

QUANTUM COMPUTING FOR COMPUTER SCIENTISTS

Noson S. Yanofsky

Brooklyn College, City University of New York

and

Mirco A. Mannucci

HoloMathics, LLC



CAMBRIDGE
UNIVERSITY PRESS

Complex Numbers

You, have you really understood all that stuff?

What?

The story of imaginary numbers?

Robert Musil, *The Confusions of Young
Törless* (1907)¹

Complex numbers lie at the very core of quantum mechanics and are therefore absolutely essential to a basic understanding of quantum computation. In this chapter we present this important system of numbers from both the algebraic and the geometric standpoints. Section 1.1 presents some motivation and the basic definitions. The algebraic structure and operations on complex numbers are given in Section 1.2. The chapter concludes with Section 1.3, where complex numbers are presented from a geometric point of view and advanced topics are discussed. Our hope is that this chapter will help you get a little closer to what Sir Roger Penrose has very aptly called the “magic of complex numbers” (Penrose, 2005).

Reader Tip. Many readers will find that they are already familiar with some of the material presented in this chapter. The reader who feels confident in her comprehension of the fundamental knowledge of complex numbers, the basic operations, and their properties can safely move on to later chapters. We suggest, though, that you at least skim through the following pages to see what topics are covered. Return to Chapter 1 as a reference when needed (using the index to find specific topics). ♡

¹ For the German-speaking reader, here is the original text (the translation at the beginning is ours):

Du, hast du das vorhin ganz verstanden?

Was?

Die Geschichte mit den imaginären Zahlen?

Musil's *Törless* is a remarkable book. A substantial part is dedicated to the struggle of young Törless to come to grips with mathematics, as well as with his own life. Definitely recommended!

1.1 BASIC DEFINITIONS

The original motivation for the introduction of complex numbers was the theory of algebraic equations, the part of algebra that seeks solutions of polynomial equations. It became readily apparent that there are plenty of cases in which no solution among familiar numbers can be found. Here is the simplest example:

$$x^2 + 1 = 0. \quad (1.1)$$

Indeed, any possible x^2 would be positive or zero. Adding 1 ends up with some quantity to the left that is strictly positive; hence, no solution exists.

Exercise 1.1.1 Verify that the equation $x^4 + 2x^2 + 1 = 0$ has no solution among the real numbers. (Hint: Factor the polynomial.) ■

The aforementioned argument seems to dash any hope of solving Equation (1.1). But does it?

Before building any new number system, it pays to remind ourselves of other sets of numbers that we usually work with

- positive numbers, $\mathbb{P} = \{1, 2, 3, \dots\}$;
- natural numbers, $\mathbb{N} = \{0, 1, 2, 3, \dots\}$;
- integers (or whole numbers), $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$;
- rational numbers, $\mathbb{Q} = \{\frac{m}{n} | m \in \mathbb{Z}, n \in \mathbb{P}\}$;
- real numbers, $\mathbb{R} = \mathbb{Q} \cup \{\dots, \sqrt{2}, \dots, e, \dots, \pi, \dots, \frac{e}{\pi}, \dots\}$;

In none of these familiar number systems can a valid solution to Equation (1.1) be found. Mathematics often works around difficulties by simply *postulating* that such a solution, albeit unknown, is available somewhere. Let us thus boldly assume that this enigmatic solution does indeed exist and determine what it looks like: Equation (1.1) is equivalent to

$$x^2 = -1. \quad (1.2)$$

What does this state? That the solution of Equation (1.1) is a number such that its square is -1 , i.e., a number i such that

$$i^2 = -1 \quad \text{or} \quad i = \sqrt{-1}. \quad (1.3)$$

Of course we know that no such number exists among known (i.e., real) numbers, but we have already stated that this is not going to deter us. We will simply allow this new creature into the realm of well-established numbers and use it as it pleases us. Because it is *imaginary*, it is denoted i . We will impose on ourselves an important restriction: aside from its weird behavior when squared, i will behave just like an ordinary number.

Example 1.1.1 What is the value of i^3 ? We shall treat i as a legitimate number, so

$$i^3 = i \times i \times i = (i^2) \times i = -1 \times i = -i. \quad (1.4)$$

□

Exercise 1.1.2 Find the value of i^{15} . (Hint: Calculate i, i^2, i^3, i^4 , and i^5 . Find a pattern.) ■

In opening the door to our new friend i , we are now flooded with an entire universe of new numbers: to begin with, all the multiples of i by a real number, like $2 \times i$. These fellows, being akin to i , are known as **imaginary numbers**. But there is more: add a real number and an imaginary number, for instance, $3 + 5 \times i$, and you get a number that is neither a real nor an imaginary. Such a number, being a hybrid entity, is rightfully called a **complex number**.

Definition 1.1.1 A complex number is an expression

$$c = a + b \times i = a + bi, \quad (1.5)$$

where a, b are two real numbers; a is called the real part of c , whereas b is its imaginary part. The set of all complex numbers will be denoted as \mathbb{C} . When the \times is understood, we shall omit it.

Complex numbers can be added and multiplied, as shown next.

Example 1.1.2 Let $c_1 = 3 - i$ and $c_2 = 1 + 4i$. We want to compute $c_1 + c_2$ and $c_1 \times c_2$.

$$c_1 + c_2 = 3 - i + 1 + 4i = (3 + 1) + (-1 + 4)i = 4 + 3i. \quad (1.6)$$

Multiplying is not as easy. We must remember to multiply each term of the first complex number with each term of the second complex number. Also, remember that $i^2 = -1$.

$$\begin{aligned} c_1 \times c_2 &= (3 - i) \times (1 + 4i) = (3 \times 1) + (3 \times 4i) + (-i \times 1) + (-i \times 4i) \\ &= (3 + 4) + (-1 + 12)i = 7 + 11i. \end{aligned} \quad (1.7)$$

□

Exercise 1.1.3 Let $c_1 = -3 + i$ and $c_2 = 2 - 4i$. Calculate $c_1 + c_2$ and $c_1 \times c_2$. ■

With addition and multiplication we can get all polynomials. We set out to find a solution for Equation (1.1); it turns out that complex numbers are enough to provide solutions for *all* polynomial equations.

Proposition 1.1.1 (Fundamental Theorem of Algebra). Every polynomial equation of one variable with complex coefficients has a complex solution.

Exercise 1.1.4 Verify that the complex number $-1 + i$ is a solution for the polynomial equation $x^2 + 2x + 2 = 0$. ■

This nontrivial result shows that complex numbers are well worth our attention. In the next two sections, we explore the complex kingdom a little further.

Programming Drill 1.1.1 Write a program that accepts two complex numbers and outputs their sum and their product.

1.2 THE ALGEBRA OF COMPLEX NUMBERS

Admittedly, the fact that we know how to handle them does not explain away the oddity of complex numbers. What *are* they? What does it mean that i squared is equal to -1 ?

In the next section, we see that the geometrical viewpoint greatly aids our intuition. Meanwhile, we would like to convert complex numbers into more familiar objects by carefully looking at how they are built.

Definition 1.1.1 tells us *two* real numbers correspond to each complex number: its real and imaginary parts. A complex number is thus a two-pronged entity, carrying its two components along. How about *defining* a complex number as an ordered pair of reals?

$$c \mapsto (a, b). \quad (1.8)$$

Ordinary real numbers can be identified with pairs $(a, 0)$

$$a \mapsto (a, 0), \quad (1.9)$$

whereas imaginary numbers will be pairs $(0, b)$. In particular,

$$i \mapsto (0, 1). \quad (1.10)$$

Addition is rather obvious – it adds pairs componentwise:

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2). \quad (1.11)$$

Multiplication is a little trickier:

$$(a_1, b_1) \times (a_2, b_2) = (a_1, b_1)(a_2, b_2) = (a_1a_2 - b_1b_2, a_1b_2 + a_2b_1). \quad (1.12)$$

Does this work? Multiplying i by itself gives

$$i \times i = (0, 1) \times (0, 1) = (0 - 1, 0 + 0) = (-1, 0), \quad (1.13)$$

which is what we wanted.

Using addition and multiplication, we can write any complex number in the usual form:

$$c = (a, b) = (a, 0) + (0, b) = (a, 0) + (b, 0) \times (0, 1) = a + bi. \quad (1.14)$$

We have traded one oddity for another: i was previously quite mysterious, whereas now it is just $(0, 1)$. A complex number is nothing more than an ordered pair of ordinary real numbers. Multiplication, though, is rather strange: perhaps the reader would have expected a componentwise multiplication, just like addition. We shall see later that by viewing complex numbers through yet another looking glass the strangeness linked to their multiplication rule will fade away.

Example 1.2.1 Let $c_1 = (3, -2)$ and $c_2 = (1, 2)$. Let us multiply them using the aforementioned rule:

$$\begin{aligned} c_1 \times c_2 &= (3 \times 1 - (-2) \times 2, -2 \times 1 + 2 \times 3) \\ &= (3 + 4, -2 + 6) = (7, 4) = 7 + 4i. \end{aligned} \quad (1.15)$$

□

Exercise 1.2.1 Let $c_1 = (-3, -1)$ and $c_2 = (1, -2)$. Calculate their product. ■

So far, we have a set of numbers and two operations: addition and multiplication. Both operations are **commutative**, meaning that for arbitrary complex numbers c_1 and c_2 ,

$$c_1 + c_2 = c_2 + c_1 \quad (1.16)$$

and

$$c_1 \times c_2 = c_2 \times c_1. \quad (1.17)$$

Both operations are also **associative**:

$$(c_1 + c_2) + c_3 = c_1 + (c_2 + c_3) \quad (1.18)$$

and

$$(c_1 \times c_2) \times c_3 = c_1 \times (c_2 \times c_3). \quad (1.19)$$

Exercise 1.2.2 Verify that multiplication of complex numbers is associative. ■

Moreover, multiplication **distributes** over addition: for all c_1, c_2, c_3 , we have

$$c_1 \times (c_2 + c_3) = (c_1 \times c_2) + (c_1 \times c_3). \quad (1.20)$$

Let us verify this property: first we write the complex numbers as pairs $c_1 = (a_1, b_1)$, $c_2 = (a_2, b_2)$, and $c_3 = (a_3, b_3)$. Now, let us expand the left side

$$\begin{aligned} c_1 \times (c_2 + c_3) &= (a_1, b_1) \times ((a_2, b_2) + (a_3, b_3)) \\ &= (a_1, b_1) \times (a_2 + a_3, b_2 + b_3) \\ &= (a_1 \times (a_2 + a_3) - b_1 \times (b_2 + b_3), \\ &\quad a_1 \times (b_2 + b_3) + b_1 \times (a_2 + a_3)) \\ &= (a_1 \times a_2 + a_1 \times a_3 - b_1 \times b_2 - b_1 \times b_3, \\ &\quad a_1 \times b_2 + a_1 \times b_3 + b_1 \times a_2 + b_1 \times a_3). \end{aligned} \quad (1.21)$$

Turning to the right side of Equation (1.20) one piece at a time gives

$$c_1 \times c_2 = (a_1 \times a_2 - b_1 \times b_2, a_1 \times b_2 + a_2 \times b_1) \quad (1.22)$$

$$c_1 \times c_3 = (a_1 \times a_3 - b_1 \times b_3, a_1 \times b_3 + a_3 \times b_1); \quad (1.23)$$

summing them up we obtain

$$\begin{aligned} c_1 \times c_2 + c_1 \times c_3 &= (a_1 \times a_2 - b_1 \times b_2 + a_1 \times a_3 - b_1 \times b_3, \\ &\quad a_1 \times b_2 + a_2 \times b_1 + a_1 \times b_3 + a_3 \times b_1), \end{aligned} \quad (1.24)$$

which is precisely what we got in Equation (1.21).

Having addition and multiplication, we need their complementary operations: subtraction and division.

Subtraction is straightforward:

$$c_1 - c_2 = (a_1, b_1) - (a_2, b_2) = (a_1 - a_2, b_1 - b_2); \quad (1.25)$$

in other words, subtraction is defined componentwise, as expected.

As for division, we have to work a little: If

$$(x, y) = \frac{(a_1, b_1)}{(a_2, b_2)}, \quad (1.26)$$

then by definition of division as the inverse of multiplication

$$(a_1, b_1) = (x, y) \times (a_2, b_2) \quad (1.27)$$

or

$$(a_1, b_1) = (a_2x - b_2y, a_2y + b_2x). \quad (1.28)$$

So we end up with

$$(1) \quad a_1 = a_2x - b_2y, \quad (1.29)$$

$$(2) \quad b_1 = a_2y + b_2x. \quad (1.30)$$

To determine the answer, we must solve this pair of equations for x and y . Multiply both sides of (1) by a_2 and both sides of (2) by b_2 . We end up with

$$(1') \quad a_1a_2 = a_2^2x - b_2a_2y, \quad (1.31)$$

$$(2') \quad b_1b_2 = a_2b_2y + b_2^2x. \quad (1.32)$$

Now, let us add (1') and (2') to get

$$a_1a_2 + b_1b_2 = (a_2^2 + b_2^2)x. \quad (1.33)$$

Solving for x gives us

$$x = \frac{a_1a_2 + b_1b_2}{a_2^2 + b_2^2}. \quad (1.34)$$

We can perform the same trick for y by multiplying (1) and (2) by b_2 and $-a_2$, respectively, and then summing. We obtain

$$y = \frac{a_2b_1 - a_1b_2}{a_2^2 + b_2^2}. \quad (1.35)$$

In more compact notation, we can express this equation as

$$\frac{a_1 + b_1i}{a_2 + b_2i} = \frac{a_1a_2 + b_1b_2}{a_2^2 + b_2^2} + \frac{a_2b_1 - a_1b_2}{a_2^2 + b_2^2}i. \quad (1.36)$$

Notice that both x and y are calculated using the same denominator, namely, $a_2^2 + b_2^2$. We are going to see what this quantity means presently. In the meantime, here is a concrete example.

Example 1.2.2 Let $c_1 = -2 + i$ and $c_2 = 1 + 2i$. We will compute $\frac{c_1}{c_2}$. In this case, $a_1 = -2$, $b_1 = 1$, $a_2 = 1$, and $b_2 = 2$. Therefore,

$$a_2^2 + b_2^2 = 1^2 + 2^2 = 5, \quad (1.37)$$

$$a_1a_2 + b_1b_2 = -2 \times 1 + 1 \times 2 = 0, \quad (1.38)$$

$$a_2b_1 - a_1b_2 = 1 \times 1 - (-2) \times 2 = 1 + 4 = 5. \quad (1.39)$$

The answer is thus $(\frac{0}{5}, \frac{5}{5}) = (0, 1) = i$. □

Exercise 1.2.3 Let $c_1 = 3i$ and $c_2 = -1 - i$. Calculate $\frac{c_1}{c_2}$. ■

Now, let us go back to the mysterious denominator in the quotient formula in Equation (1.36). Real numbers have a unary operation, the absolute value, given by

$$|a| = +\sqrt{a^2}. \quad (1.40)$$

We can define a generalization of this operation² to the complex domain by letting

$$|c| = |a + bi| = +\sqrt{a^2 + b^2}. \quad (1.41)$$

This quantity is known as the **modulus** of a complex number.

Example 1.2.3 What is the modulus of $c = 1 - i$?

$$|c| = |1 - i| = +\sqrt{1^2 + (-1)^2} = \sqrt{2}. \quad (1.42)$$

□

The geometric meaning of the modulus is discussed in the next section. For now, we remark that the quantity in the denominator of the quotient of two complex numbers is nothing more than the modulus squared of the divisor:

$$|c|^2 = a^2 + b^2. \quad (1.43)$$

This modulus must be different from zero, which always happens unless the divisor is itself zero.

Exercise 1.2.4 Calculate the modulus of $c = 4 - 3i$. ■

Exercise 1.2.5 Verify that given two arbitrary complex numbers c_1 and c_2 , the following equality always holds:

$$|c_1||c_2| = |c_1c_2|. \quad (1.44)$$

■

Exercise 1.2.6 Prove that

$$|c_1 + c_2| \leq |c_1| + |c_2|. \quad (1.45)$$

When are they, in fact, equal? (Hint: Square both sides.) ■

Exercise 1.2.7 Show that for all $c \in \mathbb{C}$, we have $c + (0, 0) = (0, 0) + c = c$. That is, $(0, 0)$ is an additive identity. ■

² The definition given in Equation (1.40) is entirely equivalent to the more familiar one: $|a| = a$ if $a \geq 0$, and $|a| = -a$ if $a < 0$.

Exercise 1.2.8 Show that for all $c \in \mathbb{C}$ we have $c \times (1, 0) = (1, 0) \times c = c$. That is, $(1, 0)$ is a multiplicative identity. ■

In summation, we have defined a new set of numbers, \mathbb{C} , endowed with four operations, verifying the following properties:

- (i) Addition is commutative and associative.
- (ii) Multiplication is commutative and associative.
- (iii) Addition has an identity: $(0, 0)$.
- (iv) Multiplication has an identity: $(1, 0)$.
- (v) Multiplication distributes with respect to addition.
- (vi) Subtraction (i.e., the inverse of addition) is defined everywhere.
- (vii) Division (i.e., the inverse of multiplication) is defined everywhere except when the divisor is zero.

A set with operations satisfying all these properties is called a **field**. \mathbb{C} is a field, just like \mathbb{R} , the field of real numbers. In fact, via the identification that associates a real number to a complex number with 0 as the imaginary component, we can think of \mathbb{R} as a subset³ of \mathbb{C} . \mathbb{R} sits inside \mathbb{C} ; but \mathbb{C} is a vast field, so vast, indeed, that all polynomial equations with coefficients in \mathbb{C} have a solution in \mathbb{C} itself. \mathbb{R} is also a roomy field, but not enough to enjoy this last property (remember Equation (1.1)). A field that contains all solutions for any of its polynomial equations is said to be **algebraically complete**. \mathbb{C} is an algebraically complete field, whereas \mathbb{R} is not.

There is a unary operation that plays a crucial role in the complex domain. The reader is familiar with “changing signs” of real numbers. Here, however, there are *two* real numbers attached to a complex number. Therefore, there are *three* ways of changing sign: either change the sign of the real part or change the sign of the imaginary part, or both. Let us analyze them one by one.

Changing both signs of the complex number is done by multiplying by the number $-1 = (-1, 0)$.

Exercise 1.2.9 Verify that multiplication by $(-1, 0)$ changes the sign of the real and imaginary components of a complex number. ■

Changing the sign of the imaginary part only is known as **conjugation**.⁴ If $c = a + bi$ is an arbitrary complex number, then the conjugate of c is $\bar{c} = a - bi$. Two numbers related by conjugation are said to be **complex conjugates** of each other.

Changing the sign of the real part ($c \mapsto -\bar{c}$) has no particular name, at least in the algebraic context.⁵

The following exercises will guide you through conjugation’s most important properties.

³ A subset of a field that is a field in its own right is called a **subfield**: \mathbb{R} is a subfield of \mathbb{C} .

⁴ Its “geometric” name is **real-axis reflection**. The name becomes obvious in the next section.

⁵ In the geometric viewpoint, it is known as **imaginary-axis reflection**. After reading Section 1.3, we invite you to investigate this operation a bit further.

Exercise 1.2.10 Show that conjugation respects addition, i.e.,

$$\overline{c_1 + c_2} = \overline{c_1} + \overline{c_2}. \quad (1.46)$$

■

Exercise 1.2.11 Show that conjugation respects multiplication, i.e.,

$$\overline{c_1 \times c_2} = \overline{c_1} \times \overline{c_2}. \quad (1.47)$$

■

Notice that the function

$$c \mapsto \bar{c} \quad (1.48)$$

given by conjugation is **bijective**, i.e., is one-to-one and onto. Indeed, two different complex numbers are never sent to the same number by conjugation. Moreover, every number is the complex conjugate of some number. A function from a field to a field that is bijective and that respects addition and multiplication is known as a **field isomorphism**. Conjugation is thus a field isomorphism of \mathbb{C} to \mathbb{C} .

Exercise 1.2.12 Consider the operation given by flipping the sign of the real part. Is this a field isomorphism of \mathbb{C} ? If yes, prove it. Otherwise, show where it fails. ■

We cannot continue without mentioning another property of conjugation:

$$c \times \bar{c} = |c|^2. \quad (1.49)$$

In words, the modulus squared of a complex number is obtained by multiplying the number with its conjugate. For example,

$$(3 + 2i) \times (3 - 2i) = 3^2 + 2^2 = 13 = |3 + 2i|^2. \quad (1.50)$$

We have covered what we need from the algebraic perspective. We see in the next section that the geometric approach sheds some light on virtually all topics touched on here.

Programming Drill 1.2.1 *Take the program that you wrote in the last programming drill and make it also perform subtraction and division of complex numbers. In addition, let the user enter a complex number and have the computer return its modulus and conjugate.*

1.3 THE GEOMETRY OF COMPLEX NUMBERS

As far as algebra is concerned, complex numbers are an algebraically complete field, as we have described them in Section 1.2. That alone would render them invaluable as a mathematical tool. It turns out that their significance extends far beyond the algebraic domain and makes them equally useful in geometry and hence in physics. To see why this is so, we need to look at a complex number in yet another way. At the beginning of Section 1.2, we learned that a complex number is a pair of real

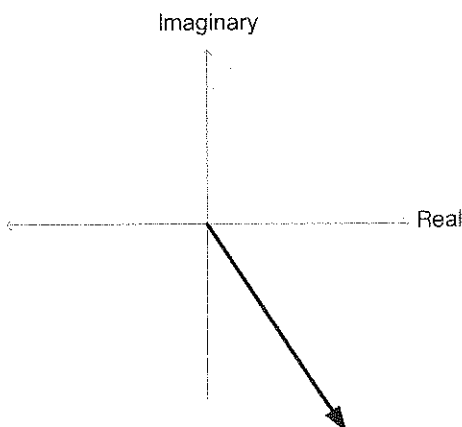


Figure 1.1. Complex plane.

numbers. This suggests a natural means of representation: real numbers are placed on the line, so pairs of reals correspond to points on the plane, or, equivalently, correspond to **vectors** starting from the origin and pointing to that point (as shown in Figure 1.1).

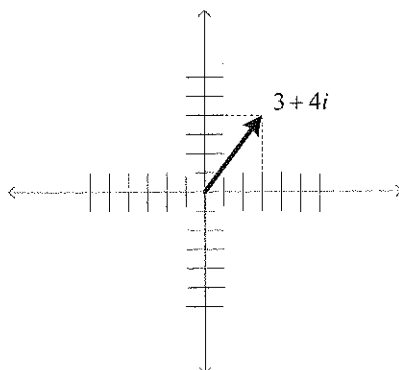
In this representation, real numbers (i.e., complex numbers with no imaginary part) sit on the horizontal axis and imaginary numbers sit on the vertical axis. This plane is known as the **complex plane** or the **Argand plane**.

Through this representation, the algebraic properties of the complex numbers can be seen in a new light. Let us start with the modulus: it is nothing more than the **length** of the vector. Indeed, the length of a vector, via Pythagoras' theorem, is the square root of the sum of the squares of its edges, which is precisely the modulus, as defined in the previous section.

Example 1.3.1 Consider the complex numbers $c = 3 + 4i$ depicted in Figure 1.2. The length of the vector is the hypotenuse of the right triangle whose edges have length 3 and 4, respectively. Pythagoras' theorem gives us the length as

$$\text{length}(c) = \sqrt{4^2 + 3^2} = \sqrt{16 + 9} = \sqrt{25} = 5. \quad (1.51)$$

This is exactly the modulus of c . □

Figure 1.2. Vector $3 + 4i$.

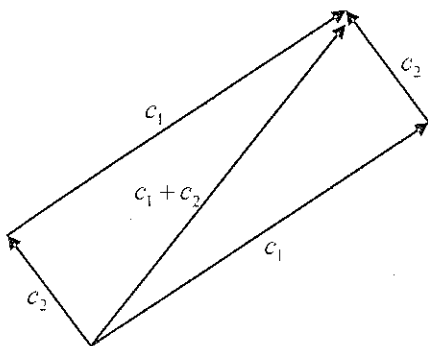


Figure 1.3. Parallelogram rule.

Next comes addition; vectors can be added using the so-called **parallelogram rule** illustrated by Figure 1.3. In words, draw the parallelogram whose parallel edges are the two vectors to be added; their sum is the diagonal.

Exercise 1.3.1 Draw the complex numbers $c_1 = 2 - i$ and $c_2 = 1 + i$ in the complex plane, and add them using the parallelogram rule. Verify that you would get the same result as adding them algebraically (the way we learned in Section 1.2). ■

Subtraction too has a clear geometric meaning: subtracting c_2 from c_1 is the same as adding the negation of c_2 , i.e., $-c_2$, to c_1 . But what is the negation of a vector? It is just the vector of the same length pointed in the opposite direction (see Figure 1.4).

Exercise 1.3.2 Let $c_1 = 2 - i$ and $c_2 = 1 + i$. Subtract c_2 from c_1 by first drawing $-c_2$ and then adding it to c_1 using the parallelogram rule. ■

To give a simple geometrical meaning to multiplication, we need to develop yet another characterization of complex numbers. We saw a moment ago that for every complex number we can draw a right triangle, whose edges' lengths are the real and imaginary parts of the number and whose hypotenuse's length is the modulus. Now, suppose someone tells us the modulus of the number what else do we need to know to draw the triangle? The answer is the angle at the origin.

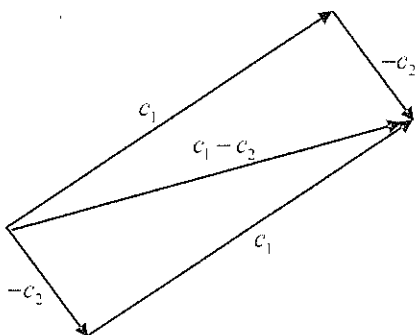


Figure 1.4. Subtraction.

The modulus ρ and the angle θ (notice: two real numbers, as before) are enough to uniquely determine the complex number.

$$(a, b) \mapsto (\rho, \theta). \quad (1.52)$$

We know how to compute ρ from a, b :

$$\rho = \sqrt{a^2 + b^2}. \quad (1.53)$$

θ is also easy, via trigonometry:

$$\theta = \tan^{-1} \left(\frac{b}{a} \right). \quad (1.54)$$

The (a, b) representation is known as the **Cartesian representation** of a complex number, whereas (ρ, θ) is the **polar representation**.

We can go back from polar to Cartesian representation, again using trigonometry:

$$a = \rho \cos(\theta), \quad b = \rho \sin(\theta). \quad (1.55)$$

Example 1.3.2 Let $c = 1 + i$. What is its polar representation?

$$\rho = \sqrt{1^2 + 1^2} = \sqrt{2} \quad (1.56)$$

$$\theta = \tan^{-1} \left(\frac{1}{1} \right) = \tan^{-1}(1) = \frac{\pi}{4} \quad (1.57)$$

c is the vector of length $\sqrt{2}$ from the origin at an angle of $\frac{\pi}{4}$ radians, or 45° . □

Exercise 1.3.3 Draw the complex number given by the polar coordinates $\rho = 3$ and $\theta = \frac{\pi}{3}$. Compute its Cartesian coordinates. ■

Programming Drill 1.3.1 Write a program that converts a complex number from its Cartesian representation to its polar representation and vice versa.

Before moving on, let us meditate a little: what kind of insight does the polar representation give us? Instead of providing a ready-made answer, let us begin with a question: how many complex numbers share exactly the same modulus? A moment's thought will tell us that for a *fixed* modulus, say, $\rho = 1$, there is an entire circle centered at the origin (as shown in Figure 1.5).

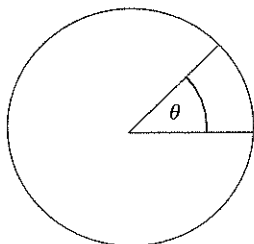


Figure 1.5. Phase θ .

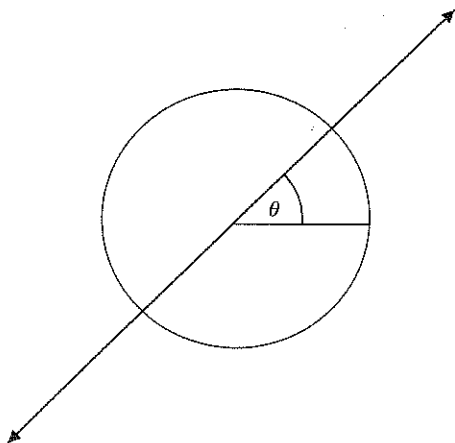


Figure 1.6. Points on a line with the same phase.

So, here comes the angle: imagine the circle as your watch, and the complex number as the needle. Angle θ tells us the “time.” The “time” is known in physics and engineering as the **phase**, whereas the length of the “needle” (i.e., the modulus) is the **magnitude** of the number.

Definition 1.3.1 A complex number is a magnitude and a phase.

The ordinary positive reals are just complex numbers such that their phase is zero. The negative reals have phase π . By the same token, imaginary numbers are numbers with constant phase equal to $\frac{\pi}{2}$ (positive imaginary) or $\frac{3\pi}{2}$ (negative imaginary).

Given a constant phase, there is an entire line of complex numbers having that phase as depicted in Figure 1.6.

Observe that a complex number has a unique polar representation only if we confine the phase between 0 and 2π :

$$0 \leq \theta < 2\pi \quad (1.58)$$

(and the $\rho \geq 0$). If we restrict θ in this fashion, though, we cannot in general add angles (the sum may be bigger than 2π). A better course is to let the angle be anything and *reduce* it modulo 2π :

$$\theta_1 = \theta_2 \quad \text{if and only if } \theta_2 = \theta_1 + 2\pi k, \text{ for some integer } k. \quad (1.59)$$

Two complex numbers in polar representations will be identical if their magnitude is the same and if the angles are the same modulo 2π , as shown by the following example.

Example 1.3.3 Are the numbers $(3, -\pi)$ and $(3, \pi)$ the same? Indeed they are: their magnitude is the same and their phases differ by $(-\pi) - \pi = -2\pi = (-1)2\pi$. \square

We are now ready for multiplication: given two complex numbers in polar coordinates, (ρ_1, θ_1) and (ρ_2, θ_2) , their product can be obtained by simply multiplying their magnitude and *adding* their phase:

$$(\rho_1, \theta_1) \times (\rho_2, \theta_2) = (\rho_1 \rho_2, \theta_1 + \theta_2). \quad (1.60)$$

Example 1.3.4 Let $c_1 = 1 + i$ and $c_2 = -1 + i$. Their product, according to the algebraic rule, is

$$c_1 c_2 = (1 + i)(-1 + i) = -2 + 0i = -2. \quad (1.61)$$

Now, let us take their polar representation

$$c_1 = \left(\sqrt{2}, \frac{\pi}{4}\right), \quad c_2 = \left(\sqrt{2}, \frac{3\pi}{4}\right). \quad (1.62)$$

(Carry out the calculations!) Therefore, their product using the rule described earlier is

$$c_1 c_2 = \left(\sqrt{2} \times \sqrt{2}, \frac{\pi}{4} + \frac{3\pi}{4}\right) = (2, \pi). \quad (1.63)$$

If we revert to its Cartesian coordinates, we get

$$(2 \times \cos(\pi), 2 \times \sin(\pi)) = (-2, 0), \quad (1.64)$$

which is precisely the answer we arrived at with the algebraic calculation in Equation (1.61).

Figure 1.7 is the graphical representation of the two numbers and their product.

As you can see, we simply rotated the first vector by an angle equal to the phase of the second vector and multiplied its length by the length of the second vector. \square

Exercise 1.3.4 Multiply $c_1 = -2 - i$ and $c_2 = -1 - 2i$ using both the algebraic and the geometric method; verify that the results are identical. \blacksquare

Reader Tip. Most of the rest of this chapter are basic ideas in complex numbers; however, they will not really be used in the text. The part on roots of unity will arise in our discussion of Shor's algorithm (Section 6.5). The rest is included for the sake

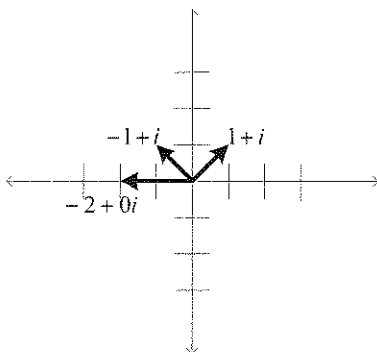


Figure 1.7. Two complex numbers and their product.

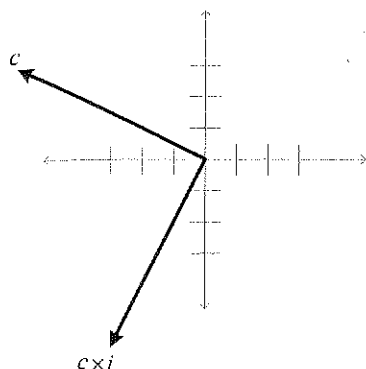


Figure 1.8. Multiplication by i .

of completeness. The restless reader can safely skim the rest of this chapter on the first reading. ♥

We have implicitly learned an important fact: multiplication in the complex domain has something to do with *rotations* of the complex plane. Indeed, observe just what happens by left or right multiplication by i :

$$c \mapsto c \times i. \quad (1.65)$$

i has modulus 1, so the magnitude of the result is exactly equal to that of the starting point. The phase of i is $\frac{\pi}{2}$, so multiplying by i has the net result of rotating the original complex number by 90° , a right angle. The same happens when we multiply any complex number; so we can safely conclude that multiplication by i is a right-angle counterclockwise rotation of the complex plane, as shown in Figure 1.8.

Exercise 1.3.5 Describe the geometric effect on the plane obtained by multiplying by a real number, i.e., the function

$$c \mapsto c \times r_0, \quad (1.66)$$

where r_0 is a fixed real number. ■

Exercise 1.3.6 Describe the geometric effect on the plane obtained by multiplying by a generic complex number, i.e., the function

$$c \mapsto c \times c_0, \quad (1.67)$$

where c_0 is a fixed complex number. ■

Programming Drill 1.3.2 If you like graphics, write a program that accepts a small drawing around the origin of the complex plane and a complex number. The program should change the drawing by multiplying every point of the diagram by a complex number.

Now that we are armed with a geometric way of looking at multiplication, we can tackle division as well. After all, division is nothing more than the inverse operation of multiplication. Assume that

$$c_1 = (\rho_1, \theta_1) \quad \text{and} \quad c_2 = (\rho_2, \theta_2), \quad (1.68)$$

are two complex numbers in polar form; what is the polar form of $\frac{c_1}{c_2}$? A moment's thought tells us that it is the number

$$\frac{c_1}{c_2} = \left(\frac{\rho_1}{\rho_2}, \theta_1 - \theta_2 \right). \quad (1.69)$$

In words, we divide the magnitudes and subtract the angles.

Example 1.3.5 Let $c_1 = -1 + 3i$ and $c_2 = -1 - 4i$. Let us calculate their polar coordinates first:

$$c_1 = \left(\sqrt{(-1)^2 + 3^2}, \tan^{-1} \left(\frac{3}{-1} \right) \right) = (\sqrt{10}, \tan^{-1}(-3)) = (3.1623, 1.8925), \quad (1.70)$$

$$c_2 = \left(\sqrt{(-1)^2 + (-4)^2}, \tan^{-1} \left(\frac{-4}{-1} \right) \right) = (\sqrt{17}, \tan^{-1}(4)) = (4.1231, -1.8158), \quad (1.71)$$

therefore, in polar coordinates the quotient is

$$\frac{c_1}{c_2} = \left(\frac{3.1623}{4.1231}, 1.8925 - (-1.8158) \right) = (0.7670, 3.7083). \quad (1.72)$$

□

Exercise 1.3.7 Divide $2 + 2i$ by $1 - i$ using both the algebraic and the geometrical method and verify that the results are the same. ■

You may have noticed that in Section 1.2, we have left out two important operations: powers and roots. The reason was that it is much easier to deal with them in the present geometric setting than from the algebraic viewpoint.

Let us begin with powers. If $c = (\rho, \theta)$ is a complex number in polar form and n a positive integer, its n th power is just

$$c^n = (\rho^n, n\theta), \quad (1.73)$$

because raising to the n th power is multiplying n times. Figure 1.9 shows a complex number and its first, second, and third powers.

Exercise 1.3.8 Let $c = 1 - i$. Convert it to polar coordinates, calculate its fifth power, and revert the answers to Cartesian coordinates. ■

What happens when the base is a number of magnitude 1? Its powers will also have magnitude 1; thus, they will stay on the same unit circle. You can think of the various powers $1, 2, \dots$ as time units, and a needle moving counterclockwise at

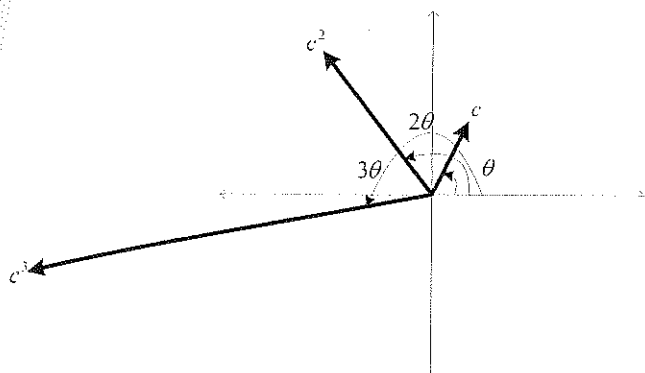


Figure 1.9. A complex number and its square and cube.

constant speed (it covers exactly θ radians per time unit, where θ is the phase of the base).

Let us move on to roots. As you know already from high-school algebra, a root is a fractional power. For instance, the square root means raising the base to the power of one-half; the cube root is raising to the power of one-third; and so forth. The same holds true here, so we may take roots of complex numbers: if $c = (\rho, \theta)$ is a complex in polar form, its n th root is

$$c^{\frac{1}{n}} = \left(\rho^{\frac{1}{n}}, \frac{1}{n}\theta \right). \quad (1.74)$$

However, things get a bit more complicated. Remember, the phase is defined only up to multiples of 2π . Therefore, we must rewrite Equation (1.74) as

$$c^{\frac{1}{n}} = \left(\sqrt[n]{\rho}, \frac{1}{n}(\theta + k2\pi) \right). \quad (1.75)$$

It appears that there are *several* roots of the same number. This fact should not surprise us: in fact, even among real numbers, roots are not always unique. Take, for instance, the number 2 and notice that there are two square roots, $\sqrt{2}$ and $-\sqrt{2}$.

How many n th roots are there? There are precisely n n th roots for a complex number. Why? Let us go back to Equation (1.75).

$$\frac{1}{n}(\theta + 2k\pi) = \frac{1}{n}\theta + \frac{k}{n}2\pi. \quad (1.76)$$

How many different solutions can we generate by varying k ? Here they are:

$k = 0$	$\frac{1}{n}\theta$	(1.77)
$k = 1$	$\frac{1}{n}\theta + \frac{1}{n}2\pi$	
\vdots	\vdots	
$k = n - 1$	$\frac{1}{n}\theta + \frac{n-1}{n}2\pi$	

That is all: when $k = n$, we obtain the first solution; when $k = n + 1$, we obtain the second solution; and so forth. (Verify this statement!)

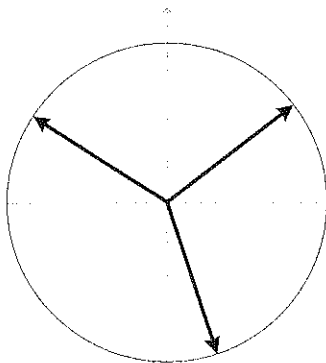


Figure 1.10. The three cube roots of unity.

To see what is happening, let us assume that $\rho = 1$; in other words, let us find n th roots of a complex number $c = (1, \theta)$ on the unit circle. The n solutions in Equation (1.77) can be interpreted in the following way: Draw the unit circle, and the vectors whose phase is $\frac{1}{n}\theta$, $\frac{1}{n}\theta$ plus an angle equal to $\frac{k}{n}$ of the entire circle, where $k = 1, \dots, n$. We get precisely the vertices of a regular polygon with n edges. Figure 1.10 is an example when $n = 3$.

Exercise 1.3.9 Find all the cube roots of $c = 1 + i$. ■

By now we should feel pretty comfortable with the polar representation: we know that any complex number, via the polar-to-Cartesian function, can be written as

$$c = \rho(\cos(\theta) + i \sin(\theta)). \quad (1.78)$$

Let us introduce yet another notation that will prove to be very handy in many situations. The starting point is the following formula, known as **Euler's formula**:

$$e^{i\theta} = \cos(\theta) + i \sin(\theta). \quad (1.79)$$

The full justification of the remarkable formula of Euler lies outside the scope of this book.⁶ However, we can at least provide some evidence that substantiates its

⁶ For the calculus-savvy reader: Use the well-known Taylor expansions.

$$e^x = 1 + x + \frac{x^2}{2} + \dots + \frac{x^n}{n!} + \dots, \quad (1.80)$$

$$\sin(x) = x - \frac{x^3}{3!} + \dots + \frac{(-1)^n}{(2n+1)!} x^{2n+1} + \dots, \quad (1.81)$$

$$\cos(x) = 1 - \frac{x^2}{2} + \dots + \frac{(-1)^n}{(2n)!} x^{2n} + \dots. \quad (1.82)$$

Assume that they hold for complex values of x . Now, formally multiply $\sin(x)$ by i and add componentwise $\cos(x)$ to obtain Euler's formula.

validity. First of all, if $\theta = 0$, we get what we expected, namely, 1. Secondly,

$$\begin{aligned}
 e^{i(\theta_1 + \theta_2)} &= \cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2) \\
 &= \cos(\theta_1) \cos(\theta_2) - \sin(\theta_1) \sin(\theta_2) \\
 &\quad + i((\sin(\theta_1) \cos(\theta_2) + \sin(\theta_2) \cos(\theta_1))) \\
 &= (\cos(\theta_1) + i \sin(\theta_1))(\cos(\theta_2) + i \sin(\theta_2)) \\
 &= e^{i\theta_1} \times e^{i\theta_2}.
 \end{aligned} \tag{1.83}$$

In other words, the exponential function takes sums into products as it does in the real case.

Exercise 1.3.10 Prove De Moivre's formula:

$$(e^{i\theta})^n = \cos(n\theta) + i \sin(n\theta). \tag{1.84}$$

(Hint: The trigonometric identities used earlier, with induction on n , will do the work.) ■

Now that we know how to take the exponential of an imaginary number, there is no problem in defining the exponential of an arbitrary complex number:

$$e^{a+bi} = e^a \times e^{bi} = e^a(\cos(b) + i \sin(b)). \tag{1.85}$$

Euler's formula enables us to rewrite Equation (1.78) in a more compact form:

$$c = \rho e^{i\theta}. \tag{1.86}$$

We shall refer to Equation (1.86) as the **exponential form** of a complex number.

Exercise 1.3.11 Write the number $c = 3 - 4i$ in exponential form. ■

The exponential notation simplifies matters when we perform multiplication:

$$c_1 c_2 = \rho_1 e^{i\theta_1} \rho_2 e^{i\theta_2} = \rho_1 \rho_2 e^{i(\theta_1 + \theta_2)}. \tag{1.87}$$

Exercise 1.3.12 Rewrite the law for dividing complex numbers in exponential form. ■

With this notation, we can look at the roots of the complex number $1 = (1, 0) = 1 + 0i$. Let n be a fixed number. There are n different **roots of unity**. Setting $c = (1, 0)$ in Equation (1.75), we get

$$c^{\frac{1}{n}} = (1, 0)^{\frac{1}{n}} = \left(\sqrt[n]{1}, \frac{1}{n}(0 + 2k\pi) = \left(1, \frac{2k\pi}{n} \right) \right). \tag{1.88}$$

By permitting $k = 0, 1, 2, \dots, n-1$, we get n different roots of unity. Notice that if we set $k = n$, we get back to the first one. The k th root of unity in exponential form is $e^{2\pi i k/n}$. We denote these n different roots of unity by

$$\omega_n^0 = 1, \omega_n^1, \omega_n^2, \dots, \omega_n^{n-1}. \tag{1.89}$$

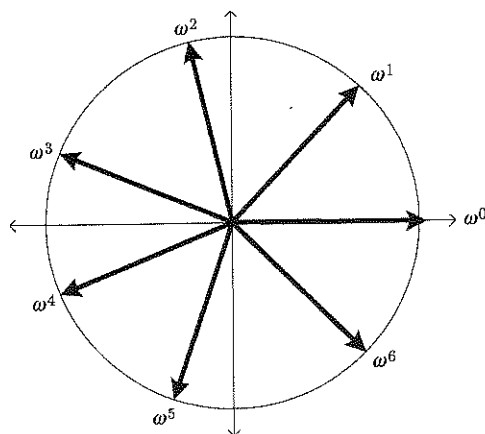


Figure 1.11. The seventh root of unity and its powers.

Geometrically these n roots of unity correspond to splitting up the unit circle into n parts where the first partition is $(1, 0)$. Figure 1.11 is a picture of the seventh root of unity and all its powers.

If we multiply two roots of unity, we get

$$\omega_n^j \omega_n^k = e^{2\pi i j/n} e^{2\pi i k/n} = e^{2\pi i (j+k)/n} = \omega_n^{j+k}. \quad (1.90)$$

Notice also that

$$\omega_n^j \omega_n^{n-j} = \omega_n^n = 1, \quad (1.91)$$

and hence

$$\overline{\omega_n^j} = \omega_n^{n-j}. \quad (1.92)$$

Exercise 1.3.13 Draw all the fifth roots of unity. ■

We are now in a position to characterize geometrically any function on complex numbers. The simplest functions one can think of beyond the elementary operations are **polynomials**. An arbitrary polynomial with complex coefficients looks like

$$P(x) = c_n x^n + c_{n-1} x^{n-1} + \cdots + c_0, \quad (1.93)$$

where c_0, c_1, \dots, c_{n-1} are in \mathbb{C} . $P(x)$ can be seen as a function from \mathbb{C} to \mathbb{C}

$$P(x) : \mathbb{C} \longrightarrow \mathbb{C}. \quad (1.94)$$

To build some geometric intuition on polynomials, you can try your hand at the following two exercises.

Exercise 1.3.14 Describe the geometric meaning of the function

$$c \longmapsto c^n \quad (1.95)$$

from \mathbb{C} to \mathbb{C} . ■

Exercise 1.3.15 Describe the geometric meaning of the function

$$c \mapsto c + c_0 \quad (1.96)$$

from \mathbb{C} to \mathbb{C} . ■

After polynomials, the next set of functions are **rational functions**, or quotients of polynomials:

$$R(x) = \frac{P_0(x)}{P_1(x)} = \frac{c_n x^n + c_{n-1} x^{n-1} + \cdots + c_0}{d_m x^m + d_{m-1} x^{m-1} + \cdots + d_0}. \quad (1.97)$$

In general, describing the action on the plane provided by a rational function is no simple matter. The simplest case, though, is relatively easy and a very important one:

$$R_{a,b,c,d}(x) = \frac{ax + b}{cx + d}, \quad (1.98)$$

where a, b, c, d are in \mathbb{C} and $ad - bc \neq 0$ is known as the **Möbius transformation**. The following set of exercises introduce some of the basic properties of Möbius transformations. (In particular, we show that the set of Möbius transformations form a group.)⁷

Exercise 1.3.16 When $a = d = 0$ and $b = c = 1$, we get $R(x) = \frac{1}{x}$. Describe the geometrical effect of this transformation. (Hint: See what happens to the points inside and outside the circle of radius 1.) ■

Exercise 1.3.17 Prove that the composition of two Möbius transformations is a Möbius transformation. In other words, if $R_{a,b,c,d}$ and $R_{a',b',c',d'}$ are two Möbius transformations, the transformation $R_{a',b',c',d'} \circ R_{a,b,c,d}$ given by

$$R_{a',b',c',d'} \circ R_{a,b,c,d}(x) = R_{a',b',c',d'}(R_{a,b,c,d}(x)) \quad (1.99)$$

is also a Möbius transformation. ■

Exercise 1.3.18 Show that the identity transformation, i.e., the transformation that leaves every point fixed, is a Möbius transformation. ■

Exercise 1.3.19 Show that each Möbius transformation has an inverse that is also a Möbius transformation, i.e., for each $R_{a,b,c,d}$ you can find $R_{a',b',c',d'}$ such that

$$R_{a',b',c',d'} \circ R_{a,b,c,d}(x) = x. \quad (1.100)$$

■

There are many more functions in the complex domain, but to introduce them one needs tools from **complex analysis**, i.e., calculus over the complex numbers. The main idea is quite simple: replace polynomials with a power series, i.e., polynomials

⁷ Möbius transformations are a truly fascinating topic, and perhaps the best entrance door to the geometry of complex numbers. We invite you to find out more about them in Schwerdtfeger (1980).

with an infinite number of terms. The functions one studies are the so-called **analytic functions**, which are functions that can be coherently pieced together from small parts, each of which is represented by a series.

Programming Drill 1.3.3 *Expand your program. Add functions for multiplication, division, and returning the polar coordinates of a number.*

We have covered the basic language of complex numbers. Before we embark on our quantum journey, we need another tool: vector spaces over the complex field.

.....
References: Most of the material found in this chapter can be found in any calculus or linear algebra textbook. References for some of the more advanced material presented at the end of the chapter can be found in, e.g., Bak and Newman (1996), Needham (1999), Schwerdtfeger (1980), and Silverman (1984).

The history of complex numbers goes back to the mid-sixteenth century during the Italian Renaissance. The story of Tartaglia, Cardano, Bombelli and their effort to solve algebraic equations is well worth reading. Some of this fascinating tale is in Nahin (1998), Mazur (2002), and several wonderful sections in Penrose (1994).

Complex Vector Spaces

Philosophy is written in that great book which continually lies open before us (I mean the Universe). But one cannot understand this book until one has learned to understand the language and to know the letters in which it is written. It is written in the language of mathematics, and the letters are triangles, circles and other geometric figures. Without these means it is impossible for mankind to understand a single word; without these means there is only vain stumbling in a dark labyrinth.¹

Galileo Galilei

Quantum theory is cast in the language of complex vector spaces. These are mathematical structures that are based on complex numbers. We learned all that we need about such numbers in Chapter 1. Armed with this knowledge, we can now tackle complex vector spaces themselves.

Section 2.1 goes through the main example of a (finite-dimensional) complex vector space at tutorial pace. Section 2.2 provides formal definitions, basic properties, and more examples. Each of Section 2.3 through Section 2.7 discusses an advanced topic.

.....
Reader Tip. The reader might find some of this chapter to be “just boring math.” If you are eager to leap into the quantum world, we suggest reading the first two or three sections before moving on to Chapter 3. Return to Chapter 2 as a reference when needed (using the index and the table of contents to find specific topics). ♡

¹ ... *La filosofia é scritta in questo grandissimo libro che continuamente ci sta aperto innanzi a gli occhi (io dico l'universo), ma non si puo intendere se prima non s'impara a intender la lingua, e conoscer i caratteri, ne' quali é scritto. Egli é scritto in lingua matematica, e i caratteri sono triangoli, cerchi, ed altre figure geometriche, senza i quali mezzi é impossibile a intenderne umanamente parola; senza questi é un aggirarsi vanamente per un'oscuro laberinto...* (Opere Il Saggiatore p. 171).

A small disclaimer is in order. The theory of complex vector spaces is a vast and beautiful subject. Lengthy textbooks have been written on this important area of mathematics. It is impossible to provide anything more than a small glimpse into the beauty and profundity of this topic in one chapter. Rather than “teaching” our reader complex vector spaces, we aim to cover the bare minimum of concepts, terminology, and notation needed in order to start quantum computing. It is our sincere hope that reading this chapter will inspire further investigation into this remarkable subject.

2.1 \mathbb{C}^n AS THE PRIMARY EXAMPLE

The primary example of a complex vector space is the set of vectors (one-dimensional arrays) of a fixed length with complex entries. These vectors will describe the states of quantum systems and quantum computers. In order to fix our ideas and to see clearly what type of structure this set has, let us carefully examine one concrete example: the set of vectors of length 4. We shall denote this set as $\mathbb{C}^4 = \mathbb{C} \times \mathbb{C} \times \mathbb{C} \times \mathbb{C}$, which reminds us that each vector is an ordered list of four complex numbers.

A typical element of \mathbb{C}^4 looks like this:

$$\begin{bmatrix} 6 - 4i \\ 7 + 3i \\ 4.2 - 8.1i \\ -3i \end{bmatrix}. \quad (2.1)$$

We might call this vector V . We denote the j th element of V as $V[j]$. The top row is row number 0 (not 1);² hence, $V[1] = 7 + 3i$.

What types of operations can we carry out with such vectors? One operation that seems obvious is to form the **addition** of two vectors. For example, given two vectors of \mathbb{C}^4

$$V = \begin{bmatrix} 6 - 4i \\ 7 + 3i \\ 4.2 - 8.1i \\ -3i \end{bmatrix} \quad \text{and} \quad W = \begin{bmatrix} 16 + 2.3i \\ -7i \\ 6 \\ -4i \end{bmatrix}, \quad (2.2)$$

² Computer scientists generally start indexing their rows and columns at 0. In contrast, mathematicians and physicists tend to start indexing at 1. The difference is irrelevant. We shall generally follow the computer science convention (after all, this is a computer science text).

we can add them to form $V + W \in \mathbb{C}^4$ by adding their respective entries:

$$\begin{bmatrix} 6 - 4i \\ 7 + 3i \\ 4.2 - 8.1i \\ -3i \end{bmatrix} + \begin{bmatrix} 16 + 2.3i \\ -7i \\ 6 \\ -4i \end{bmatrix} = \begin{bmatrix} (6 - 4i) + (16 + 2.3i) \\ (7 + 3i) + (-7i) \\ (4.2 - 8.1i) + (6) \\ (-3i) + (-4i) \end{bmatrix} = \begin{bmatrix} 22 - 1.7i \\ 7 - 4i \\ 10.2 - 8.1i \\ -7i \end{bmatrix}. \quad (2.3)$$

Formally, this operation amounts to

$$(V + W)[j] = V[j] + W[j]. \quad (2.4)$$

Exercise 2.1.1 Add the following two vectors:

$$\begin{bmatrix} 5 + 13i \\ 6 + 2i \\ 0.53 - 6i \\ 12 \end{bmatrix} + \begin{bmatrix} 7 - 8i \\ 4i \\ 2 \\ 9.4 + 3i \end{bmatrix}. \quad (2.5)$$

The addition operation satisfies certain properties. For example, because the addition of complex numbers is commutative, addition of complex vectors is also **commutative**:

$$\begin{aligned} V + W &= \begin{bmatrix} (6 - 4i) + (16 + 2.3i) \\ (7 + 3i) + (-7i) \\ (4.2 - 8.1i) + (6) \\ (-3i) + (-4i) \end{bmatrix} = \begin{bmatrix} 22 - 1.7i \\ 7 - 4i \\ 10.2 - 8.1i \\ -7i \end{bmatrix} \\ &= \begin{bmatrix} (16 + 2.3i) + (6 - 4i) \\ (-7i) + (7 + 3i) \\ (6) + (4.2 - 8.1i) \\ (-4i) + (-3i) \end{bmatrix} = W + V. \end{aligned} \quad (2.6)$$

Similarly, addition of complex vectors is also **associative**, i.e., given three vectors V , W , and X , we may add them as $(V + W) + X$ or as $V + (W + X)$. Associativity states that the resulting sums are the same:

$$(V + W) + X = V + (W + X). \quad (2.7)$$

Exercise 2.1.2 Formally prove the associativity property.

There is also a distinguished vector called **zero**:

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (2.8)$$

which satisfies the following property: for all vectors $V \in \mathbb{C}^4$, we have

$$V + \mathbf{0} = V = \mathbf{0} + V. \quad (2.9)$$

Formally, $\mathbf{0}$ is defined as $\mathbf{0}[j] = 0$.

Every vector also has an **(additive) inverse** (or **negative**). Consider

$$V = \begin{bmatrix} 6 - 4i \\ 7 + 3i \\ 4.2 - 8.1i \\ -3i \end{bmatrix}. \quad (2.10)$$

There exists in \mathbb{C}^4 another vector

$$-V = \begin{bmatrix} -6 + 4i \\ -7 - 3i \\ -4.2 + 8.1i \\ 3i \end{bmatrix} \in \mathbb{C}^4 \quad (2.11)$$

such that

$$V + (-V) = \begin{bmatrix} 6 - 4i \\ 7 + 3i \\ 4.2 - 8.1i \\ -3i \end{bmatrix} + \begin{bmatrix} -6 + 4i \\ -7 - 3i \\ -4.2 + 8.1i \\ 3i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0}. \quad (2.12)$$

In general, for every vector $W \in \mathbb{C}^4$, there exists a vector $-W \in \mathbb{C}^4$ such that $W + (-W) = (-W) + W = \mathbf{0}$. $-W$ is called the **inverse** of W . Formally,

$$(-W)[j] = -(W[j]). \quad (2.13)$$

The set \mathbb{C}^4 with the addition, inverse operations, and zero such that the addition is associative and commutative, form something called an **Abelian group**.

What other structure does our set \mathbb{C}^4 have? Take an arbitrary complex number, say, $c = 3 + 2i$. Call this number a **scalar**. Take a vector

$$V = \begin{bmatrix} 6 + 3i \\ 0 + 0i \\ 5 + 1i \\ 4 \end{bmatrix}. \quad (2.14)$$

We can **multiply an element by a scalar** by multiplying the scalar with each entry of the vector; i.e.,

$$(3 + 2i) \cdot \begin{bmatrix} 6 + 3i \\ 0 + 0i \\ 5 + 1i \\ 4 \end{bmatrix} = \begin{bmatrix} 12 + 21i \\ 0 + 0i \\ 13 + 13i \\ 12 + 8i \end{bmatrix}. \quad (2.15)$$

Formally, for a complex number c and a vector V , we form $c \cdot V$, which is defined as

$$(c \cdot V)[j] = c \times V[j], \quad (2.16)$$

where the \times is complex multiplication. We shall omit the \cdot when the scalar multiplication is understood.

Exercise 2.1.3 Scalar multiply $8 - 2i$ with $\begin{bmatrix} 16 + 2.3i \\ -7i \\ 6 \\ 5 - 4i \end{bmatrix}$. ■

Scalar multiplication satisfies the following properties: for all $c, c_1, c_2 \in \mathbb{C}$ and for all $V, W \in \mathbb{C}^4$,

- $1 \cdot V = V$,
- $c_1 \cdot (c_2 \cdot V) = (c_1 \times c_2) \cdot V$,
- $c \cdot (V + W) = c \cdot V + c \cdot W$,
- $(c_1 + c_2) \cdot V = c_1 \cdot V + c_2 \cdot V$.

Exercise 2.1.4 Formally prove that $(c_1 + c_2) \cdot V = c_1 \cdot V + c_2 \cdot V$. ■

An Abelian group with a scalar multiplication that satisfies these properties is called a **complex vector space**.

Notice that we have been working with vectors of size 4. However, everything that we have stated about vectors of size 4 is also true for vectors of arbitrary size. So the set \mathbb{C}^n for a fixed but arbitrary n also has the structure of a complex vector space. In fact, these vector spaces will be the primary examples we will be working with for the rest of the book.

Programming Drill 2.1.1 Write three functions that perform the addition, inverse, and scalar multiplication operations for \mathbb{C}^n , i.e., write a function that accepts the appropriate input for each of the operations and outputs the vector.

2.2 DEFINITIONS, PROPERTIES, AND EXAMPLES

There are many other examples of complex vector spaces. We shall need to broaden our horizon and present a formal definition of a complex vector space.

Definition 2.2.1 A **complex vector space** is a nonempty set \mathbb{V} , whose elements we shall call **vectors**, with three operations

- Addition: $+: \mathbb{V} \times \mathbb{V} \longrightarrow \mathbb{V}$
- Negation: $-: \mathbb{V} \longrightarrow \mathbb{V}$
- Scalar multiplication: $\cdot: \mathbb{C} \times \mathbb{V} \longrightarrow \mathbb{V}$

and a distinguished element called the **zero vector** $\mathbf{0} \in \mathbb{V}$ in the set. These operations and zero must satisfy the following properties: for all $V, W, X \in \mathbb{V}$ and for all $c, c_1, c_2 \in \mathbb{C}$,

- (i) Commutativity of addition: $V + W = W + V$,
- (ii) Associativity of addition: $(V + W) + X = V + (W + X)$,
- (iii) Zero is an additive identity: $V + \mathbf{0} = V = \mathbf{0} + V$,
- (iv) Every vector has an inverse: $V + (-V) = \mathbf{0} = (-V) + V$,
- (v) Scalar multiplication has a unit: $1 \cdot V = V$,
- (vi) Scalar multiplication respects complex multiplication:

$$c_1 \cdot (c_2 \cdot V) = (c_1 \times c_2) \cdot V, \quad (2.17)$$

- (vii) Scalar multiplication distributes over addition:

$$c \cdot (V + W) = c \cdot V + c \cdot W, \quad (2.18)$$

- (viii) Scalar multiplication distributes over complex addition:

$$(c_1 + c_2) \cdot V = c_1 \cdot V + c_2 \cdot V. \quad (2.19)$$

To recap, any set that has an addition operation, an inverse operation, and a zero element that satisfies Properties (i), (ii), (iii), and (iv) is called an **Abelian group**. If, furthermore, there is a scalar multiplication operation that satisfies all the properties, then the set with the operations is called a **complex vector space**.

Although our main concern is complex vector spaces, we can gain much intuition from real vector spaces.

Definition 2.2.2 A **real vector space** is a nonempty set \mathbb{V} (whose elements we shall call **vectors**), along with an addition operation and a negation operation. Most important, there is a scalar multiplication that uses \mathbb{R} and not \mathbb{C} , i.e.,

$$\cdot: \mathbb{R} \times \mathbb{V} \longrightarrow \mathbb{V}. \quad (2.20)$$

This set and these operations must satisfy the analogous properties of a complex vector space.

In plain words, a real vector space is like a complex vector space except that we only require the scalar multiplication to be defined for scalars in $\mathbb{R} \subset \mathbb{C}$. From the fact that $\mathbb{R} \subset \mathbb{C}$, it is easy to see that for every V we have $\mathbb{R} \times V \subset \mathbb{C} \times V$. If we have a given

$$\cdot : \mathbb{C} \times V \longrightarrow V, \quad (2.21)$$

then we can write

$$\mathbb{R} \times V \hookrightarrow \mathbb{C} \times V \longrightarrow V. \quad (2.22)$$

We conclude that every complex vector space can automatically be given a real vector space structure.

Let us descend from the abstract highlands and look at some concrete examples.

Example 2.2.1 \mathbb{C}^n , the set of vectors of length n with complex entries, is a complex vector space that serves as our primary example for the rest of the book. In Section 2.1, we exhibited the operations and described the properties that are satisfied. \square

Example 2.2.2 \mathbb{C}^n , the set of vectors of length n with complex entries, is also a real vector space because every complex vector space is also a real vector space. The operations are the same as those in Example 2.2.1. \square

Example 2.2.3 \mathbb{R}^n , the set of vectors of length n with real number entries, is a real vector space. Notice that there is no obvious way to make this into a complex vector space. What would the scalar multiplication of a complex number with a real vector be? \square

In Chapter 1, we discussed the geometry of $\mathbb{C} = \mathbb{C}^1$. We showed how every complex number can be thought of as a point in a two-dimensional plane. Things get more complicated for \mathbb{C}^2 . Every element of \mathbb{C}^2 involves two complex numbers or four real numbers. One could visualize this as an element of four-dimensional space. However, the human brain is not equipped to visualize four-dimensional space. The most we can deal with is three dimensions. Many times throughout this text, we shall discuss \mathbb{C}^n and then revert to \mathbb{R}^3 in order to develop an intuition for what is going on.

It pays to pause for a moment to take an in-depth look at the geometry of \mathbb{R}^3 . Every vector of \mathbb{R}^3 can be thought of as a point in three-dimensional space or equivalently, as an arrow from the origin of \mathbb{R}^3 to that point. So the vector $\begin{bmatrix} 5 \\ -7 \\ 6.3 \end{bmatrix}$ shown in Figure 2.1 is 5 units in the x direction, -7 units in the y direction, and 6.3 units in the z direction.

Given two vectors $V = \begin{bmatrix} r_0 \\ r_1 \\ r_2 \end{bmatrix}$ and $V' = \begin{bmatrix} r'_0 \\ r'_1 \\ r'_2 \end{bmatrix}$ of \mathbb{R}^3 , we may add them to form $\begin{bmatrix} r_0 + r'_0 \\ r_1 + r'_1 \\ r_2 + r'_2 \end{bmatrix}$. Addition can be seen as making a parallelogram in \mathbb{R}^3 where you attach the beginning of one arrow to the end of the other one. The result of the addition is

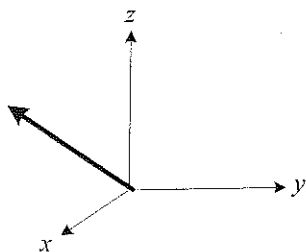


Figure 2.1. A vector in three dimensional space.

the composition of the arrows (see Figure 2.2). The reason that we can be ambiguous about which arrow comes first demonstrates the commutativity property of addition.

Given a vector $V = \begin{bmatrix} r_0 \\ r_1 \\ r_2 \end{bmatrix}$ in \mathbb{R}^3 , we form the inverse $-V = \begin{bmatrix} -r_0 \\ -r_1 \\ -r_2 \end{bmatrix}$ by looking at the arrow in the opposite direction with respect to all dimensions (as in Figure 2.3).

And finally, the scalar multiplication of a real number r and a vector $V = \begin{bmatrix} r_0 \\ r_1 \\ r_2 \end{bmatrix}$ is $r \cdot V = \begin{bmatrix} rr_0 \\ rr_1 \\ rr_2 \end{bmatrix}$, which is simply the vector V stretched or shrunk by r (as in Figure 2.4).

It is useful to look at some of the properties of a vector space from the geometric point of view. For example, consider the property $r \cdot (V + W) = r \cdot V + r \cdot W$. This corresponds to Figure 2.5.

Exercise 2.2.1 Let $r_1 = 2$, $r_2 = 3$, and $V = \begin{bmatrix} 2 \\ -4 \\ 1 \end{bmatrix}$. Verify Property (vi), i.e., calculate $r_1 \cdot (r_2 \cdot V)$ and $(r_1 \times r_2) \cdot V$ and show that they coincide. ■

Exercise 2.2.2 Draw pictures in \mathbb{R}^3 that explain Properties (vi) and (viii) of the definition of a real vector space. ■

Let us continue our list of examples.

Example 2.2.4 $\mathbb{C}^{m \times n}$, the set of all m -by- n matrices (two-dimensional arrays) with complex entries, is a complex vector space. □

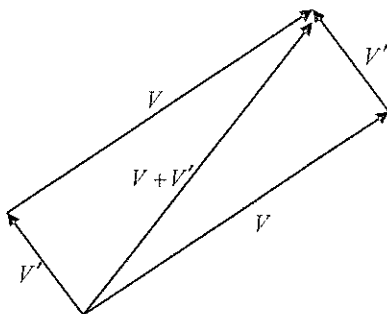


Figure 2.2. Vector addition.

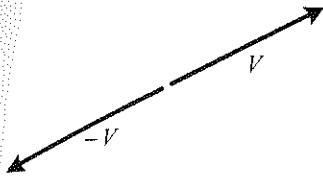


Figure 2.3. Inverse of a vector.

For a given $A \in \mathbb{C}^{m \times n}$, we denote the complex entry in the j th row and the k th column as $A[j, k]$ or $c_{j,k}$. We shall denote the j th row as $A[j, -]$ and the k th column as $A[-, k]$. Several times throughout the text we shall show the row and column numbers explicitly to the left and top of the square brackets:

$$A = \begin{matrix} & \begin{matrix} 0 & 1 & \cdots & n-1 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ \vdots \\ m-1 \end{matrix} & \begin{bmatrix} c_{0,0} & c_{0,1} & \cdots & c_{0,n-1} \\ c_{1,0} & c_{1,1} & \cdots & c_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m-1,0} & c_{m-1,1} & \cdots & c_{m-1,n-1} \end{bmatrix} \end{matrix} \quad (2.23)$$

The operations for $\mathbb{C}^{m \times n}$ are given as follows: Addition is

$$\begin{bmatrix} c_{0,0} & c_{0,1} & \cdots & c_{0,n-1} \\ c_{1,0} & c_{1,1} & \cdots & c_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m-1,0} & c_{m-1,1} & \cdots & c_{m-1,n-1} \end{bmatrix} + \begin{bmatrix} d_{0,0} & d_{0,1} & \cdots & d_{0,n-1} \\ d_{1,0} & d_{1,1} & \cdots & d_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m-1,0} & d_{m-1,1} & \cdots & d_{m-1,n-1} \end{bmatrix} = \begin{bmatrix} c_{0,0} + d_{0,0} & c_{0,1} + d_{0,1} & \cdots & c_{0,n-1} + d_{0,n-1} \\ c_{1,0} + d_{1,0} & c_{1,1} + d_{1,1} & \cdots & c_{1,n-1} + d_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m-1,0} + d_{m-1,0} & c_{m-1,1} + d_{m-1,1} & \cdots & c_{m-1,n-1} + d_{m-1,n-1} \end{bmatrix} \quad (2.24)$$

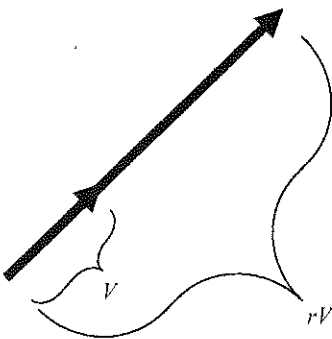


Figure 2.4. A real multiple of a vector.

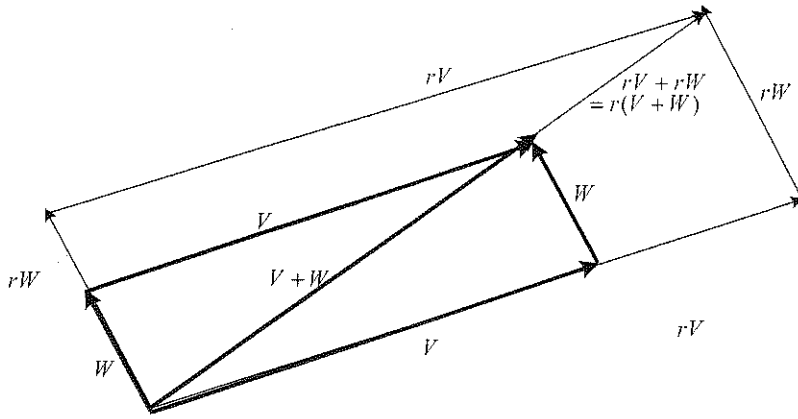


Figure 2.5. Scalar multiplication distributes over addition.

The inverse operation is given as

$$- \begin{bmatrix} c_{0,0} & c_{0,1} & \cdots & c_{0,n-1} \\ c_{1,0} & c_{1,1} & \cdots & c_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m-1,0} & c_{m-1,1} & \cdots & c_{m-1,n-1} \end{bmatrix} = \begin{bmatrix} -c_{0,0} & -c_{0,1} & \cdots & -c_{0,n-1} \\ -c_{1,0} & -c_{1,1} & \cdots & -c_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ -c_{m-1,0} & -c_{m-1,1} & \cdots & -c_{m-1,n-1} \end{bmatrix} \quad (2.25)$$

Scalar multiplication is given as

$$c \cdot \begin{bmatrix} c_{0,0} & c_{0,1} & \cdots & c_{0,n-1} \\ c_{1,0} & c_{1,1} & \cdots & c_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m-1,0} & c_{m-1,1} & \cdots & c_{m-1,n-1} \end{bmatrix} = \begin{bmatrix} c \times c_{0,0} & c \times c_{0,1} & \cdots & c \times c_{0,n-1} \\ c \times c_{1,0} & c \times c_{1,1} & \cdots & c \times c_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c \times c_{m-1,0} & c \times c_{m-1,1} & \cdots & c \times c_{m-1,n-1} \end{bmatrix} \quad (2.26)$$

Formally, these operations can be described by the following formulas:

For two matrices, $A, B \in \mathbb{C}^{m \times n}$, we add them as

$$(A + B)[j, k] = A[j, k] + B[j, k]. \quad (2.27)$$

The inverse of A is

$$(-A)[j, k] = -(A[j, k]). \quad (2.28)$$

The scalar multiplication of A with a complex number $c \in \mathbb{C}$ is

$$(c \cdot A)[j, k] = c \times A[j, k]. \quad (2.29)$$

Exercise 2.2.3 Let $c_1 = 2i$, $c_2 = 1 + 2i$, and $A = \begin{bmatrix} 1-i & 3 \\ 2+2i & 4+i \end{bmatrix}$. Verify Properties (vi) and (viii) in showing $\mathbb{C}^{2 \times 2}$ is a complex vector space. ■

Exercise 2.2.4 Show that these operations on $\mathbb{C}^{m \times n}$ satisfy Properties (v), (vi), and (viii) of being a complex vector space. ■

Programming Drill 2.2.1 Convert your functions from the last programming drill so that instead of accepting elements of \mathbb{C}^n , they accept elements of $\mathbb{C}^{m \times n}$.

When $n = 1$, the matrices $\mathbb{C}^{m \times n} = \mathbb{C}^{m \times 1} = \mathbb{C}^m$, which we dealt with in Section 2.1. Thus, we can think of vectors as special types of matrices.

When $m = n$, the vector space $\mathbb{C}^{n \times n}$ has more operations and more structure than just a complex vector space. Here are three operations that one can perform on an $A \in \mathbb{C}^{n \times n}$:

■ The **transpose** of A , denoted A^T , is defined as

$$A^T[j, k] = A[k, j]. \quad (2.30)$$

■ The **conjugate** of A , denoted \overline{A} , is the matrix in which each element is the complex conjugate of the corresponding element of the original matrix,³ i.e., $\overline{A}[j, k] = \overline{A[j, k]}$.

■ The transpose operation and the conjugate operation are combined to form the **adjoint** or **dagger** operation. The adjoint of A , denoted as A^\dagger , is defined as $A^\dagger = (\overline{A})^T = \overline{(A^T)}$ or $A^\dagger[j, k] = \overline{A[k, j]}$.

Exercise 2.2.5 Find the transpose, conjugate, and adjoint of

$$\begin{bmatrix} 6-3i & 2+12i & -19i \\ 0 & 5+2.1i & 17 \\ 1 & 2+5i & 3-4.5i \end{bmatrix}. \quad (2.31)$$

These three operations are defined even when $m \neq n$. The transpose and adjoint are both functions from $\mathbb{C}^{m \times n}$ to $\mathbb{C}^{n \times m}$.

These operations satisfy the following properties for all $c \in \mathbb{C}$ and for all $A, B \in \mathbb{C}^{m \times n}$:

- (i) Transpose is idempotent: $(A^T)^T = A$.
- (ii) Transpose respects addition: $(A + B)^T = A^T + B^T$.
- (iii) Transpose respects scalar multiplication: $(c \cdot A)^T = c \cdot A^T$.

³ This notation is overloaded. It is an operation on complex numbers and complex matrices.

- (iv) Conjugate is idempotent: $\overline{\overline{A}} = A$.
- (v) Conjugate respects addition: $\overline{A + B} = \overline{A} + \overline{B}$.
- (vi) Conjugate respects scalar multiplication: $\overline{c \cdot A} = \overline{c} \cdot \overline{A}$.
- (vii) Adjoint is idempotent: $(A^\dagger)^\dagger = A$.
- (viii) Adjoint respects addition: $(A + B)^\dagger = A^\dagger + B^\dagger$.
- (ix) Adjoint relates to scalar multiplication: $(c \cdot A)^\dagger = \overline{c} \cdot A^\dagger$.

Exercise 2.2.6 Prove that conjugation respects scalar multiplication, i.e., $\overline{c \cdot A} = \overline{c} \cdot \overline{A}$. ■

Exercise 2.2.7 Prove Properties (vii), (viii), and (ix) using Properties (i) – (vi). ■

The transpose shall be used often in the text to save space. Rather than writing

$$\begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} \quad (2.32)$$

which requires more space, we write $[c_0, c_1, \dots, c_{n-1}]^T$.

When $m = n$, there is another binary operation that is used: **matrix multiplication**. Consider the following two 3-by-3 matrices:

$$A = \begin{bmatrix} 3+2i & 0 & 5-6i \\ 1 & 4+2i & i \\ 4-i & 0 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 2-i & 6-4i \\ 0 & 4+5i & 2 \\ 7-4i & 2+7i & 0 \end{bmatrix}. \quad (2.33)$$

We form the matrix product of A and B , denoted $A \star B$. $A \star B$ will also be a 3-by-3 matrix. $(A \star B)[0, 0]$ will be found by multiplying each element of the 0th row of A with the corresponding element of the 0th column of B . We then sum the results:

$$\begin{aligned} (A \star B)[0, 0] &= ((3+2i) \times 5) + (0 \times 0) + ((5-6i) \times (7-4i)) \\ &= (15+10i) + (0) + (11-62i) = 26-52i. \end{aligned} \quad (2.34)$$

The $(A \star B)[j, k]$ entry can be found by multiplying each element of $A[j, -]$ with the appropriate element of $B[-, k]$ and summing the results. So,

$$(A \star B) = \begin{bmatrix} 26-52i & 60+24i & 26 \\ 9+7i & 1+29i & 14 \\ 48-21i & 15+22i & 20-22i \end{bmatrix}. \quad (2.35)$$

Exercise 2.2.8 Find $B \star A$. Does it equal $A \star B$? ■

Matrix multiplication is defined in a more general setting. The matrices do not have to be square. Rather, the number of columns in the first matrix must be the