This document teaches you the basics of complex numbers: arithmetic, conjugation, exponentiation, linear algebra, and their geometric meaning.

1 Arithmetic

A *complex number* is a quantity of the form $a+bi$, where a and b are real numbers and $i^2=-1$. Calculations with complex numbers are not difficult. Just use all of the usual algebraic rules, and replace i^2 with -1 whenever necessary. Here's addition:

$$(a+bi)+(c+di)=a+c+bi+di=(a+c)+(b+d)i.$$

Similarly, subtraction is

$$(a+bi)-(c+di)=(a-c)+(b-d)i.$$

Multiplication is a little more interesting:

$$(a+bi)(c+di)=ac+adi+bic+bdi^2=(ac-bd)+(ad+bc)i.$$

Division is more difficult. It's helpful to think of division as multiplication by the reciprocal:

$$\frac{a+bi}{c+di}=(a+bi)\cdot\frac{1}{c+di}.$$

But how do you compute the reciprocal? Use this trick:

$$\frac{1}{c+di}=\frac{1}{c+di}\cdot\frac{c-di}{c-di}=\frac{c-di}{c^2+d^2}=\frac{c}{c^2+d^2}-\frac{d}{c^2+d^2}i. \tag{1}$$

It doesn't work if $c^2+d^2=0$, but that happens only when you're trying to compute the reciprocal of 0. Division by zero is illegal in the complex numbers, just as it's illegal in the real numbers.

The set of complex numbers is denoted $\mathbb{C}$. When I say that the complex numbers satisfy all of the usual rules of algebra, I specifically mean these nine rules:

- Associativity of addition: $(x+y)+z=x+(y+z)$ for all $x,y,z\in\mathbb{C}$.
- Identity in addition: The complex number $0=0+0i$ satisfies $0+x=x=x+0$ for all x.
- Inverses in addition: For any complex number $x=a+bi$, there exists a complex number $-x=-a+-bi$, which satisfies $x+-x=0=-x+x$.
- Commutativity of addition: $x+y=y+x$ for all x,y.
- Associativity of multiplication: $(xy)z=x(yz)$ for all x,y,z.
- Identity in multiplication: The complex number $1=1+0i$ satisfies $1x=x=x1$ for all x.

- Inverses in multiplication: For any non-zero complex number x, there exists a complex number x^{-1}, which satisfies $xx^{-1} = 1 = x^{-1}x$. (Equation 1 tells us how to compute x^{-1}.)
- Commutativity of multiplication: $xy = yx$ for all x, y.
- Distributivity: $x(y + z) = xy + xz$ and $(x + y)z = xz + yz$ for all x, y, z.

In the mathematical jargon, we say that $\mathbb{C}$ is a *field*. The set $\mathbb{R}$ of real numbers is also a field. In one way, $\mathbb{R}$ is nicer than $\mathbb{C}$: $\mathbb{R}$ has an ordering $<$, so that we can talk about whether $x < y$, $x \geq y$, etc. for real numbers x, y. Those concepts don't exist in $\mathbb{C}$. But $\mathbb{C}$ is nicer than $\mathbb{R}$ in a different way: It is *algebraically closed*, meaning that every non-constant polynomial with coefficients in $\mathbb{C}$ has a root in $\mathbb{C}$. In contrast, there are polynomials with real coefficients that do not have any real roots. The most important example is $x^2 + 1$. Why?

The complex numbers have another operation, which has no analogue in the real numbers: conjugation. The *conjugate* of a complex number $a + bi$ is defined as

$$\overline{a + bi} = a - bi.$$

Notice that the conjugate of a real number $a = a + 0i$ is just $a - 0i = a$ again. So conjugation of real numbers is trivial, which is why we never talk about it. Notice also that

$$(a + bi) \cdot \overline{a + bi} = a^2 + b^2.$$

This was the trick that helped us compute the reciprocal in Equation 1. Conjugation also plays well with arithmetic; for example,

$$\overline{(a + bi)(c + di)} = \overline{a + bi} \cdot \overline{c + di}.$$

Exercise A: Twice now I've asserted that $(a + bi) \cdot \overline{a + bi} = a^2 + b^2$, without any justification. Prove it, by showing all steps of the algebra required.

Exercise B: Complete this exercise in the programming language of your choice (as long as it's common and readable, such as Python or C, rather than esoteric or unreadable, such as Whitespace). Define a data type for complex numbers. It should use the language's floating-point numbers to approximate the underlying real numbers. Write subroutines (functions, methods, etc.) to perform addition, multiplication, conjugation, subtraction, and division. This should all be short and simple; don't over-engineer it.

2 Geometry

Because each complex number $x = a + bi$ is made up of two real numbers a and b, it is natural to picture $\mathbb{C}$ as the two-dimensional real plane $\mathbb{R}^2$. That is, the number $a + bi$ plots at the point (a, b). The horizontal axis consists of the numbers of the form $a + 0i$ — that is, the real numbers. The vertical axis consists of the numbers of the form $0 + bi$. They are called the *imaginary* numbers.

Let's take a moment to talk about this terminology. First, the terms "real" and "imaginary" are important in the vocabulary of math, so you should learn to use them correctly. They are not antonyms. Most complex numbers are neither real nor imaginary, and the number 0 is both real and imaginary. If you want to say that a number is not real, then don't say that it's imaginary; instead, say that it's "not real" or "non-real". Second, you should ignore the non-mathematical meanings of "real" and "imaginary". You should not intuit that the real numbers actually exist and the other complex numbers actually don't exist. None of these numbers exist in our universe; they are concepts, not physical objects, and they live only in the human mind. (There is a minority view in the philosophy of science, which holds that mathematical concepts such as these are the only things that exist, and the physical universe is a phenomenon that emerges from that math. If you adopt that view, then real, imaginary, and other complex numbers are again equally existent.)

The *norm* or *magnitude* $|a + bi|$ of a complex number $a + bi$ is defined as its distance to the origin. That is,

$$|a + bi| = \sqrt{a^2 + b^2} = \sqrt{(a + bi) \cdot \overline{a + bi}}.$$

Addition has a simple geometric interpretation. Adding a complex number $x = a + bi$ to a complex number $y = c + di$ has the effect of translating y a units to the right and b units up. Scaling a complex number $c + di$ by a real number a has the effect of stretching $c + di$ away from the origin by a factor of a. In other words, if we view $\mathbb{C}$ as the vector space $\mathbb{R}^2$, then addition and real scaling have their usual geometric interpretation.

However, $\mathbb{C}$ is more than just $\mathbb{R}^2$, because it has two additional operations: conjugation and complex multiplication. Geometrically, conjugation has the effect of flipping points across the real axis. We'll explain the geometric meaning of multiplication in a moment. First it will be helpful to change coordinates.

Recall (from some calculus course) the concept of polar coordinates. Given a point (a, b) in the plane, let r be the distance from the origin to that point, and let θ be the angle, at the origin, measured counterclockwise from the positive real axis to (a, b). It is easy to convert from polar coordinates (r, θ) to Cartesian coordinates (a, b): $a = r \cos\theta$ and $b = r \sin\theta$. So

$$a + ib = r\cos\theta + ir\sin\theta = r(\cos\theta + i\sin\theta).$$

It's harder to convert from Cartesian to polar coordinates. Well, $r = \sqrt{a^2 + b^2}$ isn't hard, but computing θ requires several cases. Most programming languages offer a function, called something like `atan2`, to compute θ from a and b.

3 Exponential

In the real numbers, the exponential function is defined as the power series

$$\exp(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \cdots . \tag{2}$$

This function has many miraculous properties, the most important of which is probably

$$\exp(x) \cdot \exp(y) = \exp(x + y).$$

Let's plug our favorite real numbers into exp. First, $\exp(0) = 1$. Second,

$$\exp(1) = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots = 2.718\ldots.$$

We give the number $\exp(1)$ a special name: e. Then $\exp(2) = \exp(1+1) = \exp(1) \cdot \exp(1) = e^2$. Using induction, you can prove that $\exp(n) = e^n$ for any positive integer n, and then for all integers n. For this reason the function $\exp(x)$ is often denoted e^x. (But the function exp is more fundamental than the number e. You should view the number as an emergent phenomenon of the function.)

The same power series function definition (Equation 2) works for complex numbers x. I mean, you can plug in any complex number $a + bi$ for x, compute the required powers, divide by the required factorials, and perform the required summation (at least in principle). The complex exponential function still has that crucial sum-product property

$$e^{a+bi} e^{c+di} = e^{(a+bi)+(c+di)}.$$

You already know what exp does to complex numbers x of the form $a + 0i$, because those are just real numbers. But what about imaginary numbers $x = ib$? When one examines the power series closely, something surprising happens: $e^{ib} = \cos b + i \sin b$. The exponential function contains the trigonometric functions, even though Equation 2 seems not to be related to trigonometry at all.

Exercise C: Again using your complex number data type, write code to compute the complex exponential function. Do not use the power series expansion. Instead, derive and implement an approach that uses trigonometric functions and real exponentiation.

4 Geometry Revisited

Notice that

$$\left| e^{ib} \right| = |\cos b + i \sin b| = \cos^2 b + \sin^2 b = 1$$

for all b. So the exponential function maps the imaginary axis onto the unit circle (and the real axis onto the positive real axis). The exponential function gives us another way to view polar coordinates:

$$a + ib = r(\cos\theta + i\sin\theta) = re^{i\theta}.$$

Multiplication is especially simple in this format. If $c+di = se^{i\phi}$ is another complex number, then

$$(a + bi)(c + di) = re^{i\theta}se^{i\phi} = (rs)e^{i(\theta+\phi)}.$$

This is algebra, but we can immediately interpret it as geometry. Multipliying a complex number $x = re^{i\theta}$ by a complex number $y = se^{i\phi}$ has the effect of scaling y by r and rotating y through the angle θ about the origin.

Exercise D: Describe all complex numbers whose exponentials are real. Describe all complex numbers whose exponentials are imaginary. Describe all complex numbers whose exponentials are both real and imaginary.

5 Linear Algebra

Your linear algebra course was probably focused on vector spaces over the real numbers, meaning that all scalars were real numbers. But linear algebra works just as well over the complex numbers. If you have two matrices of complex numbers, then you can add and multiply them just as you do for real matrices. For example, if A is $M \times N$ and B is $N \times P$, then the product matrix AB is $M \times P$ and

$$(AB)_{ij} = \sum_{k=1}^{N} A_{ik}B_{kj}.$$

The usual $N \times N$ identity matrix I satisfies $AI = A = IA$ for all $N \times N$ matrices A. You can also scale complex matrices by complex numbers, just as you scale real matrices by real numbers. And you can compute determinants and eigensystems and all of that. Because $\mathbb{C}$ is algebraically closed, a complex $N \times N$ matrix always has N eigenvalues. So in some ways complex linear algebra is easier than real linear algebra.

$\mathbb{C}^N$ is the complex vector space consisting of all complex $N \times 1$ column matrices. It has a standard basis

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \vec{e}_N = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

But this is computer science, so we index from 0:

$$\vec{e}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \vec{e}_{N-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

But this is quantum theory, so we use the *ket* notation invented by physicist Paul Dirac:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad |N-1\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

This ket notation has a reputation for scaring people, but it shouldn't, because it's just notation. Take $\vec{e}_0$ vs. $|0\rangle$ for example. Instead of using "$\vec{\ }$" to signal that we're talking about a vector, we use "$|\ \ \rangle$". Instead of putting the "0" in a subscript, we put it inside the "$|\ \ \rangle$".

Because this is computer science, we often write the indices in binary. For example, in $\mathbb{C}^4$ we write $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ instead of $|0\rangle, |1\rangle, |2\rangle, |3\rangle$. Ambiguities then occasionally arise between numeric bases. For example, is $|11\rangle$ the 4th standard basis vector or the 12th? To clarify, you add a subscript inside the ket, as in $|11_2\rangle$ or $|11_{10}\rangle$. (The subscript itself is always in base 10...where 10 is ten.) But you have to use these subscripts only rarely, because the numeric base is usually clear from context.

Exercise E: In $\mathbb{C}^8$, what is $|110\rangle$?

In math notation, when we want to talk about an unspecified vector in the abstract, we might use a generic symbol such as $\vec{v}$ or $\vec{w}$. In ket notation, the analogous symbols are $|v\rangle$ and $|w\rangle$. But for historical reasons it's more common to use Greek letters than Roman letters, such as in $|\psi\rangle$ and $|\phi\rangle$.

Let's emphasize that the stuff inside the "$|\ \ \rangle$" is not the numerical value of the vector, but rather just a name for the vector. You can't figure out the vector's value from just its name. For example, in $\mathbb{C}^2$ there are two special vectors denoted $|+\rangle$ and $|-\rangle$. Is it true that

$$|+\rangle = \begin{bmatrix} + \\ + \end{bmatrix}, \quad |-\rangle = \begin{bmatrix} - \\ - \end{bmatrix}?$$

No, that doesn't make any sense at all. Rather, they are defined to be

$$|+\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad |-\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

In math notation they would be written as something like $\vec{v}_+$ and $\vec{v}_-$ probably.

6 Hermitian Inner Product

In real linear algebra, the dot product is a fundamental operation. All of Euclidean geometry in $\mathbb{R}^N$ is a consequence of the dot product. For example, the length of a vector $\vec{v}$ is $\sqrt{\vec{v}\cdot\vec{v}}$, and the angle between two unit vectors $\vec{v}$ and $\vec{w}$ is $\arccos(\vec{v}\cdot\vec{w})$. The dot product relates to transposition via the equation $\vec{v}\cdot\vec{w} = \vec{v}^\top\vec{w}$. In fact, the entire transposition operation on matrices exists because of the dot product.

The best way to extend these concepts to complex matrices is to replace transposition with conjugate transposition, as follows. For any complex matrix A, define the conjugate-transpose A^* to be $A^* = \overline{A}^\top$. That is, to compute A^*, you conjugate every entry in A and then transpose that conjugated matrix. Because conjugation plays well with arithmetic, the conjugate-transpose behaves much like the transpose; for example, $(AB)^* = B^*A^*$, just as $(AB)^\top = B^\top A^\top$.

If $|\psi\rangle$ is $N \times 1$, then there is a special notation for the conjugate transpose:

$$\langle\psi| = |\psi\rangle^* .$$

For example, in $\mathbb{C}^2$,

$$|\psi\rangle = \begin{bmatrix} 1+3i \\ 2-i \end{bmatrix} \quad \Rightarrow \quad \langle\psi| = |\psi\rangle^* = \begin{bmatrix} 1-3i & 2+i \end{bmatrix}.$$

Now we can define the complex analogue of the dot product. There are a couple of conventions. We follow the convention used by our textbook. If $|\psi\rangle, |\phi\rangle \in \mathbb{C}^N$, then their *(Hermitian) inner product* is the complex number $\langle\psi|\phi\rangle$ defined by

$$\langle\psi|\phi\rangle = \langle\psi|\,|\phi\rangle = |\psi\rangle^*\,|\phi\rangle .$$

It satisfies the following rules.

- Linearity in the second argument: $\langle\psi|c\phi + d\omega\rangle = c\langle\psi|\phi\rangle + d\langle\psi|\omega\rangle$ for all $|\psi\rangle, |\phi\rangle, |\omega\rangle \in \mathbb{C}^N$ and all $c, d \in \mathbb{C}$.
- Conjugate linearity in the first argument: $\langle c\psi + d\phi|\omega\rangle = \overline{c}\langle\psi|\omega\rangle + \overline{d}\langle\phi|\omega\rangle$.
- Conjugate symmetry: $\langle\psi|\phi\rangle = \overline{\langle\phi|\psi\rangle}$.
- Positive definiteness: $\langle\psi|\psi\rangle$ is a positive real number, except when $|\psi\rangle$ is the zero vector (in which case $\langle\psi|\psi\rangle = 0$).

With the inner product in hand, now we can define the *norm* or *magnitude* of any vector $|\psi\rangle \in \mathbb{C}^N$ to be the real number

$$\||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle} = \sqrt{\langle\psi|\,|\psi\rangle} = \sqrt{|\psi\rangle^*\,|\psi\rangle}.$$

It satisfies the following rules.

- Positivity: $\||\psi\rangle\| > 0$, unless $|\psi\rangle$ is the zero vector (in which case $\||\psi\rangle\| = 0$).
- Scaling: $\|c\,|\psi\rangle\| = |c| \cdot \||\psi\rangle\|$.
- Triangle inequality: $\||\psi\rangle + |\phi\rangle\| \leq \||\psi\rangle\| + \||\phi\rangle\|$.

The *Cauchy-Schwarz inequality* is also worth mentioning: $|\langle\psi|\phi\rangle| \leq \||\psi\rangle\| \cdot \||\phi\rangle\|$.

Exercise F: Prove that $\langle\psi|\psi\rangle$ is always real, as the definition of the norm implicitly claims.

7 Unitary Transformations

Here are a few important kinds of $N \times N$ complex matrices. A matrix H is *Hermitian* if $H^* = H$. A matrix S is *skew-Hermitian* if $S^* = -S$. The set of all $N \times N$ Hermitian matrices forms a real vector space of dimension N^2. So does the set of all $N \times N$ skew-Hermitian matrices.

Exercise G: Just for the cases $N = 2$ and $N = 4$, check that the real vector space of $N \times N$ skew-Hermitian matrices is N^2-dimensional. That is, explain why choosing a $N \times N$ skew-Hermitian matrix is tantamount to choosing N^2 real numbers.

Exercise H: Prove that if S is skew-Hermitian, then iS is Hermitian. (The converse is also true. And H is Hermitian if and only if iH is skew-Hermitian. But I'm not asking you to prove those extra theorems.)

A matrix U is *unitary* if $UU^* = I = U^*U$. The set of all $N \times N$ unitary matrices does not form a vector space. Instead it forms a *group*, meaning that:

- The identity I is a unitary matrix.
- If U is unitary, then U is invertible, and U^{-1} is also unitary.
- If U and V are unitary, then so is the product UV.

Check these three facts on your own. This next fact is so important that I want to see your verification.

Exercise I: Prove that unitary transformations preserve the Hermitian inner product. That is, if U is unitary $N \times N$ and $|\psi\rangle, |\phi\rangle \in \mathbb{C}^N$, then the inner product of $|\psi\rangle$ with $|\phi\rangle$ equals the inner product of $U|\psi\rangle$ with $U|\phi\rangle$. Then prove that $\|U|\psi\rangle\| = \||\psi\rangle\|$.

Because unitary matrices are important in our course, we want to have some ways to manufacture them. One way is via the matrix exponential function. If A is an $N \times N$ complex (or real) matrix, define

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = I + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \frac{1}{24}A^4 + \cdots.$$

Exercise J: Prove that if S is skew-Hermitian, then $\exp(S)$ is unitary. [Hint: First show that $(\exp(S))^* = \exp(-S)$. To do that, you must perform some term-by-term manipulations on the power series. If you are well trained in power series, then such manipulations should make you nervous. But this is the nicest power series in the world, so stop worrying.]

So, to make an $N \times N$ unitary matrix, one approach is: Choose N^2 real numbers. Form a skew-Hermitian matrix S from them. Then exponentiate to get a unitary $U = \exp(S)$. The exponential map is not injective, so differing choices of real numbers can sometimes produce the same U. But it's "almost injective", and this approach can make tons of unitary matrices.

8 Matrix Calculus

Let t be a real variable. Think of it as time. Suppose that $A = A(t)$ is a time-dependent complex matrix. You can think of it as a matrix, each entry of which is a function of t. You can also think of it as a curve of matrices parametrized by t. Anyway, define $\frac{d}{dt}A = A' = \dot{A}$ to be the component-wise derivative of A with respect to t. This matrix derivative obeys some reasonable rules:

- Sum rule: $\frac{d}{dt}(A + B) = \dot{A} + \dot{B}$.
- Scaling rule: $\frac{d}{dt}(cA) = c\dot{A}$ for any complex constant c.
- Product rule: $\frac{d}{dt}(AB) = \dot{A}B + A\dot{B}$. (Be careful about the order, because matrix multiplication is not commutative.)
- Conjugate-transpose rule: $\frac{d}{dt}(A^*) = (\dot{A})^*$. That is, the derivative of the conjugate-transpose is the conjugate-transpose of the derivative.

Exercise K: Prove that if $U = U(t)$ is a curve of unitary matrices, then $\dot{U}U^*$ is skew-Hermitian. [Hint: You know very little about U, so you don't have many options.]

Now suppose that $U = U(t)$ is a curve of unitary matrices, and define $|\psi\rangle = |\psi(t)\rangle$ by $|\psi\rangle = U\,|\psi(0)\rangle$, where $|\psi(0)\rangle$ is some initial vector at time $t = 0$. We are ready to show that $|\psi\rangle$ obeys a certain differential equation known as the *Schrödinger equation*.

Exercise L: Prove that there exists a Hermitian H such that $i|\dot{\psi}\rangle = H\,|\psi\rangle$. [Hint: First show that $|\dot{\psi}\rangle = S\,|\psi\rangle$ for a certain skew-Hermitian S.]