Complex numbers, linear algebra and the mathematics of quantum mechanics

Martin Laforest, PhD

Contents

Disclaimer

I would kindly ask you to not distribute, post online, copy or sell this document without my prior authorization. It is still a work in progress that might eventually be turned into a textbook.

1 Preface

1.1 A taste of quantum mechanics

The physics describing the world we experience everyday is referred to as *classical physics*. It describes how large objects (i.e., objects made of billions and billions of atoms) interact with each other. Whether it is the motion of the planets in our solar systems, the behaviour of a car when you accelerate, what happens when you play billiards, or how electronic circuits work, classical physics is a set of rules that were discovered and quantified by the likes of Galileo, Newton, Maxwell and many others.

Unfortunately, classical physics is not the ultimate description of nature. If we try to describe the behaviour of atoms and their constituents (protons, neutrons, electrons) using the laws of classical physics, it completely, and I mean, completely fails. Actually, if we described the motion of electrons around the nucleus of an atom, you can calculate that the atom would collapse within a fraction of a second. Obviously, the world we live in is made of stable atoms..., so what is going on?

Well, it turns out that classical physics is only an approximation of physics that works for large objects. In order to describe the behaviour of the building blocks of nature, we need a radically different approach that, as you will learn, leads to baffling and counter-intuitive phenomena: welcome to the world of Quantum Mechanics!

Words of warning, in order to fully appreciate the beauty of quantum mechanics, you have to give up reality as you know it, open your mind and be ready to have your mind blown! As you will learn, in the quantum world, particles behaving like waves and single, indivisible electrons being in two places at once, actually makes sense!

Quantum mechanics has the reputation of being "weird", complicated, inaccessible, mind boggling, and the list goes on and on. At the University of Waterloo's Institute for Quantum Computing (IQC), we see quantum mechanics, and certainly quantum cryptography, in a much different light.

Quantum mechanics can be baffling... yes, surprising... and certainly counter-intuitive, but claiming that it is weird would imply that the physical world we live in is weird. We prefer to think of it as fascinating! Quantum mechanics is not really that complicated. But, it does shatter our intuitive grasp of reality. As you become more familiar with the subject, you will discover that the world we experience every day is far from the whole story. Quantum mechanics opens a door to a world that may surprise you; a world where the rules of the game are different. Much different.

Developed in the first half of the 20th century by the likes of Max Planck, Erwin Schrödinger, Werner Heseinberg, Paul Dirac and many more, the theory of quantum mechanics (also called quantum theory) never ceases to amaze us, even to this day. At the time, quantum mechanics was revolutionary and controversial. Even a genius like Albert Einstein thought it could not be a serious theory. Unfortunately for him, he was wrong! Many experiments have been performed in the last few decades demonstrating the validity

of quantum theory. As a matter of fact, we can safely claim that quantum theory is the most accurate theory ever developed by mankind. Every attempt at proving it wrong has failed miserably.

1.2 Quantum technologies

Quantum mechanics has already had a tremendous impact on our lives. Not only does it tell us how the world behaves at its core – at the molecular level and beyond – but it has led to transformative technologies that have shaped, and continue to shape, the 20^{th} and 21^{st} centuries. The laser, LED lights, magnetic resonance imaging (MRI), transistors, and so much more, all exist because the world behaves quantum mechanically.

What would a world be without lasers? Well, there would be no internet. How about a world with no transistors? Well, every single piece of digital technology – computers, mp3 players, smartphones, digital cable tv – would not exist! The world would be radically different.

Speaking of digital technology, the digital world we now live in has been made possible thanks to *information theory*. All the digital technology mentioned above really boils down to one thing: information processing. Yes, their applications are vastly different from one another, but at their core, e.g., in their processor, they manipulate *bits* of information.

A second quantum revolution is underway, the *Quantum Information* revolution, where we manipulate information in a quantum mechanical fashion. This revolution is more than just an idea – small prototypes of quantum computers exist (you will even see some of them at IQC), stable quantum cryptography systems are commercially available and are currently being used by government and banks around the world, quantum sensors are bursting out of our labs and so on.

1.3 Welcome to QCSYS

This week, you will become familiar with a special type of quantum technology: quantum cryptography, or more precisely, quantum key distribution. Nowadays, when secure information is being sent around the internet (bank transactions, your password when you sign in Facebook, etc.), the privacy of the information is ensured by the fact that no computer on earth can solve, in a reasonable amount of time (e.g., hundreds to thousands of years!), a given, really hard mathematical problem. The eventual arrival of the ridiculously powerful quantum computer will render these cryptographic techniques obsolete.

Thankfully, quantum mechanics also comes to the rescue: Quantum Cryptography. By exploiting the behaviour of the quantum world, we are able to secure information in such a way that the only way for an all-evil eavesdropper to access this information would be to break the rules of physics: we are pretty confident nobody can do that. Ever!

During the Quantum Cryptography School for Young Students (QCSYS, or "cue-see-sis"), you will learn the basic concepts behind quantum cryptography: from quantum

mechanics and classical cryptography, to quantum optics and, of course, quantum cryptography.

QCSYS started in 2007 with many goals and challenges in mind. Passionate about the science and research we do at IQC, we wanted to share it with the future scientists, mathematicians and engineers (that would be you!). Also, since quantum mechanics and quantum technologies will play a key role in shaping the technological landscape of the 21^{st} century, we strongly believe it is important for the new generation to be "quantum-aware". Last, but not the least, it was a challenge we gave ourselves: can we teach quantum mechanics and quantum information to high school students? Quantum cryptography is a tiny subset of potential quantum technology, but it offers a great vehicle to teach young students about technology, information security, mathematics, quantum mechanics and even quantum computing.

Since mathematics is the language of quantum mechanics, this present document has been put together to ease you into the subject. In quantum mechanics, we use special mathematics – complex numbers and *linear algebra* (vectors and matrices). Unfortunately, most high school mathematics curriculums around the world do not teach linear algebra. It is not very complicated. It is really just a different and clever way to add and multiply numbers together, but it is a very powerful tool.

We do not claim to cover all of linear algebra in a rigorous way, nor do we claim that this is the only way to do quantum mechanics. There are different mathematical approaches, but the one described in the following pages is very well suited for quantum information and quantum cryptography, and fairly simple (we hope) to understand.

I encourage you to read through this book before the beginning of QCSYS. Do not panic if it gets a little over your head or if you are struggling with some concepts – we will spend at least five hours going through the important concepts and we will do exercises in groups. QCSYS counsellors and myself will be on hand during QCSYS to help you out.

In addition to the mathematics of quantum mechanics, we will spend another five hours exploring the "physics" of quantum mechanics – using a phenomenological approach. After which, we will try to consolidate the two so you have a good understanding of how we use mathematics to model the physical quantum world. After this introduction, we will be ready to learn about cryptography, quantum optics, quantum cryptography (of course) and even quantum hacking. We will also go in the labs and do some experiments by the end of the week – you will even have the chance to build your own quantum cryptography system!

A little note before getting into the heart of the subject: it would be unprofessional and ungrateful of me to pretend that I did all the work putting this document together. The starting point of this mathematical primer has been a condensed version of these notes put together a few years ago by a then graduate student at IQC – Jamie Sikora. Jamie was a stellar graduate student and a passionate and talented educator. His contribution to the early stages of QCSYS have been paramount and QCSYS owes him greatly.

Finally, welcome to IQC, welcome to QCSYS and we sincerely hope you will have a

great time, learn a lot and make new friendships that will last forever.

Martin Laforest, Ph. D.
Director of QCSYS
Senior Manager of Scientific Outreach
Institute for Quantum Computing
University of Waterloo

2 Complex Numbers

The number system we all know and love, like 10, -2, 0.3, $\sqrt{2}$, $\frac{22}{7}$, π and of course the answer to life, the universe and everything, 42, are known as the **real numbers**. Conventionally, we denote the family of all numbers as $\mathbb{R}$.

Unfortunately, sometimes, real numbers are not sufficient. Below is an example that might be all too familiar:

Example 2.1 If we try to solve the quadratic equation $x^2 + 5x + 10 = 0$ for x, we will get the solution:

$$x = \frac{-5 \pm \sqrt{25 - 4 \cdot 10}}{2} = \frac{-5 \pm \sqrt{-15}}{2}$$

What is the value for $\sqrt{-15}$? In other words, is there a real number a such that $a^2 = -15$? It is not hard to convince yourself that, in fact, no real numbers, being positive or negative, can satisfy this condition.

Should we just give up? Of course not! We are doing mathematics: if something doesn't exist, we invent it! This is where complex numbers come in. If you pursue your studies in virtually any field of science and/or engineering, chances are complex numbers will become your best friends. They have many applications in physics, chemistry, biology, electrical engineering, statistics and even finance and economics. As you will learn soon enough, in quantum mechanics, complex numbers are absolutely everywhere.

2.1 What is a complex number?

In some sense, we have already defined what a complex number is. In the example above, since $\sqrt{-15}$ is not real, then x is certainly not real either (adding a real number with a non-real number cannot give you something real!) So, **by definition**, we will call numbers like this one complex numbers.

But of course, being mathematicians in training, we would like to have something more concrete, better defined. Looking at x again, all the kerfuffle seems to be caused by the nasty minus sign under the square root. Let's take care of that:

Definition 2.2 (Imaginary unit number) *We define the square root of -1 as*

$$i = \sqrt{-1}.$$

It is also known as the imaginary unit number.

As a side note, the use of i to denote the imaginary unit number is used in most scientific fields, but if you end up studying electrical engineering, chances are you will know it as j, since i is a variable denoting the electrical current. But for now, let's stick to our regular convention.

Example 2.3 Using our new definition for the square root of 1, we can now write:

$$\begin{aligned}\sqrt{-15} &= \sqrt{-1 \cdot 15} \\ &= \sqrt{-1}\sqrt{15} \\ &= i\sqrt{15} \\ &\approx 3.87i\end{aligned}$$

Definition 2.4 (Imaginary Numbers) *A number is said to be imaginary if its square is negative.*

Example 2.5 $\sqrt{-15}$ is an imaginary number because $\left(\sqrt{-15}\right)^2 = -15$

Example 2.6 Similarly, but using our new notation, $i\sqrt{15}$ is, of course, an imaginary number because:

$$\begin{aligned}\left(i\sqrt{15}\right)^2 &= i^2 \cdot \left(\sqrt{15}\right)^2 \\ &= \left(\sqrt{-1}\right)^2 \cdot 15 \\ &= -1 \cdot 15 \\ &= -15\end{aligned}$$

Now that we have a definition for imaginary numbers, we can finally define the complex numbers.

Definition 2.7 (Complex Numbers) *A complex number is any number written in the form*

$$z = a + bi,$$

where a ***and*** *b* ***are real numbers****. a is known as the "real part" of z, and b as the "imaginary part". We also define $Re(z)$ and $Im(z)$ as follows:*

$$Re(z) = a$$

$$Im(z) = b$$

The family of all complex numbers in denoted by $\mathbb{C}$. Since a real number is a complex number without an imaginary part, we have $\mathbb{R} \subset \mathbb{C}$.

Example 2.8 Let's solve again the quadratic equation $x^2 + 5x + 10 = 0$ for x.

As before, we have:

$$\begin{aligned} x &= \frac{-5 \pm \sqrt{25 - 4 \cdot 10}}{2} \\ &= \frac{-5 \pm \sqrt{-15}}{2} \\ &= \frac{-5 \pm \sqrt{15}\sqrt{-1}}{2} \\ &= \frac{-5}{2} \pm \frac{\sqrt{15}}{2} i \end{aligned}$$

2.2 Properties of Complex Numbers

In this section, we will introduce some useful definitions and properties related to complex numbers. They might seem a little arbitrary at first, but as you will see soon enough, these will become very handy, especially when we will start using them in quantum mechanics.

Definition 2.9 (Complex Conjugate) *We define the complex conjugate of a complex number $z = a + bi$, denoted $\bar{z}$, as*

$$\bar{z} = a - bi.$$

Example 2.10 Let's calculate some complex conjugate:

1. If $z = 5 + 10i$, then $\bar{z} = 5 - 10i$
2. If $z = 3 - 2i$, then $\bar{z} = 3 + 2i$
3. If $z = -3$, then $\bar{z} = -3$ (i.e., the complex conjugate of a real number is itself)
4. If $z = 2i = 0 + 2i$, then $\bar{z} = -2i$ (i.e., the complex conjugate of an imaginary number is minus itself)

Observation 2.11 (The complex plane) Real numbers and imaginary numbers are exclusive (i.e., a real number has no imaginary part, and an imaginary number has no real part). This is similar to cartesian coordinates (i.e., a point on the x-axis has no y component and vice-versa). We can visualize a complex number as being a point in the complex plane, such that the x-axis represents the real part of the number and the y-axis represents the imaginary part. In other words, the real part of the complex number z is the x component of the point and the imaginary part is the y component.

Looking at Figure 1, the complex number $z = a + bi$ is represented on the cartesian plane. a, the real part of z, is the projection of that point on the "real axis", while b, the imaginary part of z, is the projection of z along the "imaginary axis".

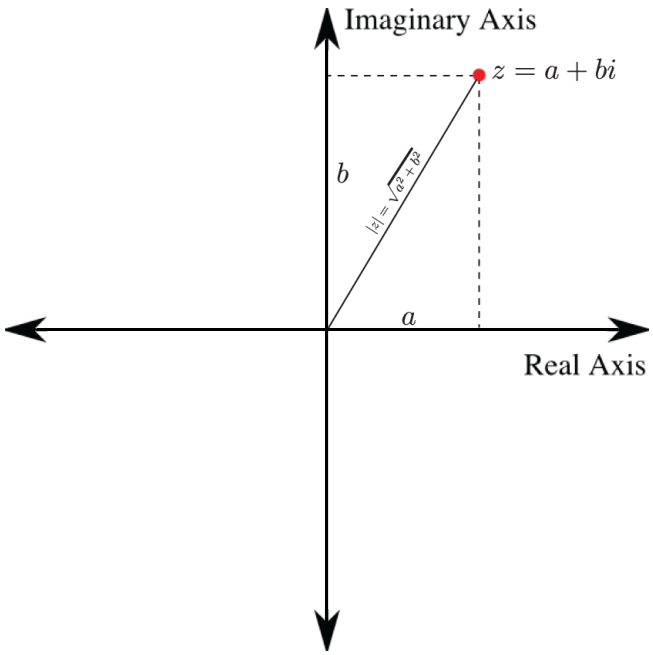


Figure 1: Using the complex plane, we can visualize any complex number z as a point on a two-dimensional plane. We can also define the modulus, or length, of a complex number as the distance between the origin $(0 + 0i)$ and the point representing z.

Definition 2.12 (Modulus, Length) *The modulus (or length) of a complex number $z = a + bi$ is given by*

$$|z| = \sqrt{a^2 + b^2}.$$

Since both a and b are real, ***the modulus of a complex number is always real and positive****.*

Observation 2.13 By looking at the representation of a complex number on the complex plane, the modulus of the number is simply the distance from the origin $0 + 0i$ the number is (hence why we also call it the "length" of the number). See Figure 1 for reference.

Note 2.14 The modulus of a complex number is similar to the absolute value of a real number.

Example 2.15 What is the modulus of the following complex numbers?

1. If $z = 5 + 10i$, then $|z| = \sqrt{5^2 + 10^2} = \sqrt{125}$
2. If $z = 3 - 2i$, then $|z| = \sqrt{3^2 + (-2)^2} = \sqrt{13}$
3. If $z = -3$, then $|z| = \sqrt{(-3)^2 + 0^2} = \sqrt{9} = 3$ (This is the absolute value!)
4. If $z = 2i$, then $|z| = \sqrt{0^2 + 2^2} = \sqrt{4} = 2$

Just like with real numbers, we can add and multiply complex numbers. (We will see later how to divide them.)

Definition 2.16 (Complex addition and multiplication) *Consider the complex numbers $z = a + bi$ and $w = c + di$. Then we can define:*

1. $z + w = (a + bi) + (c + di) = (a + c) + (b + d)i$
2. $zw = (a+bi)(c+di) = ac+adi+bci+bdi^2 = (ac-bd)+(ad+bc)i$ *(recalling $i^2 = -1$)*

Note 1: In the last step of the multiplication above, we have gathered the real part and the imaginary part together.
Note 2: The method used for the multiplication of two complex numbers is sometimes also referred to as the FOIL method (First-Outer-Inner-Last).

Example 2.17 Consider the following examples:

1. $(1 + 3i) + (-2 + i) = -1 + 4i$
2. $(1 + 3i)(-2 + i) = -2 + i - 6i + 3i^2 = -2 - 5i - 3 = -5 - 5i$

Observation 2.18 Now that we know how to multiply numbers, we can also use the complex conjugate to calculate the modulus of a complex number. Given $z = a + bi$, we see that:

$$\begin{aligned} z\bar{z} &= (a+bi)(a-bi) \\ &= a^2 + abi - abi + (-i)ib^2 \\ &= a^2 + b^2 \\ &= |z|^2 \\ \Rightarrow\ & |z| = \sqrt{z\bar{z}} \end{aligned}$$

It seems like we are in good shape to play around with complex numbers, but what if someone asked you to divide by a complex number? For example, what is the value, or the complex plane representation, of a number like:

$$\frac{1+i}{2-i}\ ?$$

First question: does it even make sense to divide by a complex number? Recall that, just like subtraction is the same thing as addition (i.e., subtraction is really just adding a negative number), division is the same thing as multiplication: i.e., dividing by x by y is really just a fancy way of saying "how many times do I have to multiply y to get x?" Therefore, since multiplication is well defined on complex numbers, so is division.

To help you visualize division by complex numbers so that you develop an intuition about it, we will use a little trick.

Observation 2.19 (Complex division) Since multiplication is well defined for complex numbers, so is division. Given a complex number $z = a + bi$, observe that:

$$\begin{aligned} \frac{1}{z} &= \frac{1}{z} \cdot \frac{\bar{z}}{\bar{z}} \quad \text{(we are just multiplying by 1!)} \\ &= \frac{\bar{z}}{|z|^2} \\ &= \frac{a-bi}{a^2+b^2} \\ &= \frac{a}{a^2+b^2} - \frac{bi}{a^2+b^2} \end{aligned}$$

Since the $a^2 + b^2$ is a real number, we found a way to express $\frac{1}{z}$ in the usual complex form $c + di$, where:

$$\begin{aligned} c &= \frac{a}{a^2+b^2}, \\ d &= \frac{-b}{a^2+b^2} \end{aligned}$$

Note that $\dfrac{1}{z}$ is also written as z^{-1}.

Example 2.20 Given the fraction $\dfrac{1+i}{2-i}$. We can clean it up as:

$$\begin{aligned}\frac{1+i}{2-i} &= \frac{(1+i)}{(2-i)} \cdot \frac{(2+i)}{(2+i)} \\ &= \frac{1+3i}{5} \\ &= \frac{1}{5} + \frac{3}{5}i\end{aligned}$$

Properties 2.21 (Summary of complex numbers properties) Below is a summary list of properties for complex numbers z and w (feel free to prove them for yourself if you are not convinced, or refer to Appendix B).

1. $z + w = w + z$ (commutativity of addition)
2. $\overline{z+w} = \overline{z} + \overline{w}$
3. $zw = wz$ (commutativity of multiplication) em $\overline{zw} = \bar{z}\bar{w}$
4. $z\bar{z} = \bar{z}z = |z|^2$
5. $\overline{\overline{z}} = z$
6. $|z| = |\overline{z}|$
7. $|zw| = |z||w|$
8. $|z + w| \leq |z| + |w|$
9. $z^{-1} = \dfrac{1}{z} = \dfrac{\bar{z}}{|z|^2}$ when $z \neq 0 + 0i$

2.3 Exponential representation of complex numbers: The polar form and Euler's formula

So far, we have explicitly written any complex number z in the form $z = a + bi$ and as you may have noticed, this form is not particularly well suited for multiplying and dividing complex numbers (which we will do a lot in quantum mechanics). Thankfully, there is a different way of handling complex numbers. In order to understand this method, you will need to be familiar with the exponential function e^x, as well as trigonometric functions using radians instead of degrees. If you are not familiar with either of those, consult Appendices C and D before continuing.

Theorem 2.22 (Euler's formula) *Euler's formula is a well-known result in complex numbers, establishing a deep relationship between trigonometric functions and complex exponential. It states that:*

$$e^{i\theta} = \cos\theta + i\sin\theta$$

where θ is ***a real number and in radians (e.g., unitless)****.*

Proof The proof of Euler's formula is not particularly complicated, but it does require the knowledge of Taylor Series. If you do have this knowledge, or are curious, I encourage you to visit Appendix E.

Observation 2.23 (Polar form notation of complex numbers) Any complex number $z = a + bi$ can be written in the form:

$$z = |z|e^{i\theta}$$

where θ is the angle (in radian) between the real axis and the complex number in the complex plane (see Figure 2). Therefore:

$$\theta = \arctan\left(\frac{b}{a}\right), \text{or } \theta = \arcsin\left(\frac{b}{|z|}\right), \text{or } \theta = \arccos\left(\frac{a}{|z|}\right)$$

θ is known as the **argument** of the complex number.

Proof By looking at the representation of $z = a + bi$ on the complex plane (Figure 2), we can see that:

$$\begin{array}{ll} a = |z|\cos\theta & \text{(projection along the real axis)} \\ b = |z|\sin\theta & \text{(projection along the imaginary axis)} \end{array}$$

By replacing a and b with these equivalent values, we can write:

$$\begin{aligned} z &= a + bi \\ &= |z|\cos\theta + i|z|\sin\theta \\ &= |z|(\cos\theta + i\sin\theta) \end{aligned}$$

By invoking Euler's formula, we thus conclude that:

$$z = |z|e^{i\theta}$$

Observation 2.24 (Periodicity) Referring to Euler's formula, the function $e^{i\theta}$ is a periodic function of θ with a period of 2π. For example:

$$\begin{aligned} e^{i(\theta\pm 2\pi)} &= \cos(\theta \pm 2\pi) + i\sin(\theta \pm 2\pi) \\ &= \cos\theta + i\sin\theta \\ &= e^{i\theta} \end{aligned}$$

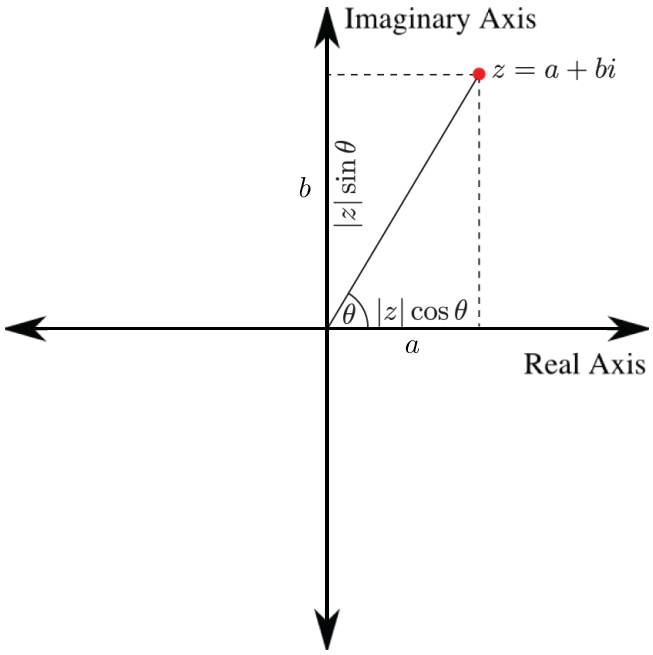


Figure 2: Using Euler's theorem, one can can represent a complex number using its modulus and its argument (the angle between the real axis and the complex number).

Given that observation, we can always assume the θ **has a value between 0 and** 2π. Similarly, other people prefer to limit the value of θ to any number between $-\pi$ and π. (Can you explain why it is the same?)

Example 2.25 What is the polar form of $z = 5 - 5i$? We first need to find the modulus of z, which is given by:

$$\begin{aligned} |z| &= \sqrt{5^2 + (-5)^2} \\ &= \sqrt{50} \end{aligned}$$

The argument is given by:

$$\begin{aligned} \theta &= \arctan(\left(\frac{5}{-5}\right) \\ &= \arctan(-1) \\ &= \frac{3\pi}{4} \text{ or } \frac{7\pi}{4}, \text{ (see Appendix D)} \end{aligned}$$

Since the the real part of z is positive and its imaginary is negative, we know that z lies in the 4th quadrant of the complex plane (refer to Figure 2 for reference), hence:

$$\theta = \frac{7\pi}{4}$$

Therefore:

$$5 - 5i = \sqrt{50}e^{i\frac{7\pi}{4}} = \sqrt{50}e^{-i\frac{\pi}{4}}$$

Observation 2.26 (Complex conjugate) What about the complex conjugate of a complex number in polar form? A little trigonometry shows us that:

$$\begin{aligned} \overline{e^{i\theta}} &= \overline{\cos\theta + i\sin\theta} \\ &= \cos\theta - i\sin\theta \\ &= \cos(-\theta) + i\sin(-\theta), \quad \text{since } \cos(-\theta) = \cos\theta, \text{ and } \sin(-\theta) = -\sin\theta \\ &= e^{-i\theta} \end{aligned}$$

Therefore, given a complex number $z = |z|e^{i\theta}$, we deduce that:

$$\begin{aligned} \bar{z} &= \overline{|z|e^{i\theta}} \\ &= \overline{|z|} \cdot \overline{e^{i\theta}} \\ &= |z|e^{-i\theta} \end{aligned}$$

The exponential notation is particularly well suited for multiplying, dividing and inverting complex numbers.

Properties 2.27 (Summary of properties) Even if the exponent of e is complex, all the basic properties of the exponential function (see Appendix C) are conserved. Given two complex numbers $z = |z|e^{i\theta}$ and $w = |w|e^{i\phi}$, below is a list of properties of complex numbers in the polar form:

1. $e^{i\theta}e^{i\phi} = e^{i(\theta+\phi)} \implies zw = (|z|e^{i\theta})(|w|e^{i\phi}) = |z||w|e^{i(\theta+\phi)}$
2. $(e^{i\theta})^n = e^{in\theta}$, for any real number n
3. From the above property $\Rightarrow \dfrac{1}{e^{i\theta}} = (e^{i\theta})^{-1} = e^{-i\theta}$
4. $|e^{i\theta}| = e^{i\theta} \cdot \overline{e^{i\theta}} = e^{i\theta}e^{-i\theta} = e^{i(\theta-\theta)} = e^0 = 1$
5. $\overline{e^{i\theta}} = e^{-i\theta}$
6. Since $e^{\pm 2\pi i} = 1$, then $e^{i(\theta \pm 2\pi)} = e^{i\theta} \cdot e^{\pm 2\pi i} = e^{i\theta}$ (this is a different way than above to show the periodicity)

3 Linear Algebra

In the previous section, we became familiar with a new family of numbers – complex numbers. As you may have noticed so far, although we have introduced new mathematical concepts, we have not really introduced new "mathematics" (e.g., integral calculus, functional analysis, etc.). We essentially only used addition and multiplication and expanded from there.

In the next section, we will introduce you to the language of Quantum Mechanics – Linear Algebra. Just like complex numbers, the type of linear algebra we will introduce here will necessitate only basic arithmetic, but used in clever ways. Welcome to the wonderful world of **vectors and matrices**.

Linear algebra is so much more than vectors and matrices, but for the purpose of QCSYS, that will be plenty! As mentioned, linear algebra is the language of Quantum Mechanics, but also the language of so many other things. Do you want to be an engineer, physicist, chemist, computer scientist? Learn linear algebra. Do you prefer to program video games? You definitely want to be an expert in linear algebra. You can even write a cookbook in the form of a matrix (more on that later).

The point is: basic linear algebra is rather simple, yet so useful for many applications. Applications aside, linear algebra is a fascinating, self-contained mathematical field on its own. But QCSYS being an multidisciplinary school in mathematics and physics, we will try to introduce the mathematical concepts of linear algebra, while making connection to potential applications (beyond Quantum Mechanics).

3.1 Vectors

Definition 3.1 (Vectors) *A vector is a column of numbers (any numbers, even complex). The amount of numbers is referred to as the dimension, or length, of the vector. For example, a 2-dimensional vector looks like:*

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

More generally, a n-dimension vector takes the form:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Notice the arrow on the v. This is a widely accepted convention to stress the fact that $\vec{v}$ refers to a vector. If we only consider the n-dimensional real vectors (i.e., v_1 and v_2 can only be real), we say the vectors lies in $\mathbb{R}^n$. If we also consider the n-dimensional complex vectors, we say they lie in $\mathbb{C}^n$.

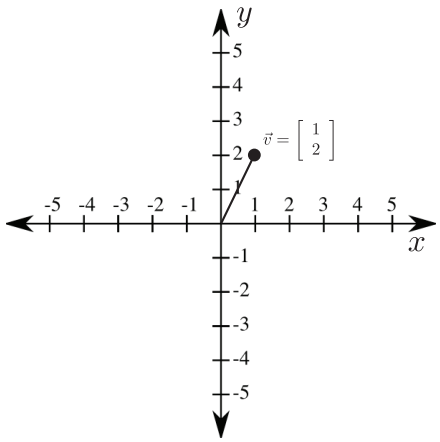


Figure 3: The cartesian coordinates or a point on a 2-dimensional surface can be written as a vector.

Even if you have never explicitly learned about vectors until now, you have already seen them. A 2-dimensional vector of real numbers is analogous to the cartesian coordinates (see Figure 3). *A* 3-dimensional vector of real numbers is analogous to the spatial coordinates in three dimensions. You can always think of a vector with n numbers as a point in n-dimensional space.

In high school physics, there is a good chance you have been told that a vector is nothing more than "a number and a direction". This is not false, but it is definitely not the whole story. It is a rather intuitive way of introducing vectors, hence why we will use this analogy extensively.

Let's try to add a little bit of abstraction: already, by using our knowledge and intuition about 2 and 3-dimensional space, it hints to the fact that each element of a vector can be used to represent quantities that are "exclusive" or "independent" of each other.

Intuition 3.2 Each number in the vector represents the value of some "property" that is unrelated and/or independent of the other properties. Too abstract? Refer to the 3-dimensional world: each number represents a position along the x, y and z-axis respectively. For example, if you are on the x-axis, you do not have any y or z coordinates. Hence those x, y and z properties are independent of each other.

For simplicity, we will generally use 2 and 3-dimensional vectors from now on, but everything explained below applies to vectors of arbitrary sizes.

Definition 3.3 (Vector Addition) *Adding vectors is easy, just add each respective en-*

try! For the vectors $\vec{v}$ and $\vec{w}$ written explicitly as:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \textit{and} \quad \vec{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

adding them gives:

$$\vec{v} + \vec{w} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \end{bmatrix}$$

Example 3.4 Adding the vectors:

$$\vec{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \vec{w} = \begin{bmatrix} 3 \\ -100 \end{bmatrix}$$

will give you:

$$\vec{v} + \vec{w} = \begin{bmatrix} 1+3 \\ 2-100 \end{bmatrix} = \begin{bmatrix} 4 \\ -98 \end{bmatrix}$$

Observation 3.5 (Geometrical representation of vector addition – parallelogram technique) If we think of a 2-dimensional vector as a point in the cartesian plane, adding two vectors can be done using the parallelogram technique (i.e., putting the second vector at the end of the first). See Figure 4 where we have used the parallelogram technique to add:

$$\vec{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Intuition 3.6 Imagine the vector $\vec{v}$ and $\vec{w}$ above represent two different movements on a 2-dimensional plane. For example, $\vec{v}$ can be thought as moving along the x-axis by 3 and along the y-axis by 1. Similarly for $\vec{w}$. Adding two vectors is essentially the same thing as saying:

> " Start at the origin. Move along x by 3, then along y by 1. After that, move along x by 1, and then along y by 2. Your final position will be at 4 along x and at 3 along y."

Note 3.7 You cannot add vectors of different length. For example, given:

$$\vec{v} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \text{ and } \vec{w} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

What is $\vec{v} + \vec{w}$? The answer, unfortunately, doesn't exist. (Or if you prefer, the answer is not defined.)

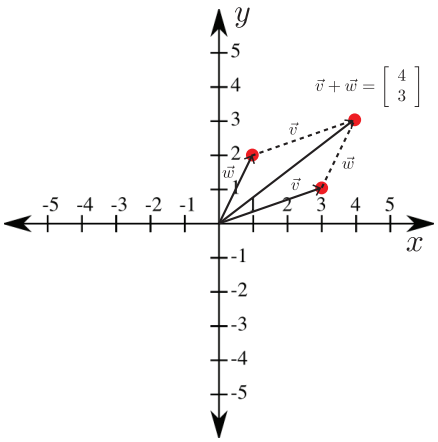


Figure 4: You can use the parallelogram technique to visualize vector addition in 2-dimensions.

Definition 3.8 (Vector Scalar Multiplication) *You can scale a vector by just multiplying each entry by the scalar (the number). If we have:*

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \text{and} \quad c \text{ (a number)}$$

then scaling gives:

$$c\vec{v} = \begin{bmatrix} cv_1 \\ cv_2 \end{bmatrix}$$

Example 3.9

$$700 \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 700 \\ -1400 \end{bmatrix}$$

This is the same as if we would have asked you to move by the vector $\vec{v}$, 700 times!

Observation 3.10 Scalar multiplication of a vector by a positive real number does not change its orientation, only its length. See Figure 5.

Observation 3.11 Scalar multiplication of a vector by a negative number inverts its direction. See Figure 6.

Note 3.12 Multiplication of vectors is not well defined, for reasons that are beyond the scope of this document. However, there is such a thing as an "inner product", which is a type of multiplication of two vectors, but the result is a scalar number... more on this later.

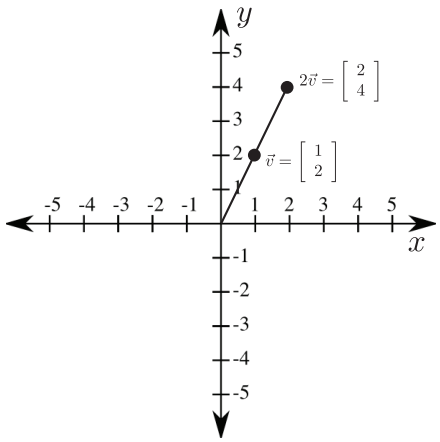


Figure 5: Scaling a vector by a positive number does not change its direction.

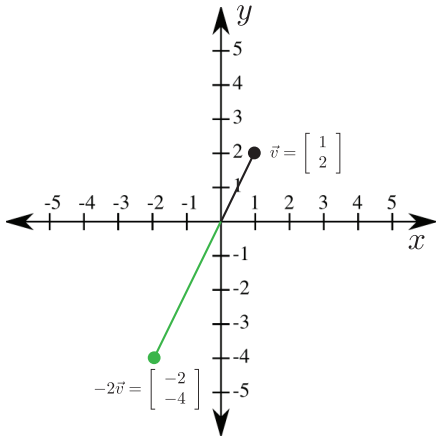


Figure 6: Scaling a vector by a negative number inverts its direction.

So far, we only used examples with real vectors and scalars. What about complex vectors and scalars? All the arithmetic is exactly the same, except that an intuitive understanding of what is going on might be of outside human perception! Just like trying to visualize a 4 or 1,300-dimensional vector might prove to be impossible! But yet, mathematically, these concepts are sound.

Properties 3.13 (Properties of vector addition and scalar multiplication) Let $\vec{v}$, $\vec{w}$ and $\vec{u}$ be vectors, c and d be scalars. Then vector addition and scalar multiplication have the following properties (again, feel free to prove them if you are not convinced):

1. $\vec{v} + \vec{w} = \vec{w} + \vec{v}$ (commutativity of vector addition)
2. $\vec{v} + (\vec{w} + \vec{u}) = (\vec{v} + \vec{w}) + \vec{u}$
3. $c(\vec{v} + \vec{w}) = c\vec{v} + c\vec{w}$
4. $(c + d)\vec{v} = c\vec{v} + d\vec{v}$

These are similar properties to real and complex numbers (numbers are actually vectors of length/dimension 1).

3.2 Matrices

Now that vectors have no more secrets for us, let's increase the complexity a little bit by introducing the concept of **matrices**. As we will see, vectors and matrices play very well with each other.

Definition 3.14 (Matrices) *A matrix is a box of numbers (real or complex). Some examples are:*

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad N = \begin{bmatrix} 3 & 3 & -1+2i \\ 1 & -3 & 0 \end{bmatrix}, \quad P = \begin{bmatrix} 2 & -1 \\ -1+2i & 3 \\ -3 & -3+i \end{bmatrix}$$

Given any matrix Q, we denote the element $Q_{i,j}$ as the number in the i^{th} row and the j^{th} column. For example, $P_{3,1} = -3$. Of course, $P_{2,3}$ is not defined as P only has 2 columns. If a matrix has m rows and n columns, we say it is an $m \times n$ matrix, or a $m \times n$ dimensional matrix. A square matrix is, as the name suggests, a matrix with the same number of rows and columns. M is an example of a square matrix.

Note 3.15 *A vector can also be thought of as an $n \times 1$ matrix.*

Definition 3.16 (Matrix addition and scalar multiplication) Similar to vector addition, matrix addition consists of adding each corresponding matrix element. Addition is only defined for matrices of the same dimensions. Similarly, scalar multiplication is simply the scaling of all the elements of the matrix.

Example 3.17

$$\begin{bmatrix} 3 & 3 & -1+2i \\ 1 & -3 & 0 \end{bmatrix} + \begin{bmatrix} 2 & -1 & i \\ 0 & 2+3i & -3 \end{bmatrix} = \begin{bmatrix} 3+2 & 3-1 & (-1+2i)+i \\ 1+0 & -3+(2+3i) & 0-3 \end{bmatrix}$$
$$= \begin{bmatrix} 5 & 2 & -1+3i \\ 1 & -1+3i & -3 \end{bmatrix}$$

As for scalar multiplication:

$$2\begin{bmatrix} 3 & 3 & -1+2i \\ 1 & -3 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 6 & -2+4i \\ 2 & -6 & 0 \end{bmatrix}$$

Before being able to help you intuitively understand matrices, we will need to define a few new concepts. Unlike vectors, it is possible to define matrix multiplication.

Definition 3.18 (Matrix multiplication) *Given 2 matrices M and N, as long as the* ***number of columns in*** M ***is the same as the number of rows in*** N*, we can define a multiplication operation on matrices as follows:*

> *The element $(MN)_{i,j}$ of the resulting matrix is given by multiplying the first number of the i^{th} row of M with the first number of the j^{th} column of N PLUS the multiplication of the second number of the i^{th} row of M with the second number of the j^{th} column of N PLUS the multiplication of the third number of the i^{th} row of M with the third number of the j^{th} column of N and so on and so forth, until you reach the end of the i^{th} row/j^{th} column (hence the requirement of the same length of the rows of M and the columns of N).*

Observation 3.19 If M is an $m \times n$ matrix and N is an $n \times l$ matrix, NM will give you a $m \times l$ matrix.

Example 3.20 Although the definition of matrix multiplication can be a little confusing, an explicit multiplication might clarify this simple task:

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} g & h \\ j & k \\ l & q \end{bmatrix} = \begin{bmatrix} ag+bj+cl & ah+bk+cq \\ dg+ej+fl & dh+ek+fq \end{bmatrix}$$

Example 3.21

$$\begin{bmatrix} 2 & 3 & i \\ 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 12 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 2\cdot 0+3\cdot 0+i\cdot 3 & 2\cdot 1+3\cdot 12+i\cdot(-2) \\ 3\cdot 0-2\cdot 0+1\cdot 3 & 3\cdot 1-2\cdot 12+1\cdot(-2) \end{bmatrix}$$
$$= \begin{bmatrix} 3i & 38-2i \\ 3 & -23 \end{bmatrix}$$

Observation 3.22 (Matrices as functions) Since a vector is a also a matrix, we can think of matrices as mapping vectors to vectors, i.e., $m \times n$ matrix maps an n-dimensional vector to an m-dimensional vector. For example, given the vector:

$$\vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

we have that P (as defined above) maps $\vec{v}$ to:

$$\begin{aligned} P\vec{v} &= \begin{bmatrix} 2 & -1 \\ -1+2i & 3 \\ -3 & -3+i \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 \cdot 2 + (-1) \cdot 1 \\ (-1+2i) \cdot 2 + 3 \cdot 1 \\ (-3) \cdot 2 + (-3+i) \cdot 1 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ 1+4i \\ -9+i \end{bmatrix} \end{aligned}$$

This is not unlike a function on real numbers, e.g., $f(x) = x^2 + 3$. A scalar function f takes a number as an input and gives you a number as an output. A $m \times n$ matrix works exactly like a scalar function, but it takes an n-dimensional vector as an input and outputs an m-dimensional vector, i.e., it is a function between an n-dimensional vector space to an m-dimensional vector space.

Example 3.23 Let $\vec{v}$ be a 2-dimensional vector (e.g., a point on the catesian plane), and consider:

$$M = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The matrix M acts as:

$$\begin{aligned} M\vec{v} &= M \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ &= \begin{bmatrix} v_1 \\ -v_2 \end{bmatrix} \end{aligned}$$

As you can see, the matrix M performs a reflection about the x-axis of the cartesian plane!

Exercise 3.24 Can you think of a 2×2 matrix that would represent a reflection about the y-axis? How about a reflection about the axis making a 45° degree angle with the x-axis? [1]

[1] The answers are $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ respectively.

When we consider a matrix as a function on vectors, we often refer to it as the matrix "operating" on the vectors. Hence, **we often refer to it as an operator**.

Example 3.25 Can you perform this multiplication $M\vec{v}$, if:

$$M = \begin{bmatrix} 1 & 3 \\ 4 & 7 \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}?$$

No. Since the matrix is a 2×2 and the vector is 3×1, this multiplication is undefined.

Observation 3.26 (Non-Commutativity of matrix multiplication) In mathematics, we say an operation is commutative if the order of operation is irrelevant. For example, if a and b are any scalar number, then $a + b = b + a$ and $ab = ba$.

Clearly, matrix addition and matrix scalar multiplication are commutative operations, but what about matrix multiplication? First, we must notice that asking the question about commutativity only makes sense for square matrices. Why? Let's look at the following example:

$$M = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}, \quad N = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$$

Multiplying in two different orders will give:

$$\begin{aligned} MN &= \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 8 & 5 \\ 14 & 10 \end{bmatrix} \\ NM &= \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 13 & 9 \\ 5 & 5 \end{bmatrix} \end{aligned}$$

These are clearly not the same matrices, therefore $MN \neq NM$. (There are cases where matrices will commute, but in general they do not.)

The fact that matrix multiplication is not commutative is what makes linear algebra radically different than number arithmetic. Without going into the details yet, we will see that one striking difference between classical and quantum physics is that "observables" (i.e., physical quantities you can measure) in quantum mechanics are also non-commutative. For example, measuring the position of a quantum particle then its momentum will not give you the same answer as if you were measuring its momentum and then its position. But we are getting ahead of ourselves. Let's stick to the math for now!

Observation 3.27 (Composition of function) Often in mathematics, we need to do a composition of function, that is, applying a function g on an input, then applying another function f using the first output as input, and so on. Abstractly, we denote the composition of function f and g as $f \circ g$ and it is defined as:

$$f \circ g(x) = f\left(g(x)\right)$$

Transposing this to matrices and vectors, say that $f(\vec{v}) = M\vec{v}$ and $g(\vec{v}) = N\vec{v}$ for any vector $\vec{v}$ of the right dimension, we have:

$$\begin{aligned} f \circ g(\vec{v}) &= f(g(\vec{v})) \\ &= f(N\vec{v}) \\ &= MN\vec{v} \end{aligned}$$

Therefore, applying the function f after the function g, is the same as applying the operator N, and then the operator M. The resulting operator will be given by:

$$f \circ g = MN$$

Example 3.28 Given the function f represented by matrix M and function g by matrix N, such that:

$$M = \begin{bmatrix} 3 & 1 & 2i \\ 1 & -3 & 0 \end{bmatrix} \quad \text{and} \quad N = \begin{bmatrix} 4 & 0 \\ -i & 5 \\ -3 & -3 \end{bmatrix}$$

First thing to notice is that since M is a 2×3 matrix and N is a 3×2 matrix, the multiplication MN is well defined. If we want to evaluate the composition of function $f \circ g(\vec{v})$ on vector:

$$\vec{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

we can use two methods:

1. We can evaluate $g(\vec{v})$ first and then $f(g(\vec{v}))$, i.e.,

$$\begin{aligned} g(\vec{v}) &= N\vec{v} \\ &= \begin{bmatrix} 4 & 0 \\ -i & 5 \\ -3 & -3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 12 \\ 10 - 3i \\ -15 \end{bmatrix} \end{aligned}$$

And then we have:

$$\begin{aligned} f(g(\vec{v})) &= Mg(\vec{v}) \\ &= \begin{bmatrix} 3 & 1 & 2i \\ 1 & -3 & 0 \end{bmatrix} \begin{bmatrix} 12 \\ 10-3i \\ -15 \end{bmatrix} \\ &= \begin{bmatrix} 46-33i \\ -18+9i \end{bmatrix} \end{aligned}$$

2. We could evaluate $f \circ g = MN$ first, and then apply this operator to $\vec{v}$:

$$\begin{aligned} MN &= \begin{bmatrix} 3 & 1 & 2i \\ 1 & -3 & 0 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ -i & 5 \\ -3 & -3 \end{bmatrix} \\ &= \begin{bmatrix} 12-7i & 5-6i \\ 4+3i & -15 \end{bmatrix} \end{aligned}$$

Therefore we have:

$$\begin{aligned} f \circ g(\vec{v}) &= MN\vec{v} \\ &= \begin{bmatrix} 12-7i & 5-6i \\ 4+3i & -15 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 46-33i \\ -18+9i \end{bmatrix} \end{aligned}$$

Remember, matrices do not commute under matrix multiplication, therefore **it is crucial to understand that matrix composition goes from right to left**, i.e., the rightmost matrix is the one being applied first.

Properties 3.29 (Properties of matrix arithmetics) *For the sake of listing properties of matrix addition and multiplication, we will assume the M, N and P are compatible for the given operations performed:*

1. $M + N = N + M$
2. $(M + N) + P = M + (N + P)$
3. $c(M + N) = cM + cN$ *for any scalar* c
4. $(c + d)M = cM + dM$, *for any scalar* c *and* d
5. $c(MN) = (cM)N = M(cN) = (MN)c$, *for any scalar* c

6. $(MN)P = M(NP)$
7. $(M+N)P = MP + NP$
8. $M(N+P) = MN + MP$
9. $MN \neq NM$, *most of the time*

Definition 3.30 (Linearity) *In mathematics, the concept of linearity plays a very important role. Mathematically, a linear function, or linear map, or linear operator f is a function that satisfies:*

1. $f(x+y) = f(x) + f(y)$, *for* ***any input*** x *and* y
2. $f(cx) = cf(x) for$ ***any input*** x *and* ***any scalar*** c

Put in words, the first condition means that the output of a function acting on a sum of inputs in just equal to the sum of the individual outputs. The second condition implies that the output of a scaled input, is just the scaled output of the original input.

Example 3.31 The world is full of linear processes. Imagine, for example, that we would charge you \$100/day to attend QCSYS. How much will it cost you if:

1. QCSYS is 5 days long? \$500.
2. QCSYS is 20 days long? \$2,000.
3. You decided to come for 5 days, but then we extended the offer to stay by 2 days (at full price)? \$700.
4. You come for 5 days, and then decide to stay twice as long (at full price)? \$1,000.

What you just did intuitively is use the concept of linearity. Let's put some mathematical rigour into this. Let f be the function calculating the cost of your stay as a function of the number of days, x you stay. You can convince yourself that:

$$f(x) = 100x$$

It is easy, from the definition of linearity, to show that f is indeed a linear function, i.e.,

- $f(x+y) = 100(x+y) = 100x + 100y = f(x) + f(y)$
- $f(cx) = 100cx = c(100x) = cf(x)$

To reconcile with our intuition, we can think of the four different scenarios above as:

1. $f(x)$, for $x = 5$

2. $f(x)$, for $x = 20$

3. $f(x + y)$, for $x = 5$ and $y = 2$

4. $f(2x)$, for $x = 5$

Example 3.32 Are matrices, when seen as a function from vectors to vectors, linear? If we define a function from n-dimensional vectors to m-dimensional vectors using an $m \times n$ matrix M, e.g.,

$$f(\vec{v}) = M\vec{v}, \text{ for } \vec{v} \in \mathbb{R}^n$$

then using the properties of matrix arithmetics listed above, it is straightforward to show that matrices can be thought as linear functions, i.e.,

- $f(\vec{v} + \vec{w}) = M(\vec{v} + \vec{w}) = M\vec{v} + M\vec{w} = f(\vec{v}) + f(\vec{w})$
- $f(c\vec{v}) = M(c\vec{v}) = cM\vec{v} = cf(\vec{v})$

Hence the term "linear" algebra!

Example 3.33 The function $f(x) = x^2$, for x being any scalar, is not linear because:

$$\begin{aligned} f(x + y) &= (x + y)^2 \\ &= x^2 + 2xy + y^2 \end{aligned}$$

but:

$$f(x) + f(y) = x^2 + y^2$$

Therefore:

$$f(x + y) \neq f(x) + f(y)$$

Similarly:

$$\begin{aligned} f(cx) &= c^2x^2 \\ &\neq cf(x), \text{ for } c \text{ being any scalar.} \end{aligned}$$

3.3 Complex conjugate, transpose and conjugate transpose

To finish this section, we will introduce a couple of other concepts related to matrices and vectors. They will become very handy soon enough!

Definition 3.34 (Matrix/vector complex conjugate) *The complex conjugate of a matrix (or vector) is defined as taking the complex conjugate on all its entries, for example if:*

$$\vec{v} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad M = \begin{bmatrix} c & d \\ f & g \end{bmatrix}$$

then:

$$\overline{\vec{v}} = \begin{bmatrix} \bar{a} \\ \bar{b} \end{bmatrix}, \quad \overline{M} = \begin{bmatrix} \bar{c} & \bar{d} \\ \bar{f} & \bar{g} \end{bmatrix}$$

Example 3.35

$$M = \begin{bmatrix} 1 & e^{-i\frac{\pi}{5}} \\ 3-i & 10 \end{bmatrix} \implies \overline{M} = \begin{bmatrix} 1 & e^{i\frac{\pi}{5}} \\ 3+i & 10 \end{bmatrix}$$

Definition 3.36 (Matrix/vector transpose) *The transpose of a matrix M, denoted M^t is such that the n^{th} row of M^t is the same as the n^{th} column of M (note: the transpose of an $m \times n$ matrix is an $n \times m$ matrix). Similarly, the transpose of a column vector $\vec{v}$, denoted $\vec{v}^t$ is just a row vector with the same entries as $\vec{v}$. For example, if:*

$$\vec{v} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad M = \begin{bmatrix} c & d \\ f & g \end{bmatrix}$$

then:

$$\vec{v}^t = \begin{bmatrix} a & b \end{bmatrix}, \quad M^t = \begin{bmatrix} c & f \\ d & g \end{bmatrix}$$

Note that the transpose of a scalar (i.e. 1×1 matrix/vector) is itself.

Example 3.37 Given the following vector and matrix:

$$\vec{v} = \begin{bmatrix} 1 \\ 3 \\ 1+3i \end{bmatrix}, \quad \text{and} \quad M = \begin{bmatrix} 2 & 1-i \\ 0 & 0 \\ e^{i\frac{\pi}{3}} & 4i \end{bmatrix}$$

then:

$$\vec{v}^t = \begin{bmatrix} 1 & 3 & 1+3i \end{bmatrix}, \quad \text{and} \quad M^t = \begin{bmatrix} 2 & 0 & e^{i\frac{\pi}{3}} \\ 1-i & 0 & 4i \end{bmatrix}$$

Definition 3.38 (Matrix/vector conjugate transpose) *The conjugate transpose of a matrix M or a vector $\vec{v}$, denoted $M^\dagger$ ("M dagger") and $\vec{v}^\dagger$ respectively, is given by taking the complex conjugate, and then the transpose. For example, if:*

$$\vec{v} = \begin{bmatrix} 1+i \\ 3 \end{bmatrix}, \quad M = \begin{bmatrix} i & 3-2i \\ -2 & 1-4i \end{bmatrix}$$

then:

$$\vec{v}^\dagger = \begin{bmatrix} 1-i & 3 \end{bmatrix}, \quad M^\dagger = \begin{bmatrix} -i & -2 \\ 3+2i & 1+4i \end{bmatrix}$$

The "dagger" notation is usually preferred by physicists, while mathematicians will often use a superscripted *. It has been a long standing debate of stubbornness from the two camps!

Note 3.39 Taking the complex conjugate and then the transpose is the same as taking the transpose then the complex conjugate. That is to say:

$$M^\dagger = (\overline{M})^t = \overline{M^t}$$

Properties 3.40 Below is a list of properties of the complex conjugate, the transpose and the conjugate transpose ("the dagger") of matrix/vectors. Feel free to prove them as an exercise.

Let M and N be any matrices/vectors compatible for multiplication, then:

1. $\overline{MN} = \overline{M}\,\overline{N}$
2. $(MN)^t = N^t M^t$ (notice the reversal of the multiplication order)
3. $(MN)^\dagger = N^\dagger M^\dagger$ (notice the reversal of the multiplication order)
4. $\overline{(\overline{M})} = M$
5. $(M^t)^t = M$
6. $(M^\dagger)^\dagger = M$

3.4 Inner Product and Norms

In this section, we will concentrate on defining some concepts that are mostly applicable for vectors. These concepts can be generalized to matrices and beyond, but the explicit forms given here are reserved for vectors.

Definition 3.41 (Inner Product) *The inner product, also known as the dot product or the scalar product, of two vectors $\vec{v}$ and $\vec{w}$, denoted $\vec{v}\bullet\vec{w}$ is a defined mathematical operation between two vectors of the* ***same dimension*** *that returns a scalar number.*

To take the inner product of two vectors, first take the complex conjugate of the first vector, and then multiply each corresponding numbers in both vectors and then add everything. Explicitly, for vectors:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad \text{and} \quad \vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

the inner product of $\vec{v}$ and $\vec{w}$ gives:

$$\vec{v} \bullet \vec{w} = \overline{v_1} \cdot w_1 + \overline{v_2} \cdot w_2 + \dots.$$

Note 3.42 The inner product is only defined for vectors of the same dimension/length.

Example 3.43 If we have:

$$\vec{v} = \begin{bmatrix} i \\ 2+i \end{bmatrix}, \quad \textit{and} \quad \vec{w} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

then:

$$\begin{aligned} \vec{v} \bullet \vec{w} &= (-i) \cdot 2 + (2-i) \cdot (-1) \\ &= -2 - i \end{aligned}$$

Observation 3.44 By looking at the definition of the inner products, we can see that taking the inner product of $\vec{v}$ and $\vec{w}$ is the equivalent of doing a matrix multiplication between $\vec{v}^\dagger$ and $\vec{w}$. That is:

$$\begin{aligned} \vec{v}^\dagger \vec{w} &= \begin{bmatrix} \overline{v_1} & \overline{v_2} & \dots & \overline{v_n} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \\ &= \overline{v_1} \cdot w_1 + \overline{v_2} \cdot w_2 + \dots \\ &= \vec{v} \bullet \vec{w} \end{aligned}$$

Properties 3.45 Notice the following properties of the inner product. Let $\vec{u}$, $\vec{v}$ and $\vec{w}$ be vectors, a and b be scalars and M a matrix of suitable dimension. Prove this as an exercise:

1. $\vec{v} \bullet (a\vec{w}) = a(\vec{v} \bullet \vec{w})$
2. $(a\vec{v}) \bullet \vec{w} = \bar{a}(\vec{v} \bullet \vec{w})$ (notice the complex conjugate of a)
3. $\vec{v} \bullet [(a+b)\vec{w}] = (a+b)\vec{v} \bullet \vec{w} = a(\vec{v} \bullet \vec{w}) + b(\vec{v} \bullet \vec{w})$
4. $[(a+b)\vec{v}] \bullet \vec{w} = (\bar{a}+\bar{b})\vec{v} \bullet \vec{w} = \bar{a}(\vec{v} \bullet \vec{w}) + \bar{b}(\vec{v} \bullet \vec{w})$
5. $(\vec{u}+\vec{v}) \bullet \vec{w} = \vec{u} \bullet \vec{w} + \vec{v} \bullet \vec{w}$
6. $\vec{u} \bullet (\vec{v}+\vec{w}) = \vec{u} \bullet \vec{v} + \vec{u} \bullet \vec{w}$

7. $\vec{v} \bullet \vec{w} = \overline{\vec{w} \bullet \vec{v}}$

8. $\vec{v} \bullet (M\vec{w}) = (M^\dagger \vec{v}) \bullet \vec{w}$

Example 3.46 Let's prove the above property 7. We need to first make a crucial observation: since $\overline{\vec{w} \bullet \vec{v}}$ is a scalar number, taking its transpose will not change anything (recall that the transpose of a scalar is itself). Therefore we have that:

$$\begin{aligned}
\overline{\vec{w} \bullet \vec{v}} &= \overline{\vec{w} \bullet \vec{v}}^{\,t} \\
&= (\vec{w} \bullet \vec{v})^\dagger, \text{ by definition of the conjugate transpose} \\
&= \left(\vec{w}^\dagger \vec{v}\right)^\dagger, \text{ recalling Observation 3.44} \\
&= \vec{v}^\dagger \left(\vec{w}^\dagger\right)^\dagger \text{ (Property 3.40-3)} \\
&= \vec{v}^\dagger \vec{w} \\
&= \vec{v} \bullet \vec{w}
\end{aligned}$$

Example 3.47 Let's now prove the last property above:

$$\begin{aligned}
\vec{v} \bullet (M\vec{w}) &= \vec{v}^\dagger (M\vec{w}), \text{ recalling Observation 3.44} \\
&= (\vec{v}^\dagger M)\vec{w} \\
&= (M^\dagger \vec{v})^\dagger \vec{w}, \text{ since } (M^\dagger \vec{v})^\dagger = \vec{v}^\dagger (M^\dagger)^\dagger = \vec{v}^\dagger M \text{ (Property 3.40-3)} \\
&= (M^\dagger \vec{v}) \bullet \vec{w}
\end{aligned}$$

Now that we have defined the inner product, we can, just like any mathematician likes to do, start expanding our back of definitions.

Definition 3.48 (Orthogonal vectors) *We say that two vectors are orthogonal, or perpendicular, if their inner product is* 0.

Example 3.49 The vectors:

$$\begin{bmatrix} i \\ i \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

are orthogonal because:

$$\begin{aligned}
\begin{bmatrix} i \\ i \end{bmatrix} \bullet \begin{bmatrix} 1 \\ -1 \end{bmatrix} &= \begin{bmatrix} -i & -i \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\
&= -i + i \\
&= 0
\end{aligned}$$

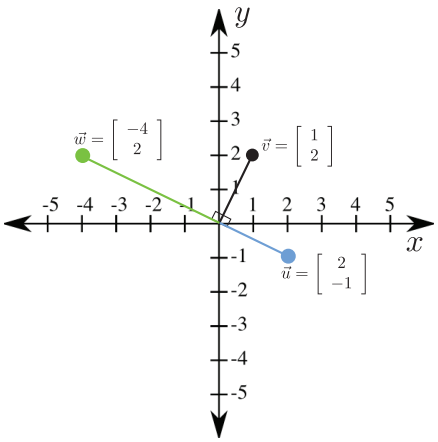


Figure 7: 2-dimensional visualization of orthogonal vectors

Observation 3.50 If we refer back to the intuitive representation of a 2-dimensional real vector as a point on the cartesian plane, we see that orthogonal vectors always have an angle of 90° between them. For example, given the vector:

$$\vec{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

then both vectors:

$$\vec{w} = \begin{bmatrix} -4 \\ 2 \end{bmatrix} \text{ and } \vec{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

are orthogonal to $\vec{v}$. See Figure 7 for visualization.

Definition 3.51 (Vector Norm) *The norm (or length) of a vector $\vec{v}$, denoted $\|\vec{v}\|$ is given by:*

$$\begin{aligned} \|\vec{v}\| &= \sqrt{\vec{v} \bullet \vec{v}} \\ &= \sqrt{\vec{v}^{\dagger}\vec{v}} \\ &= \sqrt{|v_1|^2 + |v_2|^2 + \ldots + |v_n|^2}. \end{aligned}$$

*A vector with norm 1 is called a **unit vector**. Based on this definition, given a scalar c, we also have:*

$$\|c\vec{v}\| = |c| \, \|\vec{v}\|$$

Note 3.52 Notice that we have defined the "length" of a vector in two different ways: the first time we referred it to the dimension of $\vec{v}$, while this time, it refers to its norm. Let it be clear that these are **two different concepts**. Using the same word for two different concepts is unfortunately something that happens often in mathematics (and physics for that matter). Usually, the context in which these terms are used will make the reference clear.

Example 3.53 To find how long the vector is:

$$\vec{v} = \begin{bmatrix} 1 \\ -2 \\ i \end{bmatrix}$$

we just need to plug and chug!

$$\begin{aligned} \|v\| &= \sqrt{|1|^2 + |-2|^2 + |i|^2} \\ &= \sqrt{6} \end{aligned}$$

Example 3.54 The norm of the vector:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

is:

$$\|\vec{v}\| = \sqrt{|v_1|^2 + |v_2|^2}$$

Referring back to the cartesian representation of a vector, this definition of length is intuitive and appropriate (e.g., represents the distance from the origin to that point).

Definition 3.55 (Unit Vectors, Normalizing) *By definition, unit vectors have norm equal to* 1. *Normalizing a nonzero vector* $\vec{v}$ *means to scale by* $\frac{1}{\|\vec{v}\|}$ *to make it have length* 1. *This is readily seen since:*

$$\begin{aligned} \left\| \frac{\vec{v}}{\|\vec{v}\|} \right\| &= \frac{1}{\|\vec{v}\|} \|\vec{v}\|, \text{ since } \frac{1}{\|\vec{v}\|} \text{ is a positive scalar} \\ &= 1 \end{aligned}$$

Example 3.56 Let's normalize the following vector:

$$\vec{v} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Since $\|\vec{v}\| = \sqrt{|1|^2 + |-2|^2} = \sqrt{5}$, we can scale it by $1/\sqrt{5}$ to get the unit vector

$$\frac{1}{\sqrt{5}}\vec{v} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

3.5 Bases

A very important concept in linear algebra is that of "bases". In simple words, a basis is a finite set of vectors that can be used to describe any other vectors of the same dimension. For example, any 2-dimensional vectors $\vec{v}$ can be decomposed as:

$$\begin{aligned} \vec{v} &= \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ &= v_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned}$$

The set:

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

is a basis for vectors of length 2 and v_1 and v_2 are the coefficient of each basis vector. This example clearly relates to the cartesian plane, where we can always describe any point on the plane as "how much of x" and "how much of y".

Example 3.57 We can also write any 2-dimensional vector as:

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \frac{v_1 + v_2}{\sqrt{2}} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} + \frac{v_1 - v_2}{\sqrt{2}} \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

In this example, the basis is given by the set:

$$\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

and the coefficients are $\frac{v_1 + v_2}{\sqrt{2}}$ and $\frac{v_1 - v_2}{\sqrt{2}}$

Can you think of a different basis? The fact of the matter is, there is an infinite number of them! As any decent mathematician would (let's face it, by now we are more than decent!), we will put some mathematical rigour into it.

Definition 3.58 (Linear Combination) *A linear combination is a combination of vector addition and scalar multiplication. For example, for vectors $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}$ and scalars $c_1, c_2, \ldots, c_n$, a linear combination looks like:*

$$c_1\vec{v_1} + c_2\vec{v_2} + \cdots + c_n\vec{v_n}$$

Definition 3.59 (Linearly Independent Vectors) *We say that a set of vectors is linearly dependent if one can be written as a linear combination of the others. Otherwise they are linearly independent.*

Example 3.60 *Given the three vectors:*

$$\left\{ \vec{v_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v_2} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{v_3} = \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix} \right\}$$

These vectors are linearly dependent since:

$$\vec{v_3} = 2 \cdot \vec{v_1} + \vec{v_2}$$

Note 3.61 For n-dimensional vectors, you cannot have a set of more than n linearly independent vectors.

Definition 3.62 (Basis) *Any set of n linearly independent vectors in $\mathbb{C}^n$ (or $\mathbb{R}^n$) is called a basis of $\mathbb{C}^n$ (or $\mathbb{R}^n$).*

Example 3.63 *The set given by:*

$$\left\{ \vec{v_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v_2} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{v_3} = \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix} \right\}$$

is not a basis for $\mathbb{R}^3$ since, as previously shown, they are linearly dependent.

Example 3.64 *The set given by:*

$$\left\{ \vec{v_1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v_2} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$$

is a basis for $\mathbb{R}^2$ since there are two of them and are linearly independent.

Observation 3.65 (Using a basis to write any vectors) By the very definition of a basis, we can argue that any n-dimensional vector $\vec{w}$ can be written as linear combination of the basis vector $\{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$. To see this fact, consider the set:

$$\{\vec{w}, \vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$$

Since this set has $n+1$ vectors, it cannot be a linearly independent set. By assumption, we know that $\{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ are linearly independent, therefore there must exist scalar coefficients $c_1, \ldots, c_n$ such that:

$$\vec{w} = c_1\vec{v_1} + c_2\vec{v_2} + \cdots + c_n\vec{v_n}$$

Example 3.66 Using the basis:

$$\left\{\vec{v_1} = \begin{bmatrix} 1 \\ i \end{bmatrix}, \quad \vec{v_2} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$$

we will find a way to rewrite the vector:

$$\vec{w} = \begin{bmatrix} 3 \\ 1+i \end{bmatrix}$$

as a linear superposition of $\vec{v_1}$ and $\vec{v_2}$. To do this, we will need to solve a simple system of 2 equations and 2 unknowns:

$$\begin{aligned} \begin{bmatrix} 3 \\ 1+i \end{bmatrix} &= c_1 \begin{bmatrix} 1 \\ i \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} c_1 + c_2 \\ ic_1 + 2c_2 \end{bmatrix} \\ \implies & c_1 + c_2 = 3 \qquad (1) \\ & ic_1 + 2c_2 = 1 + i \qquad (2) \end{aligned}$$

Performing $2 \times (1) - (2)$ gives us

$$\begin{aligned} (2-i)c_1 &= 5 - i \\ \implies & c_1 = \frac{5-i}{2-i} = \frac{5-i}{2-i} \cdot \frac{2+i}{2+i} = \frac{11}{5} + \frac{3i}{5} \end{aligned}$$

recalling how to do complex division from Section 2.19.

By reinserting our result for c_1 into (1), we get:

$$\begin{aligned} c_2 &= 3 - c_1 \\ &= \frac{4}{5} - \frac{3i}{5} \end{aligned}$$

You can verify for yourself that the result:

$$\vec{w} = \left(\frac{11}{5} + \frac{3i}{5}\right)\vec{v_1} + \left(\frac{4}{5} - \frac{3i}{5}\right)\vec{v_2}$$

is correct.

Observation 3.67 Because of the linearity of matrix multiplication, knowing the "action" of a matrix M on each vector of a given basis of $\mathbb{C}^n$ is enough to determine the action of M on any vectors in $\mathbb{C}^n$. To see this, let's use the basis $\{\vec{v_1}, \ldots \vec{v_n}\}$. Then any vector $\vec{w}$ can be written as:

$$\vec{w} = c_1\vec{v_1} + c_2\vec{v_2} + \cdots + c_n\vec{v_n}$$

The action of M on $\vec{w}$ can be evaluated directly since

$$\begin{aligned} M\vec{w} &= M\left(c_1\vec{v_1} + c_2\vec{v_2} + \cdots + c_n\vec{v_n}\right) \\ &= c_1 M\vec{v_1} + c_2 M\vec{v_2} + \cdots + c_n M\vec{v_n} \end{aligned}$$

Since we know the results for all $M\vec{v_i}$, we therefore know the output $M\vec{w}$.

Example 3.68 If we are given:

$$M\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} i \\ 3 \end{bmatrix}, \quad \textit{and} \quad M\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

then we can calculate that:

$$\begin{aligned} M\begin{bmatrix} 5 \\ 1+i \end{bmatrix} &= M\left(5\begin{bmatrix} 1 \\ 0 \end{bmatrix} + (1+i)\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \\ &= 5M\begin{bmatrix} 1 \\ 0 \end{bmatrix} + (1+i)M\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= 5\begin{bmatrix} i \\ 3 \end{bmatrix} + (1+i)\begin{bmatrix} 2 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 2+7i \\ 15 \end{bmatrix} \end{aligned}$$

Definition 3.69 (Orthonormal Basis) *We say a basis is orthonormal if each vector has norm 1 and each pair of vectors are orthogonal.*

Example 3.70 The three sets below are three different orthonormal bases for $\mathbb{C}^2$:

$$\left\{\vec{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \vec{v_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\},$$

$$\left\{\vec{w_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{w_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\},$$

$$\left\{\vec{u_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix}, \quad \vec{u_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}\right\}.$$

You better start liking these bases, because they will show up everywhere in quantum mechanics!

Definition 3.71 (Standard (canonical) Basis) When we explicitly write a vector, we are implicitly using the standard basis (also known as the canonical basis), i.e.,

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + v_2\begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \ldots + v_n\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

In quantum information, for reasons we will understand shortly, we often refer to the standard basis as the computational basis.

Observation 3.72 (Change of basis) Given an arbitrary vector:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

How do we write this vector as a linear combination of different basis vectors, say $\{\vec{w_1}, \vec{w_2}\}$ as defined above? There is actually a systematic way of doing this using "change of basis matrices", but turns out for small dimension, it is easier to attack the problem head-on.

Notice that if we can find how to write each of the standard basis vectors using the new basis, then we will just need to use substitution to get the job done. In our case, it is easy to find that:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}}(\vec{w_1} + \vec{w_2}) \text{ and } \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}}(\vec{w_1} - \vec{w_2})$$

By basic substitution, we get:

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \frac{x_1}{\sqrt{2}}(\vec{w_1} + \vec{w_2}) + \frac{x_2}{\sqrt{2}}(\vec{w_1} - \vec{w_2}) \\ &= \left(\frac{x_1 + x_2}{\sqrt{2}}\right)\vec{w_1} + \left(\frac{x_1 - x_2}{\sqrt{2}}\right)\vec{w_2} \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \end{bmatrix}_{\{w_1, w_2\}} \end{aligned}$$

On the last line, that notation was used to stress the fact that the vector is explicitly written in the $\{w_1, w_2\}$ basis.

Observation 3.73 (Reconstruction of a matrix) Imagine your friend just found the most amazing, perfect $n \times n$ matrix that unlocks the secrets of the universe, but they refuse to show it to you (what a friend!). On the other hand, they are willing to give you a "little taste": they agree to give you the output for n, and only n, input of your choice. Turns out, if you choose the input wisely, you can reconstruct this amazing matrix.

Given an arbitrary basis for n-dimensional vectors, and if you know the "action" of a matrix M on **every** basis vector, it is possible to explicitly reconstruct M. Without loss of generality, we only need to consider the standard basis (if you know the action of the matrix on a different basis, then we just need to perform a change of basis as described above, or use the linearity of matrix multiplication to find the action on the standard

basis). Also, for simplicity, we'll consider only the 2×2 matrix case, but the generalization is straightforward. Starting from the most general matrix:

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

we can observe that:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$$

But since we know what the output for each standard basis taken as input, we can easily deduct all the elements of M.

Example 3.74 Suppose we know that:

$$M \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ i \end{bmatrix} \quad \text{and} \quad M \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1-3i \\ 2+i \end{bmatrix}$$

then we can deduce that:

$$\begin{aligned} M \begin{bmatrix} 1 \\ 0 \end{bmatrix} &= M \left(\frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right) \\ &= \frac{1}{2} \left(M \begin{bmatrix} 1 \\ 1 \end{bmatrix} + M \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right) \\ &= \frac{1}{2} \begin{bmatrix} 3-3i \\ 2+2i \end{bmatrix} \end{aligned}$$

Similarly:

$$\begin{aligned} M \begin{bmatrix} 0 \\ 1 \end{bmatrix} &= M \left(\frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right) \\ &= \frac{1}{2} \left(M \begin{bmatrix} 1 \\ 1 \end{bmatrix} - M \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right) \\ &= \frac{1}{2} \begin{bmatrix} 1+3i \\ -2 \end{bmatrix} \end{aligned}$$

We conclude that:

$$M = \frac{1}{2} \begin{bmatrix} 3-3i & 1+3i \\ 2+2i & -2 \end{bmatrix}$$

3.6 Inner product as projection

When working with the complex plane, or with the cartesian plane for that matter, we have used the term "projection" quite often. Implicitly, "projecting" a vector $\vec{v}$ along the x-axis just means that we are interested in how much of the vector is "along" x. Similarly with projection along y-axis.

Example 3.75 Referring to Figure 3, the projection of $\vec{v}$ along x is 1 and along y, 2.

Observation 3.76 Referring to Figure 3 again, notice that:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \bullet \vec{v} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
$$= 1$$
$$= \text{projection of } \vec{v} \text{ along the } x\text{-axis,}$$

and:

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \bullet \vec{v} = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
$$= 2$$
$$= \text{projection of } \vec{v} \text{ along the } y\text{-axis.}$$

The idea of projection can be generalized for any vectors, in any dimensions. We will work our way to the formal definition step by step. For now, let's stick to the cartesian planes as they offer nice visuals. Let's use vectors:

$$\vec{v} = \begin{bmatrix} 2 \\ 2\sqrt{3} \end{bmatrix} \quad \text{and} \quad \vec{w} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

Refer to Figure 8 for a visual. You can convince yourself that $\vec{v}$ makes and angle of $\frac{\pi}{3}$ (or $60°$) with the x-axis and $\vec{w}$ makes an angle of $-\frac{\pi}{4}$ ($-45°$).

Definition 3.77 (Orthogonal Projection) The projection $P_{\vec{v},\vec{w}}$ of $\vec{v}$ onto $\vec{w}$ is given by the component of $\vec{v}$ along the direction of $\vec{w}$.

An easy trick to find that component: put a ruler perpendicular to the direction of $\vec{w}$ (that direction is given by the large dashed line in Figure 8) and move the ruler along that line until you find the point representing $\vec{v}$. Draw a line with your ruler (short dashed orange line) and the intersection of the two lines is the desired component $P_{\vec{v},\vec{w}}$. Notice that in our current example, the component we are interested in is actually along the opposite direction of $\vec{w}$, that is the direction of $-\vec{w}$. *$P_{\vec{v},\vec{w}}$ will therefore be a negative number.*

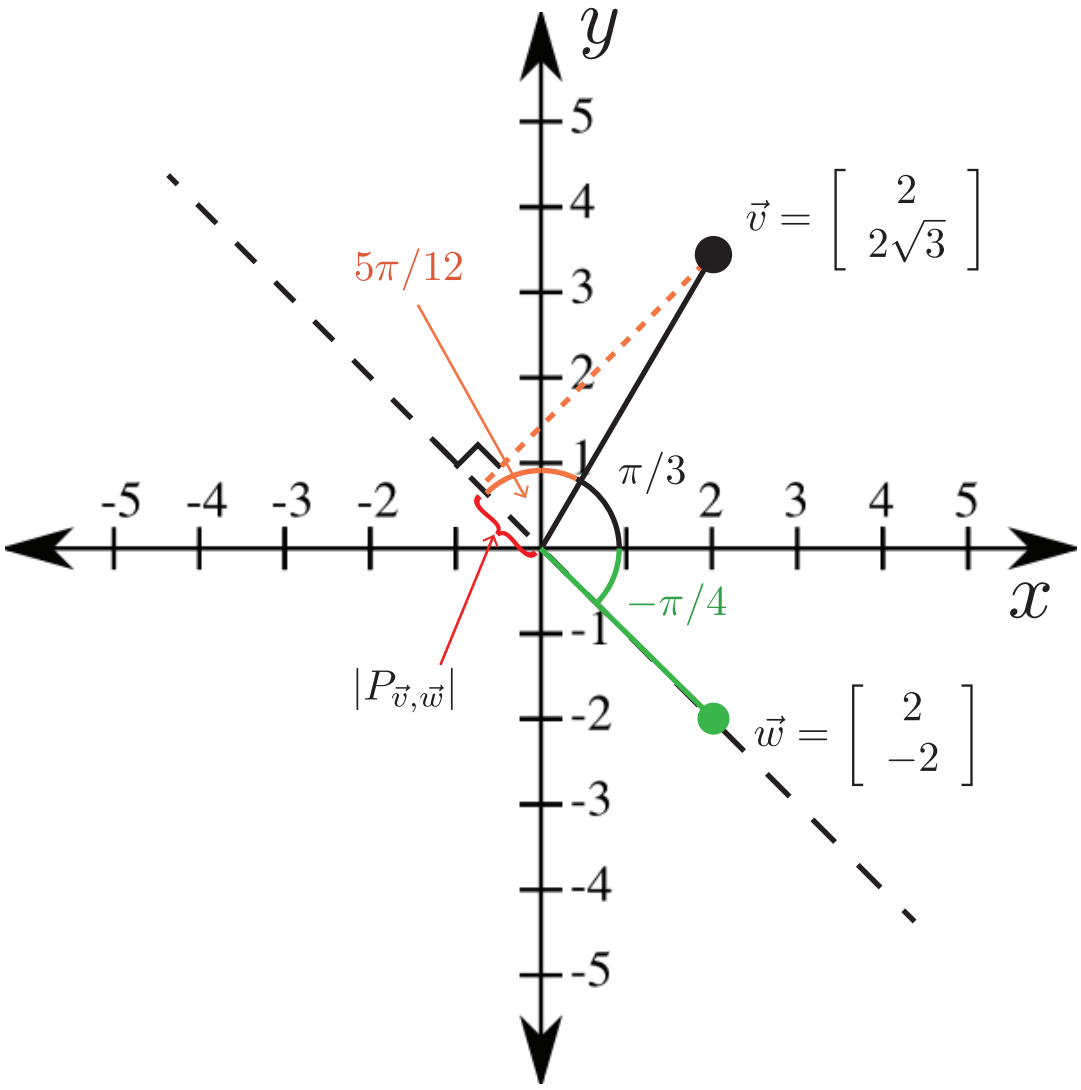


Figure 8: The projection of $\vec{v}$ onto $\vec{w}$ is given by the component of $\vec{v}$ along $\vec{w}$.

Again referring to Figure 8, basic trigonometry tells us that:

$$\begin{aligned}
\left|P_{\vec{v},\vec{w}}\right| &= ||\vec{v}|| \cos\left(\frac{5\pi}{12}\right) \\
&= ||\vec{v}|| \cos 75^\circ \quad \text{(if you prefer to work in degrees)} \\
&\approx \sqrt{16} \cdot 0.259 \\
&\approx 1.035 \\
\implies & P_{\vec{v},\vec{w}} \approx -1.035
\end{aligned}$$

Observation 3.78 The vector $\vec{w}$ above is parallel to the unit vector:

$$\vec{u} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$\vec{u}$ can be obtained simply by normalizing $\vec{w}$ (see Definition 3.55). Do you think it is a coincidence that:

$$\begin{aligned}
\vec{u} \bullet \vec{v} &= \frac{1}{||\vec{w}||} \vec{w} \bullet \vec{v} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 2\sqrt{3} \end{bmatrix} \\
&= \frac{1}{\sqrt{2}} (2 - 2\sqrt{3}) \\
&\approx -1.035?
\end{aligned}$$

The observation made above is not a coincidence at all and can readily be seen as a consequence of the properties of the inner product and the definition of basis. To see this, let's suppose that $\vec{u}_\perp$ is any unit vector that is perpendicular to $\vec{u}$, and hence $\vec{w}$). We have seen in the previous section that the set $\{\vec{u}, \vec{u}_\perp\}$ form a orthonormal basis for 2-dimensional vectors. Therefore we can write:

$$\vec{v} = c_1 \vec{u} + c_2 \vec{u}_\perp$$

where c_1 is the component of $\vec{v}$ along $\vec{u}$ (and $\vec{w}$), i.e., the projection $P_{\vec{v},\vec{w}}$ of $\vec{v}$ along $\vec{w}$, and c_2 is the component of $\vec{v}$ along $\vec{u}_\perp$, i.e., the projection of $\vec{v}$ along any vectors perpendicular to $\vec{v}$. You can therefore have:

$$\begin{aligned}
\vec{u} \bullet \vec{v} &= c_1 \vec{u} \bullet \vec{u} + c_2 \vec{u} \bullet \vec{u}_\perp \\
&= c_1 \quad \text{since } \vec{u} \bullet \vec{u} = 1 \text{ and } \vec{u} \bullet \vec{u}_\perp = 0 \\
&= P_{\vec{v},\vec{w}}
\end{aligned}$$

Using the cartesian plane, we essentially worked our way toward the formal definition of an orthogonal projection is any dimension.

Definition 3.79 Given n-dimensional vectors $\vec{v}, \vec{w}$ in $\mathbb{C}^n$, the projection of $\vec{v}$ onto $\vec{w}$, $P_{\vec{v},\vec{w}}$ is given by:

$$P_{\vec{v},\vec{w}} = \frac{1}{||\vec{w}||}\vec{w} \bullet \vec{v}$$

In other words, the projection is given by the inner product between the *unit* vector along $\vec{w}$ and $\vec{v}$.

Example 3.80 In the 3-dimension complex vector space, the projection of:

$$\vec{v} = \begin{bmatrix} 1 \\ e^{-i\frac{\pi}{4}} \\ 2 \end{bmatrix}$$

along:

$$\vec{w} = \begin{bmatrix} i \\ 2 \\ -i \end{bmatrix}$$

is given by:

$$\begin{aligned} P_{\vec{v},\vec{w}} &= \frac{1}{\sqrt{6}} \begin{bmatrix} i \\ 2 \\ -i \end{bmatrix} \bullet \begin{bmatrix} 1 \\ e^{-i\frac{\pi}{4}} \\ 2 \end{bmatrix}, \text{ since } ||\vec{w}|| = 1/\sqrt{6} \\ &= \frac{1}{\sqrt{6}} \begin{bmatrix} -i & 2 & i \end{bmatrix} \begin{bmatrix} 1 \\ e^{-i\frac{\pi}{4}} \\ 2 \end{bmatrix} \\ &\quad \text{(do not forget the complex conjugate of the first vector)} \\ &= \frac{1}{\sqrt{6}}(-i + 2e^{-i\frac{\pi}{4}} + 2i) \\ &= \frac{1}{\sqrt{6}}(i + 2\cos\frac{\pi}{4} + 2i\sin\frac{\pi}{4}), \text{ remembering Euler's formula} \\ &= \frac{1}{\sqrt{6}}\left(\sqrt{2} + i(1+\sqrt{2})\right) \end{aligned}$$

This definition of projection is actually quite intuitive. If we pick any orthonormal basis $\{u_1, \ldots, u_n\}$ such that one of the basis vectors, say, without loss of generality, u_1 is the unit vector pointing in the direction of $\vec{w}$, then just like in the 2-dimensional, real, case we can write:

$$\vec{v} = c_1\vec{u_1} + \ldots + c_n\vec{u_n}$$

such that:

$$\begin{aligned}\frac{1}{||\vec{w}||}\vec{w} \bullet \vec{v} &= \vec{u_1} \bullet \vec{v} \\ &= c_1\vec{u_1} \bullet \vec{u_1} + c_2\vec{u_1} \bullet \vec{u_2} + \ldots + \vec{u_1} \bullet \vec{u_n} \\ &= c_1, \text{ since } \vec{u_1} \bullet \vec{u_j} = 0 \text{ except if } j = 1 \\ &= P_{\vec{v},\vec{w}}\end{aligned}$$

Observation 3.81 (Linear combination revisited) In previous sections, we rewrote a given vector as a linear combination of other vectors a few times. In Example 3.60 and Observation 3.72, we found the coefficients of the combination by inspection, while in Example 3.66, we used a more systematic, yet cumbersome method.

Using the inner product now gives us a systematic and simple method to write any vector $\vec{v}$ as a linear combination of vectors belonging to an orthonormal basis $\{\vec{u_1}, \ldots, \vec{u_n}\}$. As we have seen many times already, we can always find coefficient $c_1, \ldots, c_n$ such that:

$$\begin{aligned}\vec{v} &= c_1\vec{u_1} + c_2\vec{u_2} + \ldots + c_n\vec{u_n} \\ &= (\vec{u_1} \bullet \vec{v})\vec{u_1} + (\vec{u_2} \bullet \vec{v})\vec{u_2} + \ldots + (\vec{u_n} \bullet \vec{v})\vec{u_n}\end{aligned}$$

since $c_j = \vec{u_j} \bullet \vec{v}$

Example 3.82 We will rewrite the vector:

$$\vec{v} = \begin{bmatrix} 3+i \\ 1-i \end{bmatrix}$$

as a linear combination of:

$$\left\{\vec{u_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix}, \vec{u_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}\right\}$$

First, convince yourself that $\{\vec{u_1}, \vec{u_2}\}$ is an orthonormal basis for C^2. From there, we just need to crunch the numbers:

$$\begin{aligned}\vec{u_1} \bullet \vec{v} &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -i \end{bmatrix}\begin{bmatrix} 3+i \\ 1-i \end{bmatrix} \\ &\quad \text{(again, don't forget the complex conjugate of the first vector)} \\ &= \sqrt{2} \\ \vec{u_2} \bullet \vec{v} &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \end{bmatrix}\begin{bmatrix} 3+i \\ 1-i \end{bmatrix} \\ &= 2\sqrt{2} + \sqrt{2}i \\ \implies \vec{v} &= \sqrt{2}\vec{u_1} + \left(2\sqrt{2} + \sqrt{2}i\right)\vec{u_2}\end{aligned}$$

3.7 Special Matrices

Now back to matrices. A lot "families" of matrices exist, i.e., matrices that share special properties. There are a few special types of matrices that play an important role in quantum information. Below are a few examples.

Definition 3.83 (Identity matrix) *The $n \times n$ identity matrix, denoted $\mathbb{1}$ or sometimes I, is defined such that for every $n \times n$ matrix M, and any vector $\vec{v}$ in $\mathbb{C}^n$, we have:*

$$\mathbb{1}M = M\mathbb{1} = M \quad \textbf{\textit{and}} \quad \mathbb{1}\vec{v} = \vec{v}.$$

In other words, this matrix performs no actions when operating on any vectors or matrices; the output is always the same as the input. It is similar to the number 1 in scalar multiplication. (Remember that scalar numbers are 1-dimensional vectors/matrices.)

It is easy to observe that $\mathbb{1}$ is the matrix with only 1's along the diagonal, i.e.,

$$\mathbb{1} = \begin{bmatrix} 1 & 0 & \dots \\ 0 & 1 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Example 3.84 In 3 dimensions, we see that the 3×3 matrix with only one's on the diagonal is the identity, since:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Note 3.85 The identity is only defined for square matrices. (Can you explain why, given its definition?)

Definition 3.86 (Unitary Matrices) *A unitary matrix U is a matrix that satisfies:*

$$UU^\dagger = U^\dagger U = \mathbb{1}$$

Example 3.87 The following matrices, which will become your best friends soon, are unitary:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Observation 3.88 If the same unitary matrix U is applied to any two vectors $\vec{v_1}$ and $\vec{v_2}$, the inner product is preserved. That is if:

$$\vec{w_1} = U\vec{v_1} \quad \textit{and} \quad \vec{w_2} = U\vec{v_2}$$

then:

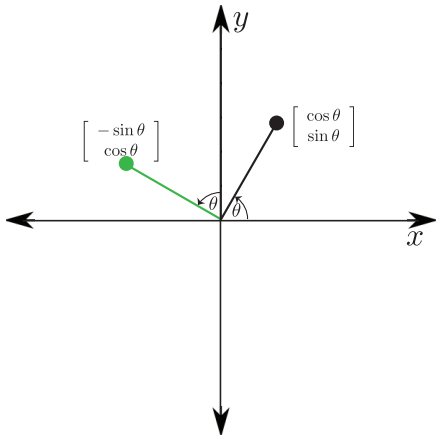
$$\begin{aligned}
\vec{w_1} \bullet \vec{w_2} &= (U\vec{v_1}) \bullet (U\vec{v_2}) \\
&= (U\vec{v_1})^\dagger U\vec{v_2} \\
&= (\vec{v_1}^\dagger U^\dagger)(U\vec{v_2}), \text{ recalling that } (MN)^\dagger = N^\dagger M^\dagger \\
&= \vec{v_1}^\dagger (U^\dagger U)\vec{v_2} \\
&= \vec{v_1}^\dagger \mathbb{1} \vec{v_2} \\
&= \vec{v_1}^\dagger \vec{v_2} \\
&= \vec{v_1} \bullet \vec{v_2}
\end{aligned}$$

Note 3.89 From the above observation, unitary matrices preserve the length (or norm) of vectors.

$$\|U\vec{v}\| = \sqrt{(U\vec{v}) \bullet (U\vec{v})} = \sqrt{\vec{v} \bullet \vec{v}} = \|\vec{v}\|$$

Observation 3.90 **(Rotation matrix)** The matrix R defined above represents a rotation in $\mathbb{R}^2$, since:

$$R \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \quad R \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$



3.8 Example of applied linear algebra – The cooking matrix

> "At the beginning of the chapter, you said that you could write a cookbook using matrices, what's up with that?"

So far we have mostly used concrete examples that we were already familiar with (cartesian and spacial coordinates). Just to show you how versatile matrices and vectors can be, we'll write a small cookbook using only numbers in a matrix.

To make things simple, we'll use a finite number of ingredients, say sugar, flour, salt, milk and water, and we will use three different dough recipes:

1. Pizza dough: 0 cup of sugar, 2 cups of flour, 1 tsp of salt, no milk, 1 cup of water
2. Cake dough: 1 cup of sugar, 1 cup of flour, no salt, 1/2 cup of milk, no water
3. Bagel dough: 1 cup of sugar, 1 cup of flour, 2 tbs salt, 1/4 cup of milk, 1/4 cup of water

Obviously, we must stress the fact that that these are not real recipes and **you should not try this at home**: it would taste terrible. But for the purpose of this example, these are our recipes.

First, we need to determine the input and output vectors and which operation the "cookbook" matrix M will represent. Our input shall be the answer to "what recipe do you want to make". The output, a list of ingredients.

Since each dough is independent of each other, we can assign each dough to a basis vector. Since there are three of them, we will need at least, and no more, than three basis vectors.

$$\overrightarrow{pizza} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \overrightarrow{cake} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \overrightarrow{bagel} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Since each ingredient is independent of each other (you would not use salt to replace water, would you?), we will also assign each ingredient a basis vector. Since there are five ingredients, we will need five basis vectors. Therefore, we can assign:

$$\overrightarrow{sugar} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \overrightarrow{flour} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \overrightarrow{salt} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \overrightarrow{milk} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \overrightarrow{water} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

From the recipe, we know the actions of the matrix M, e.g.,

$$M\overrightarrow{pizza} = \begin{bmatrix} 0 \\ 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad M\overrightarrow{cake} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0.5 \\ 0 \end{bmatrix}, \quad M\overrightarrow{bagel} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0.25 \\ 0.25 \end{bmatrix}$$

From this, we can give the explicit form of M:

$$M = \begin{bmatrix} 0 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 0 & 2 \\ 0 & 0.5 & 0.25 \\ 1 & 0 & 0.25 \end{bmatrix}$$

There is our cookbook!

Suppose, you want to do two pizzas, half a cake and two batches of bagels, here is what you would need to buy:

$$\begin{bmatrix} 0 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 0 & 2 \\ 0 & 0.5 & 0.25 \\ 1 & 0 & 0.25 \end{bmatrix} \begin{bmatrix} 2 \\ 0.5 \\ 3 \end{bmatrix} = \begin{bmatrix} 3.5 \\ 7.5 \\ 7 \\ 1 \\ 2.75 \end{bmatrix}$$

You'll need to get 3.5 cups of sugar, 7.5 cups of flour, 7 tbs of salt, 1 cup of milk and 2.75 cups of water. Isn't that the most nerd-tastic way of cooking and getting your grocery list ready?!

4 Mathematics of Quantum Mechanics

In this last section of this brief booklet (it's all relative!), we will explore how linear algebra applies to quantum mechanics. For now, we will take a purely mathematical approach, but never forget that:

> Quantum physics is not about mathematics. It describes the behaviour of atoms, molecules, photons, even nano-scale electrical circuits. But we use mathematics to quantify and model the physical phenomena of quantum mechanics.

During the week of QCSYS, we will also teach you quantum mechanics from a qualitative, phenomenological approach. By the end of the second, we will reconcile both the mathematical and the phenomenological approach. But for now, we will introduce how the different concepts of linear algebra we just learned will be used in quantum mechanics.

4.1 Mathematical postulates of quantum mechanics

As we mentioned in the Preface of this document, quantum mechanics refers to the description of the behaviour of the building blocks of nature: atoms, molecules, photons, etc. Turns out that there are five "simple" postulates that fully encapsulate the mathematical modelling of quantum mechanics. These postulates can take different equivalent forms depending to who you talk to, but for the sake of QCSYS, we will work with the postulates given by:

1. State (or wavefunction) of individual quantum systems are unit vectors living in separate complex Hilbert spaces.[2]

2. The probability of measuring a system in a given state is given by the modulus squared of the inner product of the output state and the current state of the system. Immediately after the measurement, the wavefunction collapse into that state (aka Born's rule).

3. Quantum operations are represented by unitary operators on the Hilbert space (consequence of the Schrödinger equation).

4. The Hilbert space of a composite system is given by the tensor product (aka Kronecker product) of the separate, individual Hilbert spaces.

5. Physical observable are represented by the eigenvalues of a Hermitian operators on the Hilbert Space.

[2] A loose definition of a complex Hilbert space is complex vector space with a well-defined inner product

We are just stating them here and will go into detail about each of them and explain the new terminology used. For the sake of simplicity, we will not be covering the last postulate as it is not really needed during QCSYS.

Also, note that in the first part of this booklet, we have introduced the notion of vector space using finite dimension vectors and matrices. This is a natural way to proceed. Referring to postulate 1, finite dimension vectors are a great way to represent the state of quantum systems when the physical variables only have discrete possibilities, i.e., being here or there, up or down, left or right, etc. This treatment of quantum mechanics is often referred to as "Matrix" mechanics and is really well suited for quantum information and quantum cryptography.

But what about when we want to describe physical quantities that have continuous values, such as the position of an particle along a line? In this case, we need a vector space of infinite and continuous dimension. Turns out that is it possible to define a vector space, as well as an inner product, on the set of continuous functions, e.g., $f(x) = x^3 + 2x + 1$. This is referred to as "Wave" mechanics and we will not be covering it in this booklet, although we will show you the example of a particle in a box during the QCSYS lectures. This example will require the wave mechanics treatment of quantum mechanics.

4.2 New notation: the BraKet notation

Before going into the the heart of the mathematics of quantum mechanics, we will introduce a new notation that is widely used in physics. In the Linear algebra section, we used the notation for vector and matrix notations that are widely used, but when physicists play with vectors to describe a quantum system, they like to change this around a little and use what we call the "BraKet" notation. Note that everything we have defined previously, i.e., vector addition, matrix multiplication, inner product, projections, etc, is exactly the same. The only thing that is changing is how we write down the variable.

Notation 4.1 (The "ket") *When using a vector $\vec{v}$ to represent a quantum state, we will use a different notation known as "ket", written $|v\rangle$ ("ket v"). This is a notation commonly used in quantum mechanics and does not change the nature of the vectors at all. i.e., both notations below are equivalent:*

$$|v\rangle = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \longleftrightarrow \vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

Notation 4.2 (The "bra") *The conjugate transpose of a "ket" $|v\rangle$ is denoted by:*

$$\langle v| = (|v\rangle)^\dagger$$

$\langle v|$ is called "bra v". Again, we stress the fact that this is just a notation and doesn't change the meaning of the conjugate transpose.

Notation 4.3 (The "braket") *Given two vectors $|v\rangle$ and $|w\rangle$, we use the following notation for the inner product:*

$$\langle v|w\rangle = |v\rangle \bullet |w\rangle$$

$\langle v|w\rangle$ is known as the "bracket of $|v\rangle$ and $|w\rangle$".

Note 4.4 The bracket notation follows naturally from the geometry of our notation, i.e.,

$$\begin{aligned} |v\rangle \bullet |w\rangle &= (|v\rangle)^\dagger |w\rangle \\ &= \langle v||w\rangle \\ &= \langle v|w\rangle \end{aligned}$$

Observation 4.5 Note that:

$$\begin{aligned} \overline{\langle v|w\rangle} &= \overline{|v\rangle^\dagger |w\rangle} \\ &= \overline{|v\rangle}^\dagger \overline{|w\rangle}, && \text{recalling the Property 2.2-4} \\ &= |v\rangle^t \overline{|w\rangle}, && \text{recalling the definition of the "dagger"} \\ &= \left(\overline{|w\rangle}^t |v\rangle\right)^t, && \text{recalling the Property 3.40-2} \\ &= \left(|w\rangle^\dagger |v\rangle\right)^t, \\ &= \langle w|v\rangle^t \\ &= \langle w|v\rangle, && \text{since the transpose of a scalar is a scalar.} \end{aligned}$$

Actually, this observation is the same as Property 3.45-7, but using the bracket notation.

Example 4.6 Let:

$$|v\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}, \quad \textit{and} \quad |w\rangle = \frac{i}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

We can calculate the inner product:

$$\begin{aligned} \langle w|v\rangle &= \frac{-i}{2} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} \\ &= \frac{-i}{2}(1+i) \\ &= \frac{1-i}{2} \end{aligned}$$

Similarly, we can also calculate:

$$\begin{aligned} \langle v|w\rangle &= \frac{i}{2} \begin{bmatrix} 1 & -i \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \frac{i}{2}(1-i) \\ &= \frac{1+i}{2} \end{aligned}$$

We can observe that, as proved above $\langle w|v\rangle = \overline{\langle v|w\rangle}$

4.3 Single quantum state and the qubit

Now that the new notation is out of the way, we will now look in detail Postulate 1 of quantum mechanics as defined in the previous section:

> **Postulate 1**: The state of individual quantum systems are unit vectors living in separate complex Hilbert spaces.

Definition 4.7 (Quantum States) *The collection of all relevant physical properties of a quantum system (e.g., position, momentum, spin, polarization, etc.) is known as the state of the system. The state of a quantum system can be mathematically*[3] *represented by a* ***complex vector with norm equal to*** 1 *(you'll understand why in a couple of pages!)*

Now, you may wonder how do we represent a physical state using vectors? The key point here is two understand the concept of exclusive states:

Definition 4.8 (Exclusive states) *When modelling the state of a given physical quantity, e.g., position, spin, polarization, etc., two states are said to be exclusive if the fact of being in one of the state implies that there are no chance whatsoever to be in any of the other states.*

Example 4.9 (1 particle, 3 boxes) Imagine this hypothetical situation: we have a single quantum particle, and 3 quantum boxes. The whole system behaves quantum mechanically. If I know for certain that the particle is in box 1, then it is certainly not in box 2 or 3.

Example 4.10 (Moving Particle) Imagine a particle that can move horizontally, vertically and diagonally. These three states are not exclusive since moving diagonally can be thought of a moving horizontally and vertically at the same time.

If you recall the concept of orthogonal vectors and projection, it was not unlike the concept of exclusivity (i.e., a given vector has no component along any of its orthogonal vectors), therefore it make sense to establish a connection between exclusive states and orthogonal vectors:

> **Given a quantum system with n exclusive states, each state will be represented by a vector from an orthonormal basis of a n-dimensional Hilbert space.**

[3] Actually, there is a huge debate in the quantum mechanics community as to whether the state of a system is just a mathematical description, or if it is "real". There are many arguments with no consensus, but fortunately, this is more of a philosophical question then a technical one!

Example 4.11 From the 1 particle, 3 boxes example above, the state of the electron can be in either box 1, 2, and 3 represented by the quantum state $|1\rangle, |2\rangle$ and $|3\rangle$ respectively. Explicitly, we can write:

$$|1\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad |2\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad |3\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

From now on, when speaking of a quantum system, we will use the word "vector" and "state" interchangeably. Quantum states also use the "bracket" notation to emphasize they are quantum states. See the following example.

We will introduce the abstract concept of quantum bit, or **qubits**. Qubits only have two distinct (exclusive) states, which is the starting point of everything in quantum information. But note that all the mathematical treatment we will discuss can be applied to any quantum system with any number of discrete physical states.

But before introducing the formal definition of a qubit, a quick note about the **bit**. In your cell phone, your computer, or pretty much any digital electronic devices you have, the information is treated using the simplest alphabet of all: the binary system. In binary, we only have two "letters" – 0 and 1. This basic unit of classical information is know as a bit. In your computer, each bit[4] can either be in the state 0, or in the state 1.

Definition 4.12 (The qubit) *In quantum information and quantum computing, we are using* ***quantum bits****, or* ***qubits****. Like the classical bit, a qubit only has two exclusive states, "quantum-0" and "quantum-1". But unlike the classical bit, the qubit behaves according to the laws of quantum mechanics. As we will see during the week, this new behaviour will allow us to do very amazing things.*

Since a qubit only has two different exclusive states, the state/vector representing the quantum-0 and the quantum-1 should be 2-dimensional. The vectors $|0\rangle$ and $|1\rangle$ represented by the vectors:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

This basis is known as the *computational basis*.

We typically use the computational basis $\{|0\rangle, |1\rangle\}$ to represent the two different states of a quantum system that we use as our quantum-0 and quantum-1. For example, if we use the energy of an electron in a molecule as our quantum bit, we could say that the ground state (lowest energy) is our quantum-0, and an excited state (higher energy) is our quantum-1. Since the ground and excited state are mutually exclusive, we could represent

$$\text{ground state} \leftrightarrow |0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{exited state} \leftrightarrow |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

[4]Physically implemented by a transistor in the processor, or a tiny magnet in your hard drive.

In quantum cryptography, we like to use the polarization of a photon as our qubit. By the end of the week, you will be experts on this!

Observation 4.13 Here's where the fun starts. What does the state

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

represent? If we look a little closer, we see that:

$$|\pm\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle \pm |1\rangle \right).$$

This tells us that the states $|\pm\rangle$ are linear combinations of the computational basis (i.e., the quantum-0 and quantum-1). In quantum mechanics, we call that a quantum superposition (rigorously defined below). So would that mean that my system is in both 0 and 1? Actually, yes! Is that mathematical trickery, or can you really do that in a laboratory? We can certainly do that in a lab, and you will experience it firsthand! Welcome to the world of quantum mechanics!

Definition 4.14 (Quantum superposition) *Suppose we are given the following qubit in the following state:*

$$|\psi\rangle = \begin{bmatrix} a \\ b \end{bmatrix}.$$

Notice we can write:

$$|\psi\rangle = a|0\rangle + b|1\rangle.$$

We say that $|\psi\rangle$ is in a superposition of $|0\rangle$ and $|1\rangle$ with probability amplitudes a and b. (We will see in the next section why it is called a that way.)

Note 4.15 ψ is the greek letter "psi" (silent "p") and is commonly used to denote a quantum state. We love to use Greek letters to denote quantum state. Refer to Appendix A for the complete list of Greek letters.

Example 4.16 The state $|0\rangle$ is a superposition of $|+\rangle$ and $|-\rangle$ since:

$$|0\rangle = \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle$$

So, what does it mean to be in a superposition? Well, there is still some philosophical debate about the actual meaning, but as far as we are concerned, being in a superposition means that the state of our system is not well defined. It is *as if* it was in both states at once.

4.4 Quantum Measurement

We will investigate Postulate 2 of quantum mechanics:

> **Postulate 2 (Born's rule)**: The probability of measuring a system in a given state is given by the modulus squared of the inner product of the output state and the current state of the system. Immediately after the measurement, the wavefunction collapses into that state.

As we just saw, it is possible for a quantum mechanical system to be in a superposition of exclusive states, say $|\psi\rangle = a|0\rangle + b|1\rangle$. You might might wonder: if I actually perform a measurement to see if the system is in $|0\rangle$ or in $|1\rangle$, what will I measure? Obviously, you will measure either $|0\rangle$ or $|1\rangle$, but which one? With which probability?

Definition 4.17 (Quantum Measurement) *Suppose we have a quantum state $|\psi\rangle$ and an orthonormal basis $\{|\phi_1\rangle, \ldots, |\phi_n\rangle\}$. Then we can measure $|\psi\rangle$ with respect to this orthonormal basis, i.e., we "ask" the quantum system which one of these states it is in. The probability of measuring the state $|\phi_i\rangle$, $P(\phi_i)$, is given by*

$$P(\phi_i) = |\langle\phi_i|\psi\rangle|^2$$

After the measurement is performed the original state "collapses" in the measured state, that is, we are left with one of the states $|\phi_1\rangle, \ldots, |\phi_n\rangle$. This is known as the ***Born's rule.***

Example 4.18 Suppose $|\psi\rangle = |+\rangle$ and we measure it in the orthonormal basis $\{|\phi_0\rangle = |0\rangle, |\phi_1\rangle = |1\rangle\}$. Then the state collapses to:

$$\begin{cases} |0\rangle & \text{with probability } |\langle 0|+\rangle|^2 = 1/2, \\ |1\rangle & \text{with probability } |\langle 1|+\rangle|^2 = 1/2. \end{cases}$$

This is like flipping a coin. If we decide to measure in the basis $\{|+\rangle, |-\rangle\}$, then the outcome state will be:

$$\begin{cases} |+\rangle & \text{with probability } |\langle +|+\rangle|^2 = 1, \\ |-\rangle & \text{with probability } |\langle -|+\rangle|^2 = 0. \end{cases}$$

Example 4.19 If we treat the general case of a qubit in an unknown quantum state $|\psi\rangle = a|0\rangle + b|1\rangle$ and we measure in the computational basis, the outcome will be:

$$\begin{cases} |0\rangle & \text{with probability } |\langle 0|\psi\rangle|^2 = |a|^2, \\ |1\rangle & \text{with probability } |\langle 1|\psi\rangle|^2 = |b|^2. \end{cases}$$

Example 4.20 What if we measure the system on the basis $\{|+\rangle, |-\rangle\}$? Note that:

$$\begin{aligned}\langle \pm|\psi\rangle &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & \pm 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} \\ &= \frac{a \pm b}{\sqrt{2}}\end{aligned}$$

which implies that the outcome will be the state:

$$|\pm\rangle \quad \text{with probability } \frac{|a \pm b|^2}{2}$$

Note 4.21 In the example above, we used the the explicit representation of the vector to perform the inner product, but we could have done everything as symbolic, i.e.,

$$\begin{aligned}\langle \pm|\psi\rangle &= \frac{1}{\sqrt{2}}\left(\langle 0| \pm \langle 1|\right)\left(a|0\rangle + b|1\rangle\right) \\ &= \frac{1}{\sqrt{2}}\left(a\langle 0|0\rangle + b\langle 0|1\rangle \pm a\langle 1|0\rangle \pm b\langle 1|1\rangle\right) \\ &= \frac{a \pm b}{\sqrt{2}}\end{aligned}$$

Observation 4.22 In the last two examples, we saw that measuring in two different basis yield probabilistic measure. You can actually show that the only time we will get a deterministic result is when we measure on a basis that includes the state $|\psi\rangle$. But since $|\psi\rangle$ was an unknown state to begin with, we conclude measuring a unknown quantum state will always yield a random result.

What Born's rule is actually telling us is that *quantum mechanics is inherently random.* When you have an unknown quantum superposition, it is impossible to predict precisely which outcome you will measure. You can only predict the probability of the outcome.

Example 4.23 To relate the examples above to physics, one could think of the following scenario: say we prepare an atom in a superposition of the ground and excited states (recall our assignment of the computational basis for the ground and excited state of an atom on page 57). Say the way we carried out the experiment leads to a superposition which mathematical representation is given by $|+\rangle$, i.e., the atom is in a superposition of ground and excited states. If we measure the atom to investigate its state, there is a probability of $1/2$ measuring the atom in its ground state, and a probability of $1/2$ measuring it in its excited state. If the measurement gives us the ground state for example, then right after the measurement, the state has "collapsed" into the ground state.

Measuring in the $\{|0\rangle, |1\rangle\}$ orthonormal basis is a way to transform a qubit into a regular bit. In a sense, we are destroying the quantum superposition, so make your measurement count!

Definition 4.24 (Quantum Measurement (alternative)) *Suppose we have a quantum state $|\psi\rangle$ and an orthonormal basis $\{|\phi_1\rangle, \ldots, |\phi_n\rangle\}$. We can explicitly write the state using the orthonormal basis, i.e.,*

$$|\psi\rangle = c_1|\phi_1\rangle + c_2|\phi_2\rangle + \ldots + c_n|\phi_n\rangle,$$

then the probability of measuring each state $|\phi_i\rangle$ is given by:

$$P(\phi_i) = |c_i|^2$$

After the measurement, the state of the system collapses into $|\phi_i\rangle$.

Example 4.25 This is a rather simple exanoke to show that the two definitions of quantum measurement are equivalent: according to the first definition:

$$\begin{aligned} P(\phi_i) &= |\langle\phi_i|\psi\rangle|^2 \\ &= |\langle\phi_i| \left(c_1|\phi_1\rangle + c_2|\phi_2\rangle + \ldots + c_n|\phi_n\rangle\right)|^2 \\ &= |c_1\langle\phi_i|\phi_1\rangle + c_2\langle\phi_i|\phi_2\rangle + \ldots + c_n\langle\phi_i|\phi_n\rangle|^2 \end{aligned}$$

Since $\{|\phi_1\rangle, \ldots, |\phi_n\rangle\}$ is an orthonormal basis, all the inner products are 0, except with $|\phi_i\rangle$, therefore:

$$P(\phi_i) = |c_i|^2$$

Example 4.26 Given the quantum state:

$$|\psi\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix},$$

what is the probability of measuring it in the state $|+\rangle$? We can use either method to find it.

1. First method:

$$\begin{aligned} P(+) &= |\langle+|\psi\rangle|^2 \\ &= \left|\frac{1}{2}\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix}\right|^2 \\ &= \frac{1}{4}|1+i|^2 \\ &= \frac{1}{4}(1+i)(1-i), \text{ recalling that } |z|^2 = z\bar{z} \text{ for complex numbers} \\ &= \frac{1}{2} \end{aligned}$$

2. Second method: we already know that the states $|+\rangle$ and $|-\rangle$ form an orthonormal basis, therefore we need to write $|\psi\rangle$ in this basis. Moreover, observe that:

$$|0\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle + |-\rangle\right), \quad \textit{and} \quad |1\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle - |-\rangle\right)$$

therefore:

$$\begin{aligned} |\psi\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle + i|1\rangle\right) \\ &= \frac{1}{2}\left(|+\rangle + |-\rangle + i|+\rangle - i|-\rangle\right) \\ &= \frac{1+i}{2}|+\rangle + \frac{1-i}{2}|-\rangle \end{aligned}$$

The probability of measuring + is thus given by:

$$\begin{aligned} P(+) &= \frac{1}{4}|1+i|^2 \\ &= \frac{1}{4}(1+i)(1-i) \\ &= \frac{1}{2} \end{aligned}$$

Observation 4.27 **(The need of unit vectors)** At this point, it becomes clear why quantum states must be unit vectors. If we write our quantum state using any orthonormal basis:

$$|\psi\rangle = c_1|\phi_1\rangle + c_2|\phi_2\rangle + \ldots + c_n|\phi_n\rangle$$

then the modulus of $|\psi\rangle$ is given by:

$$|||\psi\rangle|| = \sqrt{|c_1|^2 + |c_2|^2 + \ldots + |c_n|^2}$$

Since $|c_i|^2$ is the probability of measuring the state $|\psi\rangle$ in the state $|\phi_i\rangle$, then the norm of a quantum state is simply the sum of the probability of measure each $|\phi_i\rangle$. Since our measurement must given us something, then the sum of the probabilities must be 1. Hence the need of a unit vector.

4.5 Quantum Operations

So far, we have learned how to calculate the measurement probability given the state of a system. But how do we create that state? In other words, is it possible, given a known initial state, to transform it into any other states? This answer is ''yes" and the mathematical representation is known as a **quantum operation**.

Postulate 3: Quantum operations are represented by unitary operators on the Hilbert space.

Definition 4.28 (Quantum Operations) *A quantum operation transforms a quantum state to another quantum state, therefore, we must have that the norm of the vector is preserved (recall, quantum states must have norm 1). Therefore,* ***the mathematical representation of any quantum operations is represented using a unitary matrix****. Similarly,* ***any unitary matrix represents a possible quantum operation****.*

Refer to Section 3.7 for the properties of unitary matrices.

Example 4.29 The following are some popular quantum operations:

$$\mathbb{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

These are called the **Pauli matrices**. Another important quantum operation is the **Hadamard matrix** defined as:

$$H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Example 4.30 What would be the final state if we perform:

1. $\mathbb{1}$ on the state $|0\rangle$?
2. X on the state $|1\rangle$?
3. Z on the state $|+\rangle$?
4. H on the state $|1\rangle$?

To find the final state, we just need to carry the matrix multiplication:

1. $I|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle$
2. $X|1\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle$
3. $Z|+\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = |-\rangle$
4. $H|1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = |-\rangle$

Observation 4.31 If one implements many quantum operations one after another, say $U_1, U_2, \ldots U_m$ (where the U_i's are labelled in chronological order) then the matrix representation of the combined quantum operation U_{tot}, is given by:

$$U_{tot} \quad = \quad U_m \ldots U_2 U_1$$

Notice that the order of the multiplication goes **right to left** in chronological order. The reason for the reversal is quite simple. After the first operation, the state is now $U_1|\psi\rangle$, so the second operation will be applied on this state to give $U_2 U_1|\psi\rangle$, and so on, and so forth.

In the lab, quantum operations are performed in various ways, depending on the quantum system you are trying to manipulate. If our qubit is represented by the polarization of a photon, we use quarter and half wave plates (essentially a piece of glass), if we use the ground and excited states of a atom, we can use a laser pulse. There are many different qubits and many different ways to perform operations. During QCSYS, you will have the opportunity for hands-on learning about how we do it with photons, but you will also learn how we do it with other type of qubits.

4.6 Multiple Quantum States

Playing around with one qubit is fun, but wouldn't it be even nicer to play with multiple qubits, i.e., a composite system? We will need more fancy mathematics to talk about multiple quantum states.

> **Postulate 4**: The Hilbert space of a composite system is given by the tensor product (aka Kronecker product) of the separate, individual Hilbert spaces.

Before giving you the explicit representation of the tensor/kronecker product, let's discuss some physical properties of composite quantum systems. This should lead us to some abstract properties about the mathematical operation needed to treat the behaviour of multiple quantum systems as one, bigger quantum system. For simplicity, we will consider a composite system of two qubits, but the generalization to multiple quantum systems of different dimensions is straightforward.

Notation 4.32 Let's use the symbol $\otimes$ to denote the abstract mathematical operation of joining the two separate Hilbert spaces of qubit 1 and qubit 2.

Property 1: Dimensions

The first observation to make is that we should be able to see a composite system made of two qubits (2 dimensions each) as a single quantum system with 4 dimensions. This follows from the fact that since each qubit has two distinct states each ($|0\rangle_2$ and $|1\rangle_2$), then

the full system will have 4 distinct states namely:

$$
\begin{aligned}
|00\rangle_4 &= |0\rangle_2 \otimes |0\rangle_2 \longleftrightarrow \text{qubit 1 in } |0\rangle_2 \text{ and qubit 2 in } |0\rangle_2 \\
|01\rangle_4 &= |0\rangle_2 \otimes |1\rangle_2 \longleftrightarrow \text{qubit 1 in } |0\rangle_2 \text{ and qubit 2 in } |1\rangle_2 \\
|10\rangle_4 &= |1\rangle_2 \otimes |0\rangle_2 \longleftrightarrow \text{qubit 1 in } |1_2\rangle \text{ and qubit 2 in } |0\rangle_2 \\
|11\rangle_4 &= |1\rangle_2 \otimes |1\rangle_2 \longleftrightarrow \text{qubit 1 in } |1\rangle_2 \text{ and qubit 2 in } |1\rangle_2
\end{aligned}
$$

We have added the extra subscript 2 and 4 to explicitly denote the fact that they are vectors of dimensions 2 and 4 respectively.

Using the same argument as in Section 4.3, it would make sense to explicitly have:

$$
|00\rangle_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad |01\rangle_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad |10\rangle_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad |11\rangle_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.
$$

Property 2: Measurement probabilities

Now, we will introduce some arguments about probability. If qubit 1 is in $|\psi_1\rangle_2$ and qubit 2 in is $|\psi_2\rangle_2$, such that the joint, 4-dimensional state is abstractly given by $|\Psi\rangle_4 = |\psi_1\psi_2\rangle_4 = |\psi_1\rangle_2 \otimes |\psi_2\rangle_2$. Since the joint system can be seen as a single, larger system of higher dimension, **Born's rule still applies**. Therefore, the probability of measuring qubit 1 in state $|\phi_1\rangle$ and qubit 2 in $|\phi_2\rangle$, i.e., measuring the joint system in state $|\Phi\rangle_4 = |\phi_1\phi_2\rangle_4 = |\phi_1\rangle_2 \otimes |\phi_2\rangle_2$ is given by:

$$
\begin{aligned}
P(\Phi) &= |\langle\Phi|\Psi\rangle_4|^2 \\
&= |\langle\phi_1\phi_2|\psi_1\psi_2\rangle_4|^2 \\
&= \left|\left[\langle\phi_1|_2 \otimes \langle\phi_2|_2\right]\left[|\phi_1\rangle_2 \otimes |\phi_2\rangle_2\right]\right|^2
\end{aligned}
$$

But, if we think of each qubit as thier own separate system, then the probability of measuring qubit 1 in $|\phi_1\rangle_2$ and the probability of measuring qubit 2 in $|\phi\rangle_2$ is given by $P(\phi_1) = |\langle\phi_1|\psi_1\rangle|^2$ and $P(\phi_2) = |\langle\phi_2|\psi_2\rangle|^2$, respectively. Basic probability theory tells us that the probability of two independent things happening is given by the product of the individual probability[5], then we must have that:

$$
P(\Phi) = P(\phi_1)P(\phi_2)
$$

[5] E.g., if the probability of me eating a sandwich tomorrow is 1/4 and the probability of the Toronto Blue Jays winning tomorrow's game is 1/10 (they are not very good!), then since my eating a sandwich has nothing to do with the result of the baseball game, the probability of me eating a sandwich and the Jay to win the game is 1/40.

which essentially means that we must have:

$$|\langle\phi_1\phi_2|\psi_1\psi_2\rangle_4|^2 = |\langle\phi_1|\psi_1\rangle|^2|\langle\phi_2|\psi_2\rangle|^2$$

Property 3: Joint quantum operations

A final physical argument has to do with quantum operations. If U_1 is a unitary operator on qubit 1 and U_2 is a unitary operator on qubit 2, then the join operation $U_1 \otimes U_2$ must have the property that:

$$\begin{aligned}\big[U_1 \otimes U_2\big]|\psi_1\psi_2\rangle_4 &= \big[U_1 \otimes U_2\big]\big[|\psi_1\rangle_2 \otimes |\psi_2\rangle_2\big] \\ &= \big[U_1|\psi_1\rangle_2\big] \otimes \big[U_2|\psi_2\rangle_2\big]\end{aligned}$$

In other words, the resulting joint state after applying the joint quantum operation, $\big[U_1 \otimes U_2\big]|\psi_1\psi_2\rangle_4$ must be equal to the joint state of the individual state after the individual operations, $U_1|\psi_1\rangle_2$ and $U_2|\psi_2\rangle_2$ respectively.

Now we would like to have an explicit representation of the operator $\otimes$, so that we can have an explicit representation of states like $|\psi_1\psi_2\rangle$ and operators $U_1 \otimes U_2$. This is where the Kronecker, or tensor, product comes in.

Definition 4.33 (Kronecker Product for vectors) *The Kronecker product (also referred as the tensor product) is a special way to multiply vectors together to make bigger vectors. Suppose we have the two vectors:*

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \textit{and} \quad \vec{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

The Kronecker product is defined as:

$$\vec{v} \otimes \vec{w} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \otimes \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} v_1 \cdot \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \\ v_2 \cdot \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} v_1 w_1 \\ v_1 w_2 \\ v_2 w_1 \\ v_2 w_2 \end{bmatrix}$$

From the definition of the Kronecker product, it is relatively easy to prove the Property 1 and 2 are satisfied.

Example 4.34 The Kronecker (tensor) product of the following two vectors:

$$\vec{v} = \begin{bmatrix} 2 \\ -3 \end{bmatrix} \quad \text{and} \quad \vec{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

is given by:

$$\begin{aligned}
\vec{v} \otimes \vec{w} &= \begin{bmatrix} 2 \\ -3 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\
&= \begin{bmatrix} 2 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ -3 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} 2 \\ 4 \\ -3 \\ -6 \end{bmatrix}
\end{aligned}$$

From now on, we will drop the subscript of the bras and kets representing the dimensionality of the vectors, as this dimensionality will be obvious from the context.

Definition 4.35 (Multiple Qubits) *If we have two qubits with individual states $|\psi\rangle$ and $|\phi\rangle$, their joint quantum state $|\Psi\rangle$ is given by:*

$$|\Psi\rangle = |\psi\rangle \otimes |\phi\rangle$$

*where $\otimes$ represent the Kronecker product. Notice that the state of two qubits is given by a 4-dimensional vector. In general, any **4-dimensional unit vector can represent the state of 2 qubits**. Note that this definition naturally extends to any number of qubits, e.g., if we have n qubits with individual states $|\psi_1\rangle, \ldots, |\psi_n\rangle$, the joint state will be given by $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots |\psi_n\rangle$. The final vector will have a dimension of 2^n. Similarly, any 2^n-dimensional unit vector can be seen as a n qubit state.*

Example 4.36 Using the explicit definition of the Kronecker product, we see that:

$$
\begin{aligned}
|0\rangle \otimes |0\rangle &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ 0 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \\
|0\rangle \otimes |1\rangle &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ 0 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \\
|1\rangle \otimes |0\rangle &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ 1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \\
|1\rangle \otimes |1\rangle &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ 1 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}
\end{aligned}
$$

which is in agreement with our intuitive argument in Property 1.

Example 4.37 If Alice has the quantum state $|\psi\rangle = |0\rangle$ and Bob has the quantum state $|\phi\rangle = |+\rangle$ then their combined state is given by:

$$
\begin{aligned}
|\psi\rangle \otimes |\phi\rangle &= |0\rangle \otimes |+\rangle \\
&= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}
\end{aligned}
$$

Notation 4.38 For simplicity, we often use the shorthand notation:

$$|\psi\rangle \otimes |\phi\rangle \longleftrightarrow |\phi\rangle|\psi\rangle \longleftrightarrow |\psi\phi\rangle$$

Example 4.39 If $|\Psi\rangle = |01\rangle$ and $|\Phi\rangle = |1-\rangle$, we can calculate the outcome probability in three ways. Let's first write those vectors explicitly:

$$
|\Psi\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \textit{and} \quad |\Phi\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}
$$

1. We will compute directly:

$$\begin{aligned} P(\Phi) &= |\langle\Phi|\Psi\rangle|^2 \\ &= \frac{1}{2}\left|\begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}\right|^2 \\ &= 0 \end{aligned}$$

2. We can use the inner product property of the Kronecker product (Property 2):

$$\begin{aligned} P(\Phi) &= |\langle 01|1-\rangle)|^2 \\ &= |\langle 0|1\rangle|^2 \cdot |\langle 1|-\rangle|^2 \\ &= \left|\begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right|^2 \cdot \frac{1}{2}\left|\begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right|^2 \\ &= 0 \cdot \frac{1}{2} \\ &= 0 \end{aligned}$$

3. From the second method, we did not have to explicitly carry on the inner product. If fact, we already know that $\langle 1|0\rangle = 0$ since $\{|0\rangle, |1\rangle\}$ is an orthornormal basis (recall the definition of orthonormal basis on page 41. Actually, the bracket notation is particularly well suited to carry out inner products without having to explicitly do the calculations.

Observation 4.40 Not every 4-dimensional vector can be written as a Kronecker product of two 2-dimension vectors, e.g., you can have a two qubit state $|\Psi\rangle$ such that:

$$|\Psi\rangle \neq |\psi\rangle|\phi\rangle, \text{ for any one qubit state } |\psi\rangle \text{ and } |\phi\rangle.$$

These type of states (called entangled states) are very intriguing and play a fundamental role in quantum mechanics. Next section is dedicated to them.

Definition 4.41 (Separable state) *If a two qubit state can be written as:*

$$|\Psi\rangle = |\psi\rangle|\phi\rangle, \textit{ for some one qubit state } |\psi\rangle \textit{ and } |\phi\rangle,$$

then we say that $|\Psi\rangle$ is a separable state. Otherwise, we call them entangled states.

Definition 4.42 (Quantum Entanglement) *If a two qubit quantum state $|\Psi\rangle$ cannot be written as $|\psi\rangle \otimes |\phi\rangle$ for any choice of $|\psi\rangle$ and $|\phi\rangle$, then $|\Psi\rangle$ is said to be entangled.*

Entanglement is a fascinating property of quantum mechanics that is completely counter-intuitive. Physically, entangled qubits have a well-defined joint state, yet they do not have a well-defined individual state! It is widely accepted that the phenomenon of quantum entanglement is what makes quantum physics intrinsically different than classical physics. You will see later that entangled qubits have very strong correlation between them and show very interesting properties. During QCSYS, we will investigate some of these amazing properties and we will even use entanglement to do quantum cryptography.

Example 4.43 Are the states $|\Phi\rangle = \frac{1}{2}(|00\rangle + |01\rangle - |10\rangle - |11\rangle)$ and $|\Psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ separable?

1. Lets do the case of $|\Phi\rangle$ first. Assume that $|\Phi\rangle$ is separable, that is there exist two vectors:
$$|\psi\rangle = \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \quad \text{and} \quad |\phi\rangle = \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}$$
such that $|\Phi\rangle = |\psi\rangle|\phi\rangle$. By explicitly writing the vectors, we find the equality:
$$\frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} \psi_1\phi_1 \\ \psi_1\phi_2 \\ \psi_2\phi_1 \\ \psi_2\phi_2 \end{bmatrix}$$
A quick look at the situation tells us the that $\psi_1 = \phi_1 = \phi_2 = 1/\sqrt{2}$ and $\psi_2 = -1/\sqrt{2}$ is a solution. Therefore:
$$\begin{aligned} |\Phi\rangle &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \otimes \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= |-+\rangle \end{aligned}$$
and $|\Phi\rangle$ is thus separable.

2. We'll apply the same technique for the case of $|\Psi\rangle$. Assume that $|\Psi\rangle$ is separable, therefore:
$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \psi_1\phi_1 \\ \psi_1\phi_2 \\ \psi_2\phi_1 \\ \psi_2\phi_2 \end{bmatrix}$$
A quick look at the situation tells us the that $\psi_1 = \phi_1 = \phi_2 = 1/\sqrt{2}$ and $\psi_2 = -1/\sqrt{2}$ is a solution. This tells us that $\psi_1\phi_2 = 0$ which imply that either $\psi_1 = 0$ or $\phi_2 = 0$. If $\psi_1 = 0$ then, $\psi_1\phi_1 = 0$, which is a contradiction. Therefore we must have $\phi_2 = 0$, which would imply that $\psi_2\phi_2 = 0$, which is also a contradiction. We thus conclude that $|\Psi\rangle$ is not separable.

Example 4.44 Given the state $|\Psi\rangle = \frac{1}{2}(|00\rangle + |01\rangle - |10\rangle + |11\rangle)$ and $|\Phi\rangle = \frac{1}{\sqrt{6}}(|00\rangle + i|01\rangle + 2|10\rangle)$, what is the probability of measuring the system in $|\Phi\rangle$ given that it starts in $|\Psi\rangle$?

You can verify for yourself that neither of these states are separable. Therefore, we will not be able to use the second and third methods of Example 4.39 directly, but we can still evaluate the probability without explicitly writing down the vectors.

1. Let's do it explicitly first:

$$|\Psi\rangle = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \quad and \quad |\Phi\rangle = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ i \\ 2 \\ 0 \end{bmatrix}.$$

If our quantum system is prepared in $|\Psi\rangle$, then the probability of measuring it in the state $|\Phi\rangle$ is given, as before:

$$\begin{aligned} \langle\Phi|\Psi\rangle &= \frac{1}{2\sqrt{6}}\begin{bmatrix} 1 & -i & 2 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \\ &= \frac{1 - i - 2}{2\sqrt{6}} \\ &= \frac{-1 - i}{2\sqrt{6}} \\ \implies P(\Phi) &= |\langle\Phi|\Psi\rangle|^2 = \frac{1}{12} \end{aligned}$$

2. For the explicit method, we just carry on the abstract multiplication:

$$\langle\Phi|\Psi\rangle = \frac{1}{2\sqrt{6}}\big[\langle 00| + i\langle 01| + 2\langle 10|\big]\big[|00\rangle + |01\rangle - |10\rangle + |11\rangle\big]$$

The key observation here is that the inner product is written using an orthormal basis, so we only need to multiply the coefficient of the same terms , i.e., the coefficient of $\langle 00|$ with the coefficient of $|00\rangle$, then the coefficient of $\langle 01|$ with the coefficient of $|01\rangle$

and so on. In a blink of an eye, we conclude that:

$$\begin{aligned}\langle\Phi|\Psi\rangle &= \frac{(1-i-2)}{2\sqrt{6}} \\ &= \frac{-1-i}{2\sqrt{6}} \\ \implies & P(\Phi) = |\langle\Phi|\Psi\rangle|^2 = \frac{1}{12}\end{aligned}$$

The second method used above is actually much quicker than writing things down explicitly. With a little bit of practice, of which you will have plenty during QCSYS, you will be able to evaluate measurement probability without writing anything down!

Notation 4.45 (Bell's States) The following two qubit states are known as the Bell's states. They represent an orthonormal, entangled basis for two qubits:

$$\begin{aligned}|\Phi^+\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\ |\Phi^-\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \\ |\Psi^+\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \\ |\Psi^-\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)\end{aligned}$$

Definition 4.46 (Unitary matrices acting on two qubits) *Suppose we have two unitary matrices U_1 and U_2. Then*

$$U = U_1 \otimes U_2$$

is a bigger matrix which satisfies the following rule (Property 3 above):

$$U(|\psi\rangle \otimes |\phi\rangle) = (U_1 \otimes U_2)(|\psi\rangle \otimes |\phi\rangle) = U_1|\psi\rangle \otimes U_2|\phi\rangle.$$

Example 4.47 Recall the matrices defined in Example 4.29. If we apply the unitary matrix $U = X \otimes Z$ to the state $|0+\rangle$, we get:

$$\begin{aligned}U|0+\rangle &= (X \otimes Z)(|0\rangle \otimes |+\rangle) \\ &= X|0\rangle \otimes Z|+\rangle \\ &= |1\rangle \otimes |-\rangle \text{ (verify that for yourself)} \\ &= \frac{1}{\sqrt{2}}\begin{bmatrix}0\\0\\1\\-1\end{bmatrix}\end{aligned}$$

Definition 4.48 (Kronecker Product for matrices) *We can generalize the Kronecker product to matrices the following way. Suppose we have the two matrices:*

$$M = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix} \quad \text{and} \quad N = \begin{bmatrix} n_1 & n_2 \\ n_3 & n_4 \end{bmatrix}$$

The Kronecker product is defined as:

$$\begin{aligned} M \otimes N &= \begin{bmatrix} m_1 \begin{bmatrix} n_1 & n_2 \\ n_3 & n_4 \end{bmatrix} & m_2 \begin{bmatrix} n_1 & n_2 \\ n_3 & n_4 \end{bmatrix} \\ m_3 \begin{bmatrix} n_1 & n_2 \\ n_3 & n_4 \end{bmatrix} & m_4 \begin{bmatrix} n_1 & n_2 \\ n_3 & n_4 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} m_1n_1 & m_1n_2 & m_2n_1 & m_2n_2 \\ m_1n_3 & m_1n_4 & m_2n_3 & m_2n_4 \\ m_3n_1 & m_3n_2 & m_4n_1 & m_4n_2 \\ m_3n_3 & m_3n_4 & m_4n_3 & m_4n_4 \end{bmatrix} \end{aligned}$$

Example 4.49 Using the above definition for the Kronecker product of two matrices, we can redo the previous example explicitly:

$$\begin{aligned} X \otimes Z &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 0 \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} & 1 \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ 1 \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} & 0 \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \end{aligned}$$

Applying this unitary operation on the state:

$$|0+\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

yields:

$$
\begin{aligned}
X \otimes Z|0+\rangle &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} \\
&= |1-\rangle
\end{aligned}
$$

Observation 4.50 (Arbitrary two qubit operation) Similar to the state representation of multiple qubits, not every 4×4 unitary matrix can be written as the Kronecker product of two 2×2 matrices. On the other hand, any 4×4 unitary matrix can be thought of as a quantum operation of two qubits. If the operation cannot be written as a Kronecker product, we say that this is an **entangling operation**.

More generally, any $2^n \times 2^n$ matrix can be seen as a quantum operation on n qubits.

A Greek letters

Lower	Upper	Name	Lower	Upper	Name
α	A	Alpha	ν	N	Nu
β	B	Beta	ξ	Ξ	Xi
γ	Γ	Gamma	o	O	Omicron
δ	Δ	Delta	π	Π	Pi
ϵ	E	Epsilon	ρ	P	Rho
ζ	Z	Zeta	σ	Σ	Sigma
η	H	Eta	τ	T	Tau
θ	Θ	Theta	υ	Υ	Upsilon
ι	I	Iota	ϕ	Φ	Phi
κ	K	Kappa	χ	X	Chi
λ	Λ	Lambda	ψ	Ψ	Psi
μ	M	Mu	ω	Ω	Omega

B Properties of complex numbers: Proofs

In section 2.2, we stated the following properties of complex numbers:

1. $z + w = w + z$ (commutativity of addition)
2. $\overline{z + w} = \overline{z} + \overline{w}$
3. $zw = wz$ (commutativity of multiplication)
4. $\overline{zw} = \overline{z}\,\overline{w}$
5. $z\bar{z} = \bar{z}z = |z|^2$
6. $\overline{\overline{z}} = z$
7. $|z| = |\overline{z}|$
8. $|zw| = |z||w|$
9. $|z + w| \leq |z| + |w|$
10. $z^{-1} = \dfrac{1}{z} = \dfrac{\bar{z}}{|z|^2}$ when $z \neq 0 + 0i$

We will prove each of them, but we strongly encourage you to do it yourself first. Assume that $z = a + bi$ and $w = c + di$.

Property B.1 $z + w = w + z$

Proof

$$\begin{aligned} z + w &= (a + bi) + (c + di) \\ &= a + bi + c + di \\ &= (c + di) + (a + bi), \text{by commutativity of addition of real numbers} \\ &= (c + di) + (a + bi) \\ &= = w + z \end{aligned}$$

□

Property B.2 $\overline{z + w} = \overline{z} + \overline{w}$

Proof

$$\begin{aligned}
\overline{z+w} &= \overline{(a+bi)+(c+di)} \\
&= \overline{(a+c)+(b+d)i} \\
&= (a+c)-(b+d)i \\
&= (a-bi)+(c-di) \\
&= \overline{z}+\overline{w}
\end{aligned}$$

□

Property B.3 *zw=wz*

Proof On one hand, we have:

$$\begin{aligned}
zw &= (a+bi)(c+di) \\
&= ac+adi+bci+bdi^2 \\
&= (ac-bd)+(ad+bc)i, \text{ since } i^2=-1
\end{aligned}$$

On the other hand, we have:

$$\begin{aligned}
wz &= (c+di)(a+bi) \\
&= ca+cbi+dai+bdi^2 \\
&= (ca-db)+(da+bc)i \\
&= (ac-bd)+(ad+bc)i, \text{ using commutativity of multiplication of real numbers}
\end{aligned}$$

Since both results are the same, that concludes our proof.

□

Property B.4 $\overline{zw} = \overline{z}\,\overline{w}$

Proof On one hand, we have:

$$\begin{aligned}
\overline{zw} &= \overline{(a+bi)(c+di)} \\
&= \overline{(ac-bd)+(ad+bc)i} \\
&= (ac-bd)-(ad+bc)i
\end{aligned}$$

On the other hand, we have:

$$\begin{aligned}
\overline{z}\,\overline{w} &= \overline{(a+bi)}\,\overline{(c+di)} \\
&= (a-bi)(c-di) \\
&= ac-adi-bci+bdi^2 \\
&= (ac-bd)-(ad+bc)i
\end{aligned}$$

Since both results are the same, that concludes our proof.

□

Property B.5 $z\overline{z} = \overline{z}z = |z|^2$

Proof

$$\begin{aligned} z\overline{z} &= (a+bi)(a-bi) \\ &= a^2 - abi + bai - b^2 i^i \\ &= a^2 - abi + abi + b^2 \\ &= a^2 + b^2 \\ &= |z|^2, \text{ since } |z| = \sqrt{a^2+b^2} \end{aligned}$$

We already know that $z\overline{z} = \overline{z}z$ by property 3, so we are done.

□

Property B.6 $\overline{\overline{z}} = z$

Proof

$$\begin{aligned} \overline{\overline{z}} &= \overline{\overline{a+bi}} \\ &= \overline{a-bi} \\ &= a+bi \\ &= z \end{aligned}$$

□

Property B.7 $|z| = |\overline{z}|$

Proof

$$\begin{aligned} |\overline{z}| &= |a-bi| \\ &= \sqrt{a^2+(-b)^2} \\ &= \sqrt{a^2+b^2} \\ &= |z| \end{aligned}$$

□

Property B.8 $|zw| = |z||w|$

Proof

$$\begin{aligned}
|zw|^2 &= |(a+bi)(c+di)| \\
&= |(ac-bd)+(ad+bc)i| \\
&= (ac-bd)^2+(ad+bc)^2 \\
&= a^2c^2-2abcd+b^2d^2+a^2d^2+2abcd+b^2c^2 \\
&= a^2c^2+b^2d^2+a^2d^2+b^2c^2 \\
&= a^2c^2+a^2d^2+b^2c^2+b^2d^2, \text{ (just moving things around)} \\
&= a^2(c^2+d^2)+b^2(c^2+d^2), \text{ by factoring out } a^2 \text{ and } b^2 \\
&= (a^2+b^2)(c^2+d^2) \text{ by factoring out } (c^2+b^2) \\
&= |z|^2|w|^2 \\
\Longrightarrow\ & |zw| = |z||w|
\end{aligned}$$

□

Property B.9 $|z+w| \leq |z|+|w|$

Proof That one is a little trickier, but let's do it.

$$\begin{aligned}
|z+w|^2 &= |(a+c)+(b+d)i|^2 \\
&= (a+c)^2+(b+d)^2 \\
&= a^2+2ac+c^2+b^2+2bd+b^2 \\
&= (a^2+b^2)+2(ac+bd)+(c^2+d^2), \text{ (just moving things around)} \\
&= |z|^2+2(ac+bd)+|w|^2
\end{aligned}$$

Now, the critical observation is that $(ac+bd)$ is actually the real part of $z\overline{w}$ since $z\overline{w} = (ac+bd)+(-ad+bc)i$, i.e., $(ac+bd) = Re(z\overline{w})$. The other critical observation is that the real part of a complex number, will always be smaller that the length, or modulus, of that number. To see this, take for example $Re(z) = a \leq \sqrt{a^2+b^2}$.

Therefore, we have that:

$$\begin{aligned}
Re(z\overline{w}) &\leq |z\overline{w}| \\
&= |z||\overline{w}|, \text{ using property 8} \\
&= |z||w|, \text{ using property 7}
\end{aligned}$$

If we put our two results together, we have that:

$$\begin{aligned}
|z+w|^2 &= |z|^2+2(ac+bd)+|w|^2 \\
&\leq |z|^2+2|z||w|+|w|^2 \\
&= (|z|+|w|)^2 \\
\Longrightarrow\ & |z+w| \leq |z|+|w|
\end{aligned}$$

□

Property B.10 $z^{-1} = \dfrac{1}{z} = \dfrac{\bar{z}}{|z|^2}$ *when* $z \neq 0 + 0i$

Proof

$$\begin{aligned} \frac{1}{z} &= \frac{1}{z} \cdot \frac{\bar{z}}{\bar{z}} \quad \text{(we are just multiplying by 1!)} \\ &= \frac{\bar{z}}{|z|^2} \end{aligned}$$

□

C Euler number and exponential functions

Euler numbers e and the exponential functions e^x plays a central role in mathematics. In order to fully grasp the beauty of these two concepts, some knowledge of derivative and integral calculus is needed. Since we do not assume you know calculus, we will not go into the details of what the function is, but focus instead on how we can use it.

In general, an exponential function refer to *any* function of the form $f(x) = a^x$ where a is a constant which can be any numbers (real or complex) and x is a real or a complex variable. The notation a^x refers to multiplying a by itself x time.

Example C.1 Let's see a few examples where the exponent is an integer

1. $2^4 = 2 \cdot 2 \cdot 2 \cdot 2 = 8$
2. $(-3.1)^3 = (-3.1) \cdot (-3.1) \cdot (-3.1) = 29.79$
3. $(2i)^5 = 2i \cdot 2i \cdot 2i \cdot 2i \cdot 2i = 32(i^5) = 32i$

Although we can explicitly and exactly evaluate an exponential function when the exponent is a real integer, we have methods to compute the exponential function for any real and complex exponent. (You will see an example below.)

Exponential functions are particularly well suited to model mathematical situation when a constant change in the independent variable x makes the dependant variable $f(x)$ undergo a proportional change (i.e., its value gets multiplied by a fix amount). Take $f(x) = 2^x$ for example. Every time x increase by 1, the value of $f(x)$ doubles.

Properties C.2 *Here are some properties, without proofs, of exponential functions:*

1. $a^0 = 1$
2. $a^x a^y = a^{x+y}$
3. $a^{xy} = (a^x)^y$
4. $\left(\frac{a}{b}\right)^x = \frac{a^x}{b^x}$
5. $a^{-1} = \frac{1}{a}$. *This is actually a consequence of Property 1 and 2.*
6. $a^{-x} = (a^{-1})^x = \left(\frac{1}{a}\right)^x = \frac{1}{a^x}$

It is now time to introduce Euler's number e, which has the value:

$$e = 2.71828182845904523536028747135 27\ldots$$

The value of e seems to come out of nowhere, but the function $f(x) = e^x$ has some very nice properties. First, the rate of change of that function at any point x is actually e^x. If you are familiar with the concept of the "slope" of a curve, then the slope of e^x is e^x. If you are familiar with calculus, then it means that:

$$\frac{d}{dx}e^x = e^x$$

Moreover, e^x has a very nice series expansion that can be used to evaluate its value for any x:

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \\ &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!}\ldots \end{aligned}$$

Since this series converge very rapidly, we can evaluate e^x for any real or complex x only by adding a few terms.

D Radians

When you learned trigonometric functions (cosine, sine, tangent), chances are you learned to evaluate them using angles given in degrees. The fact that the full circle is split into 360 "degrees" or slices, is actually a completely arbitrary system of units that has nothing to do with anything and is more of a nuisance than anything else!

Again, as mathematicians in training, we are attracted to mathematical beauty, simplicity and logic. Therefore we are going to define a new way to split the circle: the *radian*. We know that the circumference of a circle with radius r is given by $2\pi r$ and we will try to define our angle units in a way that relates to the circumference.

Definition D.1 *Refer to Figure 9. The angle subtended at the centre of a circle, measured in radians, is equal to the ratio of the length of the enclosed arc to the length of the circle's radius. In other words,*

$$\theta_{rad} = \frac{l}{r}$$

where r is the circle radius and l is the arc subtended by the angle.

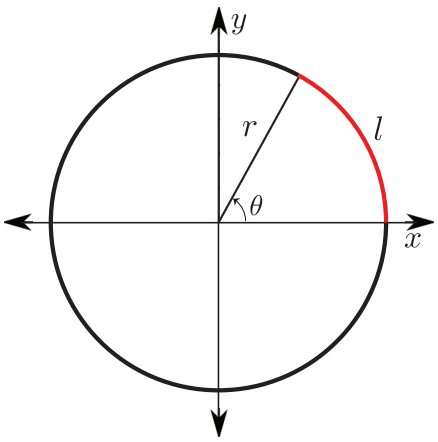


Figure 9: Angle in radian can be defined as the ratio of the the length of the subtended arc l and the radius r.

This might be a little abstract, but at least the definition depends on properties of the circle, and not some arbitrary numbers. From the above definition and the value of the circumference of the circle, it follows that:

$$\text{full circle} \leftrightarrow 2\pi\, rad$$

Note that even though *rad* is a unit of measure, since it is defined as a ratio of two lengths, it is dimensionless. Therefore, we usually don't use the "*rad*".

Since we have the equivalence $360^\circ \leftrightarrow 2\pi$, we can convert an angle given in degrees, say θ° into its corresponding radian equivalent, θ using this simple proportionality relation:

$$\theta_{rad} \quad = \quad \frac{\theta^\circ}{360^\circ} \times 2\pi$$

The first term above is the fraction of the full circle the angle subtend and since we know that the full circle is 2π radian, that we multiply that ratio by 2π to get the radian angle equivalent. From this you can calculate that:

$$1\,rad \approx 57.3^\circ$$

Here is a quick reference table to help you:

Degrees	**Radians**	$\cos\theta$	$\sin\theta$	$\tan\theta$
0	0	1	0	0
30	$\frac{\pi}{6}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\frac{1}{\sqrt{3}}$
45	$\frac{\pi}{4}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	1
60	$\frac{\pi}{3}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\sqrt{3}$
90	$\frac{\pi}{2}$	0	1	∞
135	$\frac{3\pi}{4}$	$-\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	-1
180	π	-1	0	0
225	$\frac{5\pi}{4}$	$-\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$	1
270	$\frac{3\pi}{2}$	0	-1	$-\infty$
315	$\frac{7\pi}{4}$	$\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$	-1

E Proof of Euler's theorem

Theorem E.1 (Euler's formula) *We would like to prove that:*

$$e^{i\theta} = \cos\theta + i\sin\theta$$

*where θ is in **radians**.*

Proof The Taylor series of basic functions are:

$$\begin{aligned}
\sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}\dots \\
\cos x &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!}\dots \\
e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!}\dots
\end{aligned}$$

Note that we can break the Taylor series of the exponential functions into even and odd exponents:

$$\begin{aligned}
e^x &= \sum_{n\,even} \frac{x^n}{n!} + \sum_{n\,odd} \frac{x^n}{n!} \\
&= \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}
\end{aligned}$$

By observing the pattern: $i^2 = -1, i^3 = -i, i^4 = 1, i^5 = i\dots$, we see that:

$$\begin{aligned}
i^{2n} &= (-1)^n, \text{ for } n = 0, 1, 2, \dots \\
i^{2n+1} &= i(-1)^n, \text{ for } n = 0, 1, 2, \dots
\end{aligned}$$

Putting all this together, we have:

$$\begin{aligned}
e^{i\theta} &= \sum_{n=0}^{\infty} \frac{(i\theta)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(i\theta)^{2n+1}}{(2n+1)!} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{i(-1)^n \theta^{2n+1}}{(2n+1)!} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n}}{(2n)!} + i\sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n+1}}{(2n+1)!} \\
&= \cos\theta + i\sin\theta
\end{aligned}$$