THE MATHEMATICS OF QUANTUM MECHANICS

Martin Laforest, PhD

UNIVERSITY OF WATERLOO | IQC Institute for Quantum Computing

QCSYS
Quantum Cryptography School for Young Students

Contents

Preface 5

0.1 A taste of quantum mechanics 6

0.2 Quantum technologies 7

0.3 Welcome to QCSYS 8

0.4 About the author 9

0.5 How to read this book 10

Complex numbers 13

1.1 What is a complex number? 14

1.2 Doing math with complex numbers 17

1.3 Euler's formula and the polar form 21

1.4 Summary 25

Linear algebra 27

2.1 Vectors 28

2.2 Matrices 35

2.3 Complex conjugate, transpose and conjugate transpose 47

2.4 Inner product and norms 50

2.5 Basis 55

2.6 Inner product as projection 61

2.7 Special matrices 66

2.8 The cooking matrix 67

2.9 Summary 70

Quantum mechanics 71

3.1 Mathematical postulates of quantum mechanics 72

3.2 New notation: the braket notation 73

3.3 Single quantum state and the qubit 74

3.4 Quantum measurement 77

3.5 Quantum operations 81

3.6 Multiple quantum states 83

3.7 Observables 98

3.8 Summary 100

Appendices 101

A.1 Greek letters 102

A.2 Properties of complex numbers: Proofs 103

A.3 Euler number and exponential functions 107

A.4 Radians 109

A.5 Proof of Euler's theorem 111

Preface

"*If quantum mechanics hasn't profoundly shocked you, you haven't understood it yet.*"
NIELS BOHR

"*The nineteenth century was known as the machine age, the twentieth century will go down in history as the information age. I believe the twenty-first century will be the quantum age.*"
PAUL DAVIES

0.1 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's 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 and many others.

Classical physics is a tremendously successful theory of nature and has led to astonishing human feats. We put a man on the moon, built bridges, skyscrapers, supersonic jets, developed wireless communication, etc. The fascinating part is that classical physics is not the ultimate description of nature – there is much more to reality than what we see and experience.

If we try to describe the behaviour of atoms and their constituents (e.g., 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 using classical principles, you can calculate that any atom would collapse within a tiny fraction of a second. Obviously, the world we live in is made of stable atoms. . . so what's 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'll learn, leads to surprising and fascinating new phenomena: welcome to the beautiful world of quantum mechanics!

Particles behave like waves and waves behave like particles. Electrons tunnel through barriers. It's impossible to perform a measurement without perturbing the environment. Quantum entangled photons are so strongly interconnected that they behave as one, no matter how far apart they are. These are all part of everyday life in the quantum world.

Quantum mechanics can be baffling, yes, surprising, definitely, and certainly counter-intuitive. That's because quantum mechanics lives outside of our everyday lives and any attempt to explain quantum phenomena using classical physics fails. Quantum mechanics just is, and it's awesome!

Turns out that quantum mechanics isn't really that complicated, we just need to experience it and build an intuition about it. 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 20^{th} century by the likes of Max Planck, Erwin Schrödinger, Werner Heseinberg, Paul Dirac and many others, 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 couldn't be a serious theory. Unfortunately for him, he was wrong!

An astonishing amount of 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 to prove it wrong has failed miserably.

You may have already heard about wave-particle duality. It's one of the most quintessential phenomena of quantum. Sometimes an electron (or any other quantum object) behaves like a particle, sometimes it behaves like a wave. How do you know when you should treat the electron like a wave, and when you should treat it like a particle? Part of the beauty of quantum mechanics is that we don't need to make that distinction - it's all contained within the theory.

A final note: Not only does quantum mechanics accurately describe the behaviour of atoms and molecules, it can actually describe nature at any scale. The only reason we don't really need quantum mechanics to describe large objects is because the quantum effects play such a small role that they can be ignored, and classical physics represent an accurate approximation. Though that doesn't mean quantum effects can't be observed in larger objects, such as superconductors (material conducting electricity with zero resistance), nano-size electrical circuits and transistors, just to name a few.

0.2 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 atomic level and beyond – but it has led to transformative technologies that have shaped, and continue to shape, the 20th and 21st centuries. The laser, LEDs, magnetic resonance imaging (MRI), transistors and so much more, all exist because the world behaves according to the rules of quantum mechanics.

What would a world be like 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 – wouldn't exist! The world would be radically different.

Speaking of digital technology, the digital world we now live in has been made possible thanks largely 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, 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'll even see some of them at the Institute for Quantum Computing (IQC), stable quantum cryptography systems are commercially available used by government and banks around the world, quantum sensors are bursting out of our labs and used in medicine, material science, resource exploration and other fields.

0.3 Welcome to QCSYS

During the Quantum Cryptography School for Young Students (QCSYS or "cue-see-sis"), you'll become familiar with a special type of quantum technology: quantum cryptography, or more precisely, Quantum Key Distribution (QKD). Nowadays, when secure information is being sent over the internet (bank transactions, your password when you log in to your favourite social media website, etc.) your information remains private. 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 difficult 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 can secure information such that the only way for an all-evil eavesdropper to access this information would be to break the rules of physics. We're pretty confident nobody can do that. Ever!

During QCSYS, you'll 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 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 21st century, we strongly believe it's 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 technologies, but it offers a great vehicle to teach young students about technology, information security, mathematics, quantum mechanics and quantum computing.

We'll repeat it many times: quantum physics isn't about mathematics, it's about the behaviour of nature at its core. But since mathematics is the language of nature, it's required to quantify the prediction of quantum mechanics. This present document has been put together to ease you into the mathematics of quantum mechanics. We'll use special mathematics – complex numbers and linear algebra (vectors and matrices). Unfortunately, most high school mathematics curricula around the world do not teach linear algebra. It's not very complicated. It's really just a different and clever way to add and multiply numbers together, but it's a very powerful tool.

We don't 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 here 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 you come to Waterloo for QCSYS. Do not panic if it feels like it's over your head or you're struggling with some of the concepts. We'll spend at least five hours going through the key sections of the book and work through exercises in groups. QCSYS counsellors, graduate students and I will be on hand during QCSYS to help you out.

In addition to the mathematics of quantum mechanics, we'll spend another five hours exploring the "physics" of quantum mechanics. We'll first explain the behaviour of quantum particles without quantifying it. We'll then consolidate the two so you have a good understanding of how we use mathematics to model the physical quantum world. After this introduction, we'll be ready to learn about cryptography, quantum optics, quantum cryptography (of course) and even quantum hacking. We'll also go in the labs and do some experiments. You'll even have the chance to build your own quantum cryptography system!

A little note before getting into the heart of the subject: I would like to thank the people who helped me put this document together. The starting point of this mathematical primer was a set of class notes put together a few years ago by a then graduate student at IQC – Jamie Sikora. Jamie was one of the first teachers of QCSYS and we owe him greatly. Thanks also to Juan-Miguel Arrazola for providing great feedback on earlier versions of this book and Jodi Szimanski and Kathryn Fedy for proofreading and editing.

Finally, welcome to QCSYS, welcome to the Institute for Quantum Computing and welcome to Waterloo. We sincerely hope you'll have a great time, learn a lot and build new friendships that will last forever.

0.4 About the author

Martin Laforest is the Senior Manager, Scientific Outreach at the Institute for Quantum Computing, University of Waterloo, Canada. Officially, Martin's role is to bring science out of the lab and into peoples lives. Unofficially, he's the talkative guy who's passionate about quantum science and technology and likes to share it with young minds. Martin leads IQC's youth outreach programs including the Undergraduate School on

Experimental Quantum Information Processing (USEQIP) and the Quantum Cryptography School for Young Students (QCSYS).

Martin has always been fascinated by trying to understand how the world works. That led him to earn an undergraduate degree in Mathematics and Physics at McGill University and later, a PhD in quantum physics from the University of Waterloo. Before starting his current position at IQC, Martin spent two years doing research at the Delft University of Technology in the Netherlands.

0.5 How to read this book

We understand that a lot of material in this book will be new to you. In order to make it as easy to read as possible, we came up with a series of notations to help you. The material in each chapter has been classified using six different categories. Notice how some of them are numbered for easy reference. We'll list them below and use the concept of mathematical sets (which isn't something we need to know to understand the rest of the book) to give examples.

Definitions
Since mathematics is a man-made construction, we'll introduce each new mathematical concept by highlighting them in a box as follows:

> **DEFINITION 0.5.1: Mathematical sets.** In mathematics, a **set** is a collection of distinct objects. A set is also considered an object in its own right. If we have objects a, b and c, the set containing these three objects, denoted S, would be written using the following notation:
>
> $$S = \{a, b, c\}$$

Examples
This book includes loads of examples. We believe concrete examples can help solidify your understanding of new mathematical concepts. Some of them are rather trivial, some of them are quite hard – designed to make you think a little. Examples are also highlighted using a box as follows:

> **EXAMPLE 0.1:** Suppose we have the number 42, a cat and planet Mars, then the collection
>
> $$S = \{42, \text{cat}, \text{Mars}\}$$
>
> is a set containing 3 objects. Since S is also an object, the collection
>
> $$S' = \{42, \text{cat}, \text{Mars}, S\}$$
>
> is a set containing 4 objects.

You'll also notice that there aren't any exercises. This is by design. We'll spend a lot of time during QCSYS solving problems in group. You're therefore strongly encouraged to try to solve the examples before completely reading through them.

Observations
We use this category to stress intuitive conclusions that can be derived from a definition or an example. They'll help further your understanding of certain applications of new concepts.

Observation 0.5.2: A set must be seen as a collection of object, therefore it doesn't matter in which order the objects are listed. That is to say, the sets $\{42, \text{cat}, \text{Mars}\}$ and $\{\text{cat}, 42, \text{Mars}\}$ are the exact same set.

Notes
We use these to bring your attention to something we want to make sure you didn't miss or to give you a little extra information. Notes are written in the outer margin of the page as you can see on the right.

Food for thought
These are designed to be brain teasers, to make you think a little harder, a little deeper, maybe even make you obsess. See if you can answer them. Just like notes, they're also found in the outer margin as you can see on the right.

Trivia facts
These will give you interesting facts about some concepts you've just learned. They are complementary to the rest of the material and aren't necessarily needed to understand the upcoming material. They're also found in the outer margins.

Finally, one thing you'll notice is that there are a lot of equations. It's a mathematics book after all. Some of the equations will be boxed, for example:

$$\boxed{E = mc^2}$$

These boxed equations are particularly important, so pay close attention to them. They're also summarized at the end of each chapter.

Now, get to the first chapter and have a good read!

Note 0.5.3 Objects in a set don't have to be of the same nature. For example, the set S in Example 0.1 contains a number, a planet inhabited solely by robots and the best/worst house pet there is.

Food for thought We saw that, by definition, a set is an object on its own right. Therefore, does the "set of all sets" exist? Explain why.

Trivia fact Set theory is the modern study of sets initiated by Georg Cantor and Richard Dedekind in the 1870s. It's commonly employed as a foundational system for mathematics, that is to say a series of definition and rules (known as axioms, or postulates) designed to study the logical basis of mathematics.

Chapter 1:
Complex numbers

"I tell you, with complex numbers you can do anything."
JOHN DERBYSHIRE

"Mathematics is the tool specially suited for dealing with abstract concepts of any kind and there is no limit to its power in this field."
PAUL DIRAC

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 $\mathbb{R}$.

But, sometimes, real numbers aren't sufficient. Below is an example that might be all too familiar.

EXAMPLE 1.1: Recall that the solution to any quadratic equation of the form $ax^2+bx+c=0$, is given by:

$$x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$$

If we try to solve the quadratic equation $x^2+5x+10=0$ for x, we'll get:

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

What's the value for $\sqrt{-15}$? In other words, is there a real number a such that $a^2 = -15$? It's 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're 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'll learn soon enough, in quantum mechanics, complex numbers are absolutely everywhere.

1.1 What is a complex number?

In some sense, we've 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'll call numbers like this complex numbers.

But of course, being mathematicians-in-training, we'd 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 1.1.1: Imaginary unit number. We define the **imaginary unit number** i as the **square root of -1**, that is:

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

This definition might look arbitrary at first, but mathematics is an abstract subject and as long as we're consistent, we're free to define anything we want!

Trivia fact

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'll know it as j, since i is a variable denoting the electrical current. But for now, let's stick to our regular convention.

EXAMPLE 1.2: Using our new definition for the square root of -1, it's now possible to write:

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

DEFINITION 1.1.2: Imaginary numbers. A number is said to be **imaginary** if its square is negative.

EXAMPLE 1.3: $\sqrt{-15}$ is an imaginary number because $\left(\sqrt{-15}\right)^2 = -15$

EXAMPLE 1.4: Let's look at the same example 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've defined imaginary numbers, we finally can define complex numbers.

DEFINITION 1.1.3: 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:

$$\begin{aligned}Re(z) &= a \\ Im(z) &= b\end{aligned}$$

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

Trivia fact

Italian mathematician Gerolamo Cardano is the first known person to have defined and used complex numbers. That was waaaay back in the 16th century. Cardano was trying to solutions to equations of the form $x^3 + ax + c = 0$.

EXAMPLE 1.5: Let's solve again the quadratic equation $x^2 + 5x + 10 = 0$ for x.

As before, we have:

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

There's a very nice way to graphically represent complex numbers as long as you realize that real numbers and imaginary numbers are exclusive. That is to say, 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). For this reason we can use the cartesian plane to represent and visualize a complex number:

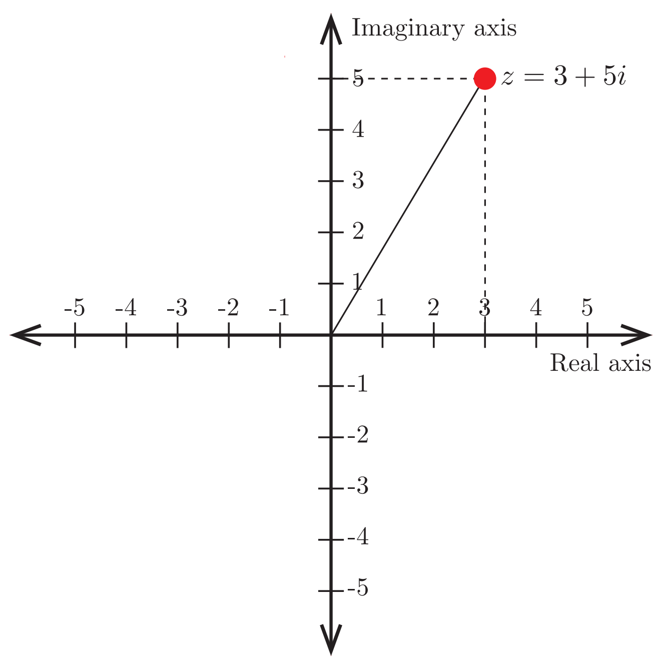

Figure 1.1: Using the **complex plane**, we can visualize any complex number z as a point on a two-dimensional plane. Represented here is the number $z = 3 + 5i$.

DEFINITION 1.1.4: Complex plane. 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.

EXAMPLE 1.6: Looking at Figure 1.1 on previous page, the number $z = 3 + 5i$ is represented on the cartesian plane. The real part of z, 3, is the projection of that point on the "real axis", while the imaginary part of z, 5, is the projection of z along the "imaginary axis".

1.2 Doing math with complex numbers

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

DEFINITION 1.2.1: Complex addition and multiplication. Consider the complex numbers $z = a + bi$ and $w = c + di$, where a, b, c and d are real numbers. Then we can define:

1. Complex addition:

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

2. Complex multiplication:

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

Note 1: In the last step of the multiplication above, we've 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 1.7: Consider the following examples using the complex numbers $z = 1 + 3i$ and $w = -2 + i$:

$$\begin{aligned} z + w &= (1 + 3i) + (-2 + i) \\ &= -1 + 4i \\ zw &= (1 + 3i)(-2 + i) \\ &= -2 + i - 6i + 3i^2 \end{aligned}$$

$$= -2 - 5i - 3$$
$$= -5 - 5i$$

Now that we know how to add and multiply complex numbers, we'll introduce some useful definitions and properties. They might seem a little arbitrary at first. But as you'll soon see, these will become very handy, especially when we start using them in quantum mechanics.

DEFINITION 1.2.2: Complex conjugate. We define the **complex conjugate** of a complex number $z = a + bi$, denoted $\bar{z}$ (notice the over bar), as:

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

EXAMPLE 1.8: Let's calculate some complex conjugates:

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

Food for thought

Referring to the complex plane, what kind of geometric operation (e.g., translation, rotation, etc.) does taking the complex conjugate of a number represent?

DEFINITION 1.2.3: Modulus. The **modulus** (or length) of a complex number $z = a + bi$ is given by:

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

Since both a and b are real, the modulus is always real and positive.

EXAMPLE 1.9: Let's calculate the moduli 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 $|z| = \sqrt{0^2 + 2^2} = \sqrt{4} = 2$

Note 1.2.4

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

Observation 1.2.5: Modulus in the complex plane. By looking at the representation of a complex number on the complex plane (see Figure 1.2 on page 19), the modulus of the number is simply the distance from the origin $0 + 0i$ to the number. (Hence why we also call it the **length** of the number.)

Observation 1.2.6: Complex conjugate and modulus. Notice the very useful application of the complex conjugate: Given $z = a + bi$, we see that:

$$z\bar{z} = (a + bi)(a - bi)$$
$$= a^2 + abi - abi + (-i)ib^2$$

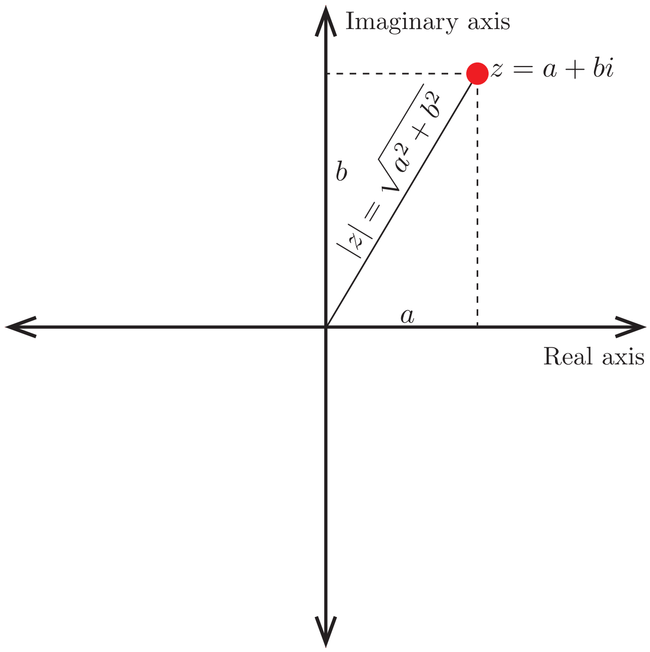

Figure 1.2: The **modulus**, or length, of a complex number is the distance between the origin $(0+0i)$ and the point representing z.

$$
\begin{aligned}
&= a^2 + b^2 \\
&= |z|^2 \\
&\Rightarrow |z| = \sqrt{z\bar{z}}
\end{aligned}
$$

It seems like we're in good shape to play around with complex numbers, but what if someone asked you to divide by a complex number? For example, what's the value 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. Dividing x by y is really just a fancy way of saying, "how many times do I have to multiply y to get x?" Since multiplication is well defined for complex numbers, so is division.

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

Observation 1.2.7: Complex division. Since multiplication is well defined for complex numbers, so is **complex division**. Given a complex number $z = a + bi$, as long as $z \neq 0+0i$, observe that:

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

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 $\frac{1}{z}$ is also written as z^{-1}.

EXAMPLE 1.10: We can clean up the fraction $\frac{1+i}{2-i}$ as:

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

Trivia fact

We just learned that complex numbers are essentially an extension of real numbers. A natural question to ask: Is there an extension to complex numbers? Turns out yes! The **Quaternions** – otherwise known as $\mathbb{H}$. Quaternions are a number system extending complex numbers with the important distinction that they're non-commutative, that is, if a and b are arbitrary quaternion, ab isn't necessarily equal to ba. Just like a complex number can be visualized as being a point of a plane (i.e., a 2-dimensional surface), quaternions are points in a 4-dimensional "hyperspace". Quaternions are very useful in applied mathematics, physics, computer graphics and computer vision.
Is there an extension to quaternions? Of course there is! The **Octonions**, or $\mathbb{O}$. Not only are octonions non-commutative, they're are also non-associative, that is if a, b and c are octonions, then $(a+b)+c$ is not necessarily equal to $a+(b+c)$. Octonions can be seen as a point in a 8-dimensional hyperspace. Although not as studied as quaternions, octonions do have some applications in string theory and special relativity.
How about an extension to octonions? Sadly, it doesn't exist.

Properties 1.2.8: Summary of complex numbers properties. Below is a summary list of properties for complex numbers. Feel free to prove them for yourself if you're not convinced, or refer to appendix A.2 on page 103. Let z and w be any complex numbers:

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

1.3 Euler's formula and the polar form

So far, we've explicitly written any complex number z in the form $z = a + bi$. As you may have noticed, this form is not particularly well suited for multiplying and dividing complex numbers (which we do a lot in quantum mechanics). Thankfully, there's a different way of handling complex numbers. In order to understand this method, you'll need to be familiar with the exponential function e^x, as well as trigonometric functions using radians instead of degrees. If you're not familiar with either of those, consult Appendices A.3 and A.4, on pages 107 and 109 respectively, before continuing.

DEFINITION 1.3.1: Euler's formula. Euler's formula is a well-known result for complex numbers, establishing a deep relationship between trigonometric functions and complex exponentials. It states that:

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

where θ is a real number and in radians (i.e., unitless). The proof of Euler's formula is not particularly complicated, but it does require the knowledge of Taylor Series. If you don't have this knowledge, or you're curious, visit Appendix A.5 on page 111.

Observation 1.3.2: Polar form of complex numbers. Any complex number $z = a + bi$ can be written in the form:

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

where $|z|$ is the modulus of z as previously defined on page 18 and θ is the angle (in radian) between the real axis and the complex number in the complex plane (see Figure 1.3 on page 22). Therefore:

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

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

You can convince yourself that the polar form notation is equivalent by looking at the representation of $z = a + bi$ on the complex plane (Figure 1.3 on page 22). We can see that:

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

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

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

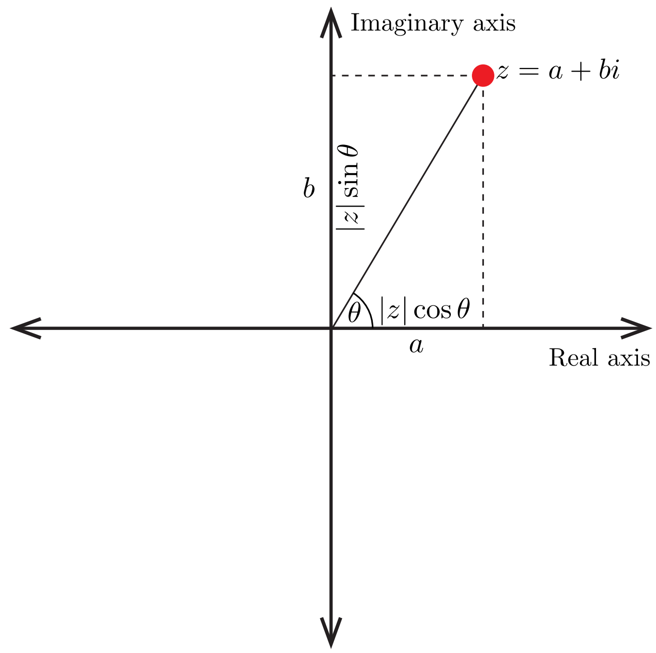

Figure 1.3: Using Euler's theorem, you can represent a complex number using its modulus and its argument (the angle between the real axis and the complex number).

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

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

EXAMPLE 1.11: What's the polar form of $z = 5 - 5i$? We first need to find the modulus of z, which is given by:

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

The argument is given by:

$$\theta = \arctan\left(\frac{5}{-5}\right)$$
$$= \arctan(-1)$$
$$= \frac{3\pi}{4} \text{ or } \frac{7\pi}{4}$$

(see Appendix A.4 on page 109)

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 1.3 for reference), hence:

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

Trivia fact

If you use $\theta = \pi$ in Euler's formula, you obtain the famous **Euler's identity**:

$$e^{i\pi} + 1 = 0$$

Euler's identity is seen as one of the most deeply beautiful mathematical equations. Think about it: it contains one occurrence of three of the most important arithmetic operations (multiplication, addition and exponentiation) and it links five of the most fundamental and celebrated mathematical constants:

- The number 0, the additive identity
- The number 1, the multiplicative identity
- The number π, because we all love π and because it's found everywhere in nature
- The number e, the base of the natural logarithm
- The number i, our new best friend and unit of imaginary numbers

Therefore:

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

Observation 1.3.3: Periodicity. Referring to Euler's formula, the function $e^{i\theta}$ is a **periodic function** of θ with a period of 2π, that is:

$$\boxed{e^{i(\theta\pm 2\pi)}=e^{i\theta}}$$

This can be readily proven since:

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

Given that observation, we can always assume that θ has a <u>value between 0 and 2π</u>.

EXAMPLE 1.12: Given the periodicity of complex numbers, the results of the previous example could also have been written as:

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

Food for thought

Some people prefer to limit the value of θ to any number between $-\pi$ and π. Can you explain why it's equivalent?

Observation 1.3.4: Complex conjugate in polar form. Calculating the complex conjugate 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:

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

Even if the exponent of e is complex, all the basic properties of the exponential function (see Appendix A.3 on page 107) are conserved. The exponential notation is particularly well suited for multiplying, dividing and inverting complex numbers.

Properties 1.3.5: Summary of properties of the polar form. Given two complex numbers $z = |z|e^{i\theta}$ and $w = |w|e^{i\phi}$, below is a list of properties of 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 number n (i.e., n could be complex!)
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 $e^{\pm 2\pi i} = \cos(\pm 2\pi) + i\sin(\pm 2\pi) = 1$, then $e^{i(\theta\pm 2\pi)} = e^{i\theta} \cdot e^{\pm 2\pi i} = e^{i\theta}$

EXAMPLE 1.13: Here are a few examples to get used to this new notation:

1. If $z = e^{-i\frac{\pi}{2}}$ and $w = 4e^{i\frac{\pi}{4}}$, then:

$$\begin{aligned} zw &= 4e^{-i\frac{\pi}{2}}e^{i\frac{\pi}{4}} \\ &= 4e^{-i\frac{\pi}{4}} \\ &= 4\left(\cos\frac{\pi}{4} - i\sin\frac{\pi}{4}\right) \\ &= 4\left(\frac{1}{\sqrt{2}} - i\frac{1}{\sqrt{2}}\right) \\ &= 2\sqrt{2}(1-i) \end{aligned}$$

2. If $z = 2e^{i\frac{\pi}{3}}$, then:

$$\begin{aligned} z^4 &= 2^4\left(e^{i\frac{\pi}{3}}\right)^4 \\ &= 16e^{i\frac{4\pi}{3}} \\ &= 16\left(\cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3}\right) \\ &= 16\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) \\ &= -8\left(1 + i\sqrt{3}\right) \end{aligned}$$

3. If $z = e^{-i\frac{\pi}{2}}$ and $w = 4e^{i\frac{\pi}{4}}$, then:

$$\begin{aligned} \frac{w}{z} &= \frac{4e^{i\frac{\pi}{4}}}{e^{-i\frac{\pi}{2}}} \\ &= 4e^{i\frac{\pi}{4}}e^{i\frac{\pi}{2}} \\ &= 4e^{i\frac{3\pi}{4}} \\ &= 4\left(\cos\frac{3\pi}{4} + i\sin\frac{3\pi}{4}\right) \\ &= 4\left(-\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}\right) \\ &= -2\sqrt{2}(1-i) \end{aligned}$$

Chapter 1 Summary

Below we've listed the key formulas and concepts we've learned so far. Let's assume that $z = a + bi$ and $w = c + di$.

Imaginary unit number: $\boxed{i = \sqrt{-1}}$

Complex numbers addition: $\boxed{(a+bi)+(c+di) = (a+c)+i(b+d)}$

Complex numbers multiplication: $\boxed{(a+bi)(c+di) = (ac-bd)+i(ad+bc)}$

Complex conjugate: $\boxed{\bar{z} = a - bi}$

Modulus of a complex number: $\boxed{|z| = \sqrt{z\bar{z}} = \sqrt{a^2+b^2}}$

Euler's formula: $\boxed{e^{i\theta} = \cos\theta + i\sin\theta}$

Polar form: $\boxed{z = |z|e^{i\theta}}$

Periodicity of complex numbers: $\boxed{e^{i\theta \pm 2\pi} = e^{i\theta}}$

Chapter 2:
Linear algebra

"How can it be that mathematics, a product of human thought independent of experience, is so admirably adapted to the objects of reality?"
ALBERT EINSTEIN

"Algebra is generous; she often gives more than is asked of her."
D'ALEMBERT

In the previous section, we became familiar with a new family of numbers – complex numbers. As you may have noticed so far, although we've introduced new mathematical **concepts**, we haven't 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'll introduce you to the language of quantum mechanics – linear algebra. Just like complex numbers, the type of linear algebra we'll introduce here will necessitate only basic arithmetic, but we'll use it 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 already 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 or computer scientist? Learn linear algebra. Do you want 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 this: Basic linear algebra is rather simple. Yet it's so useful for so many applications. Applications aside, linear algebra is a fascinating, self-contained mathematical field on its own. But since QCSYS is a multidisciplinary school in mathematics and physics, we'll introduce you to the mathematical concepts of linear algebra, while connecting it to potential applications beyond quantum mechanics.

2.1 Vectors

The starting point of linear algebra is the concept of **vectors**. In high school physics, chances are you've already seen that concept as being nothing more than "a number and a direction". This isn't false, but it's definitely not the whole story. It's a rather intuitive way of introducing vectors, hence why we'll use this analogy extensively.

From a mathematical perspective, vectors are just a way of stacking numbers together in a column or a row. For example:

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} i \\ 3 \\ -3 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 2.1 \end{bmatrix}, \quad \begin{bmatrix} -3 & \frac{1}{2} & 4 \end{bmatrix}$$

The first two vector examples are naturally referred to as **column vectors** and the last two as **row vectors**. Unless otherwise specified, when we refer to a vector, we're referring to a column vector.

Since we're mathematicians-in-training, we want to come up with a rigorous definition for this new concept.

DEFINITION 2.1.1: Vectors. A **vector** is a column of numbers (any numbers, even complex). The amount of numbers is referred to as the **dimension** 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:

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

v_1 is referred to as the first component of $\vec{v}$, v_2 as the second component, and so on and so forth.

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 lie in $\mathbb{R}^n$. If we also consider the complex vectors of n-dimensions, we say they lie in $\mathbb{C}^n$.

Even if you've never explicitly learned about vectors until now, you've already seen them. A 2-dimensional vector of real numbers is analogous to the cartesian coordinates. If you refer to Figure 2.1 on page 30, the vector:

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

is equivalent to the (1,2) coordinates. It's also the mathematical way of representing the point on the cartesian plane you would end up at if you were starting from the origin, then move 1 along the x-axis and then 2 along the y-axis.

Observation 2.1.2: Spatial coordinates. *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 "hyperspace".

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 component of a vector can be used to represent quantities that are "exclusive" or "independent" of each other.

Observation 2.1.3: Vector components. Each component in a 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 component of a 3-dimensional vector represents a position along the x, y and z-axis respectively. For example, if you're on the x-axis, you don't have any y or z coordinates. If you're on

» Food for thought

What vector would you use to represent the origin of the cartesian plane?

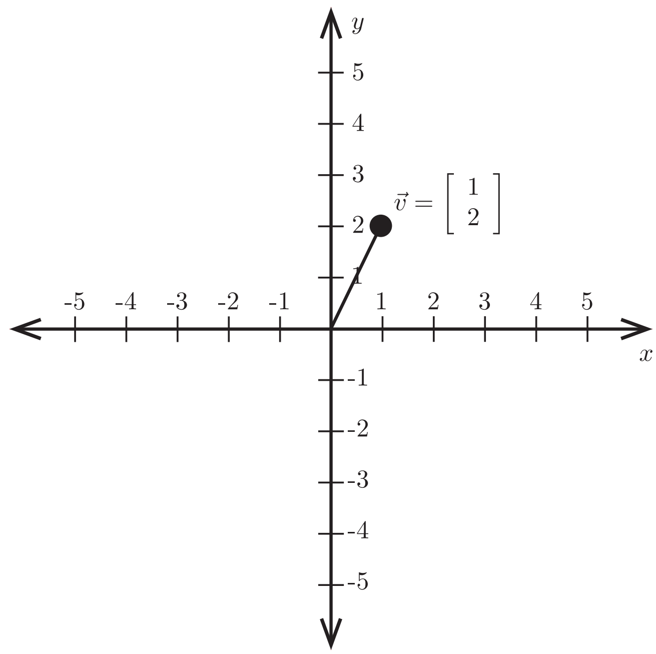

Figure 2.1: The cartesian coordinates or a point on a 2-dimensional surface can be written as a 2-dimensional vector.

the y-axis, you don't have any x or z coordinates, etc. Hence those x, y and z properties are independent of each other.

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

In the coming pages, we'll extensively use the cartesian plane as an analogy to develop our intuition about vectors, but we'll then rigorously define these new concepts so they can be applied to complex vectors of arbitrary dimensions.

Since we just defined a new mathematical concept, it's natural to ask the question: can we add and multiply vectors? The question of multiplying vectors is quite subtle and we'll discuss it a little later. But when it comes to adding vectors, there's something intuitive to it. Let's use the two following vectors:

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

Observation 2.1.4: Adding vectors. Imagine the vectors $\vec{v}$ 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.

Observation 2.1.5: Parallelogram technique. If we added a little more mathematical rigour to our intuition, adding two vectors can be done geometrically by using the **parallelogram technique** (i.e., putting the second vector at the end of the first). See Figure 2.2 where we've used the parallelogram technique to add $\vec{v}$ and $\vec{w}$ as defined above.

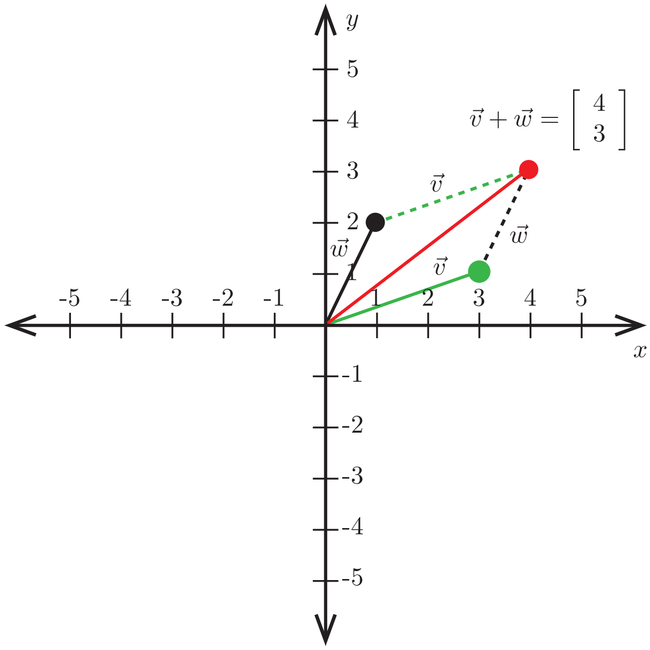

Figure 2.2: You can use the parallelogram technique to visualize vector addition in 2-dimensions.

In some sense, we've already defined vector addition using our intuition. We'll give it a rigorous definition so that it holds for vectors of any dimensions, including complex vectors.

DEFINITION 2.1.6: Vector Addition. Adding vectors is easy, just add each corresponding component! If $\vec{v}$ and $\vec{w}$ are complex vectors written explicitly as:

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

adding them gives:

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

Even more generally, if $\vec{v}$ and $\vec{w}$ are arbitrary n-dimensional vectors, the j^{th} component of $\vec{v}+\vec{w}$, denoted $(\vec{v}+\vec{w})_j$ is given by:

$$\boxed{(\vec{v}+\vec{w})_j = v_j + w_j}$$

Note 2.1.7

You cannot add vectors of different dimensions. 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 doesn't exist. (Or if you prefer, the answer is not defined.)

EXAMPLE 2.1: Adding the vectors:

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

will give you:

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

Earlier, we mentioned that vector multiplication is a little tricky. Multiplying vectors together is, but multiplying vectors by a scalar (i.e., just a number) is also intuitive.

Observation 2.1.8: Multiplying a vector by a number. As before, suppose you start at the origin of the cartesian plane and move 1 along the x-axis and 2 along the y-axis. You find yourself in position:

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

What if you repeat this procedure twice? Then you'll be in:

$$\begin{bmatrix} 2 \\ 4 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 2\vec{v}$$

Three times? You'll end up at:

$$\begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 3\vec{v}$$

Refer to Figure 2.3 on page 33 for visuals.

Based on the intuitive description above, we can generalize vector scalar multiplication so it applies to complex vectors of any dimensions:

DEFINITION 2.1.9: Vector Scalar Multiplication. You can scale a vector by just multiplying each entry by the scalar (the number). If we have:

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

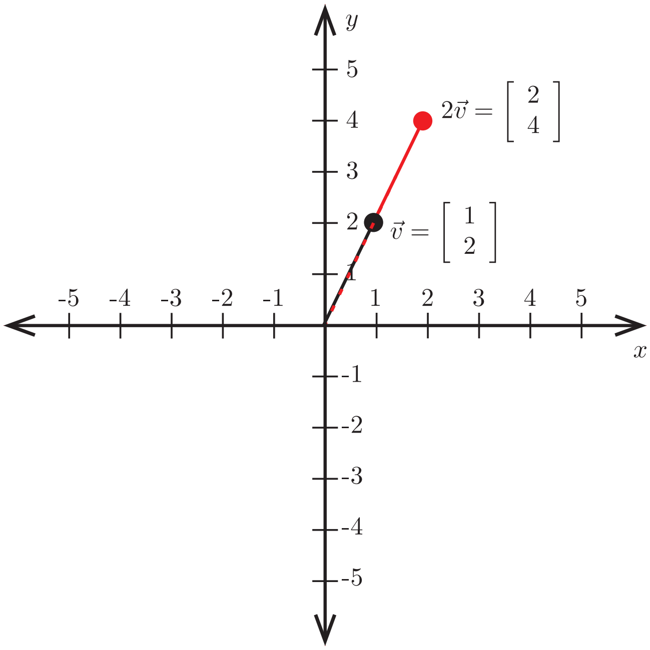

Figure 2.3: Scaling a vector by a positive number doesn't change its direction.

then the scalar multiplication of $\vec{v}$ and c gives:

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

In other words, for vectors of any dimensions:

$$\boxed{(c\vec{v})_j = cv_j}$$

EXAMPLE 2.2:

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

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

Observation 2.1.10: Scalar multiplication of a vector by a positive real number doesn't change its orientation, only its length. See Figure 2.3 above.

Observation 2.1.11: Scalar multiplication of a vector by a negative number inverts its direction. See Figure 2.4 on page 34.

So far, we've mainly 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's going on might be outside of human perception!

Trivia fact

We saw that adding two vectors together was very intuitive when we think about vectors describing position on the cartesian plane. Multiplying a vector by a number is also intuitive. But what does it mean to multiply a vector by another vector? There's not much intuition about it. But again, we're doing math, so if something doesn't exist, we invent it! There's a mathematical operation out there called the **cross product**, also known as **vector product**, denoted $\vec{v} \times \vec{w}$ or $\vec{v} \wedge \vec{w}$ which is only defined for 3-dimensional vectors. It's an operation that takes two vectors as inputs and outputs a third one that is perpendicular to the plane form by the input vectors. It sounds a little arbitrary, but it turns out that the cross product is extremely useful in physics, engineering, and of course, mathematics.

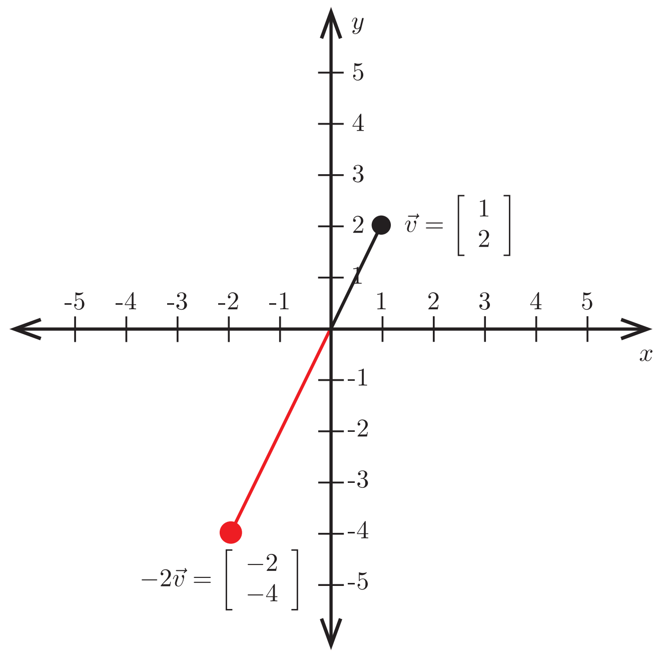

Figure 2.4: Scaling a vector by a negative number inverts its direction.

Note 2.1.12

Soon, we'll learn an operation called the **inner product**, also known as the scalar product or dot product. It's a type of multiplication of two vectors, but the result is a scalar number. You'll see that this concept is very useful, especially in quantum mechanics.

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

1. $\vec{v}+\vec{w}=\vec{w}+\vec{v}$ (commutativity)
2. $\vec{v}+(\vec{w}+\vec{u})=(\vec{v}+\vec{w})+\vec{u}$ (associativity)
3. $c(\vec{v}+\vec{w})=c\vec{v}+c\vec{w}$ (distributivity of scalar multiplication)
4. $(c+d)\vec{v}=c\vec{v}+d\vec{v}$ (distributivity of scalar addition)
5. There exists a unique additive zero, denoted $\vec{0}$ such that $\vec{v}+\vec{0}=\vec{v}$ for any vector $\vec{v}$.
6. For any vector $\vec{v}$, there exists an additive inverse $-\vec{v}$ such that $\vec{v}+(-\vec{v})=\vec{0}$.

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

The last two properties might seem obvious. We added them because we're just about to define the notion of **vector space**, which doesn't only apply to vectors, but potentially any other king of mathematical objects such as functions.

DEFINITION 2.1.14: Vector space. The less abstract definition: The collection of all the complex vectors of a given dimension with vector addition and scalar multiplication, is called a **vector space**. If we use the set of all n-dimensional vectors, we can call it a n-dimensional vector space.

The abstract definition: Take a collection of mathematical objects (known as a set) with a well-defined addition and scalar multiplication. If:

1. The set is closed under addition and scalar multiplication, that's the result of adding two arbitrary objects from the set, or the scalar multiplication of any objects, is also in the set;
2. The set, the addition and scalar multiplication follow all the properties listed in Properties 2.1.13 above.

We call such a set a **vector space**. Not using this definition, vector space can be of an infinite dimension (but this is beyond the scope of this book).

Food for thought

Is the set of all polynomials of degree 2, that is any function of the form $P(x) = ax^2 + bx + c$, using usual arithmetic addition and scalar multiplication, a vector space? Explain why.

Trivia fact

You may have heard before that Einstein developed the idea that we don't live in a 3-dimensional world, but in a 4-dimensional world (time being the 4th dimension.) Turns out that in the Theory of Relativity, time must be treated on the same footing as space (and vice versa) and we need to use a 4-dimensional vector space to describe how measurements of space and time by two observers are related.

2.2 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 you'll see, vectors and matrices play very well with each other.

DEFINITION 2.2.1: 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_{ij} as the number in the i^{th} row and the j^{th} column. If a matrix has m rows and n columns, the convention is to say that 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 2.2.2

A vector can also be thought of as a $n \times 1$ matrix.

EXAMPLE 2.3: Referring to the matrix N and P we've explicitly written above, $N_{12} = 3$ and $P_{31} = -3$. Of course, P_{23} is not defined as P only has 2 columns.

At first glance, there doesn't seem to be much intuition about matrices. As we'll see soon enough, a matrix can be thought of as a mathematical object "acting on vectors". So bear with us for a couple pages so we can introduce some concepts, and then we'll build some intuition.

DEFINITION 2.2.3: Matrix addition and scalar multiplication. Similar to vector addition, adding two matrices, as long as both matrices have the same dimensions, consists of adding each corresponding matrix element. Similarly, scalar multiplication is simply the scaling of all the elements of the matrix. Addition is only defined for matrices of the same dimensions. In other words, given matrices M,N and scalar c, we define:

$$\boxed{(M+N)_{ij} = M_{ij} + N_{ij}}$$

$$\boxed{(cM)_{ij} = c(M_{ij})}$$

EXAMPLE 2.4:

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

As for scalar multiplication:

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

We discussed that vector multiplication was not very well defined, but how about matrix multiplication? Turns out we can define a way to multiply matrices that's suitable. It'll sound a little bit arbitrary at first, but we'll then work through some examples. We won't necessarily try to explain where the definition comes from, but we'll justify that it leads to things that make sense!

We'll first start by giving the rigorous definition of matrix multiplication and then give a visual, step-by-step example. First: Look at the definition. Try to make some sense of it. Look at the example. Then come back to the definition and try to fully understand what's going on.

DEFINITION 2.2.4: 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. Let's assume that the number of columns in M and the number of rows in N is n. The i,j component of MN,

$(MN)_{ij}$ is given by:

$$\begin{aligned}(MN)_{ij} &= \sum_{k=1}^{n} M_{ik}N_{kj}\\ &= M_{i1}N_{1j} + M_{i2}N_{2j} + \ldots + M_{in}N_{nj}\end{aligned}$$

Observation 2.2.5: Suppose M is an $m \times n$ matrix and N is an $n \times l$ matrix, that is i only goes from 1 to m, k from 1 to n and j from 1 to l then from the definition above, NM will give you a $m \times l$ matrix.

EXAMPLE 2.5: Although the definition of matrix multiplication can be a little confusing, an explicit multiplication might clarify this simple task. We'll multiply the two following matrices:

$$M = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}, \quad N = \begin{bmatrix} g & h \\ j & k \\ l & q \end{bmatrix}$$

1. Since A is a 2×3 matrix and B is a 3×2 matrix, we expect the result of the multiplication to be a 2×2 matrix, that is:

$$\begin{bmatrix} \text{a} & b & \text{c} \\ d & e & f \end{bmatrix} \begin{bmatrix} \text{g} & h \\ j & k \\ 1 & q \end{bmatrix} = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}$$

2. Find the $(MN)_{11}$ element of the resulting matrix the following way:

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} g & h \\ j & k \\ l & q \end{bmatrix} = \begin{bmatrix} ag+bj+cl & \quad \\ \quad & \quad \end{bmatrix}$$

Explicitly, we've done:

$$\begin{aligned}(MN)_{11} &= M_{11}N_{11} + M_{12}N_{21} + M_{13}N_{31}\\ &= ag+bj+cl\end{aligned}$$

3. Similarly, the $(MN)_{12}$ element is:

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

4. And so on:

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} g & h \\ j & k \\ l & q \end{bmatrix} = \left[\begin{array}{c|c} ag+bj+cl & ah+bk+cq \\ \hline dg+ej+fl & \end{array}\right]$$

5. And so forth:

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

Food for thought

Can we perform the multiplication NM? Explain why.

EXAMPLE 2.6:

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

$$= \begin{bmatrix} 3i & 38-2i \\ 3 & -23 \end{bmatrix}$$

Observation 2.2.7: Matrices as functions on vectors. Since a vector is also a matrix that has only one column, we can think of matrices as **functions on vectors**, e.g., $m \times n$ matrix maps an n-dimensional vector to an m-dimensional vector. For example, given the vector and matrix:

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

Since P is a 3×2 matrix and $\vec{v}$ is a 2×1 vector (matrix) we have that P maps $\vec{v}$ to:

$$P\vec{v} = \begin{bmatrix} 2 & -1 \\ -1+2i & 3 \\ -3 & -3+i \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

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

$$= \begin{bmatrix} 3 \\ 1+4i \\ -9+i \end{bmatrix}$$

Note 2.2.6

When we multiply a vector by a matrix, we often refer to this as the "matrix acting on the vector", or "applying the matrix to the vector" and also as "the matrix operating on the vector". The latter explains why we often refer to matrices as **operators**.

which is a 3×1 vector (matrix), as expected. 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's a function between an n-dimensional vector space to an m-dimensional vector space.

EXAMPLE 2.7: Consider:

$$\vec{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad \text{and} \quad M = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The matrix M acts as:

$$\begin{aligned} M\vec{v} &= M\begin{bmatrix} 3 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ -2 \end{bmatrix} \end{aligned}$$

If you visualize $\vec{v}$ as a vector in the cartesian plane (see Figure 2.5 below), the matrix M performs a reflection of the x-axis of the cartesian plane.

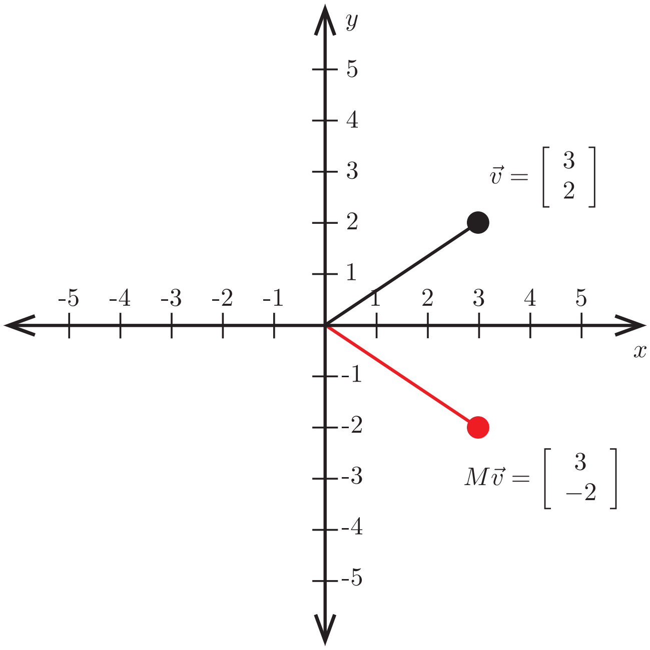

Figure 2.5: The matrix M as defined in Example 2.7 acting on a vector in the cartesian plane performs a reflection about the x-axis.

EXAMPLE 2.8: Similarly, consider:

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

The matrix N acts as:

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

The matrix N performs a reflection about the y-axis of the cartesian plane!

» Food for thought

Can you think of a 2×2 matrix that would represent a reflection of the axis making a $45°$ angle with the x-axis?

EXAMPLE 2.9: Given:

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

You can't perform the multiplication $M\vec{v}$ since the matrix is a 2×2 and the vector is 3×1. This multiplication is undefined.

Observation 2.2.8: Non-commutativity of matrix multiplication. An interesting property of matrix multiplication is its **non-commutativity**. In mathematics, we say an operation is commutative if the order of operations 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. Let's look at the following example:

» Food for thought

Why does investigating commutativity of matrix multiplication only make sense when we consider two square matrices of the same dimensions?

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

Multiplying in two different orders will give:

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

$$NM = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 13 & 9 \\ 5 & 5 \end{bmatrix}$$

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

Observation 2.2.9: Composition of functions. Often in mathematics, we need to do a **composition of function**, that is, applying a function g on an input, then apply 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's defined as:

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

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 2.10: Given the function f represented by matrix M and function g by matrix N, where:

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

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 $f(g(\vec{v}))$:

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

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 matrix 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's crucial to understand that matrix composition goes from right to left, that is, the rightmost matrix is the one being applied first.

Now we're in a position to verify whether or not matrix multiplication as previously defined makes sense from an intuitive perspective. We investigate one example. Of course, that's not enough to prove that the definition is rigorous, but at least it'll tell us whether we're on the right path.

EXAMPLE 2.11: Recall Example 2.7 and 2.8 on page 39 where we've showed that matrices:

$$M = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad N = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

represent the reflection of any vector $\vec{v}$ in the cartesian plane about the x- and y-axis respectively. Let's use the vector:

$$\vec{v} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

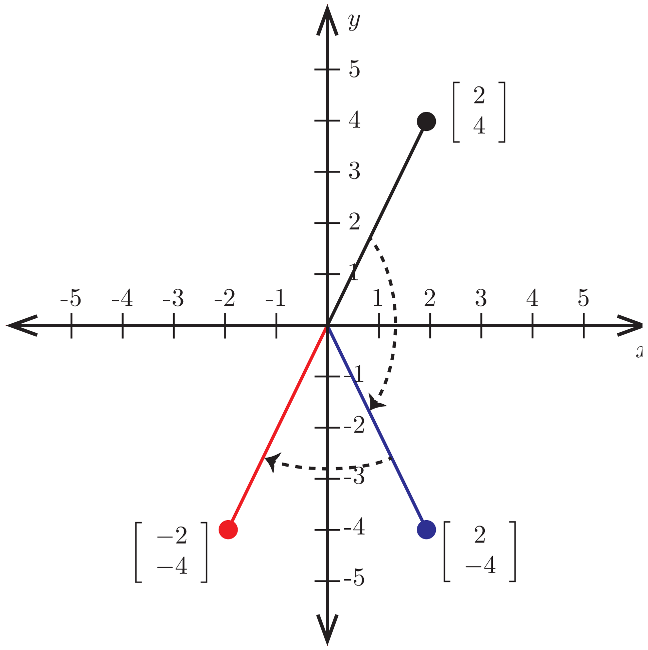

Figure 2.6: Performing a reflection about the x-axis and then the y-axis.

If we refer to Figure 2.6, you can convince yourself that performing a reflection about the x-axis and then about y-axis will give you:

$$-\vec{v} = \begin{bmatrix} -2 \\ -4 \end{bmatrix}$$

Let's see now if the math holds up. As we did for the composition of function, there are two ways of calculating the final vector:

1. Perform the first reflection, then the second, that is:

$$\begin{aligned} M\vec{v} &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ -4 \end{bmatrix} \\ N \begin{bmatrix} 2 \\ -4 \end{bmatrix} &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -4 \end{bmatrix} \\ &= \begin{bmatrix} -2 \\ -4 \end{bmatrix} \\ &= -\vec{v} \end{aligned}$$

2. Find the composition of the functions, NM (recall, since we do the x reflection first, M has to be the rightmost matrix), then apply it to $\vec{v}$, that is:

$$NM = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$
$$= \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$
$$\implies \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \end{bmatrix}$$

Both methods yield the same results, so our definition of matrix multiplication is on the right path.

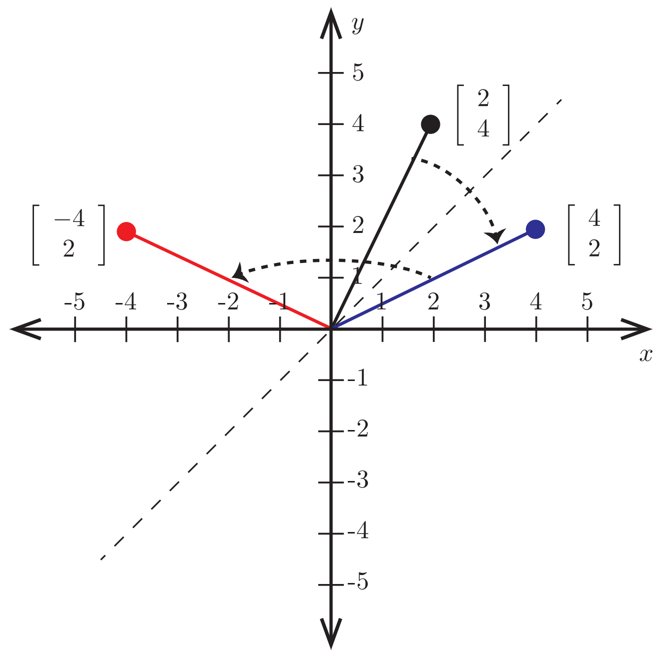

Figure 2.7: Performing a reflection about the 45°-axis and then the y-axis.

EXAMPLE 2.12: Let's kick it up a notch. A few moments ago you (hopefully) figured out that a reflection about the axis making a 45° angle with the x-axis is given by:

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

We'll do the same exercise as above, but we'll first do a reflection about this axis, then a reflection about the y axis. Let's use our trusted vector $\vec{v}$ as defined above again. Refer to Figure 2.7 above to see the geometric approach. After these two operations, we expect the final vector $\vec{w}$ to be:

$$\vec{w} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

Let's check if the math holds up again:

1. Perform the first reflection about the 45° axis, then about the y axis. We get:

$$\begin{aligned} P\vec{v} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} && \text{(reflection about the 45° axis)} \\ &= \begin{bmatrix} 4 \\ 2 \end{bmatrix} \\ N\begin{bmatrix} 4 \\ 2 \end{bmatrix} &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \end{bmatrix} && \text{(reflection about the } y \text{ axis)} \\ &= \begin{bmatrix} -4 \\ 2 \end{bmatrix} \end{aligned}$$

2. Find the composition of the function NP, then apply it to $\vec{v}$, that is:

$$\begin{aligned} NM &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\ &\implies \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix} \end{aligned}$$

There you go. Our definition of matrix multiplication seems to be holding up quite well!

Food for thought

In this case, what would have happened if we did the matrix multiplication in the wrong order?

Properties 2.2.10: Properties of matrix arithmetic. For the sake of listing properties of matrix addition and multiplication, we'll assume the dimensions of 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. $MN \neq NM$, in most cases

Food for thought

Do all the $m \times n$ matrices form a vector space if we use matrix addition and scalar multiplication?

To finish this section, we'd like to highlight a very important, yet very simple, mathematical concept that applies to vectors and matrices which makes them so useful and practical. This is the concept of linearity.

DEFINITION 2.2.11: **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 into words, the first condition means that the output of a function acting on a sum of inputs is 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. Linearity is found everywhere, as the example below shows.

EXAMPLE 2.13: Imagine that we charged you $100/day to attend QCSYS (but we don't, because we're nice that way!) 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 2 extra 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 to use the concept of linearity. Let's put some mathematical rigour into this. Let f be the function giving 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's easy, from the definition of linearity, to show that f is indeed a linear function, e.g.,

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

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

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 2.14: 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's straightforward to show that matrices can be thought of as linear functions, e.g.,

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

Hence the term "linear" algebra!

EXAMPLE 2.15: The function $f(x) = x^2$, for x being any scalar real or complex, isn't 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:

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

Food for thought

Is the function $f(x) = 3x+10$ linear? Why?

Why are we teaching you about linearity you may ask? Turns out that quantum mechanics is driven by linear processes and as we'll see in the next chapter, it will give rise to, let's just say, surprising results!

2.3 Complex conjugate, transpose and conjugate transpose

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

DEFINITION 2.3.1: **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:

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

The general definition, for vectors and matrices of any dimensions, would be:

$$\boxed{\left(\overline{\vec{v}}\right)_i = \overline{v_i}, \quad \left(\overline{M}\right)_{ij} = \overline{M_{ij}}}$$

EXAMPLE 2.16:

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

DEFINITION 2.3.2: **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 that the transpose of an $m \times n$ matrix is an $n \times m$ matrix. It follows that 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:

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

The general definition, for vectors and matrices of any dimensions, would be:

$$\boxed{\left(\vec{v}^t\right)_i = v_i, \quad \left(M^t\right)_{ij} = M_{ji}}$$

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

EXAMPLE 2.17: Given the following vector and matrix:

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

then:

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

In even more detail, here's the step-by-step solution to find M^t:

1. Since M is a 3×2 matrix, then M^t must be 2×3:

$$M = \begin{bmatrix} 2 & i \\ 0 & 0 \\ e^{i\frac{\pi}{3}} & 4i \end{bmatrix}, \quad M^t = \begin{bmatrix} \, & \, & \, \\ \, & \, & \, \end{bmatrix}$$

2. The first column of M becomes the first row of M^t:

$$M = \begin{bmatrix} 2 & i \\ 0 & 0 \\ e^{i\frac{\pi}{3}} & 4i \end{bmatrix}, \quad M^t = \begin{bmatrix} 2 & 0 & e^{i\frac{\pi}{3}} \\ \, & \, & \, \end{bmatrix}$$

3. The second column of M becomes the second row of M^t:

$$M = \begin{bmatrix} 2 & i \\ 0 & 0 \\ e^{i\frac{\pi}{3}} & 4i \end{bmatrix}, \quad M^t = \begin{bmatrix} 2 & 0 & e^{i\frac{\pi}{3}} \\ i & 0 & 4i \end{bmatrix}$$

DEFINITION 2.3.3: 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. That is:

$$\left(M^\dagger\right)_{ij} = \overline{M}_{ji}$$

The dagger subscript is usually preferred by physicists, while mathematicians will often use a superscripted *. It's been a long-standing debate between the two camps!

EXAMPLE 2.18: If:

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

then:

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

Note 2.3.4

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 2.3.5: Summary of properties. Following 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. $(M^\dagger)^\dagger = M$

2.4 Inner product and norms

In this section, we'll concentrate on defining some concepts that are mostly applicable for vectors. These concepts can be generalized to matrices, functions and beyond, but we'll restrict ourselves to the explicit definition applied to vectors.

In the previous section, we've briefly discussed the notation of multiplying vectors. As we'll soon see, the inner product is a very useful definition in linear algebra.

DEFINITION 2.4.1: Inner product. The **inner product**, also known as the dot product or the scalar product, of two vectors $\vec{v}$ and $\vec{w}$ is a mathematical operation between two vectors of the same dimension that returns a scalar number. It is denoted $\vec{v} \bullet \vec{w}$.

To take the inner product of two vectors, first take the complex conjugate of the first vector, then multiply each of the corresponding numbers in both vectors and then add everything. That is to say, if $\vec{v}$ and $\vec{w}$ are n-dimensional vectors,

$$\boxed{\vec{v} \bullet \vec{w} = \sum_{j=1}^{n} \overline{v_j} w_j}$$

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:

$$\vec{v} \bullet \vec{w} = \overline{v_1} w_1 + \overline{v_2} w_2 + \ldots + \overline{v_n} w_n$$

Note 2.4.2

The inner product is only defined for vectors of the same dimension.

EXAMPLE 2.19: If we have:

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

then:

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

Observation 2.4.3: Inner product as matrix multiplication. Taking the inner product of $\vec{v}$ and $\vec{w}$ is the same as doing a matrix multiplication between $\vec{v}^\dagger$ and $\vec{w}$:

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

In other words:

$$\boxed{\vec{v} \bullet \vec{w} = \vec{v}^\dagger \vec{w}}$$

Properties 2.4.4: Properties of the inner product. Let $\vec{u}$, $\vec{v}$ and $\vec{w}$ be vectors of the same dimension, a and b be scalars and M a matrix of suitable dimensions. The inner product has the following properties (feel free to prove these 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 2.20: 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:

$$\begin{aligned} \overline{\vec{w} \bullet \vec{v}} &= \overline{\vec{w} \bullet \vec{v}}^t \\ &= (\vec{w} \bullet \vec{v})^\dagger && \text{(by definition of the conjugate transpose)} \\ &= \left(\vec{w}^\dagger \vec{v}\right)^\dagger && \text{(recall Observation 2.4.3)} \\ &= \vec{v}^\dagger \left(\vec{w}^\dagger\right)^\dagger && \text{(recall Property 2.3.5-3)} \end{aligned}$$

$$= \vec{v}^\dagger \vec{w}$$
$$= \vec{v} \bullet \vec{w}$$

EXAMPLE 2.21: Let's now prove the above Property 8:

$$\begin{aligned}
\vec{v} \bullet (M\vec{w}) &= \vec{v}^\dagger (M\vec{w}) && \text{(recall Observation 2.4.3)}\\
&= (\vec{v}^\dagger M)\vec{w}\\
&= \left(M^\dagger \vec{v}\right)^\dagger \vec{w} && \text{(Property 2.3.5-3))}\\
&= (M^\dagger \vec{v}) \bullet \vec{w}
\end{aligned}$$

DEFINITION 2.4.5: Hilbert space. Recall the definition of vector space on page 35. A vector space with a well-defined inner product is called a **Hilbert space**. Therefore, the collection of all n-dimensional vectors with the inner product defined above form a Hilbert space.

Now that we've defined the inner product, we can, just as any mathematician likes to do, start expanding our bag of definitions.

DEFINITION 2.4.6: Orthogonal vectors. We say that two vectors are **orthogonal**, or perpendicular, if their inner product is 0.

EXAMPLE 2.22: The vectors:

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

are orthogonal because:

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

Observation 2.4.7: Angle between orthogonal vectors. 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:

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

are orthogonal to $\vec{v}$. See Figure 2.8 below for visualization. Looking at it from a mathematical approach, we have that:

$$\begin{aligned} \vec{v} \bullet \vec{w} &= \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} -4 \\ 2 \end{bmatrix} \\ &= (-1) \cdot 4 + 2 \cdot 2 \\ &= 0 \end{aligned}$$

$$\begin{aligned} \vec{v} \bullet \vec{u} &= \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \\ &= 1 \cdot 2 + 2 \cdot (-1) \\ &= 0 \end{aligned}$$

From visualizing vectors as part of the cartesian plane, the concept of length is rather intuitive. That is, some vectors are longer than others. Let's define that concept properly:

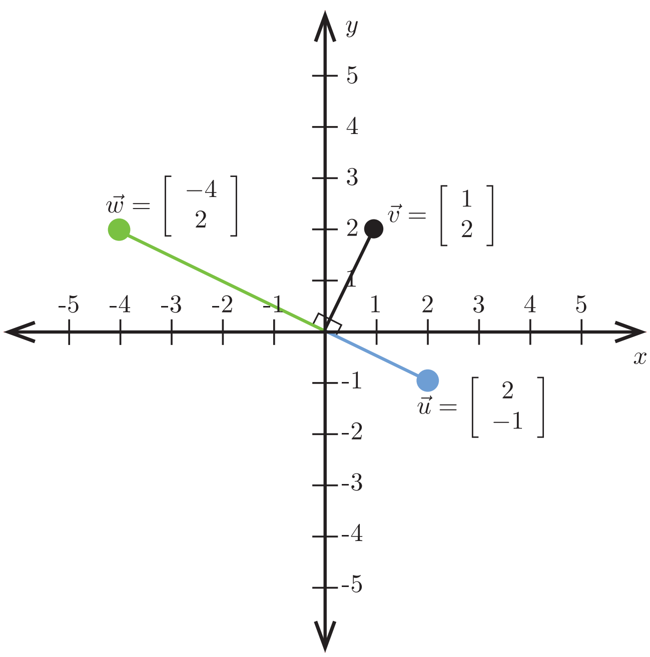

Figure 2.8: 2-dimensional visualization of orthogonal vectors.

DEFINITION 2.4.8: Vector Norm. The **norm** (or length) of a n-dimensional 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:

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

EXAMPLE 2.23: 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 2.24: 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 (i.e., it represents the distance from the origin to that point).

DEFINITION 2.4.9: 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 unit length. This is readily seen since:

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

EXAMPLE 2.25: 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:

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

Trivia fact

If you feel up to the task, you can prove that for any n-dimensional vector $\vec{v}$ and $\vec{w}$, the following inequality is always true:

$$|\vec{v} \bullet \vec{w}| \leq ||\vec{v}|| \cdot ||\vec{w}||$$

This result is known as the *Cauchy-Schwartz inequality* and is considered to be one of the most important inequalities in all of mathematics. Not only has this result had a huge impact on mathematics, it turns out that the famous Heisenberg uncertainty principle, $\Delta x \Delta p \leq \frac{\hbar}{2}$, a very physical phenomena, is derived from the Cauchy-Schwartz inequality.

2.5 Basis

A very important concept in linear algebra is that of **basis** (bases in plural). 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 dimension 2 and v_1 and v_2 are the **coefficients** 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 2.26: 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's an infinite number of them! As any decent mathematician would (and let's face it, by now we're more than decent!), we'll put some mathematical rigour into it, but first, let's introduce a couple of new concepts.

Note 2.5.2

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

DEFINITION 2.5.1: Linear combination. A **linear combination** is a combination of any number of vectors using vector addition and scalar multiplication. For example, if we use the n-dimensional vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ and scalars $c_1, c_2, \ldots, c_k$, a linear combination looks like:

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k$$

DEFINITION 2.5.3: Linearly dependent/independent vectors. We say that a set of vectors is **linearly dependent** if at least one of the vectors can be written as a linear combination of the others. Otherwise they're **linearly independent**.

EXAMPLE 2.27: 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:

$$\vec{v}_3 = 2\vec{v}_1 + \vec{v}_2$$

Food for thought

There is a rigorous proof for the statement in the Note above. Can you think of how we would write this proof?

DEFINITION 2.5.4: 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 2.28: 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're linearly dependent.

EXAMPLE 2.29: 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 they're linearly independent.

Observation 2.5.5: Basis vectors to describe any vectors. Let $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ be a basis for n-dimensional vectors. By the very definition of a basis, we can argue that any n-dimensional vector $\vec{w}$ can be written as a linear combination of the basis vectors

$\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. 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 (recall Note 2.5.2 on page 56). 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:

$$\vec{w} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$$

Since $\vec{w}$ was any arbitrary vector, we conclude that $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ can be used to describe any other n-dimensional vector by using appropriate coefficients in a linear combination.

EXAMPLE 2.30: 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'll 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'll need to solve a simple system of two equations and two unknowns:

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

Performing $2 \times (\triangle) - (\square)$ gives us $(2-i)c_1 = 5 - i$, which implies that

$$\begin{aligned} c_1 &= \frac{5-i}{2-i} = \frac{5-i}{2-i} \cdot \frac{2+i}{2+i} \\ &= \frac{11}{5} + \frac{3i}{5} \quad \text{(recall how to divide by complex numbers)} \end{aligned}$$

By reinserting our result for c_1 into $(\triangle)$, 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 2.5.6: Action of matrices on basis vectors. 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:

$$\begin{aligned} M\vec{w} &= M(c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n) \\ &= c_1 M\vec{v}_1 + c_2 M\vec{v}_2 + \cdots + c_n M\vec{v}_n \end{aligned}$$

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

EXAMPLE 2.31: If we're given:

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

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 2.5.7: Orthonormal basis. We say a basis is **orthonormal** if each vector has norm 1 and each pair of vectors are orthogonal.

EXAMPLE 2.32: The 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'll show up everywhere in quantum mechanics!

DEFINITION 2.5.8: Standard (canonical) basis. When we explicitly write a vector, we're implicitly using the **standard basis** (also known as the canonical basis), e.g.,

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

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

Observation 2.5.9: Change of basis. Given an arbitrary vector $\vec{x}$:

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

As we just explained, using x_1 and x_2 to explicitly write $\vec{x}$ as a column vector is the same as saying that $\vec{x}$ was written as a linear combination of the standard basis using coefficients x_1 and x_2. How do we write this vector as a linear combination of different basis vectors, say $\{\vec{w}_1, \vec{w}_2\}$ as defined above? This is referred to as a **change of basis**. There's actually a systematic way of doing this using a **change of basis matrix**, It turns out that for small dimensions, it's easier and much quicker 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'll just need to use substitution to get the job done. In our case, it's easy to find that:

$$\begin{bmatrix}1\\0\end{bmatrix} = \frac{1}{\sqrt{2}}(\vec{w}_1 + \vec{w}_2) \quad \text{and} \quad \begin{bmatrix}0\\1\end{bmatrix} = \frac{1}{\sqrt{2}}(\vec{w}_1 - \vec{w}_2)$$

By basic substitution, we get:

$$\begin{aligned}\begin{bmatrix}x_1\\x_2\end{bmatrix} &= x_1\begin{bmatrix}1\\0\end{bmatrix} + x_2\begin{bmatrix}0\\1\end{bmatrix}\\ &= \frac{x_1}{\sqrt{2}}(\vec{w}_1 + \vec{w}_2) + \frac{x_2}{\sqrt{2}}(\vec{w}_1 - \vec{w}_2)\end{aligned}$$

$$= \left(\frac{x_1+x_2}{\sqrt{2}}\right)\vec{w}_1 + \left(\frac{x_1-x_2}{\sqrt{2}}\right)\vec{w}_2$$
$$= \frac{1}{\sqrt{2}}\begin{bmatrix} x_1+x_2 \\ x_1-x_2 \end{bmatrix}_{\{w_1,w_2\}}$$

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 2.5.10: Reconstructing a matrix. Imagine your friend just found the most amazing, perfect $n \times n$ matrix that unlocks the secrets of the universe, but your friend refuses to show it to you (what a friend!). On the other hand, they're willing to give you a little taste: they agree to apply the matrix for n, and only n, input vectors of your choice and give you the corresponding outputs. Turns out, if you choose the input vectors 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 vector of a given basis, it's 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 to any dimension is easy. Starting from the most general matrix:

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

we can observe that:

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

Since we can ask our friend about the output for each standard basis, we can therefore deduce all the elements of M. Just to make things even more obvious, note that the first column of M is the vector that we'll get if we apply M to the first standard basis vector and the second column of M is the vector that we'll get if we apply M to the second standard basis. This observation is valid for any $n \times n$ matrix.

EXAMPLE 2.33: 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

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

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

Similarly:

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

We conclude that:

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

2.6 Inner product as projection

We already saw that the inner product can be used to calculate the norm of a vector, but it also has a lot of other applications, one of which is to perform the **projection** of a vector onto another vector.

EXAMPLE 2.34: Referring to Figure 2.1 on page 30, the projection of $\vec{v}$ along x is 1 and along y, 2.

Observation 2.6.2: Inner product vs. projection. Referring to Figure 2.1 again, notice that:

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

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

Note 2.6.1

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

The idea of projection can be generalized for any vector, in any dimension. We'll work our way to the formal definition step-by-step. For now, let's stick to the cartesian planes as they offer a nice visual. 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 2.9 below for a visual. You can convince yourself that $\vec{v}$ makes an angle of $\frac{\pi}{3}$ (or 60°) with the x-axis and $\vec{w}$ makes an angle of $-\frac{\pi}{4}$ (−45°).

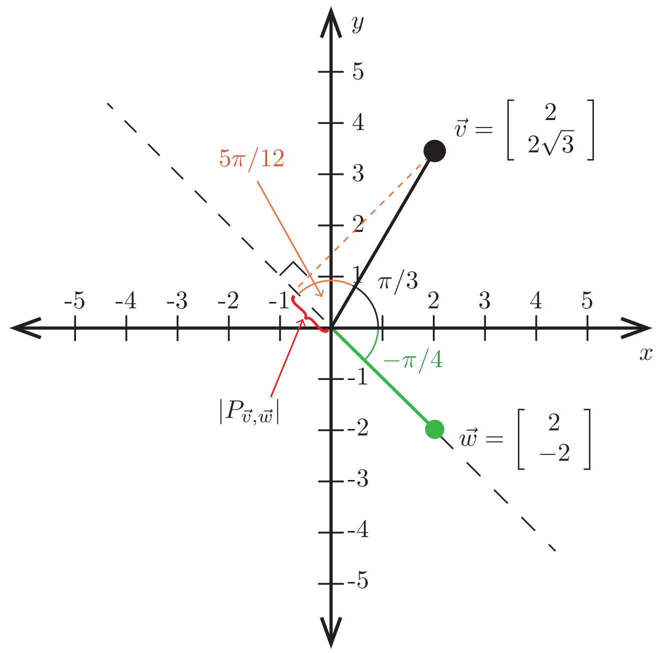

Figure 2.9: The projection of $\vec{v}$ onto $\vec{w}$ is given by the component of $\vec{v}$ along $\vec{w}$.

DEFINITION 2.6.3: 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 $\vec{w}$. Since we define it as the component along $\vec{w}$, $P_{\vec{v},\vec{w}}$ is a **scalar number**.

Here's an easy trick to find that component: place a ruler perpendicular to the direction of $\vec{w}$ (that direction is given by the long, dashed black line in Figure 2.9 above) 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're interested in is actually along the opposite direction of $\vec{w}$ – that is the direction of $-\vec{w}$. $P_{\vec{v},\vec{w}}$ is therefore a negative number.

Again referring to Figure 2.9 above, basic trigonometry tells us that:

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

$$\approx \sqrt{16} \cdot 0.259$$
$$\approx 1.035$$
$$\implies P_{\vec{v},\vec{w}} \approx -1.035$$

Observation 2.6.4: Inner product vs. projection revisited. The vector $\vec{w}$ above is parallel to the unit vector:

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

$\vec{u}$ can be obtained simply by normalizing $\vec{w}$ (see Definition 2.4.9 on page 54). Do you think it's a coincidence that:

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

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 a basis. To see this, let's suppose that $\vec{u}_\perp$ is any unit vector that is perpendicular to $\vec{u}$, and hence perpendicular to $\vec{w}$. We've seen in the previous section that the set $\{\vec{u}, \vec{u}_\perp\}$ forms an 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 $\vec{v}$. We therefore have:

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

Food for thought

Can you find an explicit representation for $\vec{u}_\perp$?

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

DEFINITION 2.6.5: Orthogonal projection revisited. 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:

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

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

EXAMPLE 2.35: In the 3-dimensional 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:

$$P_{\vec{v},\vec{w}} = \frac{1}{\sqrt{6}} \begin{bmatrix} i \\ 2 \\ -i \end{bmatrix} \bullet \begin{bmatrix} 1 \\ e^{-i\frac{\pi}{4}} \\ 2 \end{bmatrix}, \text{ since } ||\vec{w}|| = \sqrt{6}$$

$$= \frac{1}{\sqrt{6}} \begin{bmatrix} -i & 2 & i \end{bmatrix} \begin{bmatrix} 1 \\ e^{-i\frac{\pi}{4}} \\ 2 \end{bmatrix}$$

(do not forget the complex conjugate of the first vector)

$$= \frac{1}{\sqrt{6}}(-i + 2e^{-i\frac{\pi}{4}} + 2i)$$

$$= \frac{1}{\sqrt{6}}(i + 2\cos\frac{\pi}{4} - 2i\sin\frac{\pi}{4})$$

(remembering Euler's formula)

$$= \frac{1}{\sqrt{6}}\left(\sqrt{2} + i(1-\sqrt{2})\right)$$

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, u_1 is the unit vector pointing in the direction of $\vec{w}$, then 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 2.6.6: Writing a vector using projection. In previous sections, we rewrote a given vector as a linear combination of other vectors a few times. In Example 2.27 and Observation 2.5.9 on pages 56 and 59 respectively, we found the coefficients of the combination by inspection, while in Example 2.30 on page 57, we used a more systematic, yet cumbersome method.

The inner product now gives us a systematic and simple method to write any n-dimensional vector $\vec{v}$ as a linear combination of vectors belonging to an orthonormal basis $\{\vec{u}_1, \ldots, \vec{u}_n\}$. As we've seen many times already, we can always find coefficient $c_1, \ldots, c_n$ such that:

$$\vec{v} = c_1\vec{u}_1 + c_2\vec{u}_2 + \ldots + c_n\vec{u}_n$$

If we perform the inner product of $\vec{u}_j$ with $\vec{v}$, we have:

$$\begin{aligned}\vec{u}_j \bullet \vec{v} &= c_1\vec{u}_j \bullet \vec{u}_1 + c_2\vec{u}_j \bullet \vec{u}_2 + \ldots + c_j\vec{u}_j \bullet \vec{u}_j + \ldots + c_n\vec{u}_j \bullet \vec{u}_n \\ &= c_1 \cdot 0 + c_2 \cdot 0 + \ldots + c_j \cdot 1 + \ldots + \ldots c_n \cdot 0 \\ &\quad (\text{since } \vec{u}_k \bullet \vec{u}_j = 0 \text{ except if } k = j) \\ &= c_j\end{aligned}$$

If we perform this inner product for each basis vector, we can easily find all the coefficients of the linear combinations, and we conclude that:

$$\vec{v} = (\vec{u}_1 \bullet \vec{v})\vec{u}_1 + (\vec{u}_2 \bullet \vec{v})\vec{u}_2 + \ldots + (\vec{u}_n \bullet \vec{v})\vec{u}_n$$

since $c_j = \vec{u}_j \bullet \vec{v}$

EXAMPLE 2.36: We'll 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 $\mathbb{C}^2$. From there, we just need to crunch the numbers:

$$\begin{aligned}\vec{u}_1 \bullet \vec{v} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \end{bmatrix} \begin{bmatrix} 3+i \\ 1-i \end{bmatrix} \\ &\quad \text{(again, don't forget the complex conjugate of the first vector)} \\ &= \sqrt{2} \\ \vec{u}_2 \bullet \vec{v} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \end{bmatrix} \begin{bmatrix} 3+i \\ 1-i \end{bmatrix} \\ &= 2\sqrt{2} + \sqrt{2}i\end{aligned}$$

We therefore conclude that:

$$\vec{v} = \sqrt{2}\vec{u}_1 + \left(2\sqrt{2} + \sqrt{2}i\right)\vec{u}_2$$

2.7 Special matrices

Now that we know pretty much everything we need to know about vectors, let's go back to matrices for a moment. A lot of "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 2.7.1: Identity matrix. The $n \times n$ **identity matrix**, often denoted by $\mathbb{1}$ and sometimes by I, is defined such that for every $n \times n$ matrix M, and any vector $\vec{v}$ in $\mathbb{C}^n$, we have:

$$\mathbb{1}M = M\mathbb{1} = M \quad \underline{\text{and}} \quad \mathbb{1}\vec{v} = \vec{v}$$

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

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

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

Note 2.7.2

The identity is only defined for square matrices.

EXAMPLE 2.37: In 3 dimensions, the 3×3 matrix with only 1s on the diagonal is the identity, since:

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

and:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$$

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$$

Food for thought

Can you explain the note above, given the definition of identity?

DEFINITION 2.7.3: Unitary matrices. A **unitary matrix**, U, is a matrix that satisfies:

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

EXAMPLE 2.38: The following matrices, which will soon become your best friends, 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 2.7.4: Unitary matrices preserve the inner product. 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:

$$\vec{w}_1 = U\vec{v}_1 \quad \text{and} \quad \vec{w}_2 = U\vec{v}_2$$

then:

$$\begin{aligned} \vec{w}_1 \bullet \vec{w}_2 &= (U\vec{v}_1) \bullet (U\vec{v}_2) \\ &= (U\vec{v}_1)^\dagger U\vec{v}_2 \\ &= (\vec{v}_1^{\,\dagger} U^\dagger)(U\vec{v}_2) \qquad \text{(recalling that } (MN)^\dagger = N^\dagger M^\dagger\text{)} \\ &= \vec{v}_1^{\,\dagger}(U^\dagger U)\vec{v}_2 \\ &= \vec{v}_1^{\,\dagger}\mathbb{1}\vec{v}_2 \\ &= \vec{v}_1^{\,\dagger}\vec{v}_2 \\ &= \vec{v}_1 \bullet \vec{v}_2 \end{aligned}$$

Note 2.7.5

From Observation 2.7.4, unitary matrices also preserve the length (or norm) of vectors.

$$\begin{aligned} \|U\vec{v}\| &= \sqrt{(U\vec{v}) \bullet (U\vec{v})} \\ &= \sqrt{\vec{v} \bullet \vec{v}} \\ &= \|\vec{v}\| \end{aligned}$$

Observation 2.7.6: Rotation matrix. The matrix R defined above represents a **rotation** in $\mathbb{R}^2$. Notice:

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

Since both basis vectors are rotated by an angle θ, then any vector in $\mathbb{R}^2$ will be rotated by θ. See Figure 2.7 on page 68 for visuals.

Trivia fact

Unitary matrices are very useful in physics and many other fields as they represent **reversible** processes. If you think of the unitary matrix U as operating on a vector $\vec{v}$, then $U^\dagger$ undoes what U did (since $U^\dagger U = \mathbb{1}$). We'll learn soon enough that quantum mechanical processes are reversible, and hence unitary matrices will be very important.

2.8 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've mostly used concrete examples that we were already familiar with (cartesian and spacial coordinates). Just to show you how versatile matrices and

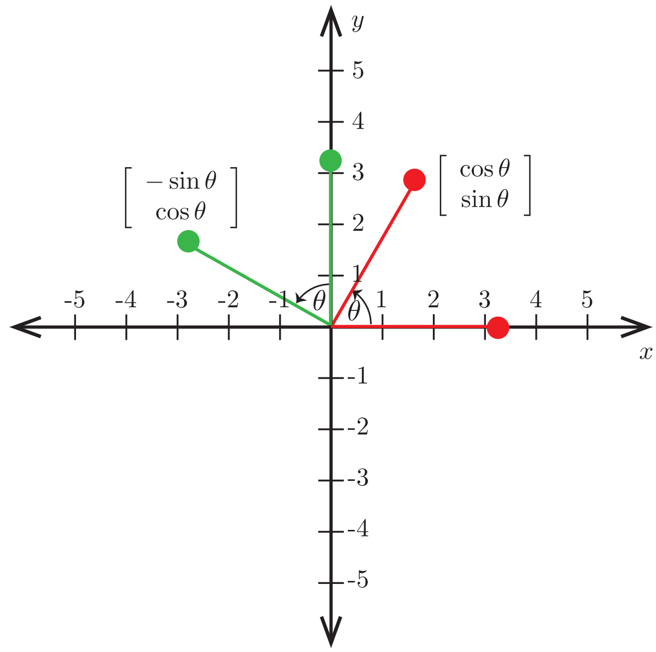

Figure 2.10: The matrix R, as defined in Observation 2.7.6 on page 67, rotates any vectors by an angle θ.

vectors can be, we'll write a small cookbook using only numbers in a matrix to show you an example of applied linear algebra.

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

	Sugar (cup)	**Flour** (cup)	**Salt** (tsp)	**Milk** (cup)	**Water** (cup)
Pizza	0	2	1	0	1
Cake	1	1	0	0.5	0
Bagel	1	1	2	0.25	0.25

Obviously, we must stress the fact that that these are not real recipes and you should not try this at home. They 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 recipes do you want to make?". The output, a list of ingredients.

Each dough is independent of each other, so we can assign each dough to a basis vector. Since there are three of them, we'll need at least, and no more, than three basis vectors. A 3-dimensional vector space will work just fine and we'll assign:

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

Since each ingredient is independent of each other (you wouldn't use salt to replace water, would you?), we'll also assign each ingredient a basis vector. Since there are five ingredients, we'll 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 recipes, we know the actions of the cookbook 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, if you remember Observation 2.5.10 on page 60, we can give the explicit form of M:

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

This is our cookbook!

Suppose, you want to do two pizzas, half a cake and two bagels, here's what you'd 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?!

Chapter 2 Summary

Below we summarize the most important concepts in linear algebra as a quick reference. Let $\vec{v}$, $\vec{w}$, M and N be vectors and matrices of suitable dimensions and c a scalar:

Vector addition:

$$(\vec{v}+\vec{w})_j = v_j + w_j$$

Vector scalar multiplication:

$$(c\vec{v})_j = cv_j$$

Matrix addition:

$$(M+N)_{ij} = M_{ij} + N_{ij}$$

Matrix scalar multiplication:

$$(cM)_{ij} = c(M_{ij})$$

Matrix multiplication:

$$(MN)_{ij} = \sum_k M_{ik}N_{kj}$$

Matrix/vector complex conjugate:

$$\left(\overline{M}\right)_{ij} = \overline{M_{ij}}$$

Matrix/vector transpose:

$$\left(M^t\right)_{ij} = M_{ji}$$

Matrix/vector conjugate transposed:

$$\left(M^\dagger\right)_{ij} = \overline{M_{ji}}$$

Inner/Scalar/Dot product:

$$\vec{v} \bullet \vec{w} = \vec{v}^\dagger \vec{w} = \sum_{j=1}^{n} \overline{v_j} w_j$$

Norm of a vector:

$$||\vec{v}|| = \sqrt{\vec{v} \bullet \vec{v}}$$

Projection of $\vec{v}$ onto $\vec{w}$:

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

Unitary matrices:

Any matrices U such that $U^\dagger U = UU^\dagger = \mathbb{1}$

Chapter 3: Quantum mechanics

"Quantum physics thus reveals a basic oneness of the universe."
ERWIN SCHRÖDINGER

"The "paradox" is only a conflict between reality and your feeling of what reality ought to be."
RICHARD FEYNMAN

Now that we've learned and mastered the beautiful concepts of complex numbers and linear algebra, it's time to explore how it applies to quantum mechanics. For now, we'll take a purely mathematical approach, but never forget that:

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

During QCSYS, we'll also teach you quantum mechanics from a qualitative, phenomenological approach. You'll even do a few experiments to investigate some of the quintessential quantum phenomena, such as the wave-particle duality and the Heisenberg uncertainty principle. By the end of these lectures, we'll reconcile both the mathematical and the phenomenological approach. But for now, we'll introduce how the different concepts of linear algebra we just learned will be used to quantify quantum mechanics.

3.1 Mathematical postulates of quantum mechanics

As we mentioned in the preface of this book, 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. We're just stating them here and will go into detail about each of them and explain the new terminology used. These postulates can take different equivalent forms depending on who you talk to, but for the sake of QCSYS, we'll work with the postulates below:

Postulates of Quantum Mechanics

1. The state (or wavefunction) of individual quantum systems is described by unit vectors living in separate complex Hilbert spaces. (Recall the definition of Hilbert space on page 52.)
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. This is known as **Born's rule**. Immediately after the measurement, the wavefunction **collapses** into that state.
3. Quantum operations are represented by unitary operators on the Hilbert space (a 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 observables are represented by the eigenvalues of a Hermitian operator on the Hilbert Space.

For the sake of simplicity, we won't cover the last postulate as it's not really needed during QCSYS. To do so, we'd need to introduce a few more mathematical concepts and let's be honest, we've learned quite enough so far! If you're curious about it, we'll be more than happy to discuss it with you during the mentoring sessions.

Also, note that in the first part of this book, we've 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, e.g., 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 very well suited to quantum information and quantum cryptography.

But what about when we want to describe physical quantities that have continuous values, such as the position of a particle along a line? In this case, we need a vector space of infinite and continuous dimension. Turns out that it's possible to define a Hilbert space on the set of continuous functions, e.g., $f(x) = x^3 + 2x + 1$. This is referred to as "wave mechanics" and we won't cover it in this book. However, we'll 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.

3.2 New notation: the braket notation

Before getting into the the heart of the mathematics of quantum mechanics, we'll introduce a new notation widely used in physics. In Chapter 2, we used the vector and matrix notations that are widely used in mathematics, but when physicists play with vectors to describe a quantum system, they like to change it around a little and use what we call the **braket notation**. Note that everything we've defined so far, including vector addition, matrix multiplication, inner product, projections, is exactly the same. The only thing that changes is how we write down the variable.

DEFINITION 3.2.1: The "ket". When using a vector $\vec{v}$ to represent a quantum state, we'll use a different notation known as "**ket**", written $|v\rangle$ ("ket v"). This is a notation commonly used in quantum mechanics and doesn't change the nature of the vectors at all. That is, both notations below are equivalent:

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

DEFINITION 3.2.2: 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.

DEFINITION 3.2.3: 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 braket of $|v\rangle$ and $|w\rangle$.

Note 3.2.4

The braket notation follows naturally from the geometry of our notation, e.g.,

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

Observation 3.2.5: Complex conjugate of a braket. If you recall one of the properties of the inner product, namely, $\vec{v} \bullet \vec{w} = \overline{\vec{w} \bullet \vec{v}}$, you can readily see that:

$$\overline{\langle w|v\rangle} = \langle v|w\rangle$$

EXAMPLE 3.1: Let:

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

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, $\langle w|v\rangle = \overline{\langle v|w\rangle}$

Trivia fact

The braket notation was introduce by Paul Dirac in 1939 and is also known as the **Dirac notation**. Dirac is notoriously known in history for having an odd and peculiar personality. One story: when Dirac first met the young Richard Feynman, he said after a long silence, "I have an equation. Do you have one too?"

3.3 Single quantum state and the qubit

Now that the new notation is out of the way, we'll now look in detail at Postulate 1 of quantum mechanics as stated at the beginning of this chapter:

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

First and foremost, we need to explain what we actually mean by **state**.

DEFINITION 3.3.1: Quantum states. The collection of all relevant physical properties of a quantum system (e.g., position, momentum, spin, polarization) is known as the **state** of the system.

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

DEFINITION 3.3.2: Exclusive states. When modelling the state of a given physical quantity (position, spin, polarization), two states are said to be exclusive if the fact of being in one of the states with certainty implies that there are no chances whatsoever of being in any of the other states.

EXAMPLE 3.2: 1 particle, 3 boxes. Imagine this hypothetical situation: we have a single quantum particle, and three quantum boxes. The whole system behaves quantum mechanically. If I know for certain that the particle is in box 1, then it's certainly not in box 2 or 3.

EXAMPLE 3.3: Moving particle. Imagine a particle that can move horizontally, vertically and diagonally. These three states are definitely not exclusive since moving diagonally can be thought of as moving horizontally and vertically at the same time.

If you recall the concept of orthogonal vectors and projection (see Definitions 2.4.6 and 2.6.3 on pages 52 and 62 respectively), it wasn't unlike the concept of exclusivity (i.e., a given vector has no component along any of its orthogonal vectors). Therefore it makes 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.

EXAMPLE 3.4: 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 words "vector" and "state" interchangeably. Quantum states also use the **braket notation** to emphasize they're quantum states.

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

But before we introduce the formal definition of a qubit, a quick note about the **bit**. In your cell phone, your computer, or pretty much any digital electronic device 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. Physically, a bit is implemented by a transistor in a processor, or a tiny magnet in your hard drive. In your computer, each bit is either in the state 0, or in the state 1.

Just like the bit is the basic unit of classical information, the qubit is the basic unit of quantum information.

DEFINITION 3.3.3: Qubit. In quantum information, we're 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'll see during QCSYS, 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 $|1\rangle$ are conventionally 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 exclusive 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 an atom 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 states are mutually exclusive, we could represent:

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

Note 3.3.4

For the rest of this book, we'll always assume that the quantum system we speak of is a qubit, i.e., it has only two exclusive states. But bear in mind that everything we'll discuss extends to quantum systems of any dimensions in a straightforward manner.

During QCSYS, we'll see how the states of qubits are used to perform quantum cryptography. For this particular case, we like to use the polarization of a photon as our qubit. By the end of the QCSYS, you'll be experts at this!

Observation 3.3.5: Both states at once. Here's where the fun starts. Let's define the state $|+\rangle$ and $|-\rangle$ with the vectors:

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

If we look a little closer, we see that:

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

The above 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 (kind of)! Is that mathematical trickery, or can we really do that in a lab? We can certainly do that in a lab, and you'll experience it firsthand! Welcome to the world of quantum mechanics!

DEFINITION 3.3.6: Quantum superposition principle. If a quantum system can be in the state $|0\rangle$, and can also be in $|1\rangle$, then quantum mechanics allows the system to be in any arbitrary state

$$\begin{aligned}|\psi\rangle &= a|0\rangle + b|1\rangle \\ &= \begin{bmatrix} a \\ b \end{bmatrix}\end{aligned}$$

We say that $|\psi\rangle$ is in a **superposition** of $|0\rangle$ and $|1\rangle$ with **probability amplitudes** a and b. We'll see in the next section why it's called an amplitude.

EXAMPLE 3.5: The state $|0\rangle$ is a superposition of $|+\rangle$ and $|-\rangle$ since:

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

3.4 Quantum measurement

Now that we know how to represent the state of a given quantum system, we'll investigate Postulate 2 of quantum mechanics:

> **Postulate 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 (Born's rule). Immediately after the measurement, the wavefunction collapses into that state.

As we just saw, it's 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 wonder: if I actually perform a measurement to see if the system is in $|0\rangle$ or in $|1\rangle$, what will I measure? You'll measure either $|0\rangle$ or $|1\rangle$, but which one? With what probability?

DEFINITION 3.4.1: Born's Rule. 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

Note 3.3.7

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 states. Refer to Appendix A.1 for the complete list of Greek letters.

Trivia fact

Even if Erwin Schrödinger was one of the fathers of the mathematical formulation of quantum mechanics, the superposition principle left him baffled. To show the absurdity of such a concept, he applied it to an everyday object and asked: can a cat be both dead and alive? Hence was born the famous Schrödinger's Cat.

orthonormal basis, i.e., we "ask" the quantum system which one of these states it's in. The probability of measuring the state $|\phi_i\rangle$, $P(\phi_i)$, is given by:

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

DEFINITION 3.4.2: Wave collapse. After the measurement is performed the original state collapses in the measured state, i.e., we're left with one of the states $|\phi_1\rangle, \ldots, |\phi_n\rangle$.

DEFINITION 3.4.3: Quantum measurement. From a mathematical point of view, applying Born's rule and then the wave collapse correspond to a **quantum measurement**.

EXAMPLE 3.6: Suppose we have $|\psi\rangle = |+\rangle$ (as defined in Observation 3.3.5 on page 76) and we measure it in the orthonormal basis $\{|\phi_0\rangle = |0\rangle, |\phi_1\rangle = |1\rangle\}$. Then the state would collapse 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 3.7: 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 3.8: What if we measure the system in the basis $\{|+\rangle, |-\rangle\}$? Note that:

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

Note 3.4.4

If you recall the definition of the superposition principle on page 77, a and b were called **probability amplitudes**. The reason for such a name is that they don't quite represent a probability, but their modulus squared does give us the probability.

which implies that the outcome will be the state:

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

Observation 3.4.6: Probabilism. In the last two examples, we saw that measuring in two different bases yields probabilistic measure. You can actually show that the only time we'll get a deterministic result is when we measure in a basis that includes the state $|\psi\rangle$. But since $|\psi\rangle$ was an arbitrary (unknown) state to begin with, we conclude measuring an unknown quantum state will always yield a random result.

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

EXAMPLE 3.9: To relate Example 3.6 on page 78 to physics, think about the following scenario: 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 76.) 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 50% probability of measuring the atom in its ground state, and a 50% probability of 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.

DEFINITION 3.4.7: Quantum measurement (alternative). Suppose we have a quantum state $|\psi\rangle$ and an orthonormal basis $\{|\phi_1\rangle, \dots, |\phi_n\rangle\}$. We can explicitly write the state using the orthonormal basis, i.e.,

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

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 3.10: Let's show that our two definitions of quantum measurements are equivalent. Starting with the fact that we can always write any state $|\psi\rangle$ using the orthonormal basis $\{|\phi_1\rangle, \dots, |\phi_n\rangle\}$, i.e.,

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

Note 3.4.5

In Example 3.8 above, we used the explicit representation of the vector to perform the inner product, but we could have done everything in braket notation, e.g.,

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

Food for thought

In Observation 3.4.6, we said that "the only time we'll get a deterministic result is when we measure in a basis that includes the state $|\psi\rangle$". Can you explain why?

we can then apply Born's rule and get that the probability of measuring the state $|\phi_1\rangle$ is given by:

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

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

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

which shows that both our definitions are equivalent.

Trivia fact

Born's rule was formulated in 1926 by German physicist Max Born. Although incredibly successful at predicting experimental results, we had to wait until 2010 for a group of experimentalists (led by Raymond Laflamme, Thomas Jennewein and Gregor Weihs from the Institute for Quantum Computing. Go IQC!) to directly put it to the test. They showed that Born's rule is accurate to at least 1%. Since then, further experiments have shown the accuracy to be within a whopping 0.001%.

EXAMPLE 3.11: Given the quantum state:

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

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

1. First method:

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

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

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

therefore:

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

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

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 3.4.8: Requirement for unit norm. 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 give us something, then the sum of the probabilities must be 1. Hence the need for a unit vector.

3.5 Quantum operations

So far, we've 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 3.5.1: Quantum operations. A **quantum operation** transforms a quantum state to another quantum state, therefore, we must have it so that the norm of the vector is preserved (recall, quantum states must have norm 1). If you recall Note 2.7.5 on page 67, the mathematical representation of any quantum operation can therefore be represented by a unitary matrix. Similarly, any unitary matrix represents a possible quantum operation.

Refer back to Section 2.7 on page 66 for the properties of unitary matrices.

EXAMPLE 3.12: 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 3.13: 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. $\mathbb{1}|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle$
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. $H|1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = |-\rangle$

EXAMPLE 3.14: The Hadamard matrix defined above is very useful in quantum information. Let's see how it acts on some on our favourite quantum states:

1. $H|0\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = |+\rangle$
2. $H|1\rangle = |-\rangle$, as seen above
3. $H|+\rangle = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle$
4. $H|-\rangle = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle$

You should memorize these four results as we'll use them over, and over, and over, and over again!

Observation 3.5.2: The power of symbolic notation. Sometimes, calculating the results of a quantum operation using only symbolic notation can provide us with some insight. Let's look at the third calculation above:

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

Notice the notations highlighted above in colour. The signs in front on both $|0\rangle$ are the same, while the signs in front of both $|1\rangle$ are opposite. This is an example of **quantum interference**. $|0\rangle$ undergoes **constructive** interference and $|1\rangle$ undergoes **destructive** interference. Interference is a property of waves and what you just saw is an example of the **wave-particle duality** inherent to quantum mechanics.

Observation 3.5.3: If one implements several quantum operations one after the other, 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:

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

Notice that the order of the multiplication goes from 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 $U_2U_1|\psi\rangle$, and so on, and so forth.

In the lab, quantum operations are performed in various ways, depending on the quantum system you're 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 an atom, we can use a laser pulse. There are many different types of qubits and many different ways to perform operations. During QCSYS, you'll have the opportunity to learn how we do it with photons in a hands-on experiment, but you'll also learn how we do it with other type of qubits.

Food for thought

We've already seen in Example 3.6 on page 78 that measuring the state $|+\rangle$, and similarly $|-\rangle$ in the $\{|0\rangle, |1\rangle\}$ basis leads to a 50% probability of measuring each state. Therefore, you can think of the Hadamard operation on the states $|0\rangle$ of $|1\rangle$ as the quantum version of a coin flip. Say you only have the ability to perform a measurement in the $\{|0\rangle, |1\rangle\}$ basis, but you can perform any quantum operation you desire. How is a quantum coin flip different than a classical coin flip?

3.6 Multiple quantum states

Playing around with one qubit is fun. As you'll see during QCSYS, manipulating and measuring one qubit at a time allows us to perform quantum cryptography. But what if we want to build a quantum computer? A one-qubit quantum computer is not really powerful, if powerful at all! In this section, we'll start playing with multiple qubits at once, i.e., a composite system. We'll need more fancy mathematics to talk about multiple quantum states, but nothing you can't handle!

> **Postulate 4:** The Hilbert space of a composite system is given by the tensor product (also known as the 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'll consider a composite system of two qubits, but the generalization to multiple quantum systems of different dimensions is straightforward.

Properties 3.6.2: Joint quantum systems must follow the following properties:

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 exclusive states each ($|0\rangle_2$ and $|1\rangle_2$), then the full system will have 4 distinct states namely:

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

   Note: We've added the extra subscript 2 and 4 to explicitly denote the fact that they're vectors of dimensions 2 and 4 respectively.
   Using the same argument as in Section 3.3 on page 74, it would make sense to explicitly have:

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

2. **Measurement probabilities**
   Now, we'll introduce some arguments about probability. Let's assume that 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 $|\phi_1\rangle$ and qubit 2 in $|\phi_2\rangle$, i.e., measuring the joint system in $|\Phi\rangle_4 = |\phi_1\phi_2\rangle_4 = |\phi_1\rangle_2 \otimes |\phi_2\rangle_2$ will be given by:

Note 3.6.1

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.

$$\begin{aligned} P(\Phi) &= |\langle\Phi|_4|\Psi\rangle_4|^2 \\ &= |\langle\phi_1\phi_2|_4|\psi_1\psi_2\rangle_4|^2 \\ &= \left|\left[\langle\phi_1|_2\otimes\langle\phi_2|_2\right]\left[|\phi_1\rangle_2\otimes|\phi_2\rangle_2\right]\right|^2 \end{aligned}$$

But, if we think of each qubit as their 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[1], then we must have that:

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

which essentially means that we must have:

$$\boxed{\left|\left[\langle\phi_1|_2\otimes\langle\phi_2|_2\right]\left[|\phi_1\rangle_2\otimes|\phi_2\rangle_2\right]\right|^2 = |\langle\phi_1|\psi_1\rangle|^2|\langle\phi_2|\psi_2\rangle|^2}$$

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 joint operation $U_1\otimes U_2$ must have the property that:

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

   In other words, the resulting joint state after applying the joint quantum operation, $\left[U_1\otimes U_2\right]|\psi_1\psi_2\rangle_4$ 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'd 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 3.6.3: Kronecker product for vectors. The **Kronecker product** (also referred as the tensor product) is a special way to join vectors together to make bigger vectors. Suppose we have the two vectors:

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

[1]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're not very good!), then since me eating a sandwich has nothing to do with the result of the baseball game, the probability of me eating a sandwich and the Jays to win the game is 1/40.

The Kronecker product is defined as:

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

From the definition of the Kronecker product, it's straightforward to show that Property 1 is satisfied. Property 2 is also easy to show, but a little more tedious. Feel free to do it yourself. We'll see an example in a few pages showing that it's satisfied.

EXAMPLE 3.15: 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:

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

EXAMPLE 3.16: Similarly, the Kronecker (tensor) product can readily be extended to vectors of any dimension. For example, the Kronecker product of the following two vectors:

$$\vec{v} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \quad \text{and} \quad \vec{w} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

is given by:

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

$$= \begin{bmatrix} 1 \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} \\ -1 \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} \\ 2 \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} 3 \\ 4 \\ -3 \\ -4 \\ 6 \\ 8 \end{bmatrix}$$

Notice that the final vector is 6-dimensional, as expected from Property 1.

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

DEFINITION 3.6.4: Multiple qubits. If we have two qubits with individual states $|\psi\rangle$ and $|\phi\rangle$, their joint quantum state $|\Psi\rangle$ is given by:

$$\boxed{|\Psi\rangle = |\psi\rangle \otimes |\phi\rangle}$$

where $\otimes$ represents 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. For example, if we have n qubits with individual states $|\psi_1\rangle, \ldots, |\psi_n\rangle$, the joint state will be given by $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots \otimes |\psi_n\rangle$. 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.

Trivia fact

As you can see from Definition 3.6.4, every time a qubit is added to a system, the dimension of the vector doubles. This grows really fast, e.g., the dimension of a 10-qubit system is $2^{10} = 1024$, that of a 20-qubit system is $2^{20} = 1048576$, and so on. This shows that quantum mechanics is extremely hard to simulate on a regular computer, as there isn't enough memory available. As a matter of fact, if we wanted to just write down the state of roughly 235 qubits, there would not be enough atoms in the universe to do so!

EXAMPLE 3.17: Using the explicit definition of the Kronecker product, we see that:

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

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

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

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

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

Before going any further, let's introduce a couple of friends we'll become very familiar with. Please meet Alice and Bob! Who are they? Well, often in quantum information, when we talk about two qubits, we often say the first qubit belongs to party A and the second belongs to B. But since we're friendly with our qubits, we gave them names: Alice and Bob! So let's take a look at a few examples with Alice and Bob.

EXAMPLE 3.18: If Alice has the quantum state $|\psi\rangle = |0\rangle$ and Bob has the state $|\phi\rangle = |+\rangle$ then their combined state is given by:

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

EXAMPLE 3.19: We're going to introduce a new qubit whom we'll call Charlie. Suppose Alice is in the state $|\psi\rangle = |+\rangle$, Bob in the state $|\phi\rangle = |+\rangle$ and their new pal, Charlie, is in $|\chi\rangle = |-\rangle$ then their combined state is given by:

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

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

EXAMPLE 3.20: If $|\Psi\rangle = |+1\rangle$ and $|\Phi\rangle = |1-\rangle$, we can calculate the outcome probability of measuring the state $|\Psi\rangle$ in the state $|\Phi\rangle$ in two ways. Let's first write those vectors explicitly:

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

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

1. We will compute directly using Born's rule:

$$\begin{aligned} P(\Phi) &= |\langle \Phi | \Psi \rangle|^2 \\ &= \left| \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right|^2 \\ &= \left| -\frac{1}{2} \right|^2 = \frac{1}{4} \end{aligned}$$

2. Using the inner product property of the Kronecker product (Property 2):

$$\begin{aligned} P(\Phi) &= |\langle +1 | 1- \rangle|^2 \\ &= |\langle + | 1 \rangle|^2 \cdot |\langle 1 | - \rangle|^2 \\ &= \left| \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right|^2 \cdot \left| \tfrac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right|^2 \\ &= \left| \frac{1}{\sqrt{2}} \right|^2 \cdot \left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{4} \end{aligned}$$

EXAMPLE 3.21: Similar example than the previous one, but this time suppose $|\Psi\rangle = |01\rangle$ and $|\Phi\rangle = |1-\rangle$. We can conclude right away that $P(\Phi) = 0$. Why? If we calculate the probability using the second method above, we don't have to explicitly carry out the inner product. Since:

$$P(\Phi) = |\langle 01|1-\rangle\rangle|^2 = |\langle 0|1\rangle|^2 \cdot |\langle 1|-\rangle|^2$$

and we already know that $\langle 1|0\rangle = 0$, then the joint probability must be 0.

Actually, the braket notation is particularly well suited to carry out inner products without having to explicitly do the calculations.

Observation 3.6.5: Interesting two-qubit states. 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:

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

These types of states (called entangled states) are very intriguing and play a fundamental role in quantum mechanics.

DEFINITION 3.6.6: Separable state. If a two-qubit state can be written as:

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

then we say that $|\Psi\rangle$ is a **separable state**. In any other case, we call them entangled states.

Trivia fact The concept of entanglement baffled Einstein so much that he questioned the validity and completeness of quantum mechanics. In a highly regarded article he published in 1935 with his colleagues Podolsky and Rosen, he presented a thought experiment which revealed consequences of quantum mechanics that seemed unreasonable and unphysical. This is known as the "EPR paradox". Sadly for Einstein, although the thought experiment was correct, his conclusions were wrong! Subsequent work by theorist John Bell in the 1960s and experiments by Alain Aspect in the 1980s showed that entanglement was indeed a real physical phenomena, hence shattering forever our concept of physical reality.

DEFINITION 3.6.7: Quantum entanglement. If a two-qubit quantum state $|\Psi\rangle$ cannot be written as $|\psi\rangle \otimes |\phi\rangle$ for any possible choice of $|\psi\rangle$ and $|\phi\rangle$, then $|\Psi\rangle$ is said to be **entangled**.

Entanglement is a fascinating property of quantum mechanics that's completely counter-intuitive. Physically, entangled qubits have a well-defined joint state, yet they do not have a well-defined individual state! It's widely accepted that the phenomenon of quantum entanglement is what makes quantum physics intrinsically different than classical physics. During QCSYS, we'll investigate some of these amazing properties and we'll even use entanglement to do quantum cryptography.

EXAMPLE 3.22: Are $|\Phi\rangle = \frac{1}{2}(|00\rangle + |01\rangle - |10\rangle - |11\rangle)$ and $|\Psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ separable quantum states?

1. Let's do the case of $|\Phi\rangle$ first. Assume that $|\Phi\rangle$ is separable; therefore there must 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 $|\Phi\rangle = |\psi\rangle|\phi\rangle$. Explicitly writing the vectorsgives us 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:

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \psi_1\phi_1 \\ \psi_1\phi_2 \\ \psi_2\phi_1 \\ \psi_2\phi_2 \end{bmatrix}$$

This tells us that $\psi_1\phi_2 = 0$ which implies 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 3.23: We could have done the example above without ever looking at the explicit representation of the state. The case of $|\Psi\rangle$, for example. We want to know if two states exist, $|\psi\rangle = \psi_1|0\rangle + \psi_2|1\rangle$ and $|\phi\rangle = \phi_1|0\rangle + \phi_2|1\rangle$, such that $|\psi\rangle \otimes |\phi\rangle = |\Psi\rangle$. Can we solve:

$$\begin{aligned} \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) &= (\psi_1|0\rangle + \psi_2|1\rangle) \otimes (|\phi_1|0\rangle + \phi_2|1\rangle\rangle) \\ &= \psi_1\phi_1|00\rangle + \psi_1\phi_2|01\rangle + \psi_2\phi_1|10\rangle + \psi_2\phi_2|11\rangle? \end{aligned}$$

This leads us to the same four equations to be satisfied, that is:

$$\begin{aligned} \psi_1\phi_1 &= \frac{1}{\sqrt{2}} \\ \psi_1\phi_2 &= 0 \\ \psi_2\phi_1 &= 0 \\ \psi_2\phi_2 &= \frac{1}{\sqrt{2}} \end{aligned}$$

EXAMPLE 3.24: Given the two-qubit 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 is originally in $|\Psi\rangle$?

You can verify for yourself that neither of these states are separable. Therefore, we won't be able to use the second method in Example 3.20 on page 89 directly. We can still evaluate the probability without explicitly writing down the vectors.

1. Let's do it explicitly first:

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

   If our quantum system is prepared in $|\Psi\rangle$, then Born's rule tells us that the probability of measuring it in the state $|\Phi\rangle$ is given, by $P(\Phi) = |\langle\Phi|\Psi\rangle|^2$:

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

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

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

   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 the blink of an eye, we conclude that:

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

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

DEFINITION 3.6.8: Bell's states. The following two-qubit states are known as the **Bell's states**. They represent an orthonormal, entangled basis for two qubits:

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

Now that we've covered how to represent the state of multiple qubits and how to predict the measurement probabilities, let's investigate how to represent a quantum operation on multiple qubits.

DEFINITION 3.6.9: 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 on page 85):

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

EXAMPLE 3.25: Recall the matrices defined in Example 3.12 on page 82. If we apply the unitary matrix $U = X \otimes Z$ to the state $|0+\rangle$, we get:

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

DEFINITION 3.6.10: Kronecker product for matrices. We can generalize the Kronecker product to matrices. 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:

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

$$= \begin{bmatrix} m_1n_1 & m_1n_2 & m_2n_1 & m_2n_2 \\ m_1n_3 & m_1n_4 & m_2n_3 & m_2n_4 \\ m_3n_1 & m_3n_2 & m_4n_1 & m_4n_2 \\ m_3n_3 & m_3n_4 & m_4n_3 & m_4n_4 \end{bmatrix}$$

EXAMPLE 3.26: Using the above definition for the Kronecker product of two matrices, we could find the same result by performing the calculation explicitly:

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

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

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

Applying this unitary operation on the state:

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

yields:

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

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$$

$$= |1-\rangle$$

Observation 3.6.11: 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 of two matries, 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.

EXAMPLE 3.27: Take the two-qubit quantum operation represented by the following matrix:

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

You can verify for yourself that U is a unitary matrix and therefore represents a valid quantum operation (recall Definition 3.5.1 on page 81). You can prove for yourself that no 2×2 matrices V_1 and V_2 such that $U = V_1 \otimes V_2$ exist.

Imagine the following situation: Start with two qubits in the joint state $|\psi_0\rangle = |00\rangle$. First, you apply a Hadamard operation on the first qubit (recall

the Hadamard matrix from Example 3.12 on page 82) and then you applied the operation U. What would be the final state? There are several ways of finding out the final state, so we'll investigate a few of them.

1. Explicit calculation: The Hadamard is a one-qubit operation, but we want to apply it to a two-qubit state, how is that done? First, you have to realize that applying the Hadamard to the first qubit only is the exact same thing as saying "apply the Hadamard to the first qubit and nothing to the second qubit". Recall from Definition 2.7.1 on page 66 that the identity matrix applied to any vectors always returns the same vector. That is to say, from a quantum operation perspective, the identity operation is just like doing nothing! Therefore, applying a Hadamard on only the first qubit is equivalent to the quantum operation on two qubits given by:

$$\begin{aligned} H \otimes \mathbb{1} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & 1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ 1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & -1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{bmatrix} \end{aligned}$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

Therefore, right after that operation, the qubits will be in the state:

$$\begin{aligned} |\psi_1\rangle &= (H \otimes \mathbb{1})\,|00\rangle \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

After applying the operation U, the final state will be given by:

$$\begin{aligned} |\psi_2\rangle &= U|\psi_1\rangle \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\ &= |\Phi^+\rangle \quad \text{(recalling the notation for Bell's states)} \end{aligned}$$

2. Explicit calculation revisited: Instead of calculating the state at each stage, we could instead calculate the matrix that represents both operations first, and then apply it to the initial state. From Observation 3.5.3 on page 83, recall that if we apply the operator U_1 followed by the operator U_2, the quantum operator describing the overall operation is represented by $U_{tot} = U_2 U_1$ (notice the order of the indices). In our case, $U_1 = H \otimes \mathbb{1}$ and $U_2 = U$, therefore:

$$U_{tot} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{bmatrix}$$

From there, we can directly calculate the final state:

$$\begin{aligned} |\psi_2\rangle &= U(H \otimes \mathbb{1})|00\rangle \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \end{aligned}$$

3. Explicit calculation revisited again: By using the properties of the Kronecker product, we wouldn't need to explicitly find the matrix form of $H \otimes \mathbb{1}$. Instead, we could have gone this way:

$$\begin{aligned} |\psi_1\rangle &= (H \otimes \mathbb{1})\,|00\rangle \\ &= H|0\rangle \otimes \mathbb{1}|0\rangle \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

Then we can explicitly apply the operator U to find the final state.

4. A bit more implicit: Remember the property of the Hadamard matrix we saw in Example 3.14 on page 82, namely that $H|0\rangle = |+\rangle$. Using this fact, we could have performed the above calculation as:

$$\begin{aligned} |\psi_1\rangle &= (H \otimes \mathbb{1})\,|00\rangle \\ &= H|0\rangle \otimes \mathbb{1}|0\rangle \\ &= |+\rangle|0\rangle \\ &= \frac{1}{\sqrt{2}}\,(|0\rangle + |1\rangle)\,|0\rangle \\ &= \frac{1}{\sqrt{2}}\,(|00\rangle + |10\rangle) \end{aligned}$$

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

5. Fully implicit: In the method above, we found the intermediate state without ever multiplying matrices and vectors. Can we implicitly apply the operator U in an implicit manner as well? Since we cannot write $U = V_1 \otimes V_2$ for any matrices V_1 and V_2, our first instinct would be to say no. But look carefully at U, specifically, look at how it operates on the computational, or standard basis states (look back at Observation 2.5.10 on page 60). You'll find that:

$$\begin{aligned} U|00\rangle &= |00\rangle \\ U|01\rangle &= |01\rangle \\ U|10\rangle &= |11\rangle \\ U|11\rangle &= |10\rangle \end{aligned}$$

Therefore, we could continue the calculation above as:

$$\begin{aligned} |\psi_2\rangle &= U|\psi_1\rangle \\ &= \frac{1}{\sqrt{2}} U\left(|00\rangle + |10\rangle\right) \\ &= \frac{1}{\sqrt{2}} \left(U|00\rangle + U|10\rangle\right) \\ &= \frac{1}{\sqrt{2}} \left(|00\rangle + |11\rangle\right) \end{aligned}$$

Food for thought

In quantum information, the matrix U as defined above is referred to as a *controlled-NOT*, or *CNOT*. Can you explain why?

Food for thought

In light of Example 3.27, suppose you'd like to measure a two-qubit state in the Bell basis, but unfortunately, you can only perform a measurement in the computational basis. How would you solve that conundrum?

3.7 Observables

The last and final postulate needed to complete quantum theory has to do with the concepts on **observables**:

> **Postulate 5:** Physical observables are represented by the eigenvalues of a Hermitian operator on the Hilbert space.

As a quick side note, a **hermitian operator** is any operator, or matrix, that's its own conjugate transpose.That is, an operator V such that $V^\dagger = V$.

As mentioned at the beginning of this section, we won't go into too much detail about the notion of observables as the details are a bit beyond the scope of this book. But don't worry, it won't affect our appreciation of quantum mechanics and quantum information.

By now, we're really good at calculating the probability of finding a quantum system in a certain state (Born's rule). But what does it mean to "measure to be in a given state?" In real life, we cannot measure states, but we can measure physical quantities like magnetic fields, position, polarization, momentum, spin, etc., in other words, we're looking for an observable. The fifth postulate of quantum mechanics says that for a given physically meaningful quantity with multiple possible values, there's a hermitian operator that associates the mutually exclusive states of the system with the possible value of the measurements.

EXAMPLE 3.28: Recall that we can build a qubit using the ground state $|g\rangle$ of an atom and its excited state $|e\rangle$. In principle, $|g\rangle$ and $|e\rangle$ can be represented using any orthonormal basis vectors, not necessarily the standard basis. If the ground state has energy (the observable) E_g and the excited state has energy E_e, we can define a hermitian **energy operator** V such that:

$$V|g\rangle = E_g|g\rangle$$
$$V|e\rangle = E_e|e\rangle$$

That is to say, the operator applied to the corresponding state returns the values of the observable times the state. Another side note, E_g and E_e are referred to as the **eigenvalue** of V with corresponding **eigenvectors**, or eigenstates, $|g\rangle$ and $|e\rangle$. (More on that some other time!)

Chapter 3 Summary

Find the key formulas and concepts we have learned in this section in the list below:

Quantum states:	Any unit vectors of appropriate dimension
Quantum bit/qubit:	A two-level, or two dimensional, quantum system
Born's rule:	$P(\phi) = \lvert\langle\phi\vert\psi\rangle\rvert^2$
State collapse:	If measured in $\lvert\phi\rangle$, the system collapses in $\lvert\phi\rangle$
Quantum operator:	Any unitary matrix of appropriate dimension
Multiple qubits:	$\lvert\psi_1\psi_2\rangle = \lvert\psi_1\rangle \otimes \lvert\psi_2\rangle$
Born's rule for multiple qubits:	$P(\Phi) = \lvert\langle\phi_1\phi_2\vert\psi_1\psi_2\rangle\rvert^2 = \lvert\langle\phi_1\vert\psi_1\rangle\rvert^2\lvert\langle\phi_2\vert\psi_2\rangle\rvert^2$
Quantum entanglement:	$\lvert\Psi\rangle \neq \lvert\phi\rangle \otimes \lvert\psi\rangle$ for any $\lvert\phi\rangle$ and $\lvert\psi\rangle$

Appendices

"The development of quantum mechanics early in the twentieth century obliged physicists to change radically the concepts they used to describe the world."
ALAIN ASPECT

"As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality."
ALBERT EINSTEIN

A.1 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

A.2 Properties of complex numbers: Proofs

In Section 1.2 starting on page 17, 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. $z\overline{z}=\overline{z}z=|z|^2$
6. $\overline{\overline{z}}=z$
7. $|z|=|\overline{z}|$
8. $|zw|=|z||w|$
9. $|z+w|\leq|z|+|w|$
10. $z^{-1}=\dfrac{1}{z}=\dfrac{\overline{z}}{|z|^2}$ when $z\neq 0+0i$

We strongly encourage you to prove these properties for yourself. If you have any difficulties, refer to the proofs below.

Just assume that $z=a+bi$ and $w=c+di$

Property A.1: $z+w=w+z$

Proof :

$$\begin{aligned}
z+w &= (a+bi)+(c+di)\\
&= a+bi+c+di\\
&= (c+di)+(a+bi)\\
&= (c+di)+(a+bi)\\
&= w+z
\end{aligned}$$

Property A.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 A.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:

$$\begin{aligned} wz &= (c+di)(a+bi) \\ &= ca + cbi + dai + bdi^2 \\ &= (ca - db) + (da + bc)i \\ &= (ac - bd) + (ad + bc)i \end{aligned}$$

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

Property A.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 A.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're done.

Property A.6: $\overline{\overline{z}} = z$

Proof :

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

Property A.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 A.8: $|zw| = |z||w|$

Proof :

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

Property A.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 or equal than the length, or modulus, of that number, 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:

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

Property A.10: $z^{-1} = \frac{1}{z} = \frac{\bar{z}}{|z|^2} \quad$ when $z \neq 0 + 0i$

Proof :

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

A.3 Euler number and exponential functions

Euler number e and the exponential functions e^x play 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 don't assume you know calculus, we won't go into the details of what the function is, but focus instead on how we can use it.

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

EXAMPLE A.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'll see an example below.)

Exponential functions are particularly well suited to model mathematical situations where 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 fixed amount). Take the function $f(x) = 2^x$ for example. Every time x increases by 1, the value of $f(x)$ doubles.

Properties A.1: 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 Properties 1 and 2)
6. $a^{-x} = (a^{-1})^x = \left(\frac{1}{a}\right)^x = \frac{1}{a^x}$

It's 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're familiar with the concept of the **slope of a curve**, then the slope of e^x is e^x. If you're familiar with calculus and derivatives, 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:

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

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

A.4 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're attracted to mathematical beauty, simplicity and logic. Therefore we're 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'll try to define our angle units in a way that relates to the circumference.

Definition A.1: Radians. Refer to Figure A.1. 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.

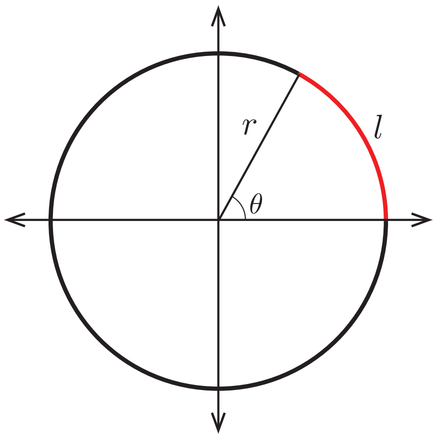

Figure A.1: An angle in radian can be defined as the ratio of the the length of the subtended arc l and the radius r.

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's defined as a ratio of two lengths, it's 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 radian equivalent, θ, using this simple proportionality relation:

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

The first term above is the fraction of the full circle the angle subtends and since we know that the full circle is 2π radian, than 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

A.5 Proof of Euler's theorem

Theorem A.1: Euler's formula. We'd 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:

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