(z-1)\,\mathrm{Log}(-\zeta)}}{e^{\zeta}-1}\, d\zeta \Big| \leq \int_0^{2\pi} \frac{2r^x e^{y(\pi-\theta)}}{r}\, d\theta \leq 4\pi e^{\pi|y|} r^{x-1} \to 0$$

when $r \to 0+$.
Therefore, if $\mathrm{Re}\, z > 1$, then

$$\lim_{r\to 0+} I_r(z) = \int_{\gamma_-} \frac{e^{(z-1)\,\mathrm{Log}_-(-\zeta)}}{e^{\zeta}-1}\, d\zeta + \int_{\gamma_+} \frac{e^{(z-1)\,\mathrm{Log}_+(-\zeta)}}{e^{\zeta}-1}\, d\zeta.$$

We finally use (8.19), that $I_r(z)$ is constant as a function of r in $(0, 2\pi)$ and also that $\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$, and we finish the proof. $\square$

From the second formula for $I_r(z)$ we see immediately that

$$I_r(n) = \oint_{C_0(r)} \frac{e^{(n-1)\,\mathrm{Log}(-\zeta)}}{e^{\zeta}-1}\, d\zeta = (-1)^{n-1} \oint_{C_0(r)} \frac{\zeta^{n-1}}{e^{\zeta}-1}\, d\zeta \tag{8.21}$$

for every $n \in \mathbb{Z}$.

Formula (8.20) holds for $\mathrm{Re}\, z > 1$. On the other hand its right side is holomorphic in all of $\mathbb{C}$ except perhaps at the points $1, 2, 3, \ldots$, which are simple poles of the function $\Gamma(1-z)$. But at the points $2, 3, \ldots$ the function ζ is holomorphic. Therefore, these points must be roots of the function I_r. (In fact, we shall see this in a more straightforward manner in a minute.) So the only possible singularity of the right side of (8.20) in $\mathbb{C}$ is the point 1. Now

$$I_r(1) = \oint_{C_0(r)} \frac{1}{e^{\zeta}-1}\, d\zeta = 2\pi i,$$

since $\zeta = 0$ is a simple pole of the function $\frac{1}{e^{\zeta}-1}$ with residue 1. Recalling that 0 is a simple pole of Γ with residue 1, we conclude that the point 1 is a simple pole of the right side of (8.20) with residue 1.

Now we use (8.20) to extend the zeta function of Riemann to all of $\mathbb{C}$, and then ζ becomes a meromorphic function in $\mathbb{C}$ with only one pole, i.e. the point 1 with residue 1. In other words, the function

$$\zeta(z) - \frac{1}{z-1}$$

is holomorphic in $\mathbb{C}$.

From (8.21) we have that

$$I_r(n) = (-1)^{n-1} 2\pi i a_{-n}, \tag{8.22}$$

where $\sum_{-\infty}^{+\infty} a_k \zeta^k$ is the Laurent expansion of the function $\frac{1}{e^{\zeta}-1}$ at the point 0. Since 0 is a simple pole of $\frac{1}{e^{\zeta}-1}$, we get that $a_n = 0$ for every $n \leq -2$ and hence $I_r(n) = 0$ for all $n \geq 2$. This we have already seen to be true. If we calculate the first terms of the Laurent expansion of $\frac{1}{e^{\zeta}-1}$ we get

$$\frac{1}{e^{\zeta}-1} = \frac{1}{\zeta} - \frac{1}{2} + \frac{\zeta}{12} + \sum_{n=2}^{+\infty} a_n \zeta^n.$$

So from (8.20) and (8.22) we get

$$\zeta(0) = -\frac{1}{2}, \quad \zeta(-1) = -\frac{1}{12}.$$

We may also observe very easily that the function $\frac{1}{e^\zeta-1}-\frac{1}{\zeta}+\frac{1}{2}$ is odd and hence $a_n=0$ for every even $n\in\mathbb{N}$. I.e.

$$\tfrac{1}{e^\zeta-1}=\tfrac{1}{\zeta}-\tfrac{1}{2}+\tfrac{\zeta}{12}+\textstyle\sum_{k=1}^{+\infty}a_{2k+1}\zeta^{2k+1}.$$

Then, again from (8.20) and (8.22) we have

$$\zeta(-2k)=0$$

for every $k\in\mathbb{N}$.

We just saw that the points $-2,-4,-6,\dots$ are roots of ζ. These are the so-called **trivial roots** of the zeta function.

If we look at exercise 5.8.11 we see that $a_{2k-1}=(-1)^{k-1}\frac{B_k}{(2k)!}$, where B_k are the *Bernoulli constants*. Thus, from (8.20) and (8.22),

$$\zeta(-(2k-1))=(-1)^k\tfrac{B_k}{2k}$$

for every $k\in\mathbb{N}$.

Proposition 8.15. *The zeta function satisfies the functional equation*

$$\zeta(z)=2^z\pi^{z-1}\sin\tfrac{\pi z}{2}\Gamma(1-z)\zeta(1-z).$$

Proof. We go back to the proof of proposition 8.14 and we consider $0<r_1<2\pi$ and $2n\pi<r_2<2(n+1)\pi$ for an arbitrary $n\in\mathbb{N}$. Now, if $0<\delta<\frac{\pi}{2}$, the closed curve $c_{r_1,r_2,\delta}$ includes the poles $\pm 2k\pi i$, $1\le k\le n$, of the function $\frac{e^{(z-1)\,\mathrm{Log}(-\zeta)}}{e^\zeta-1}$. These poles are simple with corresponding residues $(\mp 2k\pi i)^{z-1}$. So, if $0<\delta<\frac{\pi}{2}$, then

$$\oint_{c_{r_1,r_2,\delta}}\tfrac{e^{(z-1)\,\mathrm{Log}(-\zeta)}}{e^\zeta-1}\,d\zeta=2\pi i\textstyle\sum_{k=1}^n\big((2k\pi i)^{z-1}+(-2k\pi i)^{z-1}\big).$$

Taking the limit as $\delta\to 0+$, we find

$$I_{r_2}(z)-I_{r_1}(z)=2\pi i\textstyle\sum_{k=1}^n\big((2k\pi i)^{z-1}+(-2k\pi i)^{z-1}\big).$$

We consider $\mathrm{Re}\,z>1$, we set $r=r_2$ and we take the limit as $r_1\to 0+$, to find

$$\begin{aligned}I_r(z)&=-2i\sin(\pi z)\zeta(z)\Gamma(z)+2\pi i\textstyle\sum_{k=1}^n\big((2k\pi i)^{z-1}+(-2k\pi i)^{z-1}\big)\\&=-2i\sin(\pi z)\zeta(z)\Gamma(z)+4\pi i\sin\tfrac{\pi z}{2}\textstyle\sum_{k=1}^n(2k\pi)^{z-1}.\end{aligned}\tag{8.23}$$

when $2n\pi<r<2(n+1)\pi$.
We observe that $I_r(z)$ is holomorphic in $\mathbb{C}$ and that $\sin(\pi z)\zeta(z)\Gamma(z)$ is also holomorphic in $\mathbb{C}$. Indeed, the simple poles of $\zeta(z)\Gamma(z)$, i.e. the points $0,-1,-2,\dots$, are roots of $\sin(\pi z)$. So both sides of (8.23) are holomorphic in $\mathbb{C}$ and this implies that (8.23) holds not only for $\mathrm{Re}\,z>1$ but for every z.
Now we assume that $\mathrm{Re}\,z<0$. Then

$$\lim_{n\to+\infty}\textstyle\sum_{k=1}^n(2k\pi)^{z-1}=\sum_{k=1}^{+\infty}(2k\pi)^{z-1}=(2\pi)^{z-1}\zeta(1-z).$$

We write the defining formula of $I_r(z)$:

$$I_r(z)=\oint_{C_0(r)}\tfrac{e^{(z-1)\,\mathrm{Log}(-\zeta)}}{e^\zeta-1}\,d\zeta-2i\sin(\pi z)\int_r^{+\infty}\tfrac{t^{z-1}}{e^t-1}\,dt$$

and we take $r=(2n+1)\pi$.
Since $\int_1^{+\infty}\frac{t^{z-1}}{e^t-1}\,dt$ converges, we have that

$$\lim_{n\to+\infty}\int_r^{+\infty}\tfrac{t^{z-1}}{e^t-1}\,dt=0.$$

If we restrict $\frac{1}{e^\zeta - 1}$ in its period-zone $A = \{z \mid -\pi \leq \operatorname{Im} z \leq \pi\}$ we see that $\frac{1}{|e^\zeta - 1|}$ tends to 1 or to 0 as $z \to \infty$ in A. So if we exclude a small disc $D_0(\delta)$ from the period zone, then there is a constant c_δ so that $\frac{1}{|e^\zeta - 1|} \leq c_\delta$ for every $z \in A \setminus D_0(\delta)$. Of course this extends to all period-zones and, since the circle $C_0(r)$ (with $r = (2n+1)\pi$) does not intersect any of the discs $D_{2k\pi i}(\delta)$ if we select δ sufficiently small, we get that $\frac{1}{|e^\zeta - 1|} \leq c_\delta$ for every $\zeta \in C_0(r)$ and for every n. This, with $z = x + iy$ and $x < 0$, implies that

$$\Big| \oint_{C_0(r)} \frac{e^{(z-1)\operatorname{Log}(-\zeta)}}{e^\zeta - 1}\, d\zeta \Big| \leq c_\delta r^x e^{\pi|y|} \to 0$$

when $n \to +\infty$.
Now taking the limit in (8.23) as $n \to +\infty$, we find

$$0 = -2i \sin(\pi z)\zeta(z)\Gamma(z) + 2i(2\pi)^z \sin\tfrac{\pi z}{2}\zeta(1-z)$$

for $\operatorname{Re} z < 0$. Since both terms of the last sum are holomorphic in $\mathbb{C}$, the last equality holds for every z. Now we finish the proof using the $\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$. $\square$

Since Γ has no roots and ζ has no root in the half-plane $\{z \mid 1 < \operatorname{Re} z\}$, an immediate corollary of the functional equation in proposition 8.15 is that ζ has no roots in the half-plane $\{z \mid \operatorname{Re} z < 0\}$ besides the trivial roots $-2, -4, -6, \ldots$. Therefore,
All possible roots of the zeta function, besides its trivial roots, are contained in the verical zone

$$\{z \mid 0 \leq \operatorname{Re} z \leq 1\}.$$

It has been proved that there are infinitely many non-trivial roots of the zeta function and the famous Riemann Hypothesis states that *all non-trivial roots of the zeta function lie on the vertical line with equation* $\operatorname{Re} z = \frac{1}{2}$. The Riemann Hypothesis remains unsolved and all known roots of the zeta function satisfy the $\operatorname{Re} z = \frac{1}{2}$.

Exercises.

8.5.1. Prove that the series $\sum_{n=1}^{+\infty} \frac{1}{n^z}$ does not converge when $\operatorname{Re} z = 1$, and that it has bounded partial sums when $\operatorname{Re} z = 1$, $z \neq 1$.

Chapter 9

Metric spaces.

9.1 Metrics, neighborhoods, open sets, closed sets.

Let X be a non-empty set. We call **metric** on X every function $d : X \times X \to \mathbb{R}$ with the following properties:
(i) $d(x, y) \geq 0$ for every $x, y \in X$.
(ii) For every $x, y \in X$: $d(x, y) = 0$ if and only if $x = y$.
(iii) $d(x, y) = d(y, x)$ for every $x, y \in X$.
(iv) $d(x, y) \leq d(x, z) + d(z, y)$ for every $x, y, z \in X$.
We say that the pair (X, d) is a **metric space** or that "the set X is equipped with the metric d" or we just say "the set X with the metric d". The value of $d(x, y)$ is called **distance** between x, y.

A metric space consists of two things: a non-empty set X and a metric $d : X \times X \to \mathbb{R}$ which measures distances between the elements of X. When we have a non-empty set X we may talk about the *metric space* X only when there is a preassigned specific metric d on the set X.

Example 9.1.1. The cartesian product $\mathbb{R}^d = \mathbb{R} \times \cdots \times \mathbb{R}$ with $d \geq 2$ factors is the set of all ordered d-tuples $(x_1, \ldots, x_d)$ of real numbers. Using orthogonal axes, we identify $\mathbb{R}^2$ with a plane and $\mathbb{R}^3$ with the space. If $d = 1$, we consider $\mathbb{R}^1 = \mathbb{R}$ and we identify $\mathbb{R}^1$ with a line.
If for every $\mathrm{x} = (x_1, \ldots, x_d)$ we denote

$$|\mathrm{x}| = (x_1^2 + \cdots + x_d^2)^{1/2},$$

then the **euclidean distance** between $\mathrm{x} = (x_1, \ldots, x_d), \mathrm{y} = (y_1, \ldots, y_d) \in \mathbb{R}^d$ is

$$|\mathrm{x} - \mathrm{y}| = ((x_1 - y_1)^2 + \cdots + (x_d - y_d)^2)^{1/2}.$$

It is well known that the function $d : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, defined by $d(\mathrm{x}, \mathrm{y}) = |\mathrm{x} - \mathrm{y}|$, satisfies all properties of a metric and it is called **euclidean metric** on $\mathbb{R}^d$.

In everything that follows we shall consider $\mathbb{R}^d$ equipped with the euclidean metric. In case we want to use a different metric on $\mathbb{R}^d$ we shall state this explicitly and we shall give a description of the specific metric to be used.

Let (X, d) be a metric space. If $x \in X$, $r > 0$, we call r**-neighborhood** of x or neighborhood with **center** x and **radius** r the set

$$N_x(r) = \{y \in X \mid d(y, x) < r\}.$$

It is obvious that every r-neighborhood contains at least its center.

Example 9.1.2. In $\mathbb{R}^2$ (with the euclidean metric) $N_{\mathrm{x}}(r)$ is usually denoted $D_{\mathrm{x}}(r)$ and it is the *open disc* with center x and radius r: $D_{\mathrm{x}}(r) = \{\mathrm{y} \mid |\mathrm{y} - \mathrm{x}| < r\}$. The corresponding *closed disc* is $\overline{D}_{\mathrm{x}}(r) = \{\mathrm{y} \mid |\mathrm{y} - \mathrm{x}| \leq r\}$ and the corresponding *circle* is $C_{\mathrm{x}}(r) = \{\mathrm{y} \mid |\mathrm{y} - \mathrm{x}| = r\}$.
In particular, the open disc, the closed disc and the circle with center 0 and radius 1 are denoted $\mathbb{D}$, $\overline{\mathbb{D}}$ and $\mathbb{T}$, respectively.

Example 9.1.3. In $\mathbb{R}^d$ (with the euclidean metric) $N_{\mathrm{x}}(r)$ is usually denoted $B_{\mathrm{x}}(r)$ and it is the *d-dimensional open ball* with center x and radius r: $B_{\mathrm{x}}(r) = \{\mathrm{y} \,|\, |\mathrm{y} - \mathrm{x}| < r\}$. The corresponding *d-dimensional closed ball* is $\overline{B}_{\mathrm{x}}(r) = \{\mathrm{y} \,|\, |\mathrm{y} - \mathrm{x}| \le r\}$ and the corresponding $(d-1)$*-dimensional sphere* is $S_{\mathrm{x}}(r) = \{\mathrm{y} \,|\, |\mathrm{y} - \mathrm{x}| = r\}$.
The *closed ball* with center 0 and radius 1 is usually denoted $\mathbb{B}^d$ and the sphere with center 0 and radius 1 is usually denoted $\mathbb{S}^{d-1}$.
Thus, $\mathbb{B}^1 = [-1, 1]$ and $\mathbb{S}^0 = \{-1, 1\}$. Also, $\mathbb{B}^2 = \overline{\mathbb{D}}$ and $\mathbb{S}^1 = \mathbb{T}$.

Proposition 9.1. *Let (X, d) be a metric space and $x, y \in X$, $x \neq y$. Then there is $r > 0$ so that*

$$N_x(r) \cap N_y(r) = \emptyset.$$

Proof. Take $r = \frac{1}{2}\, d(x, y) > 0$. If $z \in N_x(r) \cap N_y(r)$, i.e. $d(z, x) < r$ and $d(z, y) < r$, then

$$2r = d(x, y) \le d(x, z) + d(z, y) = d(z, x) + d(z, y) < r + r = 2r$$

and we arrive at a contradiction. Therefore $N_x(r) \cap N_y(r) = \emptyset$. □

Now we define some basic notions for a metric space (X, d). Let $A \subseteq X$ and $x \in X$. We say that x is an **interior point** of A if some neighborhood of x is contained in A. We say that x is a **boundary point** of A if every neighborhood of x intersects both A and A^c. We say that x is a **limit point** of A if every neighborhood of x intersects A. We say that x is an **accumulation point** of A if every neighborhood of x intersects A at a point different from x. We also define

$$\begin{aligned} A^\circ &= \{x \in X \,|\, x \text{ is an interior point of } A\}, \\ \partial A &= \{x \in X \,|\, x \text{ is a boundary point of } A\}, \\ \overline{A} &= \{x \in X \,|\, x \text{ is a limit point of } A\}. \end{aligned}$$

The sets A°, ∂A and $\overline{A}$ are called **interior**, **boundary** and **closure** of A, respectively.
If $A \subseteq X$, the complement of A with respect to X is denoted A^c.

Proposition 9.2. *Let (X, d) be a metric space and $A \subseteq X$. Then*
(i) $\partial A = \partial(A^c)$.
(ii) $A^\circ \subseteq A \subseteq \overline{A}$.
(iii) $\overline{A} \setminus A^\circ = \partial A$.
(iv) $A^\circ = A \setminus \partial A$.
(v) $\overline{A} = A \cup \partial A$.

Proof. (i) From the definition of a boundary point it is clear that the boundary points of A are the same as the boundary points of A^c. In other words, the sets ∂A and $\partial(A^c)$ have the same elements.
(ii) If $x \in A^\circ$, then there is a neighborhood of x which is contained in A and hence $x \in A$. Also, if $x \in A$, then every neighborhood of x intersects A and hence $x \in \overline{A}$.
(iii) Let $x \in \overline{A} \setminus A^\circ$. Since $x \in \overline{A}$, every neighborhood of x intersects A. Since $x \notin A^\circ$, there is no neighborhood of x which is contained in A and so every neighborhood of x intersects A^c. Therefore, $x \in \partial A$. Conversely, let $x \in \partial A$. Then every neighborhood of x intersects A and hence $x \in \overline{A}$. Also every neighborhood of x intersects A^c which means that there is no neighborhood of x which is contained in A and hence $x \notin A^\circ$. Thus $x \in \overline{A} \setminus A^\circ$.
(iv) and (v) are straightforward corollaries of (ii) and (iii). □

Example 9.1.4. We consider $\mathbb{R}^2$ and a relatively simple curve Γ which divides the plane in three subsets: the set A_1 of the points on one side of Γ, the set A_2 of points on the other side of Γ and the set of points of Γ. For instance Γ can be a circle or an ellipse or a line or a closed polygonal line (the circumference of a rectangle, for instance). Just looking at these shapes on the plane, we understand that $A_1^\circ = A_1$, $\partial A_1 = \Gamma$ and $\overline{A_1} = A_1 \cup \Gamma$. We have analogous results for A_2 and also $\Gamma^\circ = \emptyset$, $\partial\Gamma = \Gamma$ and $\overline{\Gamma} = \Gamma$.

Example 9.1.5. Let Γ be a relatively simple surface in $\mathbb{R}^3$ which divides the space in the set A_1 of the points on one side of Γ, the set A_2 of points on the other side of Γ and the set of points of Γ. For instance Γ can be a plane or a spherical surface or the surface of a parallelopiped. Then, as in the last example, $A_1^\circ = A_1$, $\partial A_1 = \Gamma$ and $\overline{A_1} = A_1 \cup \Gamma$. There are similar results for A_2 and also $\Gamma^\circ = \emptyset$, $\partial\Gamma = \Gamma$ and $\overline{\Gamma} = \Gamma$.

Let (X, d) be a metric space and $A \subseteq X$. We say that A is **open** if it consists only of its interior points. We say that A is **closed** if it contains all its limit points.

In other words, A is open if and only if $A = A^\circ$, and A is closed if and only if $A = \overline{A}$. It is clear from proposition 9.2 that a set is open if and only if it contains none of its boundary points and that a set is closed if and only if it contains all its boundary points.

Example 9.1.6. In examples 9.1.4 and 9.1.5 the sets A_1, A_2 are open and the sets $A_1 \cup \Gamma$, $A_2 \cup \Gamma$ and Γ are closed.

Proposition 9.3. *Let (X, d) be a metric space. Every r-neighborhood is open.*

Proof. Let $x \in X$, $r > 0$. We take any $y \in N_x(r)$ and we shall prove that there is $s > 0$ so that $N_y(s) \subseteq N_x(r)$, i.e. that y is an interior point of $N_x(r)$. This will imply that $N_x(r)$ is open.
We have $d(y, x) < r$ and we take

$$s = r - d(y, x) > 0.$$

If $w \in N_y(s)$, then

$$d(w, x) \leq d(w, y) + d(y, x) < s + d(y, x) = r$$

and thus $w \in N_x(r)$. Therefore $N_y(s) \subseteq N_x(r)$. □

Proposition 9.4. *Let (X, d) be a metric space and $A \subseteq X$. Then A is closed if and only if A^c is open.*

Proof. Since A and A^c have the same boundary points, we have the following successive equivalent statements: [A is closed] ⇔ [A contains all boundary points of A] ⇔ [A contains all boundary points of A^c] ⇔ [A^c contains no boundary point of A^c] ⇔ [A^c is open]. □

The complement of the complement of a set is the set itself and hence: *A is open if and only if A^c is closed.*

Proposition 9.5. *Let (X, d) be a metric space and $A \subseteq X$. Then A° is the largest open set contained in A and $\overline{A}$ is the smallest closed set containing A.*

Proof. (i) Let $x \in A^\circ$. Then there is $r > 0$ so that $N_x(r) \subseteq A$. We take any $y \in N_x(r)$. Since $N_x(r)$ is open, there is some $s > 0$ so that $N_y(s) \subseteq N_x(r)$ and hence $N_y(s) \subseteq A$. Therefore $y \in A^\circ$. We see that $N_x(r) \subseteq A^\circ$ and so x is an interior point of A°. Thus, every point of A° is an interior point of A° and hence A° is an open set contained in A.
Now let B be an open set contained in A. If $x \in B$, then there is $r > 0$ so that $N_x(r) \subseteq B \subseteq A$ and hence $x \in A^\circ$. Therefore $B \subseteq A^\circ$.
(ii) Let x be a limit point of $\overline{A}$. We take any $r > 0$ and then $N_x(r)$ intersects $\overline{A}$. Let $y \in N_x(r) \cap \overline{A}$. Since $N_x(r)$ is open, there is some $s > 0$ so that $N_y(s) \subseteq N_x(r)$. Since $y \in \overline{A}$, $N_y(s)$ intersects A and hence $N_x(r)$ also intersects A. Therefore, every $N_x(r)$ intersects A and so $x \in \overline{A}$. We see that every limit point of $\overline{A}$ belongs to $\overline{A}$ and thus $\overline{A}$ is a closed set containing A.
Finally, let B be a closed set containing A. If $x \in \overline{A}$, then every $N_x(r)$ intersects A and hence intersects B. Therefore $x \in \overline{B}$ and, since B is closed, $x \in B$. Thus $\overline{A} \subseteq B$. □

Proposition 9.6. *Let (X, d) be a metric space.*
(i) The union of any open subsets of X is open.
(ii) The intersection of finitely many open subsets of X is open.
(iii) The intersection of any closed subsets of X is closed.
(iv) The union of finitely many closed subsets of X is closed.

Proof. (i) If x belongs to the union U of certain open sets, then x belongs to one of these sets, say A. Since A is open, there is $r > 0$ so that $N_x(r) \subseteq A \subseteq U$. Therefore every point of U is an interior point of U and then U is open.
(ii) Let $F = A_1 \cap \cdots \cap A_n$, where A_k is open for every k. If $x \in F$, then $x \in A_k$ for every k. Thus, there are $r_1, \ldots, r_n > 0$ so that $N_x(r_k) \subseteq A_k$ for every k. We take

$$r = \min\{r_1, \ldots, r_n\} > 0.$$

Then

$$N_x(r) \subseteq N_x(r_k) \subseteq A_k$$

for every k and hence $N_x(r) \subseteq F$. Therefore every point of F is an interior point of F and then F is open.
(iii) and (iv) are immediate consequences of (i) and (ii), of proposition 9.4 and of the laws of de Morgan: $(\bigcap A)^c = \bigcup A^c$ and $(\bigcup A)^c = \bigcap A^c$. □

Let X be a non-empty set and d_1, d_2 be metrics on X. We say that the two metrics are **equivalent** if the metric spaces (X, d_1) and (X, d_2) have the same open sets: every A which is open in (X, d_1) is also open in (X, d_2) and conversely.

Proposition 9.4 says that the closed sets in any netric space are the complements of the open sets. Therefore, *the metrics d_1, d_2 on X are equivalent if and only if the metric spaces (X, d_1) and (X, d_2) have the same closed sets.*

Proposition 9.7. *Let X be non-empty and d_1, d_2 be metrics on X. We denote $N_x^{d_1}(r)$ and $N_x^{d_2}(r)$ the neighborhoods of x in the metric spaces (X, d_1) and (X, d_2), respectively. The following are equivalent.*
(i) d_1, d_2 are equivalent.
(ii) For every $x \in X$ and every $\epsilon > 0$ there is $\delta > 0$ so that $N_x^{d_1}(\delta) \subseteq N_x^{d_2}(\epsilon)$ and, conversely, for every $x \in X$ and every $\epsilon > 0$ there is $\delta > 0$ so that $N_x^{d_2}(\delta) \subseteq N_x^{d_1}(\epsilon)$.

Proof. (i) $\Rightarrow$ (ii) Let $x \in X$ and $\epsilon > 0$. The neighborhood $N_x^{d_2}(\epsilon)$ is open in the metric space (X, d_2). Since (X, d_1) and (X, d_2) have the same open sets, $N_x^{d_2}(\epsilon)$ is also open in (X, d_1). Because $x \in N_x^{d_2}(\epsilon)$, there is $\delta > 0$ so that $N_x^{d_1}(\delta) \subseteq N_x^{d_2}(\epsilon)$. The converse is similar.
(ii) $\Rightarrow$ (i) Let A be open in (X, d_1). We shall prove that A is also open in (X, d_2).
We take any $x \in A$. Since A is open in (X, d_1), there is $\epsilon > 0$ so that $N_x^{d_1}(\epsilon) \subseteq A$. Then there is $\delta > 0$ so that $N_x^{d_2}(\delta) \subseteq N_x^{d_1}(\epsilon)$ and thus $N_x^{d_2}(\delta) \subseteq A$. Therefore every element of A is an interior point of A in (X, d_2) and so A is open in (X, d_2). The converse is similar. □

Exercises.

9.1.1. (i) We define three functions $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$d(x, y) = (x - y)^2, \quad d(x, y) = |x - y|^{1/2}, \quad d(x, y) = \tfrac{|x-y|}{1+|x-y|}.$$

Which of these d is a metric on $\mathbb{R}$?
(ii) For every $\mathrm{x} = (x_1, x_2), \mathrm{y} = (y_1, y_2)$ in $\mathbb{R}^2$ we set $d(\mathrm{x}, \mathrm{y}) = \big((x_1 - y_1)^2 + 4(x_2 - y_2)^2\big)^{1/2}$. Is d a metric on $\mathbb{R}^2$?
(iii) Let $d(\mathrm{x}, \mathrm{y}) = |x_1 - y_1|$ for every $\mathrm{x} = (x_1, x_2, x_3), \mathrm{y} = (y_1, y_2, y_3)$ in $\mathbb{R}^3$. Is d a metric on $\mathbb{R}^3$?

9.1.2. Which of the following are open or closed subsets of $\mathbb{R}$?

$$\mathbb{N}, \quad \mathbb{Q}, \quad \{1/n \,|\, n \in \mathbb{N}\}, \quad \{0\} \cup \{1/n \,|\, n \in \mathbb{N}\}, \quad [0,1) \cup \{1 + 1/n \,|\, n \in \mathbb{N}\}.$$

Find their interiors, their closures and their boundaries.

9.1.3. Which of the following are open or closed subsets of $\mathbb{R}^2$?

$$\{(x_1, x_2) \,|\, x_1 > 0\}, \quad \{(x_1, 0) \,|\, a \le x_1 \le b\}, \quad \{(x_1, 0) \,|\, a < x_1 < b\}, \quad \{(x_1, x_2) \,|\, x_1 x_2 \le 1\},$$

$$\{(x_1, x_2) \,|\, x_1 x_2 > 1\}, \quad \{(1/n, 0) \,|\, n \in \mathbb{N}\}, \quad [0,1] \times (\{0\} \cup \{1/n \,|\, n \in \mathbb{N}\}).$$

Find their interiors, their closures and their boundaries.

9.1.4. Which of the following are open or closed subsets of $\mathbb{R}^3$?

$$\{(x_1, x_2, x_3) \,|\, x_1 > 0\}, \quad \{(x_1, 0, 0) \,|\, a < x_1 < b\}, \quad \{(x_1, 0, 0) \,|\, a \le x_1 \le b\},$$

$$\{(x_1, x_2, 0) \,|\, a \le x_1 \le b, c \le x_2 \le d\}, \quad \{(x_1, x_2, x_3) \,|\, x_1^2 + x_2^2 < x_3\}.$$

Find their interiors, their closures and their boundaries.

9.1.5. Let $\mathrm{x} \cdot \mathrm{y} = x_1 y_1 + \cdots + x_d y_d$ be the usual euclidean inner product in $\mathbb{R}^d$. Let $\mathrm{a} \in \mathbb{R}^d$, $\mathrm{a} \neq 0$ and $a \in \mathbb{R}$. The set $\Gamma = \{\mathrm{x} \in \mathbb{R}^d \,|\, \mathrm{a} \cdot \mathrm{x} = a\}$ is a hyperplane of $\mathbb{R}^d$. The open halfspaces of $\mathbb{R}^d$ determined by Γ are $A_1 = \{\mathrm{x} \in \mathbb{R}^d \,|\, \mathrm{a} \cdot \mathrm{x} > a\}$ and $A_2 = \{\mathrm{x} \in \mathbb{R}^d \,|\, \mathrm{a} \cdot \mathrm{x} < a\}$ and the corresponding closed halfspaces are $B_1 = \{\mathrm{x} \in \mathbb{R}^d \,|\, \mathrm{a} \cdot \mathrm{x} \ge a\}$ and $B_2 = \{\mathrm{x} \in \mathbb{R}^d \,|\, \mathrm{a} \cdot \mathrm{x} \le a\}$. Find the interiors, the closures and the boundaries of Γ, A_1, A_2, B_1 and B_2.

9.1.6. In $\mathbb{R}^d$, the general open or closed orthogonal parallelepiped with edges parallel to the coordinate axes is $(a_1, b_1) \times \cdots \times (a_d, b_d)$ or $[a_1, b_1] \times \cdots \times [a_d, b_d]$, respectively. Prove that the first set is open and the second is closed.

9.1.7. Let (X, d) be a metric space.
(i) Prove that both X and $\emptyset$ are open and closed subsets of X.
(ii) If $A \subseteq X$, prove that ∂A is closed.
(iii) Prove that every finite subset of X is closed.
(iv) If $A \subseteq B \subseteq X$, prove that $A^\circ \subseteq B^\circ$ and $\overline{A} \subseteq \overline{B}$.
(v) If $A \subseteq X$ is open and $B \subseteq X$ is closed, prove that $A \setminus B$ is open and $B \setminus A$ is closed.

9.1.8. Let (X, d) be a metric space and $A, B \subseteq X$ be closed and disjoint. Prove that there are $U, V \subseteq X$ open and disjoint so that $A \subseteq U$ and $B \subseteq V$.

9.1.9. Let (X, d) be a metric space, $A \subseteq X$ and $x \in X$. We define the distance of x from A to be $d(x, A) = \inf\{d(x, y) \,|\, y \in A\}$. Prove that:
(i) $d(x, A) = d(x, \overline{A})$.
(ii) $d(x, A) = 0 \Leftrightarrow x \in \overline{A}$.
(iii) $|d(x_1, A) - d(x_2, A)| \le d(x_1, x_2)$.

9.1.10. Let X be any non-empty set and $d : X \times X \to \mathbb{R}$ be the function defined by $d(x, x) = 1$ for every $x \in X$ and by $d(x, y) = 0$ for every $x, y \in X$ with $x \neq y$.
(i) Prove that d is a metric on X. This metric is called **discrete metric**.
(ii) Prove that every $A \subseteq X$ (with the discrete metric) is open and closed. Prove that $A^\circ = \overline{A} = A$ and $\partial A = \emptyset$ for every $A \subseteq X$.

9.1.11. Let (X, d) be a metric space. We define $d' : X \times X \to \mathbb{R}$ by $d'(x, y) = \frac{d(x,y)}{d(x,y)+1}$. Prove that d' is a metric on X and that d, d' are equivalent.

9.1.12. Let X be non-empty, d_1, d_2 be equivalent metrics on X and $A \subseteq X$. Prove that in both metric spaces, (X, d_1) and (X, d_2), A has the same interior points, the same boundary points and the same limit points.

9.2 Limits and continuity of functions.

Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$, $f : A \to Y$, $x_0 \in X$ be an accumulation point of A and $y_0 \in Y$. We say that y_0 is a **limit** of f at x_0, and denote

$$y_0 = \lim_{x \to x_0} f(x),$$

if for every $\epsilon > 0$ there is $\delta > 0$ so that $f(x) \in N_{y_0}(\epsilon)$ for every $x \in N_{x_0}(\delta) \cap A$, $x \neq x_0$ or, equivalently, if for every $\epsilon > 0$ there is $\delta > 0$ so that $\rho(f(x), y_0) < \epsilon$ for every $x \in A$ with $0 < d(x, x_0) < \delta$.

This definition of the limit of a function is the direct generalization of the well known definition in case both metric spaces (X, d) and (Y, ρ) are the euclidean space $\mathbb{R}$.

Proposition 9.8. *Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$, $f : A \to Y$ and $x_0 \in X$ be an accumulation point of A. If f has a limit at x_0, then this limit is unique.*

Proof. Let

$$y_0' = \lim_{x \to x_0} f(x), \quad y_0'' = \lim_{x \to x_0} f(x),$$

where $y_0', y_0'' \in Y$. We assume $y_0' \neq y_0''$ and then proposition 9.1 implies that there is $\epsilon > 0$ so that

$$N_{y_0'}(\epsilon) \cap N_{y_0''}(\epsilon) = \emptyset.$$

Then there is $\delta > 0$ so that $f(x) \in N_{y_0'}(\epsilon)$ and $f(x) \in N_{y_0''}(\epsilon)$ for every $x \in N_{x_0}(\delta) \cap A$, $x \neq x_0$, and we arrive at a contradiction. $\square$

Proposition 9.8 allows us to talk about **the** limit of a function at a point.

Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$, $f : A \to Y$ and $x_0 \in A$. We say that f is **continuous** at x_0 if for every $\epsilon > 0$ there is $\delta > 0$ so that $f(x) \in N_{f(x_0)}(\epsilon)$ for every $x \in N_{x_0}(\delta) \cap A$ or, equivalently, if for every $\epsilon > 0$ there is $\delta > 0$ so that $\rho(f(x), f(x_0)) < \epsilon$ for every $x \in A$ with $d(x, x_0) < \delta$.

If $x_0 \in A$ is not an accumulation point of A, i.e. if it is an **isolated point** of A, then we may easily see that f is automatically continuous at x_0. On the other hand, if $x_0 \in A$ is an accumulation point of A, then f is continuous at x_0 if and only if $\lim_{x \to x_0} f(x) = f(x_0)$.

Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$ and $f : A \to Y$. We say that f is **continuous** in A if it is continuous at every point of A.

Proposition 9.9. *Let (X, d), (Y, ρ) and (Z, τ) be metric spaces, $A \subseteq X$, $B \subseteq Y$, $x_0 \in A$, $f : A \to B$ and $g : B \to Z$. If f is continuous at x_0 and g is continuous at $y_0 = f(x_0)$, then $g \circ f : A \to Z$ is continuous at x_0.*

Proof. We take $\epsilon > 0$ and then there is $\delta' > 0$ so that

$$\tau(g(y), g(y_0)) < \epsilon \tag{9.1}$$

for every $y \in B$ with $\rho(y, y_0) < \delta'$. Then there is $\delta > 0$ so that

$$\rho(f(x), y_0) = \rho(f(x), f(x_0)) < \delta' \tag{9.2}$$

for every $x \in A$ with $d(x, x_0) < \delta$. From (9.2) and from (9.1) with $y = f(x)$ we get that

$$\tau(g(f(x)), g(f(x_0))) < \epsilon$$

for every $x \in A$ with $d(x, x_0) < \delta$. Thus $g \circ f : A \to Z$ is continuous at x_0. $\square$

Proposition 9.10. *Let (X, d) be a metric space, $A \subseteq X$, $x_0 \in A$, $f, g : A \to \mathbb{R}$ be continuous at x_0 and $\lambda, \mu \in \mathbb{R}$. Then:*
(i) $\lambda f + \mu g : A \to \mathbb{R}$ and $fg : A \to \mathbb{R}$ are continuous at x_0.
(ii) If $B = \{x \in A \,|\, g(x) \neq 0\}$ and $g(x_0) \neq 0$, then $\frac{1}{g} : B \to \mathbb{R}$ is continuous at x_0.

Proof. (i) We take any $\epsilon > 0$ and then there is $\delta > 0$ so that

$$|f(x) - f(x_0)| < \tfrac{\epsilon}{2(|\lambda|+1)}, \quad |g(x) - g(x_0)| < \tfrac{\epsilon}{2(|\mu|+1)}$$

for every $x \in A$ with $d(x, x_0) < \delta$. This implies that

$$\begin{aligned} \big|(\lambda f(x) + \mu g(x)) - (\lambda f(x_0) + \mu g(x_0))\big| &\leq |\lambda||f(x) - f(x_0)| + |\mu||g(x) - g(x_0)| \\ &\leq |\lambda|\tfrac{\epsilon}{2(|\lambda|+1)} + |\mu|\tfrac{\epsilon}{2(|\mu|+1)} < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon \end{aligned}$$

for every $x \in A$ with $d(x, x_0) < \delta$ and hence $\lambda f + \mu g : A \to \mathbb{R}$ is continuous at x_0.
We then take any $\epsilon > 0$ and we set

$$\epsilon_1 = \min\left\{(\tfrac{\epsilon}{3})^{1/2}, \tfrac{\epsilon}{3(|f(x_0)|+1)}, \tfrac{\epsilon}{3(|g(x_0)|+1)}\right\} > 0.$$

Then there is $\delta > 0$ so that

$$|f(x) - f(x_0)| < \epsilon_1, \quad |g(x) - g(x_0)| < \epsilon_1$$

for every $x \in A$ with $d(x, x_0) < \delta$. This implies that

$$\begin{aligned} |f(x)g(x) - f(x_0)g(x_0)| &\leq |f(x) - f(x_0)||g(x) - g(x_0)| + |f(x_0)||g(x) - g(x_0)| \\ &\quad + |g(x_0)||f(x) - f(x_0)| \\ &\leq {\epsilon_1}^2 + |f(x_0)|\epsilon_1 + |g(x_0)|\epsilon_1 < \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} = \epsilon \end{aligned}$$

for every $x \in A$ with $d(x, x_0) < \delta$ and hence $fg : A \to \mathbb{R}$ is continuous at x_0.
(ii) We take any $\epsilon > 0$ and then there is $\delta > 0$ so that

$$|g(x) - g(x_0)| < \min\left\{\tfrac{|g(x_0)|}{2}, \tfrac{g(x_0)^2}{2}\epsilon\right\}$$

for every $x \in A$ with $d(x, x_0) < \delta$. This implies that

$$|g(x)| = |g(x_0) + (g(x) - g(x_0))| \geq |g(x_0)| - |g(x) - g(x_0)| > |g(x_0)| - \tfrac{|g(x_0)|}{2} = \tfrac{|g(x_0)|}{2}$$

and hence

$$\left|\tfrac{1}{g(x)} - \tfrac{1}{g(x_0)}\right| = \tfrac{|g(x)-g(x_0)|}{|g(x)||g(x_0)|} \leq \tfrac{2|g(x)-g(x_0)|}{g(x_0)^2} < \epsilon$$

for every $x \in B$ with $d(x, x_0) < \delta$. Therefore $\frac{1}{g} : B \to \mathbb{R}$ is continuous at x_0. ☐

The proof of proposition 9.11 is almost identical to the previous proof.

Proposition 9.11. *Let (X, d) be a metric space, $A \subseteq X$, $f, g : A \to \mathbb{R}$, $x_0 \in X$ be an accumulation point of A, $\lim_{x\to x_0} f(x) = y_0 \in \mathbb{R}$, $\lim_{x\to x_0} g(x) = z_0 \in \mathbb{R}$ and $\lambda, \mu \in \mathbb{R}$. Then:*
(i) $\lim_{x\to x_0}(\lambda f + \mu g) = \lambda y_0 + \mu z_0$ and $\lim_{x\to x_0} fg = y_0 z_0$.
(ii) If $z_0 \neq 0$, then x_0 is an accumulation point of $B = \{x \in A \,|\, g(x) \neq 0\}$ and $\lim_{x\to x_0} \frac{1}{g(x)} = \frac{1}{z_0}$.

Combining propositions 9.9 and 9.10 and starting from very simple examples of continuous functions, we can produce more complicated ones.

Example 9.2.1. In $\mathbb{R}^d$ we define the k**-projection** $\pi_k : \mathbb{R}^d \to \mathbb{R}$ by $\pi_k(\mathrm{x}) = x_k$ for every $\mathrm{x} = (x_1, \ldots, x_d)$. Every π_k is continuous, since

$$|\pi_k(\mathrm{x}) - \pi_k(\mathrm{y})| \leq |\mathrm{x} - \mathrm{y}|$$

for every $\mathrm{x}, \mathrm{y} \in \mathbb{R}^d$. Therefore, if $g : \mathbb{R} \to \mathbb{R}$ is continuous, then $f : \mathbb{R}^d \to \mathbb{R}$ defined by $f(\mathrm{x}) = g(x_k)$ for every $\mathrm{x} = (x_1, \ldots, x_d)$ is continuous, since $f = g \circ \pi_k$.
Thus, *polynomial functions*

$$p(x_1, \ldots, x_d) = Ax_1^{a_1} \cdots x_d^{a_d} + Bx_1^{b_1} \cdots x_d^{b_d} + \cdots,$$

where all exponents are non-negative integers, all coefficients are real numbers and the sum is finite, are continuous functions. *Rational functions*, i.e. quotients of polynomial functions, are also continuous (except at the points where their denominator vanishes) as well as functions which are simple combinations of exponential or trigonometric or other simple continuous functions of the coordinates.

Proposition 9.12. *Let (X,d) and (Y,ρ) be metric spaces, $A \subseteq X$ and $f : A \to Y$. Then the following are equivalent.*
(i) f is continuous in A.
(ii) For every open $W \subseteq Y$ there is an open $U \subseteq X$ so that $f^{-1}(W) = U \cap A$.
(iii) For every closed $F \subseteq Y$ there is a closed $G \subseteq X$ so that $f^{-1}(F) = G \cap A$.

Proof. (i) $\Rightarrow$ (ii) Let $x \in f^{-1}(W)$, i.e. $f(x) \in W$. Since W is open, there is $\epsilon_x > 0$ so that $N_{f(x)}(\epsilon_x) \subseteq W$. Since f is continuous, there is $\delta_x > 0$ so that

$$f(y) \in N_{f(x)}(\epsilon_x) \subseteq W,$$

and hence $y \in f^{-1}(W)$, for every $y \in N_x(\delta_x) \cap A$. Therefore,

$$N_x(\delta_x) \cap A \subseteq f^{-1}(W).$$

Now we consider the set

$$U = \bigcup_{x\in f^{-1}(W)} N_x(\delta_x).$$

Then U is a union of open sets and so it is open. We also have

$$U \cap A = \bigcup_{x\in f^{-1}(W)} \big(N_x(\delta_x) \cap A\big) \subseteq f^{-1}(W).$$

On the other hand it is clear that for every $x \in f^{-1}(W)$ we have $x \in N_x(\delta_x) \cap A$ and hence $x \in U \cap A$. Thus, $f^{-1}(W) \subseteq U \cap A$.
(ii) $\Rightarrow$ (i) Take any $x_0 \in A$ and any $\epsilon > 0$. Then $N_{f(x_0)}(\epsilon)$ is open in Y and so there is an open $U \subseteq X$ so that

$$f^{-1}\big(N_{f(x_0)}(\epsilon)\big) = U \cap A.$$

Then $x_0 \in U \cap A$ and, since U is open, there is $\delta > 0$ so that $N_{x_0}(\delta) \subseteq U$. Now, for every $x \in N_{x_0}(\delta) \cap A$ we have $x \in U \cap A$ and hence $x \in f^{-1}\big(N_{f(x_0)}(\epsilon)\big)$ i.e. $f(x) \in N_{f(x_0)}(\epsilon)$. Therefore, f is continuous at every $x_0 \in A$.
The equivalence (i) $\Leftrightarrow$ (iii) is a consequence of the equivalence (i) $\Leftrightarrow$ (ii) and of the general identity $f^{-1}(W^c) = (f^{-1}(W))^c \cap A$. $\square$

The metric spaces (X,d) and (Y,ρ) are called **homeomorphic** if there is $f : X \to Y$ which is one-to-one in X and onto Y and so that f is continuous in X and $f^{-1} : Y \to X$ is continuous in Y.

It is trivial to prove that the relation of homeomorphism between metric spaces is an equivalence relation. It is also trivial to see, based for instance on proposition 9.7, that, if d_1 and d_2 are two metrics on the non-empty set X, then the two metrics are equivalent if and only if the identity function between (X,d_1) and (X,d_2) is a homeomorphism.

Exercises.

9.2.1. Prove that $\{\mathrm{x} \in \mathbb{R}^d \setminus \{0\} \,|\, e^{-|\mathrm{x}|} + \sin|\mathrm{x}| > 0\}$ is an open subset of $\mathbb{R}^d$.
Is $\{\mathrm{x} \in \mathbb{R}^d \setminus \{0\} \,|\, |\mathrm{x}| - |\mathrm{x}|^3 \le 3\}$ a closed subset of $\mathbb{R}^d$?

9.2.2. Consider metric spaces (X,d), (Y,ρ) and $A \subseteq X$, $B \subseteq Y$ and a continuous $f : A \to Y$.
(i) If A, B are open, prove that $f^{-1}(B)$ is open.
(ii) If A, B are closed, prove that $f^{-1}(B)$ is closed.

9.2.3. Let X, Y be non-empty sets, $A \subseteq X$, $x_0 \in A$ and $f : A \to Y$. Let d_1, d_2 be equivalent metrics on X and ρ_1, ρ_2 be equivalent metrics on Y. Prove that f is continuous at x_0 with respect to d_1 and ρ_1 if and only if it is continuous at x_0 with respect to d_2 and ρ_2.

9.2.4. Let (X, d) be a non-empty set with the discrete metric (exercise 9.1.10), (Y, ρ) be any metric space, $A \subseteq X$ and $f : A \to Y$. Prove that f is continuous in A.

9.3 Sequences.

The next definition is the generalization of the analogous definition in the euclidean space $\mathbb{R}$.

Let (X, d) be a metric space, $x \in X$ and let (x_n) be a sequence in X. We say that (x_n) **converges** to x in (X, d) or that x is a **limit** of (x_n) in (X, d), and denote

$$x_n \to x \qquad \text{or} \qquad \lim_{n\to+\infty} x_n = x,$$

if for every $\epsilon > 0$ there is n_0 so that $x_n \in N_x(\epsilon)$ for every $n \geq n_0$ or, equivalently, if for every $\epsilon > 0$ there is n_0 so that $d(x_n, x) < \epsilon$ for every $n \geq n_0$.

It is clear that $x_n \to x$ in the metric space (X, d) if and only if $d(x_n, x) \to 0$ in $\mathbb{R}$.

Proposition 9.13. *Let (X, d) be a metric space and let (x_n) be a sequence in X. If (x_n) has a limit, then this limit is unique.*

Proof. Let $x_n \to x'$ and $x_n \to x''$ and assume that $x' \neq x''$. We know that there is $\epsilon > 0$ so that $N_{x'}(\epsilon) \cap N_{x''}(\epsilon) = \emptyset$. Then there is n_0 so that $x_n \in N_{x'}(\epsilon)$ and $x_n \in N_{x''}(\epsilon)$ for every $n \geq n_0$ and this is impossible. □

Because of proposition 9.13, we can talk about **the** limit of a sequence.

The next proposition reduces convergence in the euclidean space $\mathbb{R}^d$ to convergence in $\mathbb{R}$.

Proposition 9.14. *Let $\mathrm{x}_n = (x_{n,1}, \dots, x_{n,d}) \in \mathbb{R}^d$ for every n and $\mathrm{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$. The following are equivalent.*
(i) $\mathrm{x}_n \to \mathrm{x}$ in $\mathbb{R}^d$.
(ii) $x_{n,k} \to x_k$ in $\mathbb{R}$ for every $k = 1, \dots, d$.

Proof. (i) ⇒ (ii) A consequence of $|x_{n,k} - x_k| \leq |\mathrm{x}_n - \mathrm{x}|$.
(ii) ⇒ (i) A consequence of $|\mathrm{x}_n - \mathrm{x}| \leq |x_{n,1} - x_1| + \cdots + |x_{n,d} - x_d|$. □

We shall now see the close relation between the notion of convergence of sequences and certain notions we have encountered already: the notion of limit point, the notion of closed set (and, indirectly, of open set) and, finally, the notions of the limit and continuity of a function.

Proposition 9.15. *Let (X, d) be a metric space, $A \subseteq X$ and $x \in X$. Then x is a limit point of A if and only if there is a sequence (x_n) in A so that $x_n \to x$.*

Proof. Let x be a limit point of A. We take any $n \in \mathbb{N}$ and then $N_x\left(\frac{1}{n}\right)$ contains at least one point of A, i.e. there is $x_n \in A$ so that $d(x_n, x) < \frac{1}{n}$. Thus, the sequence (x_n) is in A and $x_n \to x$.
Conversely, let (x_n) be a sequence in A so that $x_n \to x$. We take any $\epsilon > 0$ and then there is n_0 so that $x_n \in N_x(\epsilon)$ for every $n \geq n_0$. Thus $N_x(\epsilon)$ intersects A and so x is a limit point of A. □

Example 9.3.1. Let us prove that the closure of the open ball $B_{\mathrm{x}_0}(r)$ in $\mathbb{R}^d$ is the corresponding closed ball $\overline{B}_{\mathrm{x}_0}(r)$.
Assume that x is a limit point of $B_{\mathrm{x}_0}(r)$. Then there is a sequence (x_n) in $B_{\mathrm{x}_0}(r)$ so that $\mathrm{x}_n \to \mathrm{x}$, i.e. $|\mathrm{x}_n - \mathrm{x}| \to 0$. Then from

$$\big||\mathrm{x}_n - \mathrm{x}_0| - |\mathrm{x} - \mathrm{x}_0|\big| \leq |\mathrm{x}_n - \mathrm{x}|$$

we find $|\mathrm{x}_n - \mathrm{x}_0| \to |\mathrm{x} - \mathrm{x}_0|$. Since $|\mathrm{x}_n - \mathrm{x}_0| < r$ for every n, we get $|\mathrm{x} - \mathrm{x}_0| \le r$ and so $\mathrm{x} \in \overline{B}_{\mathrm{x}_0}(r)$. Thus, $\overline{B_{\mathrm{x}_0}(r)} \subseteq \overline{B}_{\mathrm{x}_0}(r)$.
Conversely, take $\mathrm{x} \in \overline{B}_{\mathrm{x}_0}(r)$, i.e. $|\mathrm{x} - \mathrm{x}_0| \le r$. For each $n \in \mathbb{N}$ we consider

$$\mathrm{x}_n = \tfrac{1}{n}\,\mathrm{x}_0 + (1 - \tfrac{1}{n})\,\mathrm{x}.$$

Then

$$|\mathrm{x}_n - \mathrm{x}_0| = (1 - \tfrac{1}{n})|\mathrm{x} - \mathrm{x}_0| < r$$

and hence $\mathrm{x}_n \in B_{\mathrm{x}_0}(r)$ for every n. Also, $\mathrm{x}_n \to \mathrm{x}$ and so $\mathrm{x} \in \overline{B_{\mathrm{x}_0}(r)}$. Thus, $\overline{B}_{\mathrm{x}_0}(r) \subseteq \overline{B_{\mathrm{x}_0}(r)}$.

Proposition 9.16. *Let (X, d) be a metric space and $A \subseteq X$. The following are equivalent.*
(i) A is closed.
(ii) Every x, which is the limit of a sequence in A, belongs to A.

Proof. (i) ⇒ (ii) Take any x which is the limit of a sequence in A. Proposition 9.15 implies that x is a limit point of A and, since A is closed, $x \in A$.
(ii) ⇒ (i) Take any limit point x of A. Proposition 9.15 implies that there is a sequence in A with limit x and hence x belongs to A. Thus A contains all its limit points and so it is closed. □

Propositions 9.17 and 9.18 are generalizations of analogous propositions for $\mathbb{R}$.

Proposition 9.17. *Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$, $x_0 \in A$ and $f : A \to Y$. The following are equivalent.*
(i) f is continuous at x_0.
(ii) For every (x_n) in A with $x_n \to x_0$ we have $f(x_n) \to f(x_0)$.

Proof. (i) ⇒ (ii) Take (x_n) in A with $x_n \to x_0$. We take any $\epsilon > 0$ and then there is $\delta > 0$ so that

$$\rho(f(x), f(x_0)) < \epsilon \tag{9.3}$$

for every $x \in A$ with $d(x, x_0) < \delta$. Then there is n_0 so that

$$d(x_n, x_0) < \delta \tag{9.4}$$

for every $n \ge n_0$. Now (9.4) and (9.3) with $x = x_n$ imply that for every $n \ge n_0$ we have

$$\rho(f(x_n), f(x_0)) < \epsilon.$$

Therefore $f(x_n) \to f(x_0)$.
(ii) ⇒ (i) Assume that f *is not* continuous at x_0. Then there is $\epsilon > 0$ so that for every $\delta > 0$ there is $x \in A$ such that

$$d(x, x_0) < \delta, \quad \rho(f(x), f(x_0)) \ge \epsilon.$$

Hence for every $n \in \mathbb{N}$ there is $x_n \in A$ with

$$d(x_n, x_0) < \tfrac{1}{n}, \quad \rho(f(x_n), f(x_0)) \ge \epsilon.$$

Then (x_n) is in A and $x_n \to x_0$ but $f(x_n) \not\to f(x_0)$ and we arrived at a contradiction. □

The proof of proposition 9.18 is almost identical to the proof of proposition 9.17.

Proposition 9.18. *Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$, x_0 be an accumulation point of A, $y_0 \in Y$ and $f : A \to Y$. The following are equivalent.*
(i) $\lim_{x \to x_0} f(x) = y_0$.
(ii) For every (x_n) in $A \setminus \{x_0\}$ with $x_n \to x_0$ we have $f(x_n) \to y_0$.

Let (X, d) be a metric space and (x_n) be a sequence in X. We say that (x_n) is a **Cauchy sequence** if for every $\epsilon > 0$ there is n_0 so that $d(x_n, x_m) < \epsilon$ for every $n, m \ge n_0$.

Proposition 9.19. *Let (X,d) be a metric space and (x_n) be a sequence in X. If (x_n) converges to some element of X, then it is a Cauchy sequence.*

Proof. Let $x_n \to x$. If $\epsilon > 0$, then there is n_0 so that $d(x_n, x) < \frac{\epsilon}{2}$ for every $n \geq n_0$. Therefore,

$$d(x_n, x_m) \leq d(x_n, x) + d(x_m, x) < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon$$

for every $n, m \geq n_0$ and so (x_n) is a Cauchy sequence. □

Let (X,d) be a metric space and $A \subseteq X$. We say that A is **complete** if every Cauchy sequence in A converges to some element of A.

Proposition 9.20. *Let X be non-empty and let d_1, d_2 be metrics on X. The following are equivalent.*
(i) The metrics d_1, d_2 are equivalent.
(ii) The metric spaces (X, d_1) and (X, d_2) have the same convergent sequences.

Proof. (i) ⇒ (ii) Let $x_n \to x$ in (X, d_1). We shall prove that $x_n \to x$ also in (X, d_2).
Let $\epsilon > 0$. Proposition 9.7 implies that there is $\delta > 0$ so that $N_x^{d_1}(\delta) \subseteq N_x^{d_2}(\epsilon)$. Since $x_n \to x$ in (X, d_1), there is n_0 so that $x_n \in N_x^{d_1}(\delta)$, and hence $x_n \in N_x^{d_2}(\epsilon)$, for every $n \geq n_0$. Thus $x_n \to x$ in (X, d_2).
The converse is similar.
(ii) ⇒ (i) Let $A \subseteq X$ be closed in (X, d_1). We shall see that A is closed also in (X, d_2).
We assume that (x_n) is in A and $x_n \to x$ in (X, d_2). Then $x_n \to x$ also in (X, d_1) and, since A is closed in (X, d_1), we get $x \in A$. Thus A is closed in (X, d_2).
The converse is similar. □

Exercises.

9.3.1. Let $x_n \to x$ and $y_n \to y$ in (X, d). Prove that $d(x_n, y_n) \to d(x, y)$ in $\mathbb{R}$.

9.3.2. Let (X,d) be a metric space, $A \subseteq X$ and $x \in X$. Prove that x is a boundary point of A if and only if there are sequences (x'_n) in A and (x''_n) in A^c so that $x'_n \to x$ and $x''_n \to x$.

9.3.3. We consider sequences (x_n) and (y_n) in $\mathbb{R}^d$ and (λ_n) in $\mathbb{R}$. If $\mathrm{x}_n \to \mathrm{x}$, $\mathrm{y}_n \to \mathrm{y}$ in $\mathbb{R}^d$ and $\lambda_n \to \lambda$ in $\mathbb{R}$, prove that $\mathrm{x}_n + \mathrm{y}_n \to \mathrm{x} + \mathrm{y}$ and $\lambda_n \mathrm{x}_n \to \lambda \mathrm{x}$ in $\mathbb{R}^d$ and that $\mathrm{x}_n \cdot \mathrm{y}_n \to \mathrm{x} \cdot \mathrm{y}$ in $\mathbb{R}$.

9.3.4. Using sequences, prove that $\{\frac{1}{n} \,|\, n \in \mathbb{N}\}$ is not a closed subset of $\mathbb{R}$ while $\{0\} \cup \{\frac{1}{n} \,|\, n \in \mathbb{N}\}$ is a closed subset of $\mathbb{R}$.

9.3.5. Using sequences, prove that closed balls, hyperplanes and closed halfspaces in $\mathbb{R}^d$ are closed subsets of $\mathbb{R}^d$.

9.3.6. Let (X,d) be a non-empty set with the discrete metric (exercise 9.1.10). Prove that a sequence (x_n) in X converges if and only if it is constant after some value of its index n.

9.4 Compactness.

Let X be non-empty, $M \subseteq X$ and let Σ be a collection of subsets of X. We say that Σ is a **covering** of M if $M \subseteq \bigcup_{A \in \Sigma} A$. If, moreover, Σ is finite, we say that it is a **finite covering** of M. Now, if Σ and Σ' are coverings of M such that $\Sigma' \subseteq \Sigma$, then we say that Σ is **larger** than Σ' and that Σ' is **smaller** than Σ.

Let (X,d) be a metric space and $M \subseteq X$. If Σ is a covering of M and all $A \in \Sigma$ are open sets, then Σ is called **open covering** of M. We say that M is **compact** if for every open covering Σ of M there is a finite covering Σ' of M which is smaller than Σ.

Example 9.4.1. Let (X, d) be a metric space and $M = \{x_1, \dots, x_n\} \subseteq X$. We take any open covering Σ of M. Then every $x_k \in M$ belongs to some $A_k \in \Sigma$ and hence $M \subseteq A_1 \cup \cdots \cup A_n$. Thus, $\Sigma' = \{A_1, \dots, A_n\}$ is a finite covering of M with $\Sigma' \subseteq \Sigma$. Hence M is compact.

Let (X, d) be a metric space and $M \subseteq X$. We say that M is **bounded** if there is $x_0 \in X$ and $r > 0$ so that $M \subseteq N_{x_0}(r)$.

Example 9.4.2. A set M in $\mathbb{R}^d$ is bounded if and only if it is contained in some orthogonal parallelopiped with edges parallel to the coordinate axes.

Proposition 9.21. *Let (X, d) be a metric space and $M \subseteq X$. If M is compact, then it is bounded and closed.*

Proof. We take any $x_0 \in X$ and we consider the collection

$$\Sigma = \{N_{x_0}(n) \mid n \in \mathbb{N}\}.$$

Then Σ is an open covering of M, and so there is a covering Σ' of M which is smaller than Σ, i.e. there are $n_1, \dots, n_N$ so that

$$M \subseteq N_{x_0}(n_1) \cup \cdots \cup N_{x_0}(n_N).$$

If $r = \max\{n_1, \dots, n_N\}$, then $M \subseteq N_{x_0}(r)$ and so M is bounded.
Now we take any $x_0 \in M^c$. We consider the sets

$$A_n = \{x \in X \mid d(x, x_0) > \tfrac{1}{n}\}$$

and the collection $\Sigma = \{A_n \mid n \in \mathbb{N}\}$. Then Σ is an open covering of M, and hence there is a finite covering Σ' of M which is smaller than Σ. I.e. there are $n_1, \dots, n_N$ so that

$$M \subseteq A_{n_1} \cup \cdots \cup A_{n_N}.$$

If $n = \max\{n_1, \dots, n_N\}$, then we have $M \subseteq A_n$ and hence $N_{x_0}(\frac{1}{n}) \subseteq M^c$. We proved that every $x_0 \in M^c$ is an interior point of M^c. Thus M^c is open and so M is closed. $\square$

Proposition 9.22. *Let (X, d) be a metric space and $N \subseteq M \subseteq X$. If M is compact and N is closed, then N is compact.*

Proof. We take any open covering Σ of N. Then

$$\Sigma_1 = \{N^c\} \cup \Sigma$$

is an open covering of M. Since M is compact, there is a finite covering Σ_1' of M which is smaller than Σ_1. I.e. there are $A_1, \dots, A_n \in \Sigma_1$ so that $M \subseteq A_1 \cup \cdots \cup A_n$.
If N^c is one of $A_1, \dots, A_n$, say $N^c = A_n$, then $N \subseteq A_1 \cup \cdots \cup A_{n-1}$ and so $\Sigma' = \{A_1, \dots, A_{n-1}\}$ is a finite covering of N which is smaller than Σ.
If N^c is not one of $A_1, \dots, A_n$, then $\Sigma' = \{A_1, \dots, A_n\}$ is a finite covering of N which is smaller than Σ.
In any case there is a finite covering of N which is smaller than Σ. $\square$

Proposition 9.23. *Let (X, d) be a metric space and $M, N \subseteq X$ so that $M \cap N = \emptyset$. If M is compact and N is closed, then there is $\epsilon > 0$ so that $d(x, y) \geq \epsilon$ for every $x \in M$ and $y \in N$.*

Proof. For every $x \in M$ we have $x \in N^c$ and, since N^c is open, there is $\epsilon_x > 0$ so that

$$N_x(\epsilon_x) \subseteq N^c$$

and hence $N_x(\epsilon_x) \cap N = \emptyset$. This implies

$$d(x,y) \geq \epsilon_x \tag{9.5}$$

for every $x \in M$ and $y \in N$. The collection $\{N_x(\frac{\epsilon_x}{2}) \,|\, x \in M\}$ is an open covering of M and, since M is compact, there are $x_1, \dots, x_n \in M$ so that

$$M \subseteq N_{x_1}(\tfrac{\epsilon_{x_1}}{2}) \cup \cdots \cup N_{x_n}(\tfrac{\epsilon_{x_n}}{2}).$$

We set

$$\epsilon = \min\left\{\tfrac{\epsilon_{x_1}}{2}, \dots, \tfrac{\epsilon_{x_n}}{2}\right\} > 0.$$

If $x \in M$, there is $k = 1, \dots, n$ so that $x \in N_{x_k}(\frac{\epsilon_{x_k}}{2})$ and (9.5) implies that for every $y \in N$ we have

$$d(x,y) \geq d(y,x_k) - d(x,x_k) \geq \epsilon_{x_k} - \tfrac{\epsilon_{x_k}}{2} = \tfrac{\epsilon_{x_k}}{2} \geq \epsilon.$$

Therefore, $d(x,y) \geq \epsilon$ for every $x \in M$ and $y \in N$. □

The next theorem is a generalization of the well known result for sequences of nested closed and bounded intervales in $\mathbb{R}$: *if* $[a_1,b_1] \supseteq [a_2,b_2] \supseteq \ldots \supseteq [a_n,b_n] \supseteq \ldots$, *then there is* x *which belongs to every* $[a_n,b_n]$ *and if, moreover,* $b_n - a_n \to 0$, *then this* x *is unique.*

Let (X,d) be a metric space and $M \subseteq X$. We define the **diameter** of M to be

$$\operatorname{diam} M = \sup\{d(x,y) \,|\, x,y \in M\}.$$

Theorem 9.1. *Let* (X,d) *be a metric space and* $K_1, K_2, \dots$ *be a sequence of non-empty compact subsets of* X *so that* $K_{n+1} \subseteq K_n$ *for every* n. *Then there is some element which belongs to all* K_n. *If, moreover,* diam $K_n \to 0$, *then the common element of all* K_n *is unique.*

Proof. We assume that

$$\textstyle\bigcap_{n=1}^{+\infty} K_n = \emptyset.$$

Then the collection $\Sigma = \{K_n^c \,|\, n \in \mathbb{N}\}$ is an open covering of K_1. Since K_1 is compact, there are $n_1, \dots, n_N$ so that

$$K_1 \subseteq K_{n_1}^c \cup \cdots \cup K_{n_N}^c.$$

We take $n = \max\{n_1, \dots, n_N\}$, and then $K_1 \subseteq K_n^c$. This is wrong, because $K_n \subseteq K_1$ and $K_n \neq \emptyset$.
Now, let diam $K_n \to 0$. If x, y belong to all K_n, then

$$0 \leq d(x,y) \leq \operatorname{diam} K_n$$

for every n and hence $d(x,y) = 0$. □

The important theorem 9.2 describes the notion of compactness in terms of sequences.

Theorem 9.2. *Let* (X,d) *be a metric space and* $M \subseteq X$. *The following are equivalent.*
(i) M *is compact.*
(ii) Every sequence in M *has at least one subsequence which converges to an element of* M.

Proof. (i) ⇒ (ii) We take an arbitrary sequence (x_n) in M.
Assume that for every $x \in M$ there is a neighborhood $N_x(\epsilon_x)$ of x, which contains only finitely many terms of (x_n). Then $\Sigma = \{N_x(\epsilon_x) \,|\, x \in M\}$ is an open covering of M and hence there are $y_1, \dots, y_N \in M$ so that

$$M \subseteq N_{y_1}(\epsilon_{y_1}) \cup \cdots \cup N_{y_N}(\epsilon_{y_N}).$$

Each of these neighborhoods contains only finitely many terms of (x_n). Therefore, M also contains only finitely many terms of (x_n) and we arrive at a contradiction.

Therefore there is $x_0 \in M$ so that for every $\epsilon > 0$ the neighborhood $N_{x_0}(\epsilon)$ contains infinitely many terms of (x_n). Thus, there is $n_1 \geq 1$ so that $x_{n_1} \in N_{x_0}(1)$. Then there is $n_2 > n_1$ so that $x_{n_2} \in N_{x_0}(\frac{1}{2})$. We continue inductively and we find a subsequence (x_{n_k}) of (x_n) so that

$$x_{n_k} \in N_{x_0}(\tfrac{1}{k})$$

or, equivalently, $d(x_{n_k}, x_0) < \frac{1}{k}$ for every k. Therefore $x_{n_k} \to x_0$.
(ii) $\Rightarrow$ (i) *Step 1*. Let $\epsilon > 0$. Then there are $x_1, \dots, x_n \in M$ so that

$$M \subseteq N_{x_1}(\epsilon) \cup \cdots \cup N_{x_n}(\epsilon).$$

Assume that this is not true. We take any $x_1 \in M$. Then $M \not\subseteq N_{x_1}(\epsilon)$ and so there is $x_2 \in M$ with $x_2 \notin N_{x_1}(\epsilon)$. Then $M \not\subseteq N_{x_1}(\epsilon) \cup N_{x_2}(\epsilon)$ and so there is $x_3 \in M$ with $x_3 \notin N_{x_1}(\epsilon) \cup N_{x_2}(\epsilon)$. Then $M \not\subseteq N_{x_1}(\epsilon) \cup N_{x_2}(\epsilon) \cup N_{x_3}(\epsilon)$ and so there is $x_4 \in M$ with $x_4 \notin N_{x_1}(\epsilon) \cup N_{x_2}(\epsilon) \cup N_{x_3}(\epsilon)$. We continue inductively and we see that there is a sequence (x_n) in M so that $d(x_n, x_m) \geq \epsilon$ for every n, m with $n \neq m$. But this does not allow the existence of a convergent subsequence of (x_n) and we arrive at a contradiction.
Step 2. We take any open covering Σ of M. Then there is $\epsilon > 0$ so that for every $x \in M$ the neighborhood $N_x(\epsilon)$ is contained in some $A \in \Sigma$.
Assume that there is no $\epsilon > 0$ with this property. I.e. for every $\epsilon > 0$ there is $x \in M$ so that $N_x(\epsilon)$ is not contained in any $A \in \Sigma$. Thus, for every $n \in \mathbb{N}$ there is $x_n \in M$ so that $N_{x_n}(\frac{1}{n})$ is not contained in any $A \in \Sigma$. Now, there is a subsequence (x_{n_k}) of (x_n) so that $x_{n_k} \to x_0$ for some $x_0 \in M$. Then $x_0 \in A_0$ for some $A_0 \in \Sigma$. Since A_0 is open, there is $\delta > 0$ so that $N_{x_0}(\delta) \subseteq A_0$. We take n_k large enough so that

$$d(x_{n_k}, x_0) < \tfrac{\delta}{2}, \quad \tfrac{1}{n_k} < \tfrac{\delta}{2}.$$

Then for every $x \in N_{x_{n_k}}(\frac{1}{n_k})$ we have

$$d(x, x_0) \leq d(x, x_{n_k}) + d(x_{n_k}, x_0) < \tfrac{1}{n_k} + d(x_{n_k}, x_0) < \tfrac{\delta}{2} + \tfrac{\delta}{2} = \delta$$

and hence $N_{x_{n_k}}(\frac{1}{n_k}) \subseteq N_{x_0}(\delta) \subseteq A_0$. We arrive at a contradiction, because $N_{x_{n_k}}(\frac{1}{n_k})$ is not contained in any $A \in \Sigma$.
Step 3. We take an arbitrary open covering Σ of M. According to step 2, there is $\epsilon > 0$ so that for every $x \in M$ we have that $N_x(\epsilon)$ is contained in some $A \in \Sigma$. According to step 1, there are $x_1, \dots, x_n \in M$ so that

$$M \subseteq N_{x_1}(\epsilon) \cup \ldots \cup N_{x_n}(\epsilon).$$

Now let $N_{x_k}(\epsilon) \subseteq A_k \in \Sigma$ for each $k = 1, \dots, n$. Then

$$M \subseteq N_{x_1}(\epsilon) \cup \cdots \cup N_{x_n}(\epsilon) \subseteq A_1 \cup \cdots \cup A_n$$

and hence $\Sigma' = \{A_1, \dots, A_n\}$ is a finite covering of M which is smaller than Σ. □

Proposition 9.24. *Every closed orthogonal parallelopiped in $\mathbb{R}^d$ with edges parallel to the coordinate axes is compact.*

Proof. Let $M = [a_1, b_1] \times \cdots \times [a_d, b_d]$. We consider an arbitrary open covering Σ of M and we assume that there is no finite covering Σ' of M which is smaller than Σ.
We split every edge $[a_k, b_k]$ in the two subintervals $[a_k, \frac{a_k+b_k}{2}]$ and $[\frac{a_k+b_k}{2}, b_k]$. This induces a splitting of M in 2^d orthogonal parallelopipeds, each of which has dimensions equal to one half of the dimensions of M. We observe that for at least one of these parallelopipeds, call it M_1, there is no finite covering which is smaller than Σ. Otherwise, for each of these parallelopipeds there would exist a finite covering which is smaller than Σ, and hence the (finite) union of these finite coverings would be a finite covering of M which is smaller than Σ. Similarly, we split M_1 in 2^d orthogonal

parallelopipeds for at least one of which, call it M_2, there is no finite covering smaller than Σ. We continue inductively and we end up with a sequence (M_l) of orthogonal parallelopipeds with the following properties:
(i) For every l there is no finite covering of M_l which is smaller than Σ.
(ii) $M \supseteq M_1 \supseteq \ldots \supseteq M_{l-1} \supseteq M_l \supseteq \ldots$. This means that, if

$$M_l = [a_{l,1}, b_{l,1}] \times \cdots \times [a_{l,d}, b_{l,d}],$$

then for every $k = 1, \ldots, d$ we have

$$a_k \leq a_{1,k} \leq \ldots \leq a_{l-1,k} \leq a_{l,k} \leq \ldots \leq b_{l,k} \leq b_{l-1,k} \leq \ldots \leq b_{1,k} \leq b_k.$$

(iii) For every $k = 1, \ldots, d$ and $l \geq 1$ we have $b_{l,k} - a_{l,k} = \frac{b_k - a_k}{2^l}$ and hence

$$b_{l,k} - a_{l,k} \to 0.$$

(iv) For every $l \geq 1$ we have diam $M_l = \frac{\text{diam } M}{2^l}$ and hence

$$\text{diam } M_l \to 0.$$

From (ii) we have that for every $k = 1, \ldots, d$ the sequence$(a_{l,k})$ is increasing and bounded above and that the sequence $(b_{l,k})$ is decreasing and bounded below and hence both sequences converge to two limits which, because of (iii), coincide. We set

$$x_k = \lim_{l \to +\infty} a_{l,k} = \lim_{l \to +\infty} b_{l,k}.$$

Then $\mathrm{x} = (x_1, \ldots, x_d)$ belongs to every M_l. Since Σ is a covering of M, there is some $A_0 \in \Sigma$ so that $\mathrm{x} \in A_0$. Now, A_0 is open and hence there is $\epsilon_0 > 0$ so that

$$N_{\mathrm{x}}(\epsilon_0) \subseteq A_0.$$

Now, (iv) implies that there is l_0 so that

$$\text{diam } M_{l_0} < \epsilon_0.$$

Then, since $\mathrm{x} \in M_{l_0}$, for every $\mathrm{y} \in M_{l_0}$ we have

$$|\mathrm{y} - \mathrm{x}| \leq \text{diam } M_{l_0} < \epsilon_0$$

and hence $\mathrm{y} \in N_{\mathrm{x}}(\epsilon_0)$. Thus

$$M_{l_0} \subseteq N_{\mathrm{x}}(\epsilon_0) \subseteq A_0$$

and so $\Sigma' = \{A_0\}$ is a finite covering of M_{l_0} which is smaller than Σ and we arrive at a contradiction with (i). $\square$

Bolzano-Weierstrass theorem. *Every bounded sequence in $\mathbb{R}^d$ has at least one convergent subsequence.*

Proof. If (x_n) is any bounded sequence in $\mathbb{R}^d$, then there is a closed orthogonal parallelopiped M with edges parallel to the coordinate axes so that (x_n) is in M. Now, M is compact and hence there is a subsequence of (x_n) which converges (to an element of M). $\square$

The next theorem is the most useful result for the determination of compact subsets of $\mathbb{R}^d$.

Theorem 9.3. *Let $M \subseteq \mathbb{R}^d$. Then M is compact if and only if it bounded and closed.*

First proof. Because of proposition 9.21, we have to prove only one direction.
Let M be closed and bounded. We take any (x_n) in M. Since M is bounded, (x_n) is also bounded and the Bolzano-Weierstrass theorem implies that there is a subsequence (x_{n_k}) so that $\mathrm{x}_{n_k} \to \mathrm{x}$ for some $\mathrm{x} \in \mathbb{R}^d$. Since M is closed and (x_{n_k}) is in M, we have that $\mathrm{x} \in M$. Hence every sequence in M has a subsequence which converges to an element of M and theorem 9.2 implies that M is compact.
Second proof. Again, proposition 9.21 proves one direction.
Since M is bounded, there is a closed orthogonal parallelopiped N with edges parallel to the coordinate axes so that $M \subseteq N$. Proposition 9.24 implies that N is compact and, since M is closed, proposition 9.22 implies that M is compact. $\square$

Example 9.4.3. Every closed ball is a compact subset of $\mathbb{R}^d$.

Theorem 9.3 says that the converse of proposition 9.21 is true in $\mathbb{R}^d$. This is not the case though in an arbitrary metric space.

Theorem 9.4. *The metric space $\mathbb{R}^d$ is complete.*

Proof. Let (x_n) be a Cauchy sequence in $\mathbb{R}^d$. Then we easily see that (x_n) is bounded. Indeed, there is n_0 so that $|\mathrm{x}_n - \mathrm{x}_m| < 1$ for every $n, m \geq n_0$. This implies that $|\mathrm{x}_n - \mathrm{x}_{n_0}| < 1$ for every $n \geq n_0$ and hence $|\mathrm{x}_n| \leq |\mathrm{x}_{n_0}| + 1$ for every $n \geq n_0$. Therefore,

$$|\mathrm{x}_n| \leq \max\{|\mathrm{x}_1|, \dots, |\mathrm{x}_{n_0-1}|, |\mathrm{x}_{n_0}| + 1\}$$

for every n. Then the Bolzano-Weierstrass theorem implies that there is a subsequence (x_{n_k}) so that $\mathrm{x}_{n_k} \to \mathrm{x}$ for some x. Now, we have that $|\mathrm{x}_k - \mathrm{x}_{n_k}| \to 0$, because (x_n) is a Cauchy sequence, and hence

$$|\mathrm{x}_k - \mathrm{x}| \leq |\mathrm{x}_k - \mathrm{x}_{n_k}| + |\mathrm{x}_{n_k} - \mathrm{x}| \to 0.$$

Therefore, $\mathrm{x}_k \to \mathrm{x}$. $\square$

Proposition 9.25. *Let (X, d) and (Y, ρ) be metric spaces, $M \subseteq X$ and $f : M \to Y$. If f is continuous in M and M is compact, then $f(M)$ is compact.*

Proof. Let T be an open covering of $f(M)$. Proposition 9.12 implies that for every $B \in T$ there is an open $A_B \subseteq X$ so that

$$f^{-1}(B) = A_B \cap M. \tag{9.6}$$

Since $f(M) \subseteq \bigcup_{B \in T} B$, we have

$$M \subseteq \bigcup_{B \in T} f^{-1}(B) \subseteq \bigcup_{B \in T} A_B,$$

i.e. the collection $\Sigma = \{A_B \mid B \in T\}$ is an open covering of M. Since M is compact, there are $B_1, \dots, B_n \in T$ so that

$$M \subseteq A_{B_1} \cup \dots \cup A_{B_n}.$$

This and (9.6) imply

$$M \subseteq (A_{B_1} \cup \dots \cup A_{B_n}) \cap M = (A_{B_1} \cap M) \cup \dots \cup (A_{B_n} \cap M) = f^{-1}(B_1) \cup \dots \cup f^{-1}(B_n),$$

and hence

$$f(M) \subseteq B_1 \cup \dots \cup B_n.$$

Therefore $\{B_1, \dots, B_n\}$ is a finite covering of $f(M)$ which is smaller than T. $\square$

Proposition 9.26. *Every non-empty compact subset of $\mathbb{R}$ has a maximal and a minimal element.*

Proof. Let $M \subseteq \mathbb{R}$ be non-empty and compact. Since M is non-empty and bounded, $u = \sup M$ is in $\mathbb{R}$. Then for every $\epsilon > 0$ there is $x \in M$ so that $u - \epsilon < x \leq u$. Therefore u is a limit point of M and, since M is closed, $u \in M$. So u is the maximal element of M.
The proof for the existence of a minimal element is similar. □

Proposition 9.27 generalizes the familiar analogous proposition for continuous $f : [a, b] \to \mathbb{R}$.

Proposition 9.27. *Let (X, d) be a metric space, $M \subseteq X$ and $f : M \to \mathbb{R}$. If f is continuous on M and M is compact, then f is bounded and has a maximum and a minimum value.*

Proof. Proposition 9.25 implies that $f(M) \subseteq \mathbb{R}$ is compact. Now proposition 9.26 says that $f(M)$ is bounded and has a maximal and a minimal element. □

Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$ and $f : A \to Y$. We say that f is **uniformly continuous** in A if for every $\epsilon > 0$ there is $\delta > 0$ so that $\rho(f(x'), f(x'')) < \epsilon$ for every $x', x'' \in A$ with $d(x', x'') < \delta$.

Theorem 9.5. *Let (X, d) and (Y, ρ) be metric spaces, $M \subseteq X$ and $f : M \to Y$. If f is continuous in M and M is compact, then f is uniformly continuous in M.*

Proof. Let $\epsilon > 0$. Since f is continuous in M, for every $x \in M$ there is $\delta_x > 0$ so that

$$\rho(f(y), f(x)) < \tfrac{\epsilon}{2} \tag{9.7}$$

for every $y \in M$ with $d(y, x) < \delta_x$.
The collection $\{N_x(\frac{\delta_x}{2}) \mid x \in M\}$ is an open covering of M and, since M is compact, there are $x_1, \dots, x_n \in M$ so that

$$M \subseteq N_{x_1}(\tfrac{\delta_{x_1}}{2}) \cup \cdots \cup N_{x_n}(\tfrac{\delta_{x_n}}{2}). \tag{9.8}$$

We define

$$\delta = \min\left\{\tfrac{\delta_{x_1}}{2}, \dots, \tfrac{\delta_{x_n}}{2}\right\} > 0$$

and we take any $x', x'' \in M$ with $d(x', x'') < \delta$. Because of (9.8), there is $k = 1, \dots, n$ so that $x' \in N_{x_k}(\frac{\delta_{x_k}}{2})$ and hence

$$d(x', x_k) < \tfrac{\delta_{x_k}}{2} < \delta_{x_k}.$$

This implies that

$$d(x'', x_k) \leq d(x'', x') + d(x', x_k) < \delta + \tfrac{\delta_{x_k}}{2} \leq \delta_{x_k}$$

and from (9.7) we have

$$\rho(f(x'), f(x'')) \leq \rho(f(x'), f(x_k)) + \rho(f(x''), f(x_k)) < \epsilon.$$

We proved that for every $x', x'' \in M$ with $d(x', x'') < \delta$ we have $\rho(f(x'), f(x'')) < \epsilon$. Therefore, f is uniformly continuous in M. □

Exercises.

9.4.1. Prove that $\{(x_1, x_2) \mid x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 1\}$ is a compact subset of $\mathbb{R}^2$ and that $\{(x_1, x_2, x_3) \mid {x_1}^2 + {x_2}^2 \leq x_3 \leq 1\}$ is a compact subset of $\mathbb{R}^3$.

9.4.2. (i) Consider the subset $A = \{(x_1, x_2) \mid x_1^2 + x_2^2 \leq 1\}$ of $\mathbb{R}^2$. Does the function $f(x_1, x_2) = e^{x_1 + x_2}$ have a maximum and a minimum value in A?
(ii) Consider the subset $A = \{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 + x_2 \leq 1, |x_3| \leq 2\}$ of $\mathbb{R}^3$. Does the function $f(x_1, x_2, x_3) = e^{x_1 + x_3} \sin(x_1 x_2)$ have a maximum and a minimum value in A?

9.4.3. (i) Let $f : \mathbb{R}^d \to \mathbb{R}$ such that $f(\mathrm{x}) \to 0$ when $|\mathrm{x}| \to +\infty$. This means, by definition, that for every $\epsilon > 0$ there is $R > 0$ so that $|f(\mathrm{x})| < \epsilon$ for every $\mathrm{x} \in \mathbb{R}^d$ with $|\mathrm{x}| > R$.
If there is $\mathrm{x}_0 \in \mathbb{R}^d$ so that $f(\mathrm{x}_0) \geq 0$, prove that f has a maximum value.
(ii) Prove that $f : \mathbb{R}^2 \to \mathbb{R}$ with $f(x_1, x_2) = x_1 e^{-x_1^2 - x_2^2}$ has a maximum and a minimum value and find them.

9.4.4. Let (X, d) be a metric space, $x \in X$, (x_n) be a sequence in X so that $x_n \neq x$ for every n and $x_n \to x$. Prove that $\{x_n \,|\, n \in \mathbb{N}\}$ is not compact and that $\{x\} \cup \{x_n \,|\, n \in \mathbb{N}\}$ is compact.

9.4.5. Let (X, d) be a metric space and $M_1, \ldots, M_n \subseteq X$. If $M_1, \ldots, M_n$ are compact, prove that $M_1 \cup \cdots \cup M_n$ is compact.

9.4.6. Let (X, d) be a metric space and $A, B \subseteq X$. If A is compact and B is closed, prove that $A \cap B$ is compact.

9.4.7. Let (X, d) be a metric space, $x_0 \in X$ and M, N be non-empty compact subsets of X.
(i) Prove that there are $x', y' \in M$ so that $d(x', y') = \operatorname{diam} M$.
(ii) Prove that there is $x' \in M$ so that $d(x_0, x') = \inf\{d(x_0, x) \,|\, x \in M\}$.
(iii) Prove that there are $x' \in M$ and $y' \in N$ so that $d(x', y') = \inf\{d(x, y) \,|\, x \in M, y \in N\}$.

9.4.8. Let $\mathrm{x}_0 \in \mathbb{R}^d$, $M \subseteq \mathbb{R}^d$ be non-empty and closed and $N \subseteq \mathbb{R}^d$ be non-empty and compact.
(i) Prove that there is $\mathrm{x}' \in M$ so that $|\mathrm{x}_0 - \mathrm{x}'| = \inf\{|\mathrm{x}_0 - \mathrm{x}| \,|\, \mathrm{x} \in M\}$.
(ii) Prove that there are $\mathrm{x}' \in M$ and $\mathrm{y}' \in N$ so that $|\mathrm{x}' - \mathrm{y}'| = \inf\{|\mathrm{x} - \mathrm{y}| \,|\, \mathrm{x} \in M, \mathrm{y} \in N\}$.

9.4.9. Let M be a bounded subset of $\mathbb{R}^d$. Prove that $\overline{M}$ and ∂M are compact.

9.4.10. Let (X, d) be a metric space and $M \subseteq X$. Prove that $\operatorname{diam} M = \operatorname{diam} \overline{M}$.

9.4.11. Let (X, d) be a metric space and $M \subseteq X$. We say that M is **totally bounded** if for every $\epsilon > 0$ there are $x_1, \ldots, x_n \in M$ so that $M \subseteq N_{x_1}(\epsilon) \cup \cdots \cup N_{x_n}(\epsilon)$.
Prove that $M \subseteq X$ is compact if and only if it is complete and totally bounded.

9.4.12. Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$ and $f : A \to Y$. Assume that $\overline{A}$ is compact, Y is complete and f is continuous in A. Prove that there is a continuous $F : \overline{A} \to Y$ so that $F = f$ in A if and only if f is uniformly continuous in A.

9.4.13. Let (X, d) be a non-empty set with the discrete metric (exercise 9.1.10). Prove that $M \subseteq X$ is compact if and only if it is a finite set.

9.5 Connectedness.

Let (X, d) be a metric space and $A \subseteq X$. We say that B, C form a **decomposition** of A if (i) $B \cup C = A$, (ii) $B \neq \emptyset$, $C \neq \emptyset$, (iii) none of B, C contains a limit point of the other.

It is clear that (iii) is equivalent to $\overline{B} \cap C = \emptyset$ and $B \cap \overline{C} = \emptyset$ taken together.

Example 9.5.1. In $\mathbb{R}^2$ we consider the closed discs $B = \overline{D}_{(0,0)}(1)$, $C = \overline{D}_{(3,0)}(1)$ and their union $A = B \cup C$. It is clear that B, C form a decomposition of A.
If we consider the open discs $B = D_{(0,0)}(1)$, $C = D_{(2,0)}(1)$ and $A = B \cup C$, then the discs B, C are tangent but, again, they form a decomposition of A.
If we take the closed disc $B = \overline{D}_{(0,0)}(1)$, the open disc $C = D_{(2,0)}(1)$ and $A = B \cup C$, then the discs B, C are tangent and they *do not* form a decomposition of A. Indeed, B contains the limit point $(1, 0)$ of C.

Let (X, d) be a metric space and $A \subseteq X$. We say that A is **connected** if there is no decomposition of A, i.e. there is no pair of sets B, C with the above mentioned properties (i)-(iii).

Example 9.5.2. The first two sets A of example 9.5.1 are not connected since each of them admits a specific decomposition. But we cannot decide at this moment if the third set A of example 9.5.1 is connected or not. We know that the specific B, C related to this A do not form a decomposition of A. *To decide that A is connected we must prove that, not only the specific pair, but an arbitrary pair does not form a decomposition of A.*

Example 9.5.3. It is obvious that $\emptyset$ as well as any $\{x\}$ is a connected set.

Lemma 9.1. *Let (X, d) be a metric space and $A, B, C \subseteq X$ with $B \cap C = \emptyset$ and assume that none of B, C contains a limit point of the other. If A is connected and $A \subseteq B \cup C$, then either $A \subseteq B$ or $A \subseteq C$.*

Proof. We define

$$B_1 = A \cap B, \quad C_1 = A \cap C.$$

Clearly, $B_1 \cup C_1 = A$ and $B_1 \cap C_1 = \emptyset$.
Now let $x \in B_1$. Then $x \in B$, and x is not a limit point of C. Then there is $r > 0$ so that $N_x(r) \cap C = \emptyset$ and, since $C_1 \subseteq C$, we get $N_x(r) \cap C_1 = \emptyset$. Thus x is not a limit point of C_1. We conclude that B_1 does not contain any limit point of C_1. Similarly, C_1 does not contain any limit point of B_1.
If $B_1 \neq \emptyset$ and $C_1 \neq \emptyset$, then B_1, C_1 form a decomposition of A and this contradicts the connectedness of A. Hence, either $B_1 = \emptyset$ or $C_1 = \emptyset$ and thus either $A \subseteq C$ or $A \subseteq B$, respectively □

Proposition 9.28. *Let (X, d) be a metric space and Σ be a collection of connected subsets of X all of which have a common point. Then $\bigcup_{A \in \Sigma} A$ is connected.*

Proof. We set

$$U = \bigcup_{A \in \Sigma} A$$

and we shall prove that U is connected. Let x_0 be the common point of all $A \in \Sigma$.
We assume that U is not connected. Then there are B, C which form a decomposition of U.
Since $x_0 \in U$, we have $x_0 \in B$ or $x_0 \in C$. Assume that $x_0 \in B$ (the proof is the same if $x_0 \in C$). For every $A \in \Sigma$ we have $A \subseteq U$ and hence $A \subseteq B \cup C$. According to lemma 9.1, every $A \in \Sigma$ is contained either in B or in C. But if any $A \in \Sigma$ is contained in C, it cannot contain x_0 which is in B. Therefore every $A \in \Sigma$ is contained in B and hence $U \subseteq B$. This implies that $C = \emptyset$ and we arrived at a contradiction. □

Proposition 9.29. *Let (X, d) be a metric space and $A, D \subseteq X$ so that $A \subseteq D \subseteq \overline{A}$. If A is connected, then D is connected.*

Proof. Let D not be connected. Then there are B, C which form a decomposition of D. Since $A \subseteq D$, we have $A \subseteq B \cup C$. Lemma 9.1 implies that $A \subseteq B$ or $A \subseteq C$. Let $A \subseteq B$. (The proof is similar if $A \subseteq C$.) Now, every point of D is a limit point of A and hence a limit point of B (since $A \subseteq B$). Therefore no point of D belongs to C (since C does not contain limit points of B) and this is wrong since $C \neq \emptyset$. □

Proposition 9.30. *Let (X, d), (Y, ρ) be metric spaces, $A \subseteq X$ and $f : A \to Y$. If f is continuous in A and A is connected, then $f(A)$ is connected.*

Proof. Assume that $f(A)$ is not connected. Then there are B', C' which form a decomposition of $f(A)$. We consider the inverse images of B', C' in A, i.e. the sets

$$B = f^{-1}(B') = \{x \in A \,|\, f(x) \in B'\}, \qquad C = f^{-1}(C') = \{x \in A \,|\, f(x) \in C'\}.$$

It is clear that $B \cup C = A$, $B \neq \emptyset$, $C \neq \emptyset$.
Now, let B contain a limit point b of C. Then there is a sequence (c_n) in C so that $c_n \to b$. Since

f is continuous at b, we get $f(c_n) \to f(b)$. The sequence $(f(c_n))$ is in C' and thus $f(b)$ is a limit point of C'. But $f(b) \in B'$ and we arrive at a contradiction, because B' does not contain any limit point of C'. Hence B does not contain any limit point of C. Similarly, C does not contain a limit point of B. Thus, B, C form a decomposition of A and this is wrong since A is connected. □

Let (X, d) be a metric space, $x, y \in X$ and $r > 0$. Every finite set $\{z_0, \ldots, z_n\} \subseteq X$ with $z_0 = x$, $z_n = y$ and $d(z_{k-1}, z_k) < r$ for every $k = 1, \ldots, n$ is called r-**succession of points** which joins x, y. If, moreover, $z_k \in A$ for every $k = 0, \ldots, n$, we say that the r-succession of points *is in* A.

Theorem 9.6. *Let (X, d) be a metric space and K be a compact subset of X. Then K is connected if and only if for every $x, y \in K$ and every $r > 0$ there is an r-succession of points in K which joins x, y.*

Proof. Assume K is connected. We take any $x, y \in K$ and any $r > 0$ and let there be no r-succession of points in K which joins x, y. We define the sets

$$B = \{b \in K \mid \text{there is an } r\text{-succession of points in } K \text{ which joins } x, b\},$$
$$C = \{c \in K \mid \text{there is no } r\text{-succession of points in } K \text{ which joins } x, c\}.$$

It is clear that $B \cup C = K$, $B \neq \emptyset$ (since $x \in B$) and $C \neq \emptyset$ (since $y \in C$).
Assume that B contains a limit point b of C. Then (since $b \in B$) there is an r-succession of points in K which joins x, b and, also, (since b is a limit point of C) there is $c \in C$ so that $d(b, c) < r$. If to the r-succession of points of K which joins x, b we attach c (as a final point after b), then we get an r-succession of points in K which joins x, c. This is wrong since $c \in C$. Hence B *does not contain any limit point of* C.
Now assume that C contains a limit point c of B. Then (since c is a limit point of B) there is $b \in B$ so that $d(b, c) < r$ and (since $b \in B$) there is an r-succession of points in K which joins x, b. If to the r-succession of points in K which joins x, b we attach c (as a final point after b), then we get an r-succession of points in K which joins x, c. This is wrong since $c \in C$. Hence C *does not contain any limit point of* B.
We conclude that B, C form a decomposition of K and this is wrong since K is connected.
Therefore there is an r-succession of points in K which joins x, y.
Conversely, assume that for every $x, y \in K$ and every $r > 0$ there is an r-succession of points in K which joins x, y.
We assume that K is not connected. Then there are B, C which form a decomposition of K.
Let x be a limit point of B. Since $B \subseteq K$, x is a limit point of K and, since K is closed, we get $x \in K$. Now, $x \notin C$ (because C does not contain any limit point of B) and we get that $x \in B$. Thus B contains all its limit points and it is closed. Finally, since $B \subseteq K$ and K is compact, B is also compact. Similarly C is also compact.
Now B, C are compact and disjoint and proposition 9.23 implies that there is $r > 0$ so that

$$d(b, c) \geq r$$

for every $b \in B$ and $c \in C$. Since $B \neq \emptyset$, $C \neq \emptyset$, we consider $b' \in B$ and $c' \in C$. Then it is easy to see that there is no r-succession of points in K which joins b', c', and we arrive at a contradiction. Indeed, assume that there is an r-succession $\{z_0, \ldots, z_n\}$ in K so that $z_0 = b'$, $z_n = c'$ and

$$d(z_{k-1}, z_k) < r$$

for every $k = 1, \ldots, n$. Since $z_0 \in B$, $z_n \in C$, it is clear that there is k so that $z_{k-1} \in B$, $z_k \in C$. Then $d(z_{k-1}, z_k) < r$ contradicts that we have $d(b, c) \geq r$ for every $b \in B$, $c \in C$. □

Proposition 9.31. *A set $I \subseteq \mathbb{R}$ is connected if and only if it is an interval.*

Proof. Let I be connected. If I is not an interval, there are $x_1, x_2 \in I$ and $x \notin I$ so that $x_1 < x < x_2$. Then the sets

$$B = I \cap (-\infty, x), \quad C = I \cap (x, +\infty)$$

form a decomposition of I and we have a contradiction. Thus I is an interval.
Conversely, let I be an interval. If I has only one element, then it is connected. If $I = [a, b]$ with $a < b$, then $[a, b]$ is compact and if we take any x, y in $[a, b]$ and any $r > 0$, it is clear that we can find an r-succession of points in $[a, b]$ which joins x and y. Thus $[a, b]$ is connected. If I is an interval of any other type, we can find a sequence of intervals $I_n = [a_n, b_n]$ which increase and their union is I. Then each I_n is connected and proposition 9.28 implies that I is also connected. □

Now we have the following corollary of propositions 9.30 and 9.31.

Proposition 9.32. *Let (X, d) be a metric space, $A \subseteq X$ and $f : A \to \mathbb{R}$ be continuous on A. If A is connected, then f has the intermediate value property in A.*

Proof. $f(A)$ is a connected subset of $\mathbb{R}$ and hence it is an interval. Not let u_1, u_2 be values of f in A, i.e. u_1, u_2 belong to the interval $f(A)$. Then every u with $u_1 < u < u_2$ also belongs to the interval $f(A)$. I.e. every number between the values u_1, u_2 of f in A is also a value of f in A. □

A special case of proposition 9.32 is the well known *intermediate value theorem* which says that if $f : I \to \mathbb{R}$ is continuous in the interval $I \subseteq R$, then it has the intermediate value property in I.

Let (X, d) be a metric space, $I \subseteq \mathbb{R}$ be an interval and $\gamma : I \to X$ be continuous on I. We say that γ is a **curve** in (X, d). The set

$$\gamma^* = \gamma(I) = \{\gamma(t) \,|\, t \in I\}$$

is called **trajectory** of the curve γ. If $\gamma^* \subseteq A \subseteq X$, we say that the curve γ is in A.

Propositions 9.30 and 9.31 imply that the trajectory of any curve in (X, d) is a connected subset of X. Also, if the interval I (the domain of definition of the curve) is closed and bounded (hence compact), then proposition 9.25 implies that the trajectory of the curve is a compact subset of X.

Example 9.5.4. Every linear segment $[\mathrm{x}, \mathrm{y}]$ in $\mathbb{R}^d$ is the trajectory of the curve $\gamma : [a, b] \to \mathbb{R}^d$ given by

$$\gamma(t) = \tfrac{b-t}{b-a}\,\mathrm{x} + \tfrac{t-a}{b-a}\,\mathrm{y}$$

for $a \leq t \leq b$.
A polygonal line consisting of two successive linear segments, i.e. $[\mathrm{x}, \mathrm{y}] \cup [\mathrm{y}, \mathrm{z}]$, is also the trajectory of a curve: we may take $a < b < c$ and the continuous $\gamma : [a, c] \to \mathbb{R}^d$ given by

$$\gamma(t) = \begin{cases} \frac{b-t}{b-a}\,\mathrm{x} + \frac{t-a}{b-a}\,\mathrm{y}, & \text{if } a \leq t \leq b \\ \frac{c-t}{c-b}\,\mathrm{y} + \frac{t-b}{c-b}\,\mathrm{z}, & \text{if } b \leq t \leq c \end{cases}$$

In a similar manner we may see that a general polygonal line consisting of n successive linear segments is the trajectory of a curve.

Let (X, d) be a metric space and $A \subseteq X$. We say that A is **arcwise connected** if for every two points of A there is a curve in A which joins these two points.

Proposition 9.33. *Let (X, d) be a metric space and $A \subseteq X$. If A is arcwise connected, then it is connected.*

Proof. We fix any $x_0 \in A$. For every $x \in A$ there is a curve γ_x in A which joins x_0 and x. Then $\gamma_x^* \subseteq A$ and hence

$$\bigcup_{x \in A} \gamma_x^* \subseteq A.$$

Conversely, since every $x \in A$ is contained in the trajectory γ_x^*, we have that

$$A \subseteq \bigcup_{x \in A} \gamma_x^*.$$

Therefore $A = \bigcup_{x \in A} \gamma_x^*$. Now, every γ_x^* is connected and since all γ_x^* have the point x_0 in common, we conclude that A is connected. □

Example 9.5.5. Every ring is a connected subset of $\mathbb{R}^2$.

Example 9.5.6. Every convex set $A \subseteq \mathbb{R}^d$ is arcwise connected and hence connected. Indeed if we take any two points in A the linear segment which joins them is contained in A. For instance, balls and orthogonal parallelopipeds are connected subsets of $\mathbb{R}^d$.

Example 9.5.7. A set $A \subseteq \mathbb{R}^d$ is called **star-shaped** if there is a specific point $\mathrm{x}_0 \in A$ so that for every $\mathrm{x} \in A$ the linear segment $[\mathrm{x}_0, \mathrm{x}]$ is contained in A. Every such x_0 is called *center* of the star-shaped set A. The center of the star-shaped set A may not be unique, but this does not mean that every point of A is a center of it.
It is clear that a star-shaped A is arcwise connected and hence connected. Indeed, every two points of A can be joined with a polygonal line in A consisting of two successive linear segments: one segment from one of the points to the center x_0 and the other segment from x_0 to the second point.

Example 9.5.8. The set $A = \overline{D}_{(0,0)}(1) \cup D_{(2,0)}(1)$ in example 9.5.1 is connected, since it is star-shaped with center 1.

Theorem 9.7. *Let A be an open subset of $\mathbb{R}^d$. Then A is connected if and only if it is arcwise connected.*

Proof. If A is arcwise connected, proposition 9.33 implies that it is connected.
Conversely, let A be connected. We take $\mathrm{x}, \mathrm{y} \in A$ and we assume that there is no polygonal line in A which joins x, y. We define the sets

$$B = \{\mathrm{b} \in A \,|\, \text{there is a polygonal line in } A \text{ which joins } \mathrm{x}, \mathrm{b}\},$$
$$C = \{\mathrm{c} \in A \,|\, \text{there is no polygonal line in } A \text{ which joins } \mathrm{x}, \mathrm{c}\}.$$

It is clear that $B \cup C = A$, $B \neq \emptyset$ (since $\mathrm{x} \in B$) and $C \neq \emptyset$ (since $\mathrm{y} \in C$).
We assume that B contains some limit point b of C. Then (since $\mathrm{b} \in B$) there is a polygonal line in A which joins x, b. Since A is open, there is $r > 0$ so that $N_{\mathrm{b}}(r) \subseteq A$ and (since b is a limit point of C) there is $\mathrm{c} \in N_{\mathrm{b}}(r) \cap C$. If to the polygonal line in A which joins x, b we attach (as last) the linear segment $[\mathrm{b}, \mathrm{c}]$ (which is contained in $N_{\mathrm{b}}(r)$ and hence in A), we get a polygonal line in A which joins x, c. This is wrong, since $\mathrm{c} \in C$. Thus B *does not contain any limit point of* C.
Now we assume that C contains a limit point c of B. Since A is open, there is $r > 0$ so that $N_{\mathrm{c}}(r) \subseteq A$. Then (since c is a limit point of B) there is $\mathrm{b} \in N_{\mathrm{c}}(r) \cap B$. As before, (since $\mathrm{b} \in B$) there is a polygonal line in A which joins x, b and, if to this we attach the linear segment $[\mathrm{b}, \mathrm{c}]$ (which is contained in $N_{\mathrm{c}}(r)$ and hence in A), we get a polygonal line in A which joins x, c. This is wrong, since $\mathrm{c} \in C$. Thus C *does not contain any limit point of* B.
We conclude that B, C form a decomposition of A and we arrive at a contradiction because A is connected.
Therefore, there is a polygonal line in A which joins x, y. □

Let (X, d) be a metric space and $A \subseteq X$. We say that $C \subseteq A$ is a **connected component** of A if C is connected and has the following property: if $C \subseteq C' \subseteq A$ and C' is connected, then $C = C'$. In other words, C is a connected component of A if *it is a connected subset of A and there is no strictly larger connected subset of A.*

Let us see a characteristic property of connected components. Let C be a connected component of A and let B be any connected subset of A so that $C \cap B \neq \emptyset$. Then $C \cup B$ is connected (being the union of connected sets with a common point) and $C \subseteq C \cup B \subseteq A$. Since C is a connected component of A, we get $C \cup B = C$ and hence $B \subseteq C$. In oher words, *a connected component of A swallows every connected subset of A intersecting it*.

Let C_1, C_2 be *distinct* connected components of A and assume that $C_1 \cap C_2 \neq \emptyset$. Since C_1 is a connected subset of A and intersects the connected component C_2 of A, we get $C_1 \subseteq C_2$. Symmetrically, $C_2 \subseteq C_1$ and hence $C_1 = C_2$. We arrive at a contradiction and we conclude that $C_1 \cap C_2 = \emptyset$. Thus, *different connected components of A are disjoint*.

Proposition 9.34. *Let (X, d) be a metric space and $A \subseteq X$. Then A is the union of its (mutually disjoint) connected components.*

Proof. We shall prove that every point of A belongs to a connected component of A. We take $x \in A$ and define C_x to be the union of all connected subsets B of A which contain x. (Such a set is $\{x\}$.) I.e.

$$C_x = \bigcup\{B \,|\, B \text{ is connected } \subseteq A \text{ and } x \in B\}.$$

Now C_x is a subset of A and contains x. It is also connected, since it is the union of connected sets B with x as a common point. If $C_x \subseteq C' \subseteq A$ and C' is connected, then C' is one of the connected subsets B of A which contain x and hence $C' \subseteq C_x$. Thus $C_x = C'$. Therefore C_x is a connected component of A and contains x. □

It is obvious that A is connected if and only if A is the only connected component of A.

Example 9.5.9. In $\mathbb{R}^2$ we consider the discs $B = D_{(0,0)}(1)$ and $C = D_{(3,0)}(1)$ and the set $A = B \cup C$. The discs B, C are connected subsets of A. Lemma 9.1 implies that any connected subset of A is contained either in B or in C. I.e. there is no connected subset of A strictly larger than either B or C. Therefore the discs B and C are the connected components of A.

Example 9.5.10. We take $\mathbb{Z} \subseteq \mathbb{R}$ and any $n \in \mathbb{Z}$.
Then $\{n\}$ is a connected set. Let $\{n\} \subseteq C' \subseteq \mathbb{Z}$ and $C' \neq \{n\}$. Then

$$C' = \{n\} \cup \big(C' \setminus \{n\}\big)$$

and it is clear that the sets $\{n\}$ and $C' \setminus \{n\}$ form a decomposition of C'. Thus C' is not connected and hence $\{n\}$ is a connected component of $\mathbb{Z}$.
Therefore $\mathbb{Z}$ has infinitely many connected components, each of them being a singleton.

Proposition 9.35. *Let (X, d) be a metric space and $A \subseteq X$. If A is closed, then every connected component of A is closed.*

Proof. Let C be a connected component of A. Since $C \subseteq A$ and A is closed, we get $C \subseteq \overline{C} \subseteq A$. Proposition 9.29 implies that $\overline{C}$ is connected and, since C is a connected component of A, we get that $C = \overline{C}$. Therefore C is closed. □

Proposition 9.36. *Let A be an open subset of $\mathbb{R}^d$. Every connected component of A is open.*

Proof. Let C be a connected component of A and $\mathrm{x} \in C$. Then $\mathrm{x} \in A$ and, since A is open, there is $r > 0$ so that $N_{\mathrm{x}}(r) \subseteq A$. Since $N_{\mathrm{x}}(r)$ is a connected subset of A and intersects the connected component C of A, we see that $N_{\mathrm{x}}(r) \subseteq C$. Thus x is an interior point of C. □

Propositions 9.34 and 9.36 imply that *every open subset of $\mathbb{R}^d$ is the union of disjoint open connected sets*.

Exercises.

9.5.1. Say which of the following subsets of $\mathbb{R}^2$ are connected and find their connected components.
(i) The complement of a circle.
(ii) The complement of a linear segment.
(iii) The complement of a closed triangular line.
Also:

$$\{(\tfrac{1}{n},0)|n\in\mathbb{N}\},\quad \{(0,0)\}\cup\{(\tfrac{1}{n},0)|n\in\mathbb{N}\},\quad [(0,0),(1,0)]\cup\textstyle\bigcup_{n=1}^{+\infty}[(0,\tfrac{1}{n}),(1,\tfrac{1}{n})],$$
$$\textstyle\bigcup_{n=1}^{+\infty}\{\mathrm{x}\in\mathbb{R}^2\,|\,|\mathrm{x}|=1+\tfrac{1}{n}\},\quad \{(x,y)\,|\,x,y\in\mathbb{Q}\}.$$

9.5.2. Prove that the following subsets of $\mathbb{R}^2$ are connected:

$$\{(x,\sin x)\,|\,x\in\mathbb{R}\}\quad,\{(x,\sin\tfrac{1}{x})\,|\,0<x\leq 1\},\quad \{(x,\sin\tfrac{1}{x})\,|\,0<x\leq 1\}\cup[(0,-1),(0,1)].$$

9.5.3. (i) Find a simple example of two connected sets in $\mathbb{R}^2$ whose intersection is not connected.
(ii) Find a simple example of a connected set A in $\mathbb{R}^2$ so that ∂A is not connected.
(iii) Find a simple example of a connected set A in $\mathbb{R}^2$ so that A° is not connected.

9.5.4. Let $d\geq 2$, $U\subseteq\mathbb{R}^d$ be a connected open set and $a_1,\dots,a_n\in U$. Prove that $U\setminus\{a_1,\dots,a_n\}$ is connected and open.

9.5.5. Consider a hyperplane L in $\mathbb{R}^d$ and the two open halfspaces of $\mathbb{R}^d$ which are determined by L. If a curve γ in $\mathbb{R}^d$ joins a point of one halfspace and a point of the other halfspace, prove that the trajectory of γ intersects L.

9.5.6. Let (X,d) be a metric space, $A_n\subseteq X$ be connected and $A_n\cap A_{n+1}\neq\emptyset$ for every n. Prove that $\bigcup_{n=1}^{+\infty}A_n$ is connected.

9.5.7. Let $B\subseteq\mathbb{R}^d$. If B is open and closed prove that either $B=\emptyset$ or $B=\mathbb{R}^d$.

9.5.8. Let (X,d) be a metric space and $A\subseteq X$.
(i) If A is closed, prove that A is connected if and only if there are no closed B,C so that $B\cup C=A$, $B\cap C=\emptyset$, $B\neq\emptyset$, $C\neq\emptyset$.
(ii) If A is open, prove that A is connected if and only if there are no open B,C so that $B\cup C=A$, $B\cap C=\emptyset$, $B\neq\emptyset$, $C\neq\emptyset$.

9.5.9. Let (X,d) be a metric space and $A\subseteq X$ be connected (not necessarily compact). Prove that for every $r>0$ and every $x,y\in A$ there is an r-succession of points in A which joins x,y.

9.5.10. Let (X,d) be a metric space and $A\subseteq X$. Prove that A is connected if and only if the only continuous functions $f:A\to\mathbb{R}$ with $f(A)\subseteq\mathbb{Z}$ are the constant functions.

9.5.11. Let $A\subseteq\mathbb{R}^d$ be open and connected and let every point of $B\subseteq A$ be an isolated point of B. Prove that $A\setminus B$ is connected.

9.5.12. Let (X,d) be a metric space.
(i) Let $A_n\subseteq X$ be compact so that $A_{n+1}\subseteq A_n$ for every $n\in\mathbb{N}$ and so that every two points of A_n can be joined by some $\frac{1}{n}$-succession of points in A_n. Prove that $\bigcap_{n=1}^{+\infty}A_n$ is connected.
(ii) Let $F\subseteq X$ be compact and let $x,y\in F$ belong to different connected components of F. Prove that there is a decomposition B,C of F so that $x\in B$ and $y\in C$.

9.5.13. Let (X,d) be a metric space. We say that (X,d) is **locally connected** if for every $x\in X$ and every $r>0$ there is an open connected U so that $x\in U\subseteq N_x(r)$.
Prove that (X,d) is locally connected if and only if for every open $A\subseteq X$ all the connected components of A are open.

9.5.14. Let (X,d) be a metric space. We say that (X,d) is **locally arcwise connected** if for every $x\in X$ and every $r>0$ there is an open arcwise connected U so that $x\in U\subseteq N_x(r)$.
If (X,d) is locally arcwise connected and $A\subseteq X$ is open, prove that A is connected if and only if it is arcwise connected.

9.5.15. Let (X,d) be a non-empty set with the discrete metric (exercise 9.1.10). Prove that $M\subseteq X$ is connected if and only if it has at most one element.

9.6 Uniform convergence.

In the following we consider only complex functions, although most of the results can be stated for functions taking values in the euclidean space $\mathbb{R}^d$ or even in a more general metric space.

Let A be any non-empty set and $B(A)$ be the set of all bounded functions $f : A \to \mathbb{C}$, i.e.

$$B(A) = \{ f \mid f : A \to \mathbb{C} \text{ is bounded} \}.$$

For each $f \in B(A)$ we define the **uniform norm** of f in A to be the non-negative real number

$$\|f\|_A = \sup\{|f(x)| \mid x \in A\} = \sup_{x\in A} |f(x)|.$$

Proposition 9.37. *Let $f, g \in B(A)$ and $\lambda \in \mathbb{C}$. Then:*
(i) $\|f\|_A = 0$ if and only if $f(x) = 0$ for every $x \in A$.
(ii) $\|f + g\|_A \leq \|f\|_A + \|g\|_A$.
(iii) $\|\lambda f\|_A = |\lambda|\|f\|_A$.
(iv) $\|fg\|_A \leq \|f\|_A\|g\|_A$.

Proof. (i) is obvious.
(ii) For every $x \in A$ we have

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_A + \|g\|_A$$

and hence $\|f + g\|_A \leq \|f\|_A + \|g\|_A$.
(iii) If $\lambda = 0$, then according to (i), both sides of $\|\lambda f\|_A = |\lambda|\|f\|_A$ are equal to 0. Now let $\lambda \neq 0$. For every $x \in A$ we have

$$|\lambda f(x)| = |\lambda||f(x)| \leq |\lambda|\|f\|_A.$$

Hence $\|\lambda f\|_A \leq |\lambda|\|f\|_A$. Applying this to λf and $\frac{1}{\lambda}$, we get the opposite inequality.
(iv) For every $x \in A$ we have

$$|f(x)g(x)| = |f(x)||g(x)| \leq \|f\|_A\|g\|_A.$$

Thus, $\|fg\|_A \leq \|f\|_A\|g\|_A$. □

For every $f, g \in B(A)$ we define their **uniform distance** in A to be the non-negative real number

$$\|f - g\|_A = \sup\{|f(x) - g(x)| \mid x \in A\} = \sup_{x\in A} |f(x) - g(x)|.$$

The function $d_A : B(A) \times B(A) \to \mathbb{R}$ given by

$$d_A(f, g) = \|f - g\|_A$$

is called **uniform metric** in A or **metric of uniform convergence** in A.

Proposition 9.38 justifies the term “metric” we used for the function d_A.

Proposition 9.38. *The function $d_A : B(A) \times B(A) \to \mathbb{R}$ is a metric on $B(A)$.*

Proof. We check the four basic properties of a metric.
(i) $d_A(f, g) = \|f - g\|_A \geq 0$ is obvious.
(ii) If $d_A(f, g) = \|f - g\|_A = 0$ then proposition 9.37 implies that $f(x) = g(x)$ for every $x \in A$ and hence $f = g$.
(iii) $d_A(f, g) = \|f - g\|_A = \|g - f\|_A = d_A(g, f)$.
(iv) We have

$$d_A(f, g) = \|f - g\|_A = \|(f - h) + (h - g)\|_A \leq \|f - h\|_A + \|h - g\|_A = d_A(f, h) + d_A(h, g)$$

from proposition 9.37. □

Thus $B(A)$, equipped with the uniform metric d_A, becomes a metric space. It is called **metric space of uniform convergence** in A.

Let (f_n) be a sequence in $B(A)$ and $f \in B(A)$. We say that f_n **converges to** f **uniformly** in A if $f_n \to f$ with respect to the metric of uniform convergence in A, i.e. if

$$d_A(f_n, f) = \|f_n - f\|_A \to 0.$$

Hence f_n converges to f uniformly in A if for every $\epsilon > 0$ there is n_0 so that $\|f_n - f\|_A \leq \epsilon$ for every $n \geq 0$ or, equivalently, if for every $\epsilon > 0$ there is n_0 so that $|f_n(x) - f(x)| \leq \epsilon$ for every $x \in A$ and every $n \geq 0$.

If $f_n \to f$ uniformly in A, then for every $x \in A$ we have

$$|f_n(x) - f(x)| \leq \|f_n - f\|_A \to 0$$

and hence $f_n(x) \to f(x)$ for every $x \in A$. Therefore, uniform convergence of (f_n) to f in A implies **pointwise convergence** of (f_n) to f in A.

Proposition 9.39. *Let $\lambda, \mu \in \mathbb{C}$ and $f_n \to f$ and $g_n \to g$ uniformly in A.*
(i) If $B \subseteq A$, then $f_n \to f$ uniformly in B.
(ii) $\lambda f_n + \mu g_n \to \lambda f + \mu g$ and $f_n g_n \to fg$ uniformly in A.
(iii) If $\frac{1}{g} \in B(A)$, then $\frac{1}{g_n} \in B(A)$ after some value of the index n and $\frac{1}{g_n} \to \frac{1}{g}$ uniformly in A.

Proof. (i) We have

$$\|f_n - f\|_B = \sup_{x \in B} |f_n(x) - f(x)| \leq \sup_{x \in A} |f_n(x) - f(x)| = \|f_n - f\|_A.$$

Therefore, $\|f_n - f\|_A \to 0$ implies $\|f_n - f\|_B \to 0$.
(ii) From proposition 9.37 we have

$$\|(\lambda f_n + \mu g_n) - (\lambda f + \mu g)\|_A \leq |\lambda| \|f_n - f\|_A + |\mu| \|g_n - g\|_A$$

and hence $\|(\lambda f_n + \mu g_n) - (\lambda f + \mu g)\|_A \to 0$.
Also,

$$\|f_n g_n - fg\|_A \leq \|f_n - f\|_A \|g_n - g\|_A + \|f\|_A \|g_n - g\|_A + \|g\|_A \|f_n - f\|_A$$

and so $\|f_n g_n - fg\|_A \to 0$.
(iii) We have $\frac{1}{|g(x)|} \leq M$, and hence $|g(x)| \geq \frac{1}{M}$, for every $x \in A$. Then there is n_0 so that $\|g_n - g\|_A \leq \frac{1}{2M}$ for every $n \geq n_0$. Hence

$$|g_n(x)| \geq |g(x)| - |g_n(x) - g(x)| \geq \tfrac{1}{M} - \tfrac{1}{2M} = \tfrac{1}{2M}$$

for every $x \in A$ and every $n \geq n_0$. Therefore, $\frac{1}{|g_n(x)|} \leq 2M$ for every $x \in A$ and every $n \geq n_0$. This implies that $\frac{1}{g_n} \in B(A)$ for every $n \geq n_0$.
Moreover,

$$\Big| \tfrac{1}{g_n(x)} - \tfrac{1}{g(x)} \Big| = \tfrac{|g_n(x) - g(x)|}{|g_n(x)||g(x)|} \leq 2M^2 \|g_n - g\|_A$$

for every $x \in A$ and every $n \geq n_0$. Thus, $\big\| \frac{1}{g_n} - \frac{1}{g} \big\|_A \leq 2M^2 \|g_n - g\|_A$ for every $n \geq n_0$ and hence $\big\| \frac{1}{g_n} - \frac{1}{g} \big\|_A \to 0$. □

Theorem 9.8. *The metric space $B(A)$ with the metric of uniform convergence in A is complete.*

Proof. Let (f_n) be a Cauchy sequence in $B(A)$ with the metric of uniform convergence in A. This means that for every $\epsilon > 0$ there is n_0 so that $\|f_n - f_m\|_A \leq \epsilon$ for every $n, m \geq n_0$. In other words, we have that for every $\epsilon > 0$ there is n_0 so that

$$|f_n(x) - f_m(x)| \leq \epsilon \tag{9.9}$$

for every $x \in A$ and every $n, m \geq n_0$. Now, (9.9) implies that for every fixed $x \in A$ the sequence $(f_n(x))$ is a Cauchy sequence in $\mathbb{C}$ and hence it converges to some number. We define

$$f(x) = \lim_{n \to +\infty} f_n(x)$$

for every $x \in A$, and we get a function $f : A \to \mathbb{R}$. Taking the limit as $m \to +\infty$ in (9.9), we conclude that for every $\epsilon > 0$ there is n_0 so that

$$|f_n(x) - f(x)| \leq \epsilon \tag{9.10}$$

for every $x \in A$ and every $n \geq n_0$. Now we see that f is bounded in A, i.e. that $f \in B(A)$. Indeed, (9.10) with $n = n_0$ implies that

$$|f(x)| \leq |f_{n_0}(x) - f(x)| + |f_{n_0}(x)| \leq \epsilon + \|f_{n_0}\|_A$$

for every $x \in A$. Moreover, (9.10) says that for every $\epsilon > 0$ there is n_0 so that $\|f_n - f\|_A \leq \epsilon$ for every $n \geq n_0$. Therefore, (f_n) converges to f in the metric space $B(A)$. ☐

Proposition 9.40. *Let (X, d) be a metric space, $A \subseteq X$ and $f, f_n \in B(A)$ for every $n \in \mathbb{N}$. Let $f_n \to f$ uniformly in A and let $x \in A$. If every f_n is continuous at x, then f is continuous at x. In particular, if every f_n is continuous in A, then f is continuous in A.*

Proof. Take any $\epsilon > 0$. Then there is n_0 so that $\|f_n - f\|_A < \frac{\epsilon}{3}$ for every $n \geq n_0$ and hence

$$\|f_{n_0} - f\|_A < \tfrac{\epsilon}{3}.$$

Since f_{n_0} is continuous at x, there is $\delta > 0$ so that

$$|f_{n_0}(y) - f_{n_0}(x)| < \tfrac{\epsilon}{3}$$

for every $y \in A$ with $|y - x| < \delta$. So for every $y \in A$ with $|y - x| < \delta$ we have

$$\begin{aligned} |f(y) - f(x)| &\leq |f(y) - f_{n_0}(y)| + |f_{n_0}(y) - f_{n_0}(x)| + |f_{n_0}(x) - f(x)| \\ &\leq \|f_{n_0} - f\|_A + |f_{n_0}(y) - f_{n_0}(x)| + \|f_{n_0} - f\|_A < \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} = \epsilon \end{aligned}$$

and f is continuous at x. ☐

Let (X, d) be a metric space and $A \subseteq X$. We denote $BC(A)$ the set of all bounded and continuous functions $f : A \to \mathbb{C}$, i.e.

$$BC(A) = \{f \mid f : A \to \mathbb{C} \text{ is bounded and continuous}\}.$$

We denote $C(A)$ the set of all continuous functions $f : A \to \mathbb{C}$, i.e.

$$C(A) = \{f \mid f : A \to \mathbb{C} \text{ is continuous}\}.$$

It is obvious that $BC(A) \subseteq B(A)$. If A is a compact subset of X, then every continuous function $f : A \to \mathbb{C}$ is bounded and so, in this case, we have $BC(A) = C(A)$.

Proposition 9.41. *Let (X, d) be a metric space and $A \subseteq X$. The set $BC(A)$ is closed in $B(A)$ with respect to the uniform metric.*

Proof. This is a corollary of proposition 9.40. ☐

Theorem 9.9. *Let (X, d) be a metric space and $A \subseteq X$. The subset $BC(A)$ of $B(A)$ with the metric of uniform convergence in A is complete.*

Proof. Let (f_n) be a Cauchy sequence in $BC(A)$. Theorem 9.8 implies that there is $f \in B(A)$ so that $f_n \to f$ uniformly in A. Proposition 9.40 implies that f is continuous in A and hence $f \in BC(A)$. □

From the notion of uniform convergence of a sequence of functions we move to the notion of uniform convergence of a series of functions (through the sequence of partial sums).

Let $f_n : A \to \mathbb{C}$ for every n. We consider the partial sums $s_n : A \to \mathbb{C}$, where $s_n(x) = f_1(x) + \cdots + f_n(x)$ for every $x \in A$. Let also $s : A \to \mathbb{C}$. We say that the series of functions $\sum_{n=1}^{+\infty} f_n$ **converges** to its **sum** s **uniformly** in A if the sequence of functions (s_n) converges to the function s uniformly in A.

As in the case of a sequence of functions, we have that, if $\sum_{n=1}^{+\infty} f_n$ converges to its sum s uniformly in A, then $\sum_{n=1}^{+\infty} f_n(x) = s(x)$ for every $x \in A$, i.e. $\sum_{n=1}^{+\infty} f_n$ converges to its sum s **pointwise** in A.

Proposition 9.42. *Let (X, d) be a metric space, $A \subseteq X$ and $f_n \in B(A)$ for every $n \in \mathbb{N}$. Let $\sum_{n=1}^{+\infty} f_n$ converge to its sum s uniformly in A and let $x \in A$. If every f_n is continuous at x, then s is continuous at x. In particular, if every f_n is continuous in A, then s is continuous in A.*

Proof. We consider the partial sums $s_n = f_1 + \cdots + f_n$. Then every s_n is continuous at x and proposition 9.40 implies that s is continuous at x. □

Finally, we have a basic criterion for uniform convergence of a series of functions.

Weierstrass test. *Let $|f_n(x)| \leq M_n$ for every n and every $x \in A$. If the series (of non-negative terms) $\sum_{n=1}^{+\infty} M_n$ converges, i.e. if $\sum_{n=1}^{+\infty} M_n < +\infty$, then $\sum_{n=1}^{+\infty} f_n$ converges uniformly in A.*

Proof. For every $x \in A$ we have

$$\sum_{n=1}^{+\infty} |f_n(x)| \leq \sum_{n=1}^{+\infty} M_n < +\infty$$

and so $\sum_{n=1}^{+\infty} f_n(x)$ converges (as a series of complex numbers). Therefore, we may define the function $s : A \to \mathbb{C}$ with

$$s(x) = \sum_{n=1}^{+\infty} f_n(x)$$

for every $x \in A$. Now we consider the partial sums $s_n = f_1 + \cdots + f_n$ and then for every $x \in A$ we have

$$|s_n(x) - s(x)| = \Big| \sum_{k=n+1}^{+\infty} f_k(x) \Big| \leq \sum_{k=n+1}^{+\infty} |f_k(x)| \leq \sum_{k=n+1}^{+\infty} M_k.$$

Since this is true for every $x \in A$, we get

$$\|s_n - s\|_A \leq \sum_{k=n+1}^{+\infty} M_k \to 0$$

when $n \to +\infty$, because $\sum_{n=1}^{+\infty} M_n < +\infty$. Therefore, (s_n) converges to s uniformly in A and hence $\sum_{n=1}^{+\infty} f_n$ converges to its sum s uniformly in A. □

Let (X, d) be a metric space, $A \subseteq X$ and $\mathcal{F}$ be a family of complex functions defined in A. We say that $\mathcal{F}$ is **bounded** at some $x \in A$ if there is M so that $|f(x)| \leq M$ for every $f \in \mathcal{F}$. We say that $\mathcal{F}$ is **equicontinuous** at some $x \in A$ if for every $\epsilon > 0$ there is $\delta > 0$ so that $|f(y) - f(x)| < \epsilon$ for every $y \in A$ with $d(y, x) < \delta$ and for every $f \in \mathcal{F}$.

It is obvious that equicontinuity of $\mathcal{F}$ at x implies continuity of every $f \in \mathcal{F}$ at x. In fact the δ which corresponds to ϵ in the definition of continuity at x does not depend on the particular f: it is uniform over $f \in \mathcal{F}$.

Proposition 9.43. *Let (X, d) be a metric space, $A \subseteq X$ and (f_n) be a sequence of continuous functions in A. If $f_n \to f$ uniformly in every compact subset of A, then f is continuous in A.*

Proof. Take any $x \in A$ and a sequence (x_m) in A with $x_m \to x$. Then

$$K = \{x_m \mid m \in \mathbb{N}\} \cup \{x\}$$

is a compact subset of A and hence $f_n \to f$ uniformly in K. Since every f_n is continuous in K, we have that f is also continuous in K. Thus, $f(x_m) \to f(x)$ and so f is continuous at x. □

Let (X, d) be a metric space and $A \subseteq X$. We say that $B \subseteq A$ is **dense** in A if $A \subseteq \overline{B}$, i.e. if for every $x \in A$ and every $r > 0$ there is $y \in B$ so that $d(y, x) < r$. We say that A is **separable** if there is a countable $B \subseteq A$ which is dense in A.

The theorem of Arzela-Ascoli. *Let (X, d) be a metric space, let $A \subseteq X$ be separable, and let $\mathcal{F}$ be a collection of continuous functions in A. Then the following are equivalent:*
(i) For every sequence (f_n) in $\mathcal{F}$ there is a subsequence (f_{n_k}) and a function f continuous in A so that $f_{n_k} \to f$ uniformly in every compact subset of A.
(ii) $\mathcal{F}$ is equicontinuous and bounded at every $x \in A$.

Proof. (i) ⇒ (ii) Assume that $\mathcal{F}$ is not bounded at some $x \in A$. Then there is a sequence (f_n) in $\mathcal{F}$ so that $|f_n(x)| \to +\infty$. Now, there is a subsequence (f_{n_k}) and a function f so that $f_{n_k} \to f$ uniformly in every compact subset of A. One such compact set is $\{x\}$ and we get $f_{n_k}(x) \to f(x)$, arriving at a contradiction.
Now assume that $\mathcal{F}$ is not equicontinuous at some $x \in A$. Then there is $\epsilon > 0$ so that for every $n \in \mathbb{N}$ there is $x_n \in A$ and $f_n \in \mathcal{F}$ so that

$$d(x_n, x) < \tfrac{1}{n}, \quad |f_n(x_n) - f_n(x)| \geq \epsilon.$$

Now there is a subsequence (f_{n_k}) of (f_n) and a function f so that $f_{n_k} \to f$ uniformly in every compact subset of A. Proposition 9.43 implies that f is continuous at x. Since $x_{n_k} \to x$, the set $K = \{x_{n_k} \mid k \in \mathbb{N}\} \cup \{x\}$ is a compact subset of A and so $f_{n_k} \to f$ uniformly in K. Now

$$\begin{aligned}\epsilon \leq |f_{n_k}(x_{n_k}) - f_{n_k}(x)| &\leq |f_{n_k}(x_{n_k}) - f(x_{n_k})| + |f(x_{n_k}) - f(x)| + |f(x) - f_{n_k}(x)| \\ &\leq \|f_{n_k} - f\|_K + |f(x_{n_k}) - f(x)| + \|f_{n_k} - f\|_K.\end{aligned}$$

for every k. We arrive at a contradiction because $\|f_{n_k} - f\|_K \to 0$ and $f(x_{n_k}) \to f(x)$.
(ii) ⇒ (i) Let (f_n) be a sequence in $\mathcal{F}$. We know that there is a countable $B \subseteq A$ which is dense in A. Let

$$B = \{y_m \mid m \in \mathbb{N}\}.$$

The set $\{f_n(y_1) \mid n \in \mathbb{N}\} \subseteq \mathbb{C}$ is bounded. So there is a subsequence $(f_{1,n})$ of (f_n) such that $(f_{1,n}(y_1))$ is a Cauchy sequence in $\mathbb{C}$. Similarly, the set $\{f_{1,n}(y_2) \mid n \in \mathbb{N}\} \subseteq \mathbb{C}$ is bounded. So there is a subsequence $(f_{2,n})$ of $(f_{1,n})$ such that $(f_{2,n}(y_2))$ is a Cauchy sequence in $\mathbb{C}$. Similarly, the set $\{f_{2,n}(y_3) \mid n \in \mathbb{N}\} \subseteq \mathbb{C}$ is bounded. So there is a subsequence $(f_{3,n})$ of $(f_{2,n})$ so that $(f_{3,n}(y_3))$ is a Cauchy sequence in $\mathbb{C}$. We continue inductively and we find

$$\begin{matrix} f_{1,1} & f_{1,2} & f_{1,3} & \cdots & f_{1,n} & \cdots \\ f_{2,1} & f_{2,2} & f_{2,3} & \cdots & f_{2,n} & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \\ f_{m,1} & f_{m,2} & f_{m,3} & \cdots & f_{m,n} & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \end{matrix}$$

so that (a) the sequence in every row is a subsequence of the sequence in the previous row and hence of the original sequence (f_n) and (b) $(f_{m,n}(y_m))$ is a Cauchy sequence in $\mathbb{C}$ for every m. Now we consider the *diagonal sequence* $(f_{n,n})$. For every m, $(f_{n,n})$ is, after the value m of the index n, a subsequence of $(f_{m,n})$ and hence $(f_{n,n}(y_m))$ is a Cauchy sequence in $\mathbb{C}$. Also, $(f_{n,n})$

is a subsequence of (f_n).
Now we take any compact $K \subseteq A$ and any $\epsilon > 0$. We know that $\mathcal{F}$ is equicontinuous at every x. So for every $x \in K$ there is $\delta_x > 0$ so that

$$|f(t) - f(x)| < \tfrac{\epsilon}{5} \tag{9.11}$$

for every $t \in A$ with $d(t,x) < \delta_x$ and every $f \in \mathcal{F}$. Since K is compact, there are $x_1, \ldots, x_N \in K$ so that

$$K \subseteq \textstyle\bigcup_{k=1}^N N_{x_k}(\delta_{x_k}).$$

Since B is dense in A, for every $k = 1, \ldots, N$ there is some

$$y_{m_k} \in B \cap N_{x_k}(\delta_{x_k}).$$

Since $(f_{n,n}(y_{m_k}))$ is a Cauchy sequence in $\mathbb{C}$ for every $k = 1, \ldots, N$, there is n_0 so that

$$|f_{n',n'}(y_{m_k}) - f_{n'',n''}(y_{m_k})| < \tfrac{\epsilon}{5} \tag{9.12}$$

for every $n', n'' \geq n_0$ and every $k = 1, \ldots, N$. Now we take any $x \in K$. Then there is some $k = 1, \ldots, N$ so that $x \in N_{x_k}(\delta_{x_k})$. Then

$$d(x, x_k) < \delta_{x_k}, \quad d(y_{m_k}, x_k) < \delta_{x_k}. \tag{9.13}$$

Now, (9.11) (for $x = x_k$), (9.12) and (9.13) imply that for every $n', n'' \geq n_0$ we have

$$\begin{aligned}
|f_{n',n'}(x) - f_{n'',n''}(x)| &\leq |f_{n',n'}(x) - f_{n',n'}(x_k)| + |f_{n',n'}(x_k) - f_{n',n'}(y_{m_k})| \\
&\quad + |f_{n',n'}(y_{m_k}) - f_{n'',n''}(y_{m_k})| + |f_{n'',n''}(y_{m_k}) - f_{n'',n''}(x_k)| \\
&\quad + |f_{n'',n''}(x_k) - f_{n'',n''}(x)| \\
&< \tfrac{\epsilon}{5} + \tfrac{\epsilon}{5} + \tfrac{\epsilon}{5} + \tfrac{\epsilon}{5} + \tfrac{\epsilon}{5} = \epsilon.
\end{aligned}$$

Thus, for every $n', n'' \geq n_0$ we have

$$\|f_{n',n'} - f_{n'',n''}\|_K < \epsilon$$

and so $(f_{n,n})$ is a Cauchy sequence in $C(K)$ for every compact $K \subseteq A$. Now, theorem 9.9, applied to K, implies that $(f_{n,n})$ converges uniformly in K to some function continuous in K (the limit function depends on K), for every compact $K \subseteq A$. In particular, if $K = \{x\}$, we get that $(f_{n,n}(x))$ converges to some number in $\mathbb{C}$ and then we define

$$f(x) = \lim_{n\to+\infty} f_{n,n}(x)$$

for every $x \in A$. Since $(f_{n,n})$ converges uniformly in K to some function, say f_K, it converges also pointwise in K to f_K. But since $(f_{n,n})$ converges pointwise in A to f, we get that $f_K = f$ in K, and we proved that $(f_{n,n})$ converges to f uniformly in every compact subset of A.
Finally, proposition 9.43 says that f is continuous in A. □

Exercises.

9.6.1. Let (X,d) be a metric space and $A \subseteq X$.
(i) We define the collection of functions

$$B_c(A) = \{f \mid f : A \to \mathbb{C}, f \text{ is bounded in every compact } K \subseteq A\}.$$

Prove that $C(A) \subseteq B_c(A)$.
(ii) We say that $\mathcal{K} = (K_k)$ is an **exhausting sequence of compact sets** for A if every K_k is compact, $K_k \subseteq K_{k+1} \subseteq A$ for every k and, finally, for every compact $K \subseteq A$ there is some k so

that $K \subseteq K_k$.
If (K_k) is an exhausting sequence of compact sets for A, prove that $\bigcup_{k=1}^{+\infty} K_k = A$.
(iii) Let $F \subseteq \mathbb{R}^d$ be closed. Prove that there is an exhausting sequence of compact sets for F.
(iv) Let $\Omega \subseteq \mathbb{R}^d$ be open, $\delta > 0, R > 0$ and

$$K = \{\mathrm{x} \in \Omega \,|\, |\mathrm{x}| \leq R, |\mathrm{x} - \mathrm{y}| \geq \delta \text{ for every } \mathrm{y} \in \Omega^c\}.$$

Prove that K is a compact subset of Ω.
Prove that there is an exhausting sequence of compact sets for Ω.
(v) If $\mathcal{K} = (K_k)$ is an exhausting sequence of compact sets for A, then we define the function $d_{A,\mathcal{K}} : B_c(A) \times B_c(A) \to \mathbb{R}$ by

$$d_{A,\mathcal{K}}(f, g) = \sum_{k=1}^{+\infty} \frac{1}{2^k} \frac{\|f-g\|_{K_k}}{1+\|f-g\|_{K_k}}$$

Prove that $d_{A,\mathcal{K}}$ is a metric on $B_c(A)$.
If $f, f_n \in B_c(A)$ for every $n \in \mathbb{N}$, prove that $d_{A,\mathcal{K}}(f_n, f) \to 0$ if and only if $f_n \to f$ uniformly in every compact subset of A. Because of this result, the metric $d_{A,\mathcal{K}}$ is called **metric of uniform convergence in the compact subsets** of A.
If $\mathcal{K}' = (K_n')$ and $\mathcal{K}'' = (K_n'')$ are two exhausting sequences of compact sets for A, prove that the corresponding metrics $d_{A,\mathcal{K}'}$ and $d_{A,\mathcal{K}''}$ on $B_c(A)$ are equivalent.
If we consider $B_c(A)$ as a metric space with the metric $d_{A,\mathcal{K}}$ of uniform convergence in the compact subsets of A, prove that $C(A)$ is a closed subset of $B_c(A)$.

9.6.2. Let (X, d) be a metric space, $A \subseteq X$ and $\mathcal{F}$ be a family of complex functions defined in A. We say that $\mathcal{F}$ is **equicontinuous** in A if for every $\epsilon > 0$ there is $\delta > 0$ so that $|f(x') - f(x'')| < \epsilon$ for every $x', x'' \in A$ with $d(x', x'') < \delta$ and for every $f \in \mathcal{F}$.
If A is compact and $\mathcal{F}$ is equicontinuous at every $x \in A$, prove that $\mathcal{F}$ is equicontinuous in A.

9.6.3. Let (X, d) be a metric space. If $A \subseteq X$ is a countable union of compact sets, prove that A is separable.

9.6.4. Let (X, d) be a metric space and $A \subseteq A' \subseteq X$. If A' is separable, prove that A is separable.

9.6.5. Prove that $\mathbb{Q}^d$ is dense in $\mathbb{R}^d$ and hence $\mathbb{R}^d$ is separable. Now, exercise 9.6.4 implies that every subset of $\mathbb{R}^d$ is separable.