COMPLEX NUMBERS

TSOGTGEREL GANTUMUR

Contents

1. Brief history and introduction

The square of a real number is always nonnegative, i.e., a negative number is never a square. However, it turns out that one can extend the concept of a number to include objects other than the real numbers so that negative numbers are squares of those hypothetical objects. Roughly speaking, this is how complex numbers were discovered.

As far as the recorded history goes, Gerolamo Cardano (1501-1576) was the first person to encounter complex numbers explicitly. In his *Ars Magna* (1545), Cardano considers the equation $x(10-x)=40$, that is,

$$x^2-10x+40=0. \tag{1}$$

If we apply the usual solution formula for quadratic equations, one of the "solutions" we get is $x=5+\sqrt{25-40}=5+\sqrt{-15}$. Now, as Cardano writes, "ignoring the mental tortures involved", we can check that

$$x(10-x)=(5+\sqrt{-15})(5-\sqrt{-15})=5^2-(\sqrt{-15})^2=25-(-15)=40. \tag{2}$$

So there seems to be some sense in which $x=5+\sqrt{-15}$ is really a solution of (1). Cardano shows this calculation but dismisses it immediately by saying that it is useless.

The next step was taken by Rafael Bombelli (1526-1572) in his *Algebra* (1572). As a sort of motivation to study complex numbers, he considers the cubic equation

$$x^3=3px+2q, \tag{3}$$

with $p=5$ and $q=2$, and applies the formula

$$x=\sqrt[3]{q+\sqrt{q^2-p^3}}+\sqrt[3]{q-\sqrt{q^2-p^3}}, \tag{4}$$

for a solution of the equation (3). The formula (4), or rather an approach equivalent to it, had been described in Cardano's *Ars Magna*. Thus (4) gives

$$x=\sqrt[3]{2+\sqrt{2^2-5^3}}+\sqrt[3]{2-\sqrt{2^2-5^3}}=\sqrt[3]{2+\sqrt{-121}}+\sqrt[3]{2-\sqrt{-121}}. \tag{5}$$

Date: February 3, 2015.

Bombelli makes a guess that $\sqrt[3]{2 \pm \sqrt{-121}} = 2 \pm \sqrt{-1}$, and verifies it as

$$\begin{aligned}(2 \pm \sqrt{-1})^3 &= 2^3 \pm 3 \cdot 2^2 \cdot \sqrt{-1} + 3 \cdot 2 \cdot (\sqrt{-1})^2 \pm (\sqrt{-1})^3 \\ &= 8 \pm 12\sqrt{-1} + 6 \cdot (-1) \pm (-1\sqrt{-1}) \\ &= 2 \pm 11\sqrt{-1} = 2 \pm \sqrt{121} \cdot \sqrt{-1} = 2 \pm \sqrt{-121}.\end{aligned} \tag{6}$$

In light of this, (5) becomes

$$x = (2 + \sqrt{-1}) + (2 - \sqrt{-1}) = 4, \tag{7}$$

which is really a solution of (3), since

$$4^3 = 15 \cdot 4 + 4. \tag{8}$$

What is interesting here is that the intermediate calculations leading to the real solution $x = 4$ involve the square roots of negative numbers, and at the time there was no other way known to reach this solution. In particular, by venturing into the domain of complex numbers, Bombelli discovered a new class of solutions to cubic equations that escaped Cardano's investigations. This indicates the usefulness, and to some extent, even the necessity of complex numbers.

To fix ideas, by *complex numbers* we understand expressions of the form $a + b\sqrt{-B}$, where a, b and B are real numbers. We can restrict attention to the case $B > 0$, because if $B \leq 0$ then $a' = a + b\sqrt{-B}$ is a real number, which can be written, e.g., as $a' + 0 \cdot \sqrt{-1}$. In particular, real numbers are special cases of complex numbers. Moreover, for a positive real number B, we have

$$\sqrt{-B} = \sqrt{B \cdot (-1)} = \sqrt{B} \cdot \sqrt{-1}, \tag{9}$$

which means that any complex number can be written in the form

$$a + b\sqrt{-B} = a + b\sqrt{B} \cdot \sqrt{-1} = a + b'\sqrt{-1}, \tag{10}$$

where $b' = b\sqrt{B}$. In other words, we can always assume $B = 1$. Following Bombelli, we can give the following rules for addition and subtraction of complex numbers:

$$\begin{aligned}(a + b\sqrt{-1}) + (c + d\sqrt{-1}) &= (a + c) + (b + d)\sqrt{-1}, \\ (a + b\sqrt{-1}) - (c + d\sqrt{-1}) &= (a - c) + (b - d)\sqrt{-1}.\end{aligned} \tag{11}$$

For multiplication, also following Bombelli, we have

$$\begin{aligned}(a + b\sqrt{-1}) \cdot (c + d\sqrt{-1}) &= ac + bd(\sqrt{-1})^2 + (ad + bc)\sqrt{-1} \\ &= (ac - bd) + (ad + bc)\sqrt{-1}.\end{aligned} \tag{12}$$

Apart from the necessity in the calculation of roots of cubic polynomials, there is another, more fundamental role complex numbers play in polynomial equations, which was only beginning to be appreciated in the 17th century. This role is expressed through the *fundamental theorem of algebra*, which says that any nonconstant polynomial equation has at least one root, if we allow complex numbers to be roots. That is, if $a_0, a_1, \dots, a_n$ are real numbers such that at least one of $a_1, a_2, \dots, a_n$ is nonzero, then the equation

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = 0, \tag{13}$$

has a solution, provided x may have complex values. If $a_1 = a_2 = \dots = a_n = 0$, then the equation $p(x) = 0$ becomes $a_0 = 0$, which does not have any (complex) solution when $a_0 \neq 0$. So the condition that at least one of $a_1, a_2, \dots, a_n$ is nonzero (i.e., $p(x)$ is nonconstant) is simply to rule out this trivial case. The fundamental theorem of algebra is miraculous because complex numbers are designed to solve any quadratic equation, and it is *a priori* conceivable that we need to introduce a new kind of "number" every time we increase the degree of a polynomial equation. The first formulation of the fundamental theorem of algebra was given by Albert Girard (1595-1632) in 1629, although he did not attempt a proof. Indeed, rigorous

proofs of this theorem did not appear until the early 19th century, which incidentally marks the beginning of an era when the existence and usefulness of complex numbers were widely accepted. In the meantime, since the nature of complex numbers was unclear, and even the very status of negative numbers was somewhat shaky, most mathematicians were extremely reluctant to accept complex numbers. The father of analytic geometry, René Descartes (1596-1650) wrote that the square roots of negative numbers are "imaginary." Both inventors of calculus, Isaac Newton (1643-1727) and Gottfried Leibniz (1646-1716), never approved of the existence of complex numbers. Newton said they were "impossible numbers." Leibniz called them "an amphibian between being and not being."

Note that the polynomial $p(x)$ in (13) is initially defined only for real variable x. By allowing x to be a complex number, in effect, we have extended the polynomial $p(x)$ from a real variable to a complex variable. That is, instead of $p(x)$, we consider the polynomial

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0, \tag{14}$$

where $z = x + y\sqrt{-1}$ is now a complex variable. When we talked about complex roots of the equation $p(x) = 0$, this extension from a real to a complex variable is done implicitly and seemlessly, because given z, the computation of $p(z)$ according to (14) involves only addition and multiplication of complex numbers. In fact, we can now consider polynomials with complex coefficients (the fundamental theorem of algebra is still true for them). However, if we want to extend other functions, such as e^x and $\sin x$, to a complex variable, the situation is not completely trivial. We cannot simply replace x with z, as we have done in going from (13) to (14), because that would give "e^z" and "$\sin z$", which are the very things we are trying to define. This problem was solved by Leonhard Euler (1707-1783) in his *Introductio* (1748). First, he develops the (real) exponential function into the power series

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + \ldots, \tag{15}$$

where x is a real variable, and then simply replaces x with z to define the complex exponential

$$e^z = 1 + z + \frac{z^2}{2} + \frac{z^3}{3!} + \ldots + \frac{z^n}{n!} + \ldots, \tag{16}$$

where z is a complex variable. Of course, the principal difference between (16) and (14) is that (16) involves infinitely many terms, and so for a given z, we must ensure that the right hand side of (16) defines a complex number, which would then be the definition of the value e^z. We shall make sense of the infinite sum (or the series) in (16) as a limit. For any given complex number z and any positive integer n, the partial sum

$$S_n(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{3!} + \ldots + \frac{z^n}{n!}, \tag{17}$$

makes sense and will be a complex number. If there is a complex number w such that $S_n(z)$ gets closer and closer to w as n approaches infinity, then we say that the series in the right hand side of (16) converges to w, and we take $e^z = w$. If the series in (16) converges for every complex number z, then (16) would be a good definition of the function e^z. We will not delve into the convergence issue here, except to note that it requires the notion of "closeness" between two complex numbers. Working with infinite series, Euler discovered many fundamental identities such as

$$e^{it} = \cos t + i \sin t, \tag{18}$$

where t is a real number, and $i = \sqrt{-1}$. The notation i was introduced by Euler in 1777.

The geometric interpretation of complex numbers as points on a (two-dimensional) plane was a big step towards taking away the mystery of complex numbers. Real numbers can be represented by points on a line, and they "do not leave any gap". Then, roughly speaking,

if complex numbers really exist, in order to represent them, one needs an extra dimension. It was John Wallis (1616-1703) who first suggested a graphical representation of complex numbers in 1673, although his method had a flaw. From writings of many mathematicians such as Euler, it is clear that they were thinking of complex numbers as points on a plane, even though they do not make it explicit. The first explicit accounts of the modern approach appeared around 1800, and it is credited to Caspar Wessel (1745-1818), Carl Friedrich Gauss (1777-1855), and Jean-Robert Argand (1768-1822). In this approach, the complex number $z = a + bi$, where a and b are real numbers, is represented by the point (a, b) on the plane $\mathbb{R}^2$. Equivalently, one can think of $z = a + bi$ as the vector with the tail at $(0, 0)$ and the head at (a, b). Then the rules (11) for addition and subtraction of complex numbers coincide with the corresponding rules for vectors:

$$\begin{aligned} (a, b) + (c, d) &= (a + c, b + d), \\ (a, b) - (c, d) &= (a - c, b - d). \end{aligned} \tag{19}$$

The multiplication rule (12) applied to vectors is

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc), \tag{20}$$

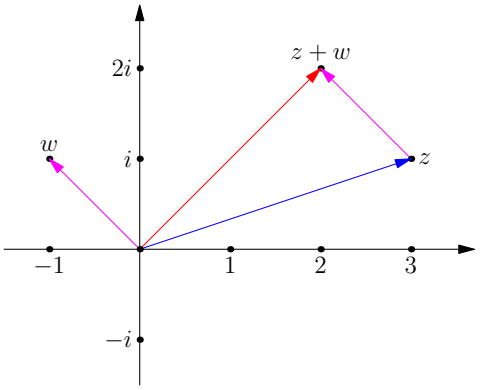
and it is not immediately clear if this can be understood in terms of common operations for vectors. An interesting special case occurs when we take $(c, d) = (0, 1)$, that is, multiplication of $a + bi$ by i:

$$(a, b) \cdot (0, 1) = (-b, a), \tag{21}$$

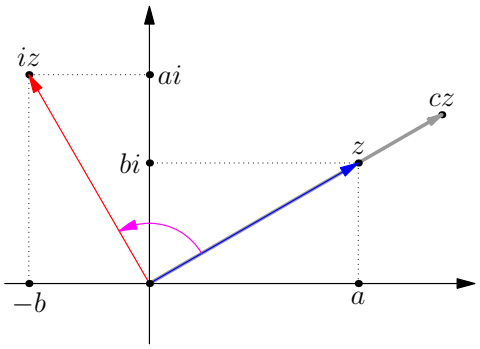
which is the vector (a, b) rotated counter-clockwise by the angle $\frac{\pi}{2}$. Another special case is when $(c, d) = (c, 0)$, that is, multiplication of $a + bi$ by a real number c:

$$(a, b) \cdot (c, 0) = (ac, bc). \tag{22}$$

This is, of course, simply the scaling of the vector (a, b) by the factor of c.



(A) Addition of complex numbers corresponds to addition of vectors.



(B) Multiplication by i corresponds to rotation by $\frac{\pi}{2}$, and multiplication by a real number corresponds to scaling.

FIGURE 1. On the so-called *Argand diagram*, the complex number $z = a + bi$ is represented by the point (a, b).

Exercise 1. Let $w = a + bi$ be a nonzero complex number, and let θ be the angle between the vector (a, b) and the positive direction of the horizontal axis, counted anticlockwise. Without using trigonometric functions, show that multiplication by w corresponds to the rotation by the angle θ, followed by the scaling with the factor of $\sqrt{a^2 + b^2}$. *Hint*: Decompose $w \cdot z$ as the sum of $a \cdot z$ and $b \cdot i \cdot z$.

Any doubts on the existence and importance of complex numbers were completely disposed of after the development of *complex analysis*, which is also known as *function theory*, or *the theory of functions of a complex variable*. The initial motivation for studying functions of a complex variable was to use them to compute (or simplify) real definite integrals, and the pioneering works in this direction were done by Euler and Joseph-Louis Lagrange (1736-1813) around 1760-1780. Their research was taken up later in the 1810's by Augustin Louis Cauchy (1789-1857), who realized by 1821 that complex functions have a rich theory of their own. Gauss reached the same understanding as early as 1811, and played a major role in popularizing complex numbers, but he did not directly contribute to the development of complex analysis. Thus roughly between 1820-1850, Cauchy singlehandedly developed all the basic results of complex analysis, perhaps with the exception of Laurent series, which first appeared in a paper submitted by Pierre Alphonse Laurent (1813-1854) in 1843. Laurent series was also known to Karl Weierstrass (1815-1897) by 1841. Weierstrass and Bernhard Riemann (1826-1866) developed complex analysis further, but the main results of their work are beyond the scope of this course. All the results we will cover in this course were known in their basic form by the year 1850, to Cauchy or Weierstrass.

2. Axioms and models of complex numbers

In this section, we will have a critical look at complex numbers and address the question if complex numbers really exist. We start with a discussion on the meaning of the symbol $\sqrt{-1}$, or i, as we have used it to define complex numbers. Certainly, we cannot think of $\sqrt{-1}$ as the result of some operation applied to -1, because it would mean that we had already defined complex numbers. What we need to do is simply *assume* the existence of an object i satisfying $i^2 = -1$. We require that the usual rules of arithmetics apply to i, and construct complex numbers as objects of the form $a + bi$, where a and b are real numbers. Now, the existence of i immediately implies the existence of another square root of -1, namely, $-i$, since $(-i)^2 = (-1 \cdot i)^2 = (-1)^2 i^2 = -1$. Let us check if there exist any other (complex) square roots of -1. So we consider the equation $(a + bi)^2 = -1$, which is equivalent to

$$a^2 - b^2 + 2abi = -1. \tag{23}$$

Since the right hand side is real, we have $ab = 0$, and so $a = 0$ or $b = 0$. If $b = 0$, then $a^2 = -1$, which is impossible. On the other hand, assuming $a = 0$ we end up with $b^2 = 1$, or $b = \pm 1$. Hence $\pm i$ are the only solutions of $z^2 = -1$. Moreover, $-i \neq i$, because $i = -i$ would imply that $2i = 0$ and hence $i = 0$. This might lead to the following confusion. Suppose that complex numbers exist as objects in some hypothetical universe. In that universe, of course, there will be 2 distinct square roots of -1. When we assume the existence of i, as we have done earlier, we are effectively picking one of the square roots of -1. However, how do we know which one we are choosing? Does the choice matter? To get out of this conundrum, we need to assume that i and $-i$ are identifiable and different, just as 1 and -1 are different. What it means is that since the relation $i^2 = -1$ cannot differentiate between i and $-i$, we do not use this relation as a definition of i, but rather, we assume that there existed an object i, and that it just happened to have the property $i^2 = -1$.

Simply assuming the existence of i may appear as a strange way to convince somebody that a square root of negative one exists. However, it is not so strange if we examine what we mean by the existence of mathematical objects. In a mathematical theory, such as Euclidean geometry or arithmetics, one starts with a few basic facts and definitions, and incrementally deduces more and more complicated facts by using the rules of logic. The basic facts and definitions one starts with are called *axioms*, and they are assumed to be self-evident[1]. There

[1]In Euclidean geometry, the axioms were used to be considered as idealizations of the geometry of the physical space. Later it was discovered that starting with a slightly different set of axioms, one can build a

can be discussions on what axioms one should choose, but once the axioms have been chosen, there is no question within the theory about the validity of the axioms. In other words, there are potentially as many mathematical theories as there are systems of axioms. The axioms of a theory inevitably introduce some basic objects, such as points and lines, which are simply assumed to exist, and describe relationships between these objects. Now, notice that even if they are assumed to exist, these objects by themselves are devoid of meaning. Only through the stated (as well as deduced) relationships between them that these objects become "alive" and they have any meaning. Thus, axioms stating that "straight lines exist", "points exist", and "a line can contain a point" do not convey much information. They simply say that there are two types of objects in the theory, and there is one relation between the two types of objects. If we have an additional axiom saying "Given any two distinct points, there is at least one line containing both of them", the notions of points and lines start to acquire some meaning. The essence of a theory is not in the objects, but in the relationships between the objects. To put it differently, since there is nothing in the theory that identifies an object except its relationship to others, the relationships *define* the objects. This makes it clear that instead of asking if particular mathematical objects exist, one should be asking if the logical relationships between them "exist." Recall that we are talking about logical relationships that are stated in and deduced from the given set of axioms. So once the focus is on the relationships, it becomes difficult to imagine when a given set of relationships does *not* exist! The only reasonable sense in which the logical relationships in a theory do *not* exist is that the theory is inconsistent, in the sense that the axioms lead to contradictory statements, such as "$0 = 1$ and $0 \neq 1$". Therefore, we identify the existence of the objects presupposed in a theory with consistency of its axioms. In particular, if we want to rigorously establish the existence of complex numbers, we need to have a clearly stated set of axioms, which we shall do now. What we want to state in the axioms is basically that complex numbers behave like real numbers as far as addition and multiplication are concerned, real numbers are special cases of complex numbers, i exists, and any complex number can be written as $x + yi$ with x and y real.

Axiom 1 (Complex numbers[2])**.** *There exists the set of complex numbers, which we denote by $\mathbb{C}$, satisfying the following properties.*

- (*a*) *The set of complex numbers contains the real numbers, i.e., $\mathbb{R} \subset \mathbb{C}$.*
- (*b*) *The addition operation for $\mathbb{R}$ extends to $\mathbb{C}$, and it satisfies the following.*
  - (*i*) *$z, w \in \mathbb{C}$ then $z + w \in \mathbb{C}$.*
  - (*ii*) *$z \in \mathbb{C}$ then $z + 0 = z$. (0 is an additive unit)*
  - (*iii*) *$z, w \in \mathbb{C}$ then $z + w = w + z$. (commutativity)*
  - (*iv*) *$z, w, s \in \mathbb{C}$ then $(z + w) + s = z + (w + s)$. (associativity)*
- (*c*) *The multiplication operation for $\mathbb{R}$ extends to $\mathbb{C}$, and it satisfies the following.*
  - (*i*) *$z, w \in \mathbb{C}$ then $z \cdot w \in \mathbb{C}$.*
  - (*ii*) *$z \in \mathbb{C}$ then $z \cdot 1 = z$. (1 is a multiplicative unit)*
  - (*iii*) *$z, w \in \mathbb{C}$ then $z \cdot w = w \cdot z$. (commutativity)*
  - (*iv*) *$z, w, s \in \mathbb{C}$ then $(z \cdot w) \cdot s = z \cdot (w \cdot s)$. (associativity)*
- (*d*) *$z, w, s \in \mathbb{C}$ then $z \cdot (w + s) = z \cdot w + z \cdot s$. (distributivity)*
- (*e*) *There exists a number $i \in \mathbb{C}$ such that $i \cdot i = -1$. (imaginary unit)*
- (*f*) *If $z \in \mathbb{C}$ then there exist $x, y \in \mathbb{R}$ such that $z = x + y \cdot i$. (real and imaginary parts)*

---

geometry, that is as valid as Euclidean geometry, in the sense that there is no way of knowing which geometry corresponds to the physical reality better without doing physical experiments. Then it is completely natural to consider the two geometries as two equally valid mathematical theories, and to leave the question of which geometry is the "physical one" to physicists.

[2]This set of axioms is a simplified version of those of Metamath and Wikiproofs.

Remark 2. That the addition and multiplication operations for $\mathbb{R}$ extend to $\mathbb{C}$ means that $z+w$ and $z\cdot w$ for $z,w\in\mathbb{C}$ coincide, respectively, with the addition and multiplication of real numbers if z and w happened to be real numbers.

In view of the discussion preceding the statement of Axiom 1, if we can show that Axiom 1 does not lead to any self-contradictory statements, it would mean that we have proved the existence of complex numbers. A common approach to deal with a consistency question is to reduce it to the consistency of a simpler theory, by building a *model* of the original theory within the simpler theory. We are going to explain it in the particular context of complex numbers. In a theorem below, we will construct a set of objects and relations by using concepts from the theory of real numbers, in such as way that the constructed set of objects and relations satisfy the axioms of complex numbers. This set is called a *model* of (the theory of) complex numbers. Now suppose that the complex number axioms were inconsistent, meaning that there is a chain of reasoning, which starts at the axioms, and ends at a self-contradictory statement. Then we would be able to express this self-contradictory statement in terms of concepts from the theory of real numbers, by using our model as a "dictionary" between complex number concepts and real number concepts. Hence we would prove that real numbers are inconsistent, which would in turn have very strong consequences. The whole argument will therefore show that complex numbers are as "real" as real numbers.

Theorem 3 (Vector model)**.** *Take the plane* $\mathbb{R}^2=\{(x,y):x,y\in\mathbb{R}\}$*, and embed* $\mathbb{R}$ *into* $\mathbb{R}^2$ *by* $x\mapsto(x,0)$*. Define*

$$(a,b)+(c,d)=(a+c,b+d), \tag{24}$$

$$(a,b)\cdot(c,d)=(ac-bd,ad+bc), \tag{25}$$

and $i=(0,1)$*. This system satisfies the complex number axioms (i.e., Axiom 1).*

Proof. Verifying Part (b) of Axiom 1 is straightforward. For example,

$$(a,b)+(c,d)=(a+c,b+d)=(c+a,d+b)=(c,d)+(a,b), \tag{26}$$

shows commutativity (b)(iii). To check that the addition defined by (24) is an extension of the addition of real numbers, we consider two arbitrary real numbers $x,y\in\mathbb{R}$. Inside $\mathbb{R}^2$, these two numbers are represented by $(x,0)$ and $(y,0)$, and their sum according to (24) is

$$(x,0)+(y,0)=(x+y,0), \tag{27}$$

which is exactly the sum $x+y\in\mathbb{R}$, considered as an element of $\mathbb{R}^2$ under the embedding rule $x\mapsto(x,0)$. So (24) coincides with the addition of real numbers, if the summands are real.

Now we turn to Part (c) of Axiom 1. We have

$$(x,0)\cdot(y,0)=(xy-0\cdot0,x\cdot0+0\cdot y)=(xy,0), \tag{28}$$

which confirms that the multiplication defined by (25) is indeed an extension of the multiplication of real numbers. The unit property (c)(ii) is easy to check, as

$$(a,b)\cdot(1,0)=(a\cdot1-b\cdot0,a\cdot0+b\cdot1)=(a,b). \tag{29}$$

Showing commutativity (c)(iii) is similar to (26), so we omit it here. To prove associativity (c)(iv), we write the vectors in $\mathbb{R}^2$ in the column form, and so, for example, we have

$$\begin{pmatrix}a\\b\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}ax-by\\ay+bx\end{pmatrix}. \tag{30}$$

The latter expression can be recognized as a matrix-vector product, giving us a way to write the multiplication (25) as a matrix-vector product:

$$\begin{pmatrix}a\\b\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}ax-by\\ay+bx\end{pmatrix}=\begin{pmatrix}a&-b\\b&a\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}. \tag{31}$$

Thus we have

$$\left(\begin{pmatrix}a\\b\end{pmatrix}\cdot\begin{pmatrix}c\\d\end{pmatrix}\right)\cdot\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}ac-bd\\ad+bc\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}ac-bd & -(ad+bc)\\ad+bc & ac-bd\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}. \tag{32}$$

On the other hand, we have

$$\begin{pmatrix}a\\b\end{pmatrix}\cdot\left(\begin{pmatrix}c\\d\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}\right)=\begin{pmatrix}a & -b\\b & a\end{pmatrix}\left(\begin{pmatrix}c & -d\\d & c\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\right)=\begin{pmatrix}a & -b\\b & a\end{pmatrix}\begin{pmatrix}c & -d\\d & c\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}, \tag{33}$$

by associativity of matrix-vector multiplication[3], and the explicit computation

$$\begin{pmatrix}a & -b\\b & a\end{pmatrix}\begin{pmatrix}c & -d\\d & c\end{pmatrix}=\begin{pmatrix}ac-bd & -ad-bc\\bc+ad & -bd+ac\end{pmatrix}, \tag{34}$$

shows the equality between (32) and (33).

Similarly, distributivity law (d) can be verified as

$$\begin{aligned}\begin{pmatrix}a\\b\end{pmatrix}\cdot\left(\begin{pmatrix}c\\d\end{pmatrix}+\begin{pmatrix}x\\y\end{pmatrix}\right)&=\begin{pmatrix}a & -b\\b & a\end{pmatrix}\left(\begin{pmatrix}c\\d\end{pmatrix}+\begin{pmatrix}x\\y\end{pmatrix}\right)=\begin{pmatrix}a & -b\\b & a\end{pmatrix}\begin{pmatrix}c\\d\end{pmatrix}+\begin{pmatrix}a & -b\\b & a\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\\&=\begin{pmatrix}a\\b\end{pmatrix}\cdot\begin{pmatrix}c\\d\end{pmatrix}+\begin{pmatrix}a\\b\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix},\end{aligned} \tag{35}$$

where we have used distributivity of matrix-vector multiplication over vector addition[4].

For (e), we have

$$i\cdot i=\begin{pmatrix}0\\1\end{pmatrix}\cdot\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}-1\\0\end{pmatrix}=-1. \tag{36}$$

To prove (f), first note that

$$i\cdot\begin{pmatrix}a\\0\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}a\\0\end{pmatrix}=\begin{pmatrix}0\\a\end{pmatrix}, \tag{37}$$

and use it in

$$\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x\\0\end{pmatrix}+\begin{pmatrix}0\\y\end{pmatrix}=\begin{pmatrix}x\\0\end{pmatrix}+i\cdot\begin{pmatrix}y\\0\end{pmatrix}=x+yi, \tag{38}$$

which completes the proof. □

In the preceding proof, we have seen that the product of (a,b) and (x,y) according to (25) can be represented by a matrix-vector product as

$$\begin{pmatrix}a\\b\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}a & -b\\b & a\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}. \tag{39}$$

This leads us to the possibility of modelling complex numbers by special 2×2 matrices.

Theorem 4 (Matrix model)**.** *We introduce the set*

$$\mathbb{CR}=\left\{\begin{pmatrix}a & b\\c & d\end{pmatrix}\in\mathbb{R}^{2\times 2}:a=d,\,b+c=0\right\}, \tag{40}$$

which we call the space of Cauchy-Riemann matrices*, and embed* $\mathbb{R}$ *into* $\mathbb{CR}$ *by* $x\mapsto\begin{pmatrix}x & 0\\0 & x\end{pmatrix}$*. Then with the usual addition and multiplication operations for matrices, and with the definition* $i=\begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}$*,* $\mathbb{CR}$ *satisfies the complex number axioms (i.e., Axiom 1).*

[3]Let $A=(a_{ij})$, $B=(b_{jk})$, and $C=(c_{k\ell})$ be matrices with compatible dimensions so that the product ABC can be formed. Then we have $((AB)C)_{i\ell}=\sum_k(\sum_j a_{ij}b_{jk})c_{k\ell}=\sum_j a_{ij}\sum_k b_{jk}c_{k\ell}=(A(BC))_{i\ell}$.

[4]Let $A=(a_{ij})$, $B=(b_{jk})$, and $C=(c_{jk})$ be matrices with compatible dimensions so that $A(B+C)$ makes sense. Then we have $(A(B+C))_{ik}=\sum_j a_{ij}(b_{jk}+c_{jk})=\sum_j a_{ij}b_{jk}+\sum_j a_{ij}c_{jk}=(AB)_{ik}+(AC)_{ik}$.

Proof. We will only check some of the axioms, and leave the others as exercises. First,

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} ac-bd & -ad-bd \\ bc+ad & -bd+ac \end{pmatrix} \in \mathbb{CR}, \tag{41}$$

shows that the Cauchy-Riemann matrices are closed under matrix multiplication, which is Axiom (c)(i). Then Axiom (e) is verified as

$$i \cdot i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -1. \tag{42}$$

For any real number y, we have

$$i \cdot y = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y & 0 \\ 0 & y \end{pmatrix} = \begin{pmatrix} 0 & -y \\ y & 0 \end{pmatrix}, \tag{43}$$

so we can write

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix} = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} + \begin{pmatrix} 0 & -y \\ y & 0 \end{pmatrix} = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y & 0 \\ 0 & y \end{pmatrix} = x + y \cdot i, \tag{44}$$

for any Cauchy-Riemann matrix. This confirms Axiom (f). □

Exercise 5. Check the remaining axioms. In particular, show that matrix multiplication is commutative in the class of Cauchy-Riemann matrices.

Remark 6. Another popular model is given in terms of polynomials, as follows. We start with the set of all polynomials with real coefficients:

$$\mathbb{R}[t] = \{a_0 + a_1 t + \ldots + a_n t^n : a_0, \ldots, a_n \in \mathbb{R},\, n \in \mathbb{N}\}, \tag{45}$$

and identify two polynomials if their difference can be divided by t^2+1. The resulting set is denoted by $\mathbb{R}[t]/(t^2+1)$. For example, t^2+2 and $(t-1)(t^2+1)+1$ represent the same element in $\mathbb{R}[t]/(t^2+1)$, because their difference is $(t-2)(t^2+1)$. Addition and multiplication in $\mathbb{R}[t]/(t^2+1)$ are defined as the usual addition and multiplication of polynomials, and we embed $\mathbb{R}$ into $\mathbb{R}[t]/(t^2+1)$ by $a \mapsto a + 0 \cdot t$, i.e., the real number a is identified with the constant polynomial a. We also set i to be the polynomial $p(t) = t$. Then in this setting, $\mathbb{R}[t]/(t^2+1)$ is a model of $\mathbb{C}$. We will not go into more details but let us verify the property $i \cdot i = -1$. We have

$$i \cdot i = t \cdot t = t^2, \tag{46}$$

and since $t^2 - (-1) = t^2 + 1$, which is divisible by t^2+1, the polynomial t^2 must be identified with -1 as an element of $\mathbb{R}[t]/(t^2+1)$.

3. Algebra and geometry of complex numbers

In this section, we will derive fundamental algebraic properties of $\mathbb{C}$ from Axiom 1, and will introduce some important geometric notions.

Lemma 7. *If $w \in \mathbb{C}$ satisfies $z + w = z$ for all $z \in \mathbb{C}$, then $w = 0$. We also have $0 \cdot i = 0$.*

Proof. Since 0 is an additive unit, we have $w = 0 + w$. Now the assumed property of w, applied with $z = 0$, gives $0 + w = 0$. Therefore, $w = 0 + w = 0$.

For the second assertion, let $z \in \mathbb{C}$. Then by Axiom (f), there are real numbers x and y such that $z = x + y \cdot i$. We have

$$z + 0 \cdot i = x + y \cdot i + 0 \cdot i = x + (y+0) \cdot i = x + y \cdot i = z, \tag{47}$$

where we have used associativity of addition and distributivity in the second step, and the real additive unit property $y + 0 = y$ in the third step. Since $z \in \mathbb{C}$ was arbitrary, by the first part of the lemma, we conclude that $0 \cdot i = 0$. □

Remark 8. From the proof, we note that the condition $0+w=0$ is sufficient to imply $w=0$.

Exercise 9. Show that if $w\in\mathbb{C}$ satisfies $zw=z$ for all $z\in\mathbb{C}$, then $w=1$.

Theorem 10 ($\mathbb{C}$ is a field)**.** *a) For each $z\in\mathbb{C}$, there is a unique $w\in\mathbb{C}$ such that $z+w=0$. We write $-z=w$.*
b) For each $z\in\mathbb{C}\setminus\{0\}$, there exists a unique $w\in\mathbb{C}$ such that $zw=1$. We write $\frac{1}{z}\equiv z^{-1}=w$.

Proof. a) Let $z\in\mathbb{C}$, and let $x,y\in\mathbb{R}$ be such that $z=x+yi$, which exist by Axiom (f). Then we define $w=(-x)+(-y)\cdot i$, and compute

$$z+w=x+y\cdot i+(-x)+(-y)\cdot i=(x+(-x))+(y+(-y))\cdot i=0+0\cdot i=0+0=0, \tag{48}$$

where we have used associativity and commutativity of addition, distributivity, the properties $x+(-x)=0$ and $y+(-y)=0$, and finally, the fact that $0\cdot i=0$. For uniqueness, suppose that $z+w=0$ and $z+u=0$. Then we get

$$u=u+0=u+(z+w)=(u+z)+w=(z+u)+w=0+w=w. \tag{49}$$

b) Let $z\in\mathbb{C}$, and let $x,y\in\mathbb{R}$ be such that $z=x+yi$. Then we define[5]

$$w=\frac{x}{x^2+y^2}+\frac{-y}{x^2+y^2}\cdot i. \tag{50}$$

Note that all divisions involved are real number operations. Note also that if $x=y=0$, then $x+y\cdot i=0+0\cdot i=0$, so $x^2+y^2\neq 0$ unless $z=0$. Now we compute

$$\begin{aligned} zw&=(x+y\cdot i)\big(\frac{x}{x^2+y^2}+\frac{-y}{x^2+y^2}\cdot i\big)=\frac{x\cdot x-y\cdot(-y)}{x^2+y^2}+\frac{x\cdot(-y)+y\cdot x}{x^2+y^2}\cdot i\\ &=1+0\cdot i=1+0=1, \end{aligned} \tag{51}$$

which confirms the existence of $\frac{1}{z}$. Uniqueness is left as an exercise. □

Definition 11. For $z,u\in\mathbb{C}$ we introduce the *difference*

$$u-z=u+(-z), \tag{52}$$

and in case $z\neq 0$, the *quotient*

$$\frac{u}{z}=uz^{-1}. \tag{53}$$

Exercise 12. Show that if $zw=zu=1$ and $z\neq 0$ then $w=u$.

Corollary 13. *a) $0\cdot z=0$ for $z\in\mathbb{C}$.*
b) $-(zw)=(-z)\cdot w$ for $z,w\in\mathbb{C}$. In particular, $-z=(-1)\cdot z$.
c) For any $z\in\mathbb{C}$, there is a unique pair $(x,y)\in\mathbb{R}^2$ such that $z=x+y\cdot i$.

Proof. a) Using the distributivity axiom, we first observe that

$$z=z\cdot 1=z\cdot(0+1)=z\cdot 0+z\cdot 1=z\cdot 0+z. \tag{54}$$

Then we infer

$$z\cdot 0=z\cdot 0+(z+(-z))=(z\cdot 0+z)+(-z)=z+(-z)=0, \tag{55}$$

where we have used (54) in the penultimate step.

b) This also follows from distributivity:

$$z\cdot w+(-z)\cdot w=(z+(-z))\cdot w=0\cdot w=0. \tag{56}$$

c) If $x+y\cdot i=x'+y'\cdot i$, then

$$(x-x')+(y-y')\cdot i=0+0\cdot i=0, \tag{57}$$

[5]The expression for w is inspired by the formal computation $\frac{1}{a+bi}=\frac{a-bi}{(a+bi)(a-bi)}=\frac{a-bi}{a^2+b^2}$.

hence it suffices to show that $a + bi = 0$ implies $a = b = 0$ for $a, b \in \mathbb{R}$. If $a + bi = 0$, then $(a + bi) \cdot z = 0$ for any $z \in \mathbb{C}$. We pick $z = a - bi$, which gives $(a + bi)(a - bi) = 0$, that is, $a^2 + b^2 + 0 \cdot i = 0$. Since $0 \cdot i = 0$, this implies that $a^2 + b^2 = 0$. □

Exercise 14. Part c) of the preceding Corollary defines a map $\phi : \mathbb{C} \to \mathbb{R}^2$. Show that ϕ is invertible, and that

- $\phi(w + z) = \phi(w) + \phi(z)$ for $w, z \in \mathbb{C}$,
- $\phi(w \cdot z) = \phi(w) \cdot \phi(z)$ for $w, z \in \mathbb{C}$,
- $\phi(0) = (0, 0)$ and $\phi(1) = (1, 0)$.

This means that ϕ is in fact a *field isomorphism* between $\mathbb{C}$ and the vector model based on $\mathbb{R}^2$ (considered in Theorem 3). In addition, show that $\phi(i) = (0, 1)$.

Exercise 15. Prove the following.
(a) $(wz)^{-1} = w^{-1}z^{-1}$ for $w, z \in \mathbb{C} \setminus \{0\}$.
(b) If $w, z \in \mathbb{C}$ satisfy $wz = 0$, then $w = 0$ or $z = 0$.

In the proofs we have just presented, the quantities such as $x - yi$ and $x^2 + y^2$ deriving from the representation $z = x + yi$ played prominent roles. Since we now know that the latter representation is unique, the aforementioned quantities become functions of z. We give names to some of those quantities.

Definition 16. For $z = x + yi$, we define its

- *complex conjugate* by $\bar{z} = x - yi$,
- *modulus* by $|z| = \sqrt{x^2 + y^2}$,
- *real part* by $\operatorname{Re} z = x$, and
- *imaginary part* by $\operatorname{Im} z = y$.

Exercise 17. Discuss the meaning of each of the aforementioned operations in the vector and matrix models. Try to write them in terms of natural vector (or matrix) operations.

Exercise 18. Prove the following.
(a) $z\bar{z} = x^2 + y^2$, hence $|z| = \sqrt{z\bar{z}}$.
(b) $|z| \geq 0$ for any $z \in \mathbb{C}$, and $|z| = 0$ if and only if $z = 0$.
(c) $z^{-1} = \frac{z}{z\bar{z}}$ and $|z^{-1}| = \frac{1}{|z|}$ for $z \neq 0$.
(d) $\overline{z + w} = \bar{z} + \bar{w}$ and $\overline{zw} = \bar{z} \cdot \bar{w}$.
(e) $\operatorname{Re} z = \frac{1}{2}(z + \bar{z})$ and $\operatorname{Im} z = \frac{1}{2i}(z - \bar{z})$

Lemma 19. *We have* $|zw| = |z||w|$ *and* $|z + w| \leq |z| + |w|$ *for* $z, w \in \mathbb{C}$.

Proof. We have $zw\overline{zw} = zw\bar{z}\bar{w} = z\bar{z} \cdot w\bar{w}$, which implies $\sqrt{zw\overline{zw}} = \sqrt{z\bar{z}} \cdot \sqrt{w\bar{w}}$. To prove the triangle inequality, we treat the case $w = 1$ first. We start with

$$|1 + z|^2 = (1 + z)(1 + \bar{z}) = 1 + z + \bar{x} + z\bar{z} = 1 + 2\operatorname{Re} z + |z|^2. \tag{58}$$

If $z = x + yi$, then $|z|^2 = x^2 + y^2 \geq x^2$, meaning that $|\operatorname{Re} z| \leq |z|^2$. Thus

$$|1 + z|^2 = 1 + 2\operatorname{Re} z + |z|^2 \leq 1 + 2|z| + |z|^2 = (1 + |z|)^2, \tag{59}$$

leading to

$$|1 + z| \leq 1 + |z|. \tag{60}$$

The case $w = 0$ is trivial, and for $w \in \mathbb{C}$ nonzero, we have

$$\begin{aligned} |w + z| &= |w + wzw^{-1}| = |w(1 + zw^{-1})| = |w||1 + zw^{-1}| \\ &\leq |w|(1 + |zw^{-1}|) = |w|(1 + |z||w^{-1}|) = |w| + |w||z||w^{-1}| = |w| + |w||z||w|^{-1} \\ &= |w| + |z|, \end{aligned} \tag{61}$$

which completes the proof. □

Exercise 20. Show that $||w| - |z|| \leq |w - z|$ for $w, z \in \mathbb{C}$.

Remark 21 (Geometry of multiplication). Identifying $\mathbb{C}$ with $\mathbb{R}^2$ through Axiom (f), we know that multiplication by $w = a + bi$ corresponds to (left) multiplication by the matrix

$$\Phi_w = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathbb{CR}. \tag{62}$$

For $w \neq 0$, by writing

$$\Phi_w = |w| \begin{pmatrix} \frac{a}{\sqrt{a^2+b^2}} & \frac{-b}{\sqrt{a^2+b^2}} \\ \frac{b}{\sqrt{a^2+b^2}} & \frac{a}{\sqrt{a^2+b^2}} \end{pmatrix}, \tag{63}$$

we realize that

$$\Phi_w = |w| \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = |w| \cdot R_\theta, \tag{64}$$

where θ is the (counterclockwise) angle between the x-axis and the vector (a, b), and R_θ is the matrix of rotation through the angle θ. Consider three points $z_1, z_2, z_3 \in \mathbb{C}$, and their images $\{\Phi_w z_1, \Phi_w z_2, \Phi_w z_3\}$. Then since $\Phi_w z_k - \Phi_w z_n = \Phi_w(z_k - z_n)$ by linearity, the angle between, e.g., the vectors $z_2 - z_1$ and $z_3 - z_1$ does not change under the mapping Φ_w. So even though absolute positions and sizes are affected by Φ_w, "general shapes" of geometric configurations are preserved. Not only the shapes, but the orientations are also preserved, in the sense that a letter p cannot be transformed into a letter q by applying a map Φ_w.

Appendix A. The real number system

For completeness, in this appendix we state (one version of) the real number axioms, and derive the most fundamental properties of real numbers from them.

Axiom 2 (Real numbers)**.** *There exists the set of real numbers, which we denote by* $\mathbb{R}$*, satisfying the following properties.*

- (*a*) *There is a binary operation* $+$*, which we call addition, satisfying the following properties.*
  - (*i*) $a, b \in \mathbb{R}$ *then* $a + b \in \mathbb{R}$.
  - (*ii*) *There exists an element* $0 \in \mathbb{R}$ *such that* $a + 0 = a$ *for each* $a \in \mathbb{R}$.
  - (*iii*) $a, b \in \mathbb{R}$ *then* $a + b = b + a$.
  - (*iv*) $a, b, c \in \mathbb{R}$ *then* $(a + b) + c = a + (b + c)$.
  - (*v*) *For any* $a \in \mathbb{R}$ *there exists* $x \in \mathbb{R}$ *such that* $x + a = 0$.
- (*b*) *There is a binary operation* $\cdot$*, called multiplication, satisfying the following properties.*
  - (*i*) $a, b \in \mathbb{R}$ *then* $a \cdot b \in \mathbb{R}$.
  - (*ii*) *There exists an element* $1 \in \mathbb{R}$ *such that* $a \cdot 1 = a$ *for each* $a \in \mathbb{R}$.
  - (*iii*) $a, b \in \mathbb{R}$ *then* $a \cdot b = b \cdot a$.
  - (*iv*) $a, b, c \in \mathbb{R}$ *then* $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
  - (*v*) *For any* $a \in \mathbb{R}$ *not equal to* 0*, there exists* $x \in \mathbb{R}$ *such that* $x \cdot a = 1$.
  - (*vi*) $a, b, c \in \mathbb{R}$ *then* $a \cdot (b + c) = a \cdot b + a \cdot b$.
- (*c*) *There is a binary relation* $<$*, satisfying the following properties.*
  - (*i*) $a, b \in \mathbb{R}$ *then one and only one of the following is true:* $a < b$*,* $a = b$*, or* $a > b$.
  - (*ii*) *If* $a, b, c \in \mathbb{R}$ *satisfy* $a < b$ *and* $b < c$ *then* $a < c$.
  - (*iii*) *If* $a, b, c \in \mathbb{R}$ *and* $a < b$ *then* $a + c < b + c$.
  - (*iv*) *If* $a, b \in \mathbb{R}$ *satisfy* $a > 0$ *and* $b > 0$ *then* $a \cdot b > 0$.
- (*d*) *If* $A \subset \mathbb{R}$ *is nonempty and there is* $b \in \mathbb{R}$ *such that* $a < b$ *for all* $a \in A$*, then there exists* $s \in \mathbb{R}$ *such that* $a \leq s$ *for all* $a \in A$*, and that for any* $c < s$ *there is* $a \in A$ *with* $a > c$.

Remark 22. The relation $a > b$ is defined as $b < a$. Similarly, $a \leq b$ means $a < b$ or $a = b$, and $a \geq b$ means $a > b$ or $a = b$. The property (d) is called the least upper bound property, and the number s is called the *least upper bound* or the *supremum* of A, which is denoted by

$$\sup A = s. \tag{65}$$

Exercise 23 (Algebraic properties). Prove the following.

a) If $a + b = a$ then $b = 0$ (uniqueness of 0).
b) If $a + b = a + c$ then $b = c$ (subtraction of a).
c) $0 \cdot a = 0$.
d) If $ab = 0$ then $a = 0$ or $b = 0$.
e) If $ab = a$ and $a \neq 0$ then $b = 1$ (uniqueness of 1).
f) If $ab = ac$ and $a \neq 0$ then $b = c$ (division by a).

Remark 24. We define subtraction $d - a$ and division $\frac{d}{a}$ as the solutions to the equations $a + x = d$ and $ax = d$. Then b) and f) of the preceding exercise guarantee that these concepts are well defined.

Exercise 25 (Order properties). Prove the following.

a) If $b < c$ and $a > 0$ then $ab < ac$.
b) If $b < c$ and $a < 0$ then $ab > ac$.
c) If $a \neq 0$ then $a \cdot a > 0$.
d) If $0 < a < b$ and $ac = bd > 0$ then $0 < d < c$.

Exercise 26 (Density of rational numbers). Prove that for any given real numbers $a \in \mathbb{R}$ and $b \in \mathbb{R}$ with $a < b$, there exists a rational number $q \in \mathbb{Q}$ such that $a < q < b$.

Definition 27. A *real number sequence* is a function $x : \mathbb{N} \to \mathbb{R}$, which is usually written as $\{x_n\} = \{x_1, x_2, \ldots\}$, with $x_n = x(n)$. We say that a sequence $\{x_n\}$ *converges* to $x \in \mathbb{R}$, if for any given $\varepsilon > 0$, there exists an index N such that

$$|x_n - x| \leq \varepsilon \qquad \text{for all } n \geq N. \tag{66}$$

If $\{x_n\}$ converges to x, we write

$$\lim_{n\to\infty} x_n = x, \qquad \text{or} \qquad x_n \to x \quad \text{as} \quad n \to \infty. \tag{67}$$

Exercise 28. Let $\lim x_n = x$ and $\lim y_n = y$. Show that the following hold.

a) $\lim(x_n \pm y_n) = x \pm y$.
b) $\lim(x_n y_n) = xy$.
c) If $x \neq 0$, then $x_n = 0$ for only finitely many indices n, and after the removal of those zero terms from the sequence $\{x_n\}$, we have $\lim \frac{1}{x_n} = \frac{1}{x}$.

Theorem 29 (Monotone convergence)**.** *Let $\{x_n\} \subset \mathbb{R}$ be a sequence that is nondecreasing and bounded from above, in the sense that*

$$x_n \leq x_{n+1} \leq M \qquad \textit{for each } n, \tag{68}$$

and with some constant $M \in \mathbb{R}$. Then there is $x \leq M$ such that $x_n \to x$ as $n \to \infty$.

Proof. Let $x = \sup\{x_n\}$, and let $\varepsilon > 0$. Then there is N such that $x - \varepsilon < x_N$. Since $\{x_n\}$ is nondecreasing, we have $x - \varepsilon < x_n \leq x$ for all $n \geq N$. This means that $\{x_n\}$ converges to x. The inequality $x \leq M$ is obvious because x is the *least* upper bound of $\{x_n\}$. □

Theorem 30 (Bolzano-Weierstrass)**.** *Let $\{x_n\} \subset \mathbb{R}$ be bounded, in the sense that there exists $M \in \mathbb{R}$ such that $|x_n| \leq M$ for all n. Then there is a subsequence $\{x_{n_k}\} \subset \{x_n\}$ that converges to some point $x \in [-M, M]$.*

Proof. Let us subdivide the interval $[-M, M]$ into two subintervals $[-M, 0]$ and $[0, M]$. Then at least one of these subintervals must contain infinitely many terms from the sequence $\{x_n\}$. Pick one such subinterval, and call it $[a_0, b_0]$. Obviously, we have $b_0 - a_0 = M$. Now we subdivide $[a_0, b_0]$ into two halves $[a_0, \frac{a_0+b_0}{2}]$ and $[\frac{a_0+b_0}{2}, b_0]$, one of which must contain infinitely many terms from $\{x_n\}$. Recall that interval $[a_1, b_1]$. Of course, we have $b_1 - a_1 = \frac{M}{2}$. We continue this process indefinitely, and obtain a sequence of intervals

$$[a_0, b_0] \supset [a_1, b_1] \supset \ldots \supset [a_m, b_m] \supset \ldots, \tag{69}$$

with each $[a_m, b_m]$ containing infinitely many terms from the sequence $\{x_n\}$, and satisfying $b_m - a_m = 2^{-m}M$. We can also write

$$a_0 \leq a_1 \leq \ldots \leq a_m < b_m \leq \ldots \leq b_1 \leq b_0, \tag{70}$$

which makes it clear that $\{a_m\}$ is nondecreasing and $\{b_m\}$ is nonincreasing. Since both of these sequences are bounded, by the monotone convergence theorem (Theorem 29), there exist a and b such that $a_m \to a$ and $b_m \to b$ as $m \to \infty$. Given any m, we have $a_m \leq a_n < b_n \leq b_m$ whenever $n \geq m$. This implies that a and b are both in the interval $[a_m, b_m]$ for any m. Since $b_m - a_m = 2^{-m}M$, we infer $a = b$, and moreover, $|a - a_m| \leq 2^{-m}M$ for all m.

For $k = 0, 1, \ldots$, let n_k be an index such that $x_{n_k} \in [a_k, b_k]$. Such n_k exists since $[a_k, b_k]$ contains infinitely many terms from $\{x_n\}$. Then we have

$$|x_{n_k} - a| \leq |x_{n_k} - a_k| + |a_k - a| \leq 2^{-k}M + 2^{-k}M, \tag{71}$$

which shows that the sequence $\{x_{n_k}\}$ converges to a. □

Theorem 31 (Cauchy's criterion)**.** *Let $\{x_n\} \subset \mathbb{R}$ be a Cauchy sequence, in the sense that*

$$|x_n - x_m| \to 0, \qquad \text{as} \quad \min\{n, m\} \to \infty. \tag{72}$$

Then $\{x_n\}$ is convergent.

Proof. Let N be such that $|x_n - x_N| \leq 1$ for all $n \geq N$. Then we have

$$|x_n| \leq |x_N| + 1 \qquad \text{for all} \quad n \geq N, \tag{73}$$

and therefore

$$|x_n| \leq \max\{|x_1|, \ldots, |x_{N-1}|, |x_N| + 1\} \qquad \text{for all} \quad n, \tag{74}$$

meaning that $\{x_n\}$ is bounded. By the Bolzano-Weierstrass theorem (Theorem 30), there is a subsequence $\{x_{n_k}\} \subset \{x_n\}$ that converges to some point $x \in \mathbb{R}$.

So far we only have shown that a subsequence of $\{x_n\}$ converges to x. Now we will show that the whole sequence $\{x_n\}$ indeed converges to x. To this end, let $\varepsilon > 0$, and let N be such that $|x_n - x_m| \leq \varepsilon$ for all $n \geq N$ and $m \geq N$. Moreover, let $k \geq N$ be large enough that $|x_{n_k} - x| \leq \varepsilon$. Then for $m \geq N$, we have

$$|x_m - x| \leq |x_m - x_{n_k}| + |x_{n_k} - x| \leq 2\varepsilon, \tag{75}$$

which shows that the entire sequence $\{x_n\}$ converges to x. □