Quantum Theory
ſſ
Quantum Computation

Seminar Notes

John Hutchinson

March 28, 2019 version

PRELIMINARY & INCOMPLETE

Introduction

Quantum Theory and Quantum Computation

The main goal is to develop an understanding of Quantum Theory and its underlying mathematical models, particularly as it applies to quantum computation.

In Quantum Mechanics (i.e. Quantum Theory in general), we consider properties such as position and momentum which take an infinite set of values. As we will indicate later, this is modelled mathematically by considering infinite dimensional Hilbert spaces and linear operators on such spaces. This leads to many (interesting and difficult) functional analysis issues.

In Quantum Theory for quantum computing we need only consider finite dimensional Hilbert Spaces, usually the n-dimensional inner product spaces $\mathbb{C}^n$.[1] The functional analysis difficulties are not present.

However, all the quantum weirdness *is* present in this setting. This includes superposition and entanglement, which imply the no-cloning theorem, superdense coding, teleportation, and non-locality of quantum effects (including the famous double-slit experiment, the Einstein–Podolsky–Rosen [EPR] paradox, and the results of John Bell showing quantum theory cannot be explained by additional information in the classical sense). We will see all this fairly soon!

In quantum computation, quantum "weirdness" is the essential resource utilised and exploited. We will see how this is done.

To reiterate, quantum theory in the finite dimensional setting is the optimal way to develop an intuition for quantum theory in general.

Mathematics

> *To understand quantum computing, you absolutely must become fluent in the mathematical model.* Michael Nielsen [MN19, First Lesson, preceding "Part I: The state of a qubit"]

> *... it is impossible to explain honestly the beauties of the laws of nature in a way that people can feel, without their having some deep understanding of mathematics. I am sorry, but this seems to be the case.* Richard Feynman [Fey, pp 39,40]

> ***DON'T PANIC*** Douglas Adams. The Hitchhiker's Guide to the Galaxy.

The background necessary to follow these notes is just some fairly basic linear algebra regarding finite dimensional complex vector spaces, and basic properties of unitary and hermitian operators. This is provided in Appendices A and B.

Structure of the Notes

In the first chapter we discuss the fundamental ideas of quantum theory, the mathematical models used, and their highly nonintuitive consequences. The discussion is fairly informal – if something is unclear it will hopefully be clarified in subsequent chapters.

[1] $\mathbb{C}^n$ is the vector space of n-tuples $a = (a_1, \ldots, a_n)$ with inner product $a \cdot b = a_1^* b_1 + \cdots + a_n^* b_n$ (where a_i^* is the complex conjugate of a_i). Note that the inner product is conjugate linear in the *first* argument and linear in the second. This convention is very useful when we later use the Dirac "bra-ket" notation, see Appendix D.

In Chapter 2 we discuss the four quantum theory postulates for closed systems, the relevant mathematical framework, and from there derive the nonintuitive consequences discussed in Chapter 1.

In Chapters 3 we discuss quantum computation, quantum circuits, non-cloning, superdense coding, quantum teleportation.

In Chapter 4 we discuss further mathematical structures used in the mathematical models and their consequences.

I use a significant number of footnotes. They provide additional information or more precise comments on definitions and discussions.

In the Appendices we develop the necessary mathematics. *I suggest you just go to the relevant Appendices as necessary.* Either there will be pointers in the main part of the text, or it should be clear from context where to go. This way you will get to very interesting quantum material very quickly!

The Future for Quantum Computation

[dW19] is an insightful essay concerning the potential impact of quantum computers.
[Pre18] discusses quantum computing in the near future.

Contents

Chapter 1

Overview

1.1 Qubits

1.1.1 Bits and Qubits

Bits: physical representations, states & mathematical model

A (classical) *bit*[1] is the basic unit of information in computing, information theory, digital communications. A bit can take one of two values, typically 0 or 1 as in computing, but also true/false, yes/no, $+/-$, etc. These values are also called the possible *states* of the bit.

A bit has various physical realisations such as a bit in computer hardware or software, the presence or absence of a hole in punched computer cards of old, a switch being on or off, an electric current having one of two distinct voltages (high or low), etc.

Mathematically, a bit is modelled by the set $\{0, 1\}$ containing two distinct elements, and the state of the bit is then either 0 or 1. Often we abuse language and

- refer to the "bit 0" or the "bit 1", rather than the state of the bit being 0 or 1,
- use the word "bit" both for the mathematical model and for various physical realisations.

Qubits: states, mathematical model & physical realisations

A *quantum bit* or *qubit*[2] is the basic unit in quantum computing. It is the quantum generalisation of a (classical binary) bit. It can take as its value/state any linear *superposition* of the form

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \in \mathbb{C}^2, \quad \text{where } |\alpha|^2 + |\beta|^2 = 1,\ \alpha, \beta \in \mathbb{C}, \tag{1.1}$$

where $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Note that $\{|0\rangle, |1\rangle\}$ is the standard/computational (orthonormal) basis for $\mathbb{C}^2$. We can think of $|0\rangle$ and $|1\rangle$ as the two values 0 and 1 of a classical bit.[3]

The coeffecients α and β are called *amplitudes*.

The notation $|\psi\rangle$ is due to Dirac. Dirac notation is discussed in Appendix D

Thus a qubit is a linear superposition of classical bits and is mathematically modelled by points on the unit sphere in $\mathbb{C}^2$.[4]

To gain an intuitive understanding it is often sufficient to take $\alpha, \beta \in \mathbb{R}$ in (1.1). In this case $-1 \leq \alpha, \beta \leq 1$ with $\alpha^2 + \beta^2 = 1$.

Treating qubits in a mathematical/abstract manner enables us to develop quantum theory in a manner independent of any particular physical realisation

[1] "Bit" is a contraction of "binary information digit".

[2] "Qubit" — pronounced "q(ueue)bit" — is a contraction of "quantum bit".

[3] This is consistent with the ket notation, where $|v\rangle$ is a vector with label v.

[4] This is not quite correct. Two vectors $|\phi\rangle$ and $|\psi\rangle$ on the unit sphere are said to be *equivalent* if $|\phi\rangle = e^{i\theta}|\psi\rangle$ for some $\theta \in \mathbb{R}$. Qubits correspond to *equivalence classes* of points on the unit sphere. In particular, $|\phi\rangle$ and $-|\phi\rangle$ denote the same qubit. We also sometimes refer to a non-zero vector $|v\rangle \in \mathbb{C}^2$ as a qubit. In this case we intend the normalised vector $\frac{|v\rangle}{\||v\rangle\|}$.

Every two-state physical quantum system is a qubit. We will discuss three such realisations: the two lowest energy levels of an electron, linear polarisation of photons, and spin 1/2 particles. Quantum computers are/will be built from these and other physical realisations!

As previously for bits, we abuse language and
- refer to the "qubit $|\psi\rangle$", rather than the state of the qubit being $|\psi\rangle$,
- use the word "qubit" both for the mathematical model and for various physical realisations.

1.1.2 State Space, Physical Reality and Measuring Qubits

Quantum Theory is truly weird, at least to our "classical" way of thinking.

1. Quantum states are "real". For any unit vector $|\psi\rangle \in \mathbb{C}^2$ (*state space* for qubits) we can *prepare* multiple qubits in the state $|\psi\rangle$. Moreover, we can physically manipulate/transform these qubits into other qubits in a manner which corresponds to applying a prescribed unitary operator in $\mathbb{C}^2$.

2. Given a classical bit, we can read/observe its state. This happens in a (classical) computer.

   But for qubits this is not the case, the situation is completely different.

   A measurement process for qubits corresponds to an orthonormal basis in $\mathbb{C}^2$, also known as the *measurement basis*.[5]

   (a) If this measurement basis is $\{|0\rangle, |1\rangle\}$ and we measure one of the prepared qubits in the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, then it will give the value 0 with probability $|\alpha|^2$ and the value 1 with probability $|\beta|^2$. Moreover, after measurement the state will change/"collapse" to the state $|0\rangle$ in the first case and the state $|1\rangle$ in the second case.

   (b) In particular from (a), if the state after measurement is $|0\rangle$ then repeated applications of the same measurement of the qubit will continue (with probability 1) to give the state $|0\rangle$. Similarly for $|1\rangle$.

   (c) The qubit "decides" whether its value is 0 or 1 at the instant of measurement, and *not* prior to measurement.[6]

   (d) More generally, suppose the measurement corresponds to the orthonormal basis $\{|u_1\rangle, |u_2\rangle\}$. After this measurement the state will change/collapse[7] to one of these basis vectors, the "value" obtained will be the "label" u_1 or u_2 for this basis vector, and if $|\psi\rangle = \alpha_1|u_1\rangle + \alpha_2|u_2\rangle$ the respective probabilities will be $|\alpha_1|^2$ or $|\alpha_2|^2$. The obvious analogues of (a), (b) and (c) also hold.

In summary, orthonormal bases of qubit states are naturally associated with measurement processes, and conversely.

1.1.3 Superposition and Ensembles

Later we consider *ensembles*, probabilistic weightings of qubits such as $\mathscr{E} = \{(|0\rangle, p), (|1\rangle, q)\}$, where $0 \leq p \leq 1$, $0 \leq q \leq 1$ and $p^2 + q^2 = 1$. In this example the qubit is in the state $|0\rangle$ with probability p and in the state $|1\rangle$ with probability q.

The qubit $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ is not the same as the ensemble $\mathscr{E} = \{(|0\rangle, |\alpha|^2), (|1\rangle, |\beta|^2)\}$. For example, let $\alpha = \beta = 1/\sqrt{2}$. Define[8]

$$|+\rangle := \frac{1}{\sqrt{2}}\,|0\rangle + \frac{1}{\sqrt{2}}\,|1\rangle\,, \quad |-\rangle := \frac{1}{\sqrt{2}}\,|0\rangle - \frac{1}{\sqrt{2}}\,|1\rangle\,. \tag{1.2}$$

See Figure 1.2. Note $|\psi\rangle = |+\rangle$. Also see (2.1).

[5] This is not quite correct, there are more general measurement processes, but for present purposes this will suffice.

[6] With two or more *entangled* qubits the situation is more complex, as we discuss in Section 1.3.

[7] I think "collapse" can be misleading terminology. Every state can serve as one of the possible outcome states for *some* measurement. But on the other hand, there is only a restricted (for us — finite) set of possible outcomes for each given measurement.

[8] These are important qubits.

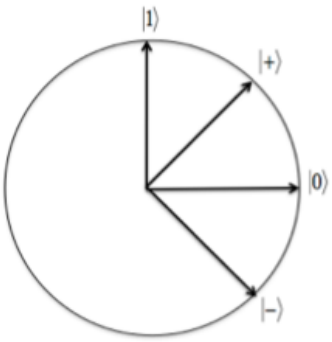

Figure 1.1: A section of $\mathbb{C}^2$. (from [Vaz12])

Measurements[9] of $|\psi\rangle$ and $\mathscr{E}$ in the $\{|0\rangle, |1\rangle\}$ basis both give $|0\rangle$ or $|1\rangle$ with probability $1/2$. But measurement of $|\psi\rangle$ in the $\{|+\rangle, |-\rangle\}$ basis gives $|+\rangle$ with certainty and $|-\rangle$ with probability 0. On the other hand, for the ensemble $\mathscr{E}$ measurement in the $\{|+\rangle, |-\rangle\}$ basis gives $|+\rangle$ and $|-\rangle$ each with probability $1/2$. *Exercise*

In the case of the ensemble $\mathscr{E}$, the state is either in state $|0\rangle$ or the state $|1\rangle$. We just do not know which. In the case of $|\psi\rangle$ the state is definitely *not* in either state $|0\rangle$ or $|1\rangle$. It is in a superposition of $|0\rangle$ and $|1\rangle$. It is also in a superposition of the states $|u_1\rangle$ and $|u_2\rangle$ for *any* orthonormal basis $\{|u_1\rangle, |u_2\rangle\}$. You will gain an intuition for this as we proceed.

1.1.4 The private world of qubits

A quantum state (or qubit) does not correspond to a physical state in the classical sense.

Qubits live in their own private $\mathbb{C}^2$ state space/world. This is a much "richer" world than the physical world it represents. The coefficients α_1 and α_2 in $|\psi\rangle = \alpha_1|u_1\rangle + \alpha_2|u_2\rangle$ contain information about the outcomes of measurement in the orthonormal basis $\{|u_1\rangle, |u_2\rangle\}$, as we saw in (d) above. But because, after measurement, the state of the qubit collapses to one of these basis vectors, the original state can no longer be "examined" by looking at that particular qubit.

But if someone sends us a large supply of identically prepared qubits $\alpha|0\rangle + \beta|1\rangle$, then we *can* obtain good estimates for $|\alpha|$ and $|\beta|$. *How?* In fact we can do considerably more than this, as we will later discuss.[10]

1.2 Young's Two-Slit Experiment

(Optional) Bonus Material!

Young's experiment is the classic quantum experiment, so a brief discussion seems mandatory.

According to Richard Feynman:

> ...it contains the *only* mystery [in quantum theory]. We cannot make the mystery go away by "explaining" how it works. We will just *tell* you how it works. In telling you how it works we will have told you about the basic peculiarities of all quantum mechanics.

The essential element in the "telling" is the superposition principle. However, we will need to move a little beyond qubits to a different state space.

[9] See the preceding discussion of measurement by projection. In the case of $\mathscr{E}$ just treat each possibility $|0\rangle$ and $|1\rangle$ separately and weight with the appropriate probabilities, here $1/2$.

[10] If $\alpha = r_1 e^{i\theta_1}, \beta = r_2 e^{i\theta_2}$ in polar form, then we estimate $r_1 = |\alpha|^2$ and $r_2 = |\beta|^2$ by a simple frequency analysis on the outcomes of measurements in the $\{|0\rangle, |1\rangle\}$ basis. Although we cannot estimate θ_1 and θ_2, we can estimate the relative phase $\theta_1 - \theta_2 \pmod{2\pi}$ by using measurements corresponding to orthonormal bases other than the standard basis. Physically, this is all that is relevant. We discuss this later.

1.2.1 Experimental Set-Up and Outcomes

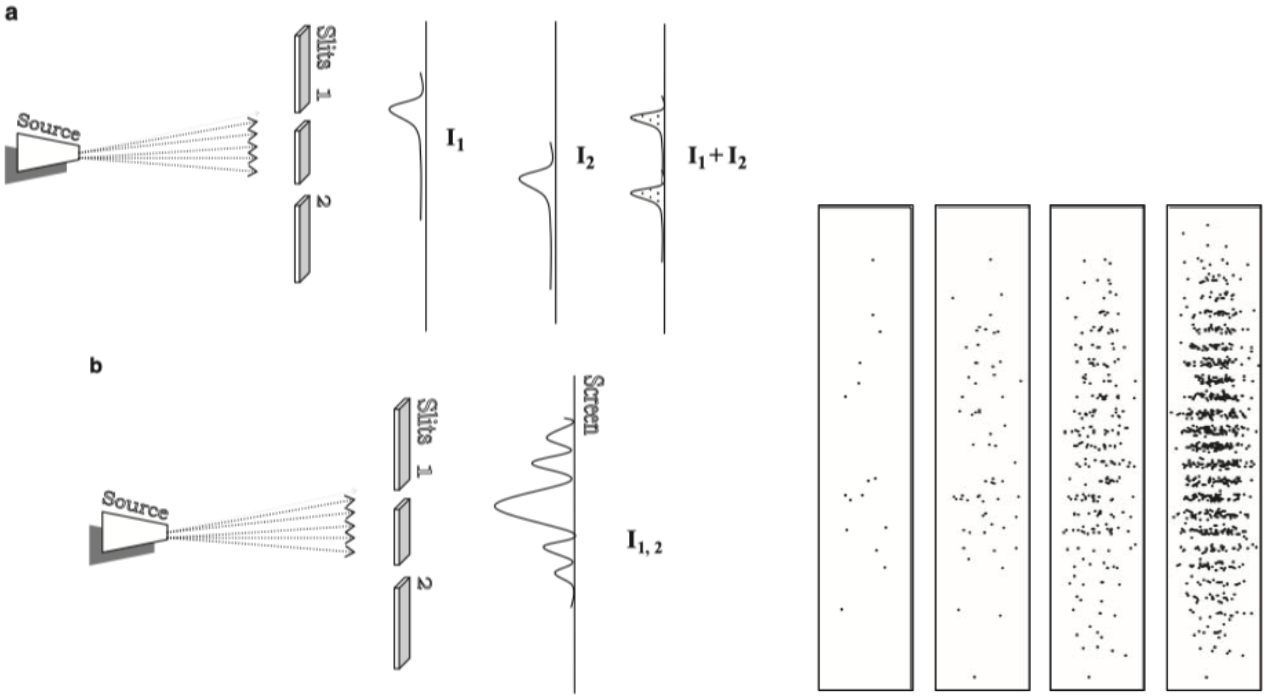

Figure 1.2: Double-slit experiment with photons/electrons or bullets. I_1 is the result *if* slit 2 is covered, I_2 is the result *if* slit 1 is covered. $I_1 + I_2$ is the result if both slits open and bullets are used, $I_{1,2}$ if photons/electrons are used ([GHW09, p 177]). The diagram on the right shows $16, 64, 256, 1024$ photons strikes respectively ([SW10, p 9])

In Diagram **a** in Figure 1.2, photons/electrons or bullets are fired at a panel with two narrow slits of width appropriate to the particles used. *If* slit 2 is covered those particles which are not absorbed by the panel, and so pass through slit 1, will concentrate on a screen behind the panel with an intensity in the vertical direction according to the approximately normal distribution I_1. The dispersion is caused by interaction with the edges of the slit. Similarly, *if* slit 1 is blocked the intensity is given by I_2.

With photons or electrons, however, the situation is quite different. If both slits are open the intensity is given by $I_{1,2}$ as in Diagram **b**. This is similar to the interference pattern obained if a steady water wave pattern is sent towards the panel, as in Figure 1.4. The peaks of the circular waves interfere constructively, as do the troughs, to give maximum intensity (squared amplitude) along the red lines. (The intensity peaks decrease as one moves away from the point on the screen opposite midway between the two slits.) In between peaks and troughs one obtains destructive interference.

Moreover, with photons or electrons one obtains the same interference pattern even if the photons/electrons are fired one at a time. Each electron/photon is exhibiting a wave pattern of self interference. This is very weird!

1.2.2 The State Space

To simplify matters, imagine the screen is divided into many small regions $R_{\boldsymbol{x}}$, each of which is denoted by some point $\boldsymbol{x} = (x, y) \in R_{\boldsymbol{x}}$. See Figure 1.4.

The possible outcomes of a particle striking the screen are given by the N possible values of $\boldsymbol{x}$.[11] Striking the screen is the measurement process. The outcomes are *mutually exclusive and regarded as exhaustive* in our set-up. For these reasons we take the state space model for a particle just prior to striking the screen to be an N-dimensional complex inner product space, not $\mathbb{C}^2$ as for qubits, with

[11] So $N = 8 \times 15 = 120$ here!

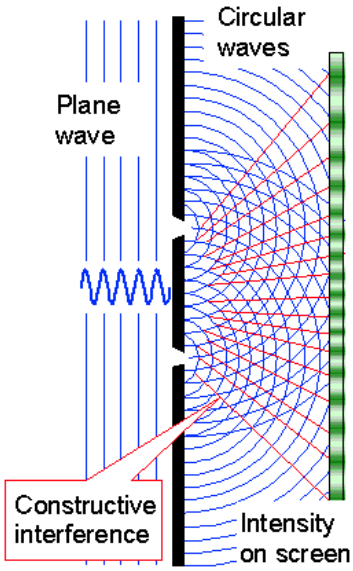

Figure 1.3: Water waves. (from https://www.tf.uni-kiel.de/matwis/amat/iss/kap_4/illustr/s4_2_2.html

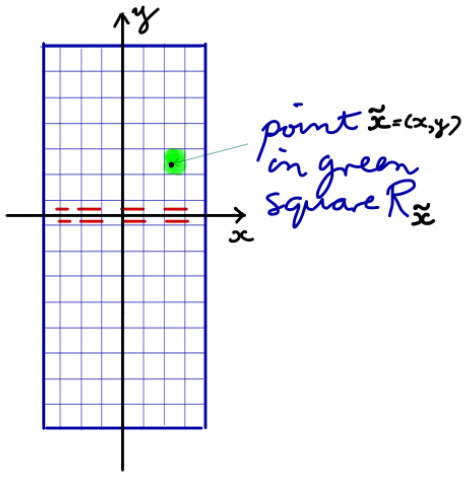

Figure 1.4: Screen as in Figure 1.2. (The dotted red lines are just to indicate that they are directly behind the two slits on the front panel.)

an orthonormal basis consisting of vectors $|\boldsymbol{x}\rangle$. Without loss of generality we can use $\mathbb{C}^N$ with the standard basis vectors.

Just as an arbitrary qubit can be written $\alpha\,|0\rangle + \beta\,|1\rangle$, here the state of an arbitrary particle just prior to impacting the screen will be a superposition

$$|\psi\rangle := \sum_{\boldsymbol{x}} \alpha(\boldsymbol{x})\,|\boldsymbol{x}\rangle\,, \qquad \sum_{\boldsymbol{x}} |\alpha(\boldsymbol{x})|^2 = 1. \tag{1.3}$$

In particular, for a particle (electron or photon) which we *know* passed through slit j (for example, if the other slit was closed), the state will be a *unit* vector $|\psi_j\rangle \in \mathbb{C}^N$ which we write as

$$|\psi_j\rangle := \sum_{\boldsymbol{x}} \alpha_j(\boldsymbol{x})\,|\boldsymbol{x}\rangle\,, \quad \alpha_j(\boldsymbol{x}) = |\alpha_j(\boldsymbol{x})|\; e^{i\phi_j(\boldsymbol{x})} \qquad \text{with } \sum_{\boldsymbol{x}} |\alpha_j(\boldsymbol{x})|^2 = 1. \tag{1.4}$$

The probability $p_j(\boldsymbol{x})$ that such a particle will arrive in $R_{\boldsymbol{x}}$ can be measured by repeated such experiments. It is given by $p_j(\boldsymbol{x}) = |\alpha_j(\boldsymbol{x})|^2$ according to the analogue of the discussion in Section 1.1.2, and thus $|\alpha_j(\boldsymbol{x})|$ can be estimated in this manner. The phase of the particle just prior to arriving at $R_{\boldsymbol{x}}$ is denoted by $\phi_j(\boldsymbol{x})$. Writing $\boldsymbol{x} = (x, y)$ in cartesian coordinates, where $y = 0$ is the vertical height corresponding to midway between the slits, $p_j\big(\boldsymbol{x} = (x, y)\big)$ is a discrete approximation to a distribution which is a product of a normal distribution in y centred at height corresponding to slit j, and a constant distribution in the x (horizontal) direction for an interval rougly the length of each slit. See I_1, I_2 in Figure 1.2, showing the y dependence. The distributions I_1, I_2 are displaced vertically from each other by the distance between the two slits.

Supose we fire electrons or photons at the panel and consider only those that are not blocked by the panel but pass through the slits and register in some way on the screen to the right in Figure 1.2. (Note that I said pass through the slits, *not* "pass through *one of* the slits"!!) Provided there was no measurement which would have enabled us to know if the particle passed through slit 1 (and equivalently in this case no measurement that would have enabled us to know if the particle passed through slit 2), the state space at any time after the particle has passed through the slits allows for both possibilities. One should *not* think: "the particle actually passed through one or the other slit but we just do not know which".

Let s be the distance between the two slits, L the distance between the panel and the screen, and λ the wave length of the particle. The state of the particle in these circumstances just prior to striking the screen, for $L \gg s$, is approximately the equal superposition[12]

$$|\psi\rangle = \frac{1}{\sqrt{2}}\big(\,|\psi_1\rangle + |\psi_2\rangle\,\big). \tag{1.5}$$

Fixing $\boldsymbol{x}$ and writing $\alpha_j(\boldsymbol{x}) = r_j e^{i\phi_j}$ in polar coordinates, the coefficient/amplitude of $|\boldsymbol{x}\rangle$ in $|\psi\rangle$ is

$$\frac{1}{\sqrt{2}}\big(\alpha_1(\boldsymbol{x}) + \alpha_2(\boldsymbol{x})\big) = \frac{1}{\sqrt{2}}\left(r_1 e^{i\phi_1} + r_2 e^{i\phi_2}\right), \tag{1.6}$$

[12]This is not surprising, but to argue it more carefully requires considering a larger state space involving the slits. This is done, for example, in [FLS65, Vol III, Chapter 2, §§3-1,3-2]. Alternatively, it is an example of a general principle for superposition of amplitudes in such situations, see for example [GHW09, p 176].

Finally, it may seem surprising that $1/\sqrt{2}$ is the correct normalisation factor to ensure $\|\,|\psi\rangle\,\| = 1$. But note $\langle\psi|\psi\rangle = 1 + \langle\psi_1|\psi_2\rangle$ and $\langle\psi_1|\psi_2\rangle \approx 0$ due to the highly oscillatory behaviour of $\cos(\phi_1 - \phi_2)$. To pursue this point a little further,

$$\begin{aligned}
\langle\psi_1|\psi_2\rangle &= \sum_{\boldsymbol{x}} r_1(\boldsymbol{x})\, r_2(\boldsymbol{x}) \cos\big(\phi_1(\boldsymbol{x}) - \phi_2(\boldsymbol{x})\big) \\
&\approx \int r_1(y)\, r_2(y) \cos\big(\phi_1(y) - \phi_2(y)\big)\, dy \quad \text{where } \boldsymbol{x} = (x, y) \text{ and ignoring "constant" } x \text{ dependence of everything} \\
&\approx \int r_1(y)\, r_2(y) \cos\left(\frac{sy}{\lambda L}\right) dy \qquad \text{from (1.8)} \\
&= -\frac{\lambda L}{s}\left(\int \frac{d}{dy}\big(r_1(y)\, r_2(y)\big) \sin\left(\frac{sy}{\lambda L}\right) dy + \text{boundary terms}\right).
\end{aligned}$$

Since $\lambda \ll s$, this last term is, in effect, bounded in absolute value by constant $\times\, \lambda$.

and so the probability of the particle landing at $\boldsymbol{x}$ is

$$\begin{aligned}\frac{1}{2}\left|r_1e^{i\phi_1}+r_2e^{i\phi_2}\right|^2 &= \frac{1}{2}\left(r_1e^{i\phi_1}+r_2e^{i\phi_2}\right)\left(r_1e^{-i\phi_1}+r_2e^{-i\phi_2}\right)\\ &= \frac{1}{2}\left(r_1^2+r_2^2\right)+r_1r_2\cos(\phi_1-\phi_2)\end{aligned} \tag{1.7}$$

The term $\frac{1}{2}\left(r_1^2+r_2^2\right)$ is the probability of arriving in $R_{\boldsymbol{x}}$ for the *mixed* state corresponding to equal probabilities of passing through slit 1 and slit 2. The second term $r_1r_2\cos(\phi_1-\phi_2)$ is the *interference term* which has a magnitude comparable to the first term but may be *additive or subtractive*.

It is not difficult to show[13] that, for $s, y \ll L$, the distance between stripes in Figure 1.2 is approximately $\lambda L/s$ and

$$\phi_1(\boldsymbol{x})-\phi_2(\boldsymbol{x}) \approx 2\pi\frac{sy}{\lambda L}, \tag{1.8}$$

This can also be used to estimate the wave length of the particles.

Remark *The fact that for the superposition of mutually exclusive events (in this case: passing through slit 1 and passing through slit 2) amplitudes add but probabilities do not, is critical to the weirdness of quantum events.*

1.3 Entangled Pairs of Qubits

1.3.1 Quantum states for pairs of qubits

We have seen how the quantum analogue of a bit with values in $\{0,1\}$ is a qubit whose state can be any linear combination, i.e. superposition, of the *orthonormal* vectors $|0\rangle$ and $|1\rangle$ in (the unit sphere for) $\mathbb{C}^2$.

Similarly, we can consider a pair of bits and the quantum analogue of a pair of qubits. For classical bits the 4 possible values of a pair of bits are $\{00,01,10,11\}$[14]. In quantum mechanics these values correspond to the following *orthonormal* unit vectors in (the unit sphere for) $\mathbb{C}^4$.

$$\{00\}\to|00\rangle := \begin{bmatrix}1\\0\\0\\0\end{bmatrix},\quad \{01\}\to|01\rangle := \begin{bmatrix}0\\1\\0\\0\end{bmatrix},\quad \{10\}\to|10\rangle := \begin{bmatrix}0\\0\\1\\0\end{bmatrix},\quad \{11\}\to 0|11\rangle := \begin{bmatrix}0\\0\\0\\1\end{bmatrix}. \tag{1.9}$$

Mathematically, the set of all possible unit length superpositions of the four basis vectors in (1.9) is the set of all vectors/states on the unit sphere in $\mathbb{C}^4$ which are of the form

$$|\psi\rangle := a_{00}|00\rangle+a_{01}|01\rangle+a_{10}|10\rangle+a_{11}|11\rangle\in\mathbb{C}^4,\quad \text{where } a_{ij}\in\mathbb{C},\ \sum|a_{ij}|^2=1. \tag{1.10}$$

It *is* physically possible to produce pairs of qubits corresponding to any such state! This has very counterintuitive consequences.

A typical measurement basis for states in $\mathbb{C}^4$ uses the standard basis $\{|00\rangle,|01\rangle,|10\rangle,|11\rangle\}$. This measurement is physically achieved by separate measurements on the first and second qubits, although as noted above the two qubits may be in highly separated locations.[15] The squared moduli of the

[13] The distance from the bottom/top slit to the point on the screen at height y, for $s, y \ll L$, is

$$\left(L^2+\left(y+\frac{s}{2}\right)^2\right)^{1/2}\approx L+\frac{1}{2}L^{-1}\left(y\pm\frac{s}{2}\right)^2,$$

and so the difference is $\approx sy/L$. The waves radiating out from the two slits are in phase at $y=0$ and the next values of y for which the two waves will be in phase are given by $sy/L=\pm\lambda$, i.e. $y=\pm L\lambda/s$. Also see [BS14, p 43, §2.3.4].

Thus the distance between bands on the right of Figure 1.2 is $\approx L\lambda/s$, and so $\phi_1(\boldsymbol{x})-\phi_2(\boldsymbol{x})\approx 2\pi\frac{sy}{\lambda L}$.

[14] It is conventional to list values in increasing binary order.

[15] Not all measurements can be achieved this way. The issue is whether or not the measurement is factorisable in the tensor sense, as we discuss in subsequent chapters.

coefficients in (1.10) give the probabilities of the possible experimental outcomes, in a way analogous to what happens for a single qubit as discussed in Section 1.1.2. The way we read off this information is straightforward but the consequences are very weird indeed. We will discuss this for the example of the Bell state (1.12).

1.3.2 Separable qubits and entangled qubits

We can easily produce pairs of qubits of the form $\alpha|0\rangle + \beta|1\rangle$ and $\alpha'|0\rangle + \beta'|1\rangle$ respectively, which leads naturally to pairs of qubits which can be written in the product form

$$\big(\alpha|0\rangle + \beta|1\rangle\big)\big(\alpha'|0\rangle + \beta'|1\rangle\big) = \alpha\alpha'|00\rangle + \alpha\beta'|01\rangle + \beta\alpha'|10\rangle + \beta\beta'|11\rangle \in \mathbb{C}^4. \tag{1.11}$$

However, not every pair of qubits in the form (1.10) can be written in the form (1.11).

If a state can be factorised as in (1.11) it is said to be *separable*, if not it is said to be *entangled*.

1.3.3 The Bell state

A particularly instructive example of an entangled pair of qubits is the *Bell state*

$$|\Phi^+\rangle := \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle \in \mathbb{C}^4. \tag{1.12}$$

Exercise: Show the Bell state is entangled.

It is possible to physically produce many examples of the Bell state, for example where the pair of qubits is a pair of entangled polarised photons. Even though the two qubits in the Bell state (1.12) are entangled, it is possible to physically separate them by over 1,200 km, and in principle by any distance. See Section 1.4.

The $\frac{1}{\sqrt{2}}$ coefficient in the Bell state for the $|00\rangle$ component and the implicit 0 coefficient for $|01\rangle$ together tell us that when the first qubit is measured in the standard basis $\{|0\rangle, |1\rangle\}$ the result is 0 with probability $(1/\sqrt{2})^2 + 0^2 = 1/2$, and in this case the first qubit moves into the state $|0\rangle$. Similarly, the measurement result is 1 with probability[16] 1/2, and in this case the first qubit moves into the state $|1\rangle$.

Suppose the measurement outcome for the first qubit is 0. The most astounding consequence is that the qubit pair instantaneously moves/collapses to the state $|00\rangle$, and as a consequence the second qubit instantaneously moves into the state $|0\rangle$.[17] In effect, this conclusion is obtained by scratching out the $|11\rangle$ term in (1.12) (since it begins with 1 but the first qubit is actually in the state $|0\rangle$) and then normalising

1.4 The EPR Paradox

1.4.1 What does it mean for quantum theory to be complete?

Einstein, Podolsky and Rosen in [EPR35] argued that quantum theory is not a complete theory. They accepted that in most quantum experiments conducted up to that time the results predicted were probabilistic, perhaps due to some random interaction with the measuring apparatus or for some other reason. This was *not* the situation they were addressing.

On the other hand, they pointed out that in certain experiments which involve entanglement (at that time thought experiments, i.e. "Gedankenexperiment", but subsequently experimentally realised in [ADR82]), the outcome *was* determined before the measurement was conducted but quantum theory did not include this information. It was in this sense that they claimed quantum theory was not complete.

I describe the [EPR35] argument in the context of entangled qubits in the Bell state (1.12) and the 2017 Micius (named after the satellite used) experiment [YCL+17] illustrated in Figure 1.5.

[16] By "probability" we just mean what occurs over many experiments in a frequency analysis.

[17] As we discuss later, using tensor notation $|\Phi^+\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle\otimes|0\rangle + |1\rangle\otimes|1\rangle\big) \in \mathbb{C}^2\otimes\mathbb{C}^2 = \mathbb{C}^4$, and orthogonal projection onto $\mathrm{span}\{|0\rangle\}\otimes\mathbb{C}^2$ gives $|00\rangle$ after normalisation. That is, just scratch out any terms beginning with $|1\rangle$ and then normalise.

1.4.2 The EPR Experiment

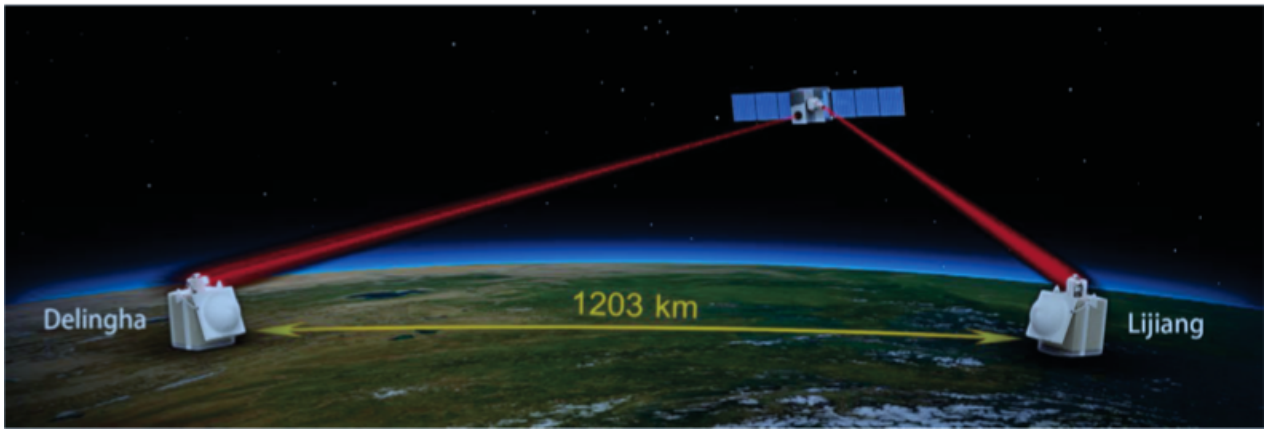

Figure 1.5: Experimental set-up of satellite-based entanglement distribution. (Arxiv version of [YCL+17])

Figure 1.6: Schematic representation of EPR set-up (from [RP11]). Alice and Bob are spacelike separated at measurement time. Alice and Bob are traditionally responsible for the measurements in such experiments. (From the list of authors and affiliations in [YCL+17] it seems unlikely either was present on this occasion.)

A pair of qubits realized as linearly polarized photons have their polarizations entangled on the satellite into the Bell state $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$.[18] After entanglement the first qubit is sent to Lijiang and the second to Delingha Lijiang. Each qubit is measured in the $\{|0\rangle, |1\rangle\}$ basis immediately upon arrival.[19] In the diagram, this means the Lijiang qubit before the Delingha qubit.

Suppose for sake of discussion that the first qubit arrives in Lijiang and is measured before the second qubit arrives in Delingha and is measured, but the second qubit arrives and is measured before there is time for any message (at the speed of light) to reach it regarding the measurement outcome of the first qubit. Each evening pass of the satellite produces about 300 such events.

From the previous discussion concerning the Bell state the probability of the first qubit being measured as 0 is 1/2, of the second qubit being measured as 0 is 1/2, and of the pair being measured as 00 is 1/2. So the measurement results of the individual qubits are not independent, since if they were then the probability of the pair being measured as 00 would be 1/4, and the state of the qubit pair would be the product state $\frac{1}{4}|00\rangle + \frac{1}{4}|01\rangle + \frac{1}{4}|10\rangle + \frac{1}{4}|11\rangle \in \mathbb{C}^4$, not the Bell state. In fact the measurement outcomes of the two qubits are *perfectly correlated*, both outcomes being 0 or both being 1.

[18] In the actual experiment the Bell state $\frac{1}{\sqrt{2}}|01\rangle - \frac{1}{\sqrt{2}}|10\rangle$ was used. This does not alter the discussion in any significant manner.

[19] A more complicated experiment was conducted with additional orthonormal bases. We return to this in later discussions.

How is this perfect correlation "enforced"? Quantum theory implies the first qubit measures as 0 with probability 1/2, as 1 with probability 1/2, but the "decision" as to *which* of these two outcomes is the case is made at the time of measurement. For sake of argument suppose the first qubit measure as 0 — how does the second qubit know that it must also measure as 0 if it is to be consistent with the quantum theory prediction of equality of outcome?

Logically there are two possibilities:

(a) *either* there is an interaction between the two qubits faster than the speed of light (what Einstein called "spooky action at a distance"), such that the second qubit "knows" the measurement outcome for the first qubit,

(b) *or* there was no such interaction. In this case information about *which* one of the two shared/common outcomes will occur for an entangled pair must have been available to each particle in the pair at the time of entanglement (the last possible time for slower than speed of light interaction before measurement).

In the case of (b) this additional information is not present in the Bell state, and so quantum theory would not be a complete theory according to [EPR35].

By way of background we remark that there is indeed a simple classical model to completely explain the outcome of the version of the EPR experiment as discussed here. All that is needed is that at the time of entanglement both qubits are produced in the state $|0\rangle$ or both in the state $|1\rangle$, with probability 1/2 for each possibility, and this choice is independent between pairs of qubits.

The difference between this model and the Bell state model shows up if we measure the individual qubits in other than the orthonormal base $\{|0\rangle, |1\rangle\}$, as we will discuss in a later section.[20]

1.4.3 Quantum theory *is* complete

[EPR35] considered (a) to not be possible, and so came to the conclusion that (b) is true and quantum theory is *not* complete. Einstein spent much of the remaining 20 years of his life looking for such a complete theory.

However, John Bell in [Bel64] designed an experiment to decide between (a) and (b). A small modification was proposed in [CHSH69] and carried out in [ADR82], with the most recent experiment being the Micius experiment [YCL+17]. This has shown convincingly that quantum theory is complete, any probabilistic or deterministic extension which includes other information (known as "hidden variables") implies correlations which are experimentally incorrect, and "spooky action at a distance" is indeed what happens.

The idea that an event can be influenced by an event outside its past light cone certainly goes against our classical intuition. It is worth remarking, however, that no *classical* information is transmitted faster that the speed of light in these cases! We will return to this later in these notes.

[20] For future reference, what we are saying here is that the Bell state is distinct from the ensemble given by $|00\rangle$ with probability 1/2 and $|11\rangle$ with probability 1/2.

Chapter 2

Quantum Postulates for Closed Systems

In this chapter we begin anew. The material in Chapter 1 may be considered as motivation for the more careful treatment here, and we refer back to it occasionally.

The quantum postulates provide the framework for all of Quantum Theory. They are essentially due to Paul Dirac (1930) [Dir58] and John Von Neumann (1932) [vN18].

We mainly discuss the case of finite dimensional state spaces. This is what is needed for quantum computation, and in addition all the counterintuitive weirdness is already present in this setting.

In this chapter we only deal with closed systems, or more precisely, systems that can be approximated by closed systems. See Remark 2 in Section 2.1.1.

The goal is to obtain an understanding of the postulates and how to apply them.

You should now read Appendix E on the Bloch sphere as it is used in the discussion of qubit measurement in Section 2.2 and provides a nice way of representing and thinking about qubits.

2.1 State Space

2.1.1 State Space Postulate

Postulate 1. *Associated to any closed physical system is a complete separable*[1] *complex inner product space V (that is, a Hilbert space) known as the* state space *of the system. The system at each time is completely described by its* state vector *at that time, which is a unit vector in the system's state space.*

Examples

In particular, $\mathbb{C}^2$ is the state space for qubits. Important examples of qubits are $|0\rangle$, $|1\rangle$ and

$$\begin{aligned} |+\rangle &:= \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\,, \qquad |-\rangle := \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle\,, \\ |i\rangle &:= \frac{1}{\sqrt{2}}|0\rangle + i\frac{1}{\sqrt{2}}|1\rangle\,, \qquad |-i\rangle := \frac{1}{\sqrt{2}}|0\rangle - i\frac{1}{\sqrt{2}}|1\rangle\,. \end{aligned} \tag{2.1}$$

A three level quantum system is represented by $\mathbb{C}^3$. State vectors in $\mathbb{C}^3$ are the standard basis vectors, usually denoted $|0\rangle$, $|1\rangle$, $|2\rangle$, and more generally

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle + \gamma\,|2\rangle \in \mathbb{C}^3, \quad \text{where } |\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1,\ \alpha,\beta,\gamma \in \mathbb{C}. \tag{2.2}$$

Similarly one uses $\mathbb{C}^n$ for n-level systems.

[1] finite dimensional spaces are always complete and separable

Remarks

1. We only consider finite dimensional state spaces V, usually $\mathbb{C}^n$ for some $n \geq 2$. This is sufficient both for quantum computation and quantum information purposes (and for understanding all the notions of "quantum weirdness").[2]

2. An *isolated* or *closed* system is one which does not exchange any information, matter or energy with any other environment. This is an idealisation and there are no closed systems, except possibly for the universe itself.[3]

3. There is no physically observable difference between the physical system described by $|\psi\rangle$ and that by $e^{i\theta}|\psi\rangle$, see Section 2.2.[4] We call $e^{i\theta}$ a *global phase factor*. The two vectors $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$ are said to be equivalent. To be more precise we should say that a state is given by an equivalence class of unit vectors.

   We later discuss *relative phase*, which is different and is physically significant. See Section 2.2.3.

4. Unit vectors in the state space are usually written in the ket form $|\psi\rangle$ using the Dirac notation, see Appendix D. It is usually convenient to think of ψ itself as a name or symbol representing the actual vector $|\psi\rangle$, as in the case $|0\rangle$ and $|1\rangle$.

5. As we discussed in Chapter 1, a state space in quantum theory describes physical reality in a very different manner than a state space in classical physics describes physical reality. See Section 2.2.

2.1.2 Three Physical Examples of Qubits

Every two-level quantum system is a qubit. Here are three examples that are useful to keep in mind when dealing more abstractly with qubits.

The following diagrams are, of course, schematic only. As with any quantum state, the energy level/polarisation/spin is not determined prior to measurement and until then the value measured/observed after measurement is only predicted probabilistically from the prior state and the particular measurement used.[5]

We give a more detailed discussion of the physical interpretation of qubit states $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ and their measurement outcomes in each of these three cases in Section 2.2.2.

Energy levels of an electron in an atom

It is often possible to effectively limit an electron in an atom to its two lowest energy levels and to produce electrons in any (complex) superposition of these levels. The quantum model is thus qubit state space $\mathbb{C}^2$. The lowest energy level state is denoted by $|0\rangle$ and the next level by $|1\rangle$.

By shining light on the atom with the appropriate energy and for the appropriate length of time, it is possible to move the electron from the $|0\rangle$ state to the $|1\rangle$ state and vice versa. By reducing the time we shine the light, an electron initially in the state $|0\rangle$ can be moved halfway into the $|+\rangle$ state.

See [NC10, p280 Box 7.1] for details of how this is physically achieved.

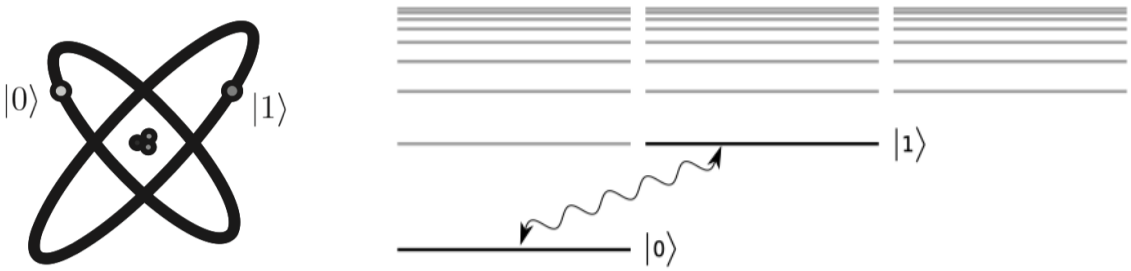

Figure 2.1: Electron energy levels of an electron in an atom. ([NC10, p14 Fig 1.2], [Vaz12, Notes Ch 1])

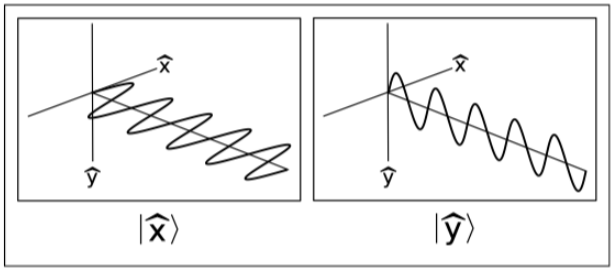

Figure 2.2: Horizontal and vertical polarisation of light. ([Vaz12, Lecture Notes Chapter 1])

Polarised photon

Individual photons have a polarisation modelled by qubit state space $\mathbb{C}^2$. The qubits $|0\rangle$ and $|1\rangle$ are conventionally assigned to horizontal and vertical linear polarisation in the plane orthogonal to the direction of propagation. See Figure 2.2.

By means of a horizontally polarised filter it is possible to produce a large supply of horizontally polarized photons which are conventionally assigned the qubit state $|0\rangle$. Similarly for $|1\rangle$. To achieve other states one passes the photons through a mixture of mirrors, phase shifters and beamsplitters. A beamsplitter is a transparent slab which shifts the phase by an amount depending on the thickness and transparency of the slab.

Spin 1/2 particle

All protons, neutrons, electrons, neutrinos, quarks, and certain atoms such as hydrogen and silver, have a property called *spin* 1/2 [6] which, after measurement in any direction in $\mathbb{R}^3$, takes one of two values, typically denoted by $1/2$ and $-1/2$,[7] or by $\uparrow$ (spin up, i.e. "in the direction of the measuring device") and $\downarrow$ (spin down, i.e. "in the opposite direction from the measuring device"). See Figure 2.3. Spin in this case is modelled by qubit state space $\mathbb{C}^2$.

[2] In quantum mechanics we generally need a separable infinite dimensional Hilbert space. In this case the analogue of the state vector, such as for the position of a particle in physical space $\mathbb{R}^3$, is a complex valued ("wave") function defined over $\mathbb{R}^3$ and typically "peaking" at some point in $\mathbb{R}^3$. Informally, the wave function is a complex superposition of Dirac measures, one at each point in $\mathbb{R}^3$.

[3] And what is *the* universe? Quantum entanglement implies we need to consider events that will not come into our past light cone until some time in our future.

[4] Thus the state space is really a complex projective space, but this is usually not the most convenient/intuitive way to think about it.

[5] If the measurement is with respect to that same state as discussed in Section 2.2 then the probability is 1 and so we do have an outcome with certainty.

[6] This is 2-state spin, which is the most common, but n states for any larger integer n are allowed in principle.

[7] The particular values $\pm 1/2$ are chosen essentially because they should sum to 0 from total spin conservation requirements, and they should differ by 1 in an appropriate system of units.

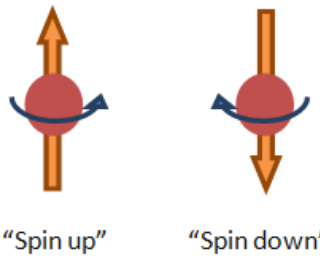

Figure 2.3: Schematic representation of spin 1/2 particle. Spin up = $|0\rangle = |\uparrow\rangle$, Spin down = $|1\rangle = |\downarrow\rangle$. (http://www.sussex.ac.uk/physics/iqt/research/researchers/simulation.html)

2.2 Measurement

2.2.1 Measurement Postulate (Preliminary Version)

Recall that we are only considering finite dimensional state spaces.

Postulate 2 (Preliminary version)**.** *Each orthonormal basis* $\{|v_1\rangle, \ldots, |v_n\rangle\}$ *for the state space* V *describes a measurement/observation as follows. If the state before measurement is*

$$|\psi\rangle = \alpha_1 |v_1\rangle + \cdots + \alpha_n |v_n\rangle$$

then the state after measurement is $|v_j\rangle$ *with probability* $|\alpha_j|^2$*. In this case we say the outcome is (the label)* v_j*.*

Remarks

1. If the orthonormal basis $\{|v_1\rangle, \ldots, |v_n\rangle\}$ is the standard/computational basis $\{e_1, \ldots, e_n\}$, which is usually written $\{|0\rangle, \cdots, |n-1\rangle\}$, then we refer to *measurement in the computational basis* or *measurement in the standard basis.*

2. In $\mathbb{C}^2$ for example, measuring $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ in the computational basis gives
   (a) outcome 0, and $|0\rangle$ as the state after measurement, with probability $|\alpha|^2$,
   (b) outcome 1, and $|1\rangle$ as the state after measurement, with probability $|\beta|^2$.

3. In the previous example, if the outcome was $|0\rangle$, then all subsequent measurements in the computational basis will give $|0\rangle$.[8] *Why?*

4. Each orthonormal basis in the state space corresponds to a particular measurement in physical space which gives *mutually exclusive and exhaustive* outcomes, one outcome for each basis vector.

   For example, with qubits and the standard identification convention of Section 2.2.2, the orthonormal basis $\{|0\rangle, |1\rangle\}$ corresponds to measurement in the z-direction for spin 1/2 particles, with outcome spin 1/2 (i.e. spin ↑) or spin -1/2 (i.e. spin ↓)

   For the two-slit experiment in Section 1.2, the set of orthonormal basis vectors $|\boldsymbol{x}\rangle$ correspond to striking the screen in the various regions $R_{\boldsymbol{x}}$.

   It may be the case that there are technological problems in realising the measurement process corresponding to a particular basis, but in principle there is no restriction. In the case of qubits, all possible measurements can in fact be realised, for example, with spin 1/2 particles and a suitably oriented Stern-Gerlach apparatus. See Appendix E.

[8] To be precise, this assumes the particle was not disturbed by the measurement process. Typically this would mean s detector, such as a screen of some sort, was set up to register if the outcome was $|1\rangle$. Any particle not detected will be in the state $|0\rangle$.

5. Since $e^{i\theta}|\psi\rangle = \sum_i (e^{i\theta}\alpha_i)|v_i\rangle$ it follows that the probability of each outcome, and the state after this outcome, are invariant under multiplication of $|\psi\rangle$ by the global phase factor $e^{i\theta}$. Similarly, the state $|v_i\rangle$ after measurement could equivalently be written $e^{i\phi}|v_i\rangle$ for any ϕ.

6. We say *a measurement is given by (or is w.r.t.[9]) the ket* $|v\rangle \in \mathbb{C}^2$ if the outcome of measuring an arbitrary $|\psi\rangle$ is $|v\rangle$ with probability $|\langle v|\psi\rangle|^2 = \cos^2\theta$ and is $|v\rangle^\perp$ with probability $1 - |\langle v|\psi\rangle|^2 = \sin^2\theta$, where θ is the angle in $\mathbb{C}^2$ between $|v\rangle$ and $|\psi\rangle$. (See Appendix B.1.) This is the same as measurement w.r.t. the orthonormal basis $\{|v\rangle, |v\rangle^\perp\}$. Thus w.l.o.g.[10] we can consider measurements given by a single ket. Of course, this is only true in $\mathbb{C}^2$.

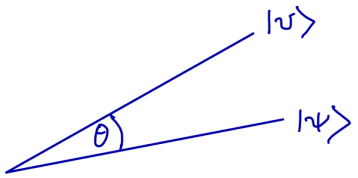

Figure 2.4: A section of $\mathbb{C}^2$. $|\psi\rangle$ is measured w.r.t. $|v\rangle$ — the outcome is $|v\rangle$ with probability $\cos^2(\theta)$.

2.2.2 Example of Qubit Measurement

We discuss the situation for spin 1/2 particles, since in this case there is a particularly nice physical interpretation of the measurement process.

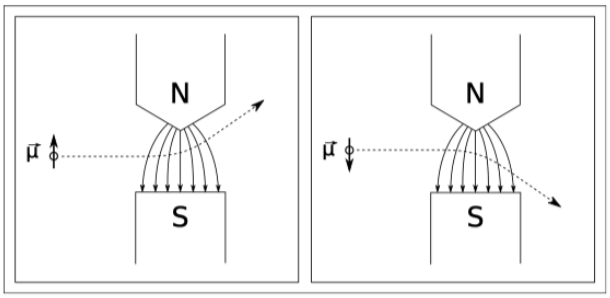

Figure 2.5: Spin 1/2 particle measured by a Stern-Gerlach apparatus ([Vaz12, Notes Chapter 10])

A Stern-Gerlach apparatus generates a non homogeneous magnetic field in some direction in $\mathbb{R}^3$, as in Figure 2.5. As is the case in this diagram, suppose the field is the vertical direction, which we will call the z-direction. When a spin 1/2 particle $|\psi\rangle$ is passed through this field from the left the particle is either deflected up or down. This corresponds to a measurement of the spin of $|\psi\rangle$ in the z direction, which by convention corresponds to a measurement of $|\psi\rangle$ w.r.t. $|0\rangle$ as discussed in Section 2.2.2 Remark 6. See the Bloch sphere representation for qubits discussed in Appendix E.

If $|v\rangle = \cos(\frac{\theta}{2})|0\rangle + \sin(\frac{\theta}{2})e^{i\phi}|1\rangle$ is an arbitrary unit vector in $\mathbb{C}^2$, then measurement of $|\psi\rangle$ w.r.t. $|v\rangle$ is physically achieved by a Stern-Gerlach apparatus oriented in the direction in $\mathbb{R}^3$ corresponding to $|v\rangle$, and hence the apparatus orientation is given by the spherical coordinates (θ, ϕ). See Appendix E.

We will not be discussing technical aspects of physical implementations in these notes, but see [NC10, Section 7.7, pp 324ff] for the physical implementation via nuclear magnetic resonance [NMR]. However, the signal to noise ratio is a major problem with this approach for quantum computation.

[9] "w.r.t." is an abbreviation for "with respect to"
[10] "w.l.o.g." is an abbreviation for "without loss of generality".

2.2.3 Distinguishing Relative Phase

We noted in Section 2.2.1 Remark 5 that measurements do not distinguish between global phase factors. However, they can distinguish relative phase factors such as $e^{i\phi}$ in

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}e^{i\phi}|1\rangle\,. \tag{2.3}$$

A measurement of $|\psi\rangle$ in the computational basis produces the $|0\rangle$ or $|1\rangle$ state with probabilities $\cos^2\frac{\theta}{2}$ and $\sin^2\frac{\theta}{2}$ respectively, which provides no information about ϕ.

But we can get some information about ϕ and estimate it accurately if we can access and measure a large number of similarly prepared qubits. For this consider measurements in the $\{|+\rangle, |-\rangle\}$ basis. From (2.1) we see

$$|0\rangle = \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle\,, \quad |1\rangle = \frac{1}{\sqrt{2}}|+\rangle - \frac{1}{\sqrt{2}}|-\rangle\,. \tag{2.4}$$

So from (2.3) and (2.4)

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(\cos\frac{\theta}{2} + \sin\frac{\theta}{2}e^{i\phi}\right)|+\rangle \;+\; \frac{1}{\sqrt{2}}\left(\cos\frac{\theta}{2} - \sin\frac{\theta}{2}e^{i\phi}\right)|-\rangle \;=:\; \alpha'|+\rangle + \beta'|-\rangle\,. \tag{2.5}$$

Therefore

$$|\alpha'|^2 = \frac{1}{2}\left(\cos\frac{\theta}{2} + \sin\frac{\theta}{2}e^{i\phi}\right)\left(\cos\frac{\theta}{2} + \sin\frac{\theta}{2}e^{-i\phi}\right) = \frac{1}{2}(1 + \sin\theta\,\cos\phi),$$
$$|\beta'|^2 = \frac{1}{2}(1 - \sin\theta\,\cos\phi).$$

So given information about θ we gain some information about the relative phase ϕ, particularly if we can measure a large number of similarly produced qubits.

2.2.4 Hermitian Operators and Measurements

Every orthonormal basis $\{v_1, \ldots, v_n\}$ in an n-dimensional state space V determines the eigenvectors of an hermitian operator $H : V \to V$, but not the eigenvalues. The eigenvalues are essentially labels, and do not normally have a canonical physical interpretation in our setting.[11]

Conversely, suppose $H : V \to V$ is hermitian with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ $(k \le n)$, corresponding orthogonal eigenspaces $E_1, \ldots, E_k$, and orthogonal projection operators $P_i : V \to E_i$. Then

$$V = E_1 \oplus \cdots \oplus E_k, \qquad H = \sum_i \lambda_i P_i, \qquad I = \sum_i P_i, \tag{2.6}$$

where $I : V \to V$ is the identity operator and $V = E_1 \oplus \cdots \oplus E_k$ just says the E_i are mutually orthogonal and every $v \in V$ has a unique decomposition $v = v_1 + \cdots + v_k$ with $v_i \in E_i$. See Figure 2.5.

The other important properties of the P_i are

$$P_i = P_i^*, \qquad P_i^2 = P_i, \qquad P_iP_j = O \quad \text{if } i \neq j, \tag{2.7}$$

where O is the zero operator sending all vectors to the zero vector.

Choosing a set of orthonormal basis vectors for each E_i and taking the union gives a basis for V which is unique up to phase factors if $k = n$, since then the E_i are all one dimensional.

The following version of Postulate 2 is indeed an extension of the preliminary version in Section 2.2.1, as we note in the subsequent Remark 3.

Postulate 2. *Suppose $H : V \to V$ is hermitian as in (2.6). Then H describes a measurement/observation as follows. If the state before measurement is $|\psi\rangle$ then the state after measurement is $P_i|\psi\rangle / \|P_i|\psi\rangle\|$ with probability $\|P_i|\psi\rangle\|^2$. In this case we say the outcome is λ_i.*

[11] However, in physics, where the state space is infinite dimensional, the eigenvectors may represent position or momentum for example, and the eigenvalues take on physical meaning. See Footnote 2

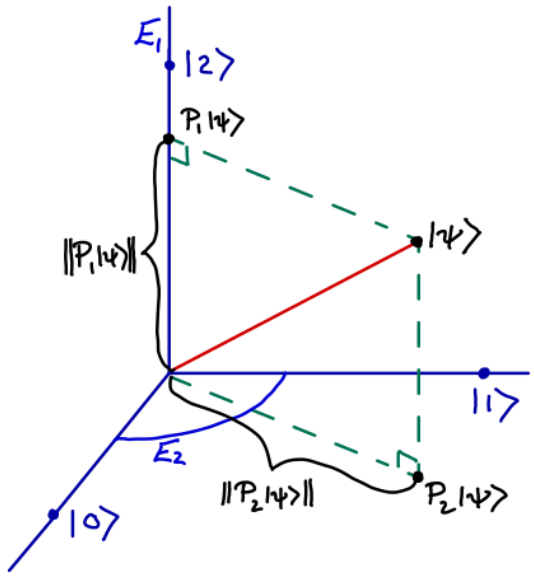

Figure 2.6: Schematic representation for (2.6) of $V = \mathbb{C}^3 = E_1 \oplus E_2$, $|\psi\rangle = P_1|\psi\rangle + P_2|\psi\rangle$, $\| |\psi\rangle \|^2 = \|P_1|\psi\rangle\|^2 + \|P_2|\psi\rangle\|^2$.

Remarks

1. It is *NOT* the case that the state after measurement by H of the state $|\psi\rangle$ is $H|\psi\rangle$.
2. The denominator $\|P_i|\psi\rangle\|$ in $P_i|\psi\rangle / \|P_i|\psi\rangle\|$ is needed for normalisation purposes.
3. Since
$$\begin{aligned} |\psi\rangle &= P_1|\psi\rangle + \cdots + P_k|\psi\rangle \qquad \text{by (2.6)} \\ &= \|P_1|\psi\rangle\| \frac{P_1|\psi\rangle}{\|P_1|\psi\rangle\|} + \cdots + \|P_k|\psi\rangle\| \frac{P_k|\psi\rangle}{\|P_k|\psi\rangle\|}, \end{aligned} \tag{2.8}$$
it is clear by taking $k = n$ that Postulate 2 is a generalisation of the previous preliminary version.
4. Let $p_{|\psi\rangle}(\lambda_i)$ be *the probability that the outcome of the measurement H is λ_i if the initial state is $|\psi\rangle$*.
Since P_i is a projection operator
$$P_i^* = P_i, \qquad P_i^2 = P_i. \tag{2.9}$$
Hence[12]
$$p_{|\psi\rangle}(\lambda_i) := \|P_i|\psi\rangle\|^2 = (P_i|\psi\rangle)^* P_i|\psi\rangle = \langle\psi|P_i^* P_i|\psi\rangle = \langle\psi|P_i|\psi\rangle\,. \tag{2.10}$$
Summarising, the important point is that
$$p_{|\psi\rangle}(\lambda_i) := \|P_i|\psi\rangle\|^2 = \langle\psi|P_i|\psi\rangle\,. \tag{2.11}$$
5. Note $\sum_i p_{|\psi\rangle}(\lambda_i) = 1$. We can see this two ways:
   - *Either*, from (2.8) and (2.11), $1 = \langle\psi|\psi\rangle = \sum_i \|P_i|\psi\rangle\|^2 = \sum_i p_{|\psi\rangle}(\lambda_i)$,
   - *or*, from (2.11) and (2.6), $\sum_i p_{|\psi\rangle}(\lambda_i) = \sum_i \langle\psi|P_i|\psi\rangle = \langle\psi|\sum_i P_i|\psi\rangle = \langle\psi|I|\psi\rangle = 1$.
6. The type of measurements we have been discussing are known as *projective measurements* or *von Neumann measurements*. In quantum computing it is sometimes convenient to deal with more general procedures known as *generalised measurements*. However, it is also possible to realise such measurements by means of projective measurements, but not so conveniently.

[12] see Appendix D.4 for the third "=".

2.3 Evolution

2.3.1 Discrete Time Evolution

Postulate 3. *The time evolution of a closed system is described by a unitary transformation U operating on its state space V. That is, if $|\psi\rangle$ is the state of the system at time t_1 then $U|\psi\rangle$ is the state at time t_2. The operator U depends on t_1 and t_2 but not on the state $|\psi\rangle$.*

Remarks

The requirement that U be unitary is a natural one. In particular, evolution should preserve superposition of states and map state vectors to state vectors. That is

$$U(\alpha|v\rangle + \beta|w\rangle) = \alpha U|v\rangle + \beta U|w\rangle\,, \qquad \|U|v\rangle\| = 1. \tag{2.12}$$

These conditions imply U is a unitary transformation. See Appendix B.5.

Examples

Please read the first few paragraphs of Appendix G for basic information about the Pauli matrices.

1. The *quantum NOT gate* is the Pauli matrix $\sigma_1 = X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. It is also called the *bit flip* matrix/gate since it takes $|0\rangle$ to $|1\rangle$ and $|1\rangle$ to $|0\rangle$.
   As part of a quantum circuit it is represented as

   $$-\boxed{X}- \qquad \text{or} \qquad \alpha|0\rangle + \beta|1\rangle \; -\boxed{X}- \; \beta|0\rangle + \alpha|1\rangle\,, \tag{2.13}$$

   with the input and output indicated in the second case.

2. The *phase flip gate* is the Pauli matrix $\sigma_3 = Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. It leaves $|0\rangle$ unchanged and changes $|1\rangle$ to $-|1\rangle$, with the factor $-1 = e^{-i\pi}$ being the phase factor change.
   As part of a quantum circuit it is represented as

   $$-\boxed{Z}- \qquad \text{or} \qquad \alpha|0\rangle + \beta|1\rangle \; -\boxed{X}- \; \alpha|0\rangle - \beta|1\rangle\,. \tag{2.14}$$

3. The *Hadamard gate* is given by the matrix $H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. It is represented by

   $$-\boxed{H}- \quad . \quad \text{Note } |0\rangle \; -\boxed{H}- \; |+\rangle\,, \quad |1\rangle \; -\boxed{H}- \; |-\rangle\,. \tag{2.15}$$

   The operator H is unitary, trace free hermitian. *Check this.* In particular $H \in \mathcal{H}_0$ and $H = \frac{1}{\sqrt{2}}(X + Z)$. See Appendix G equation (G.4).

2.3.2 Continuous Time Evolution

Often we require a version of this postulate concerning evolution in continuous time.

Postulate 3. *The time evolution of a closed system is described by the Schrödinger equation*

$$\frac{d}{dt}|\psi\rangle = -iH|\psi\rangle\,. \tag{2.16}$$

H is a time-independent hermitian operator known as the Hamiltonian *of the system.*

Remarks

1. The Schrödinger equation is usually written in the form $i\hbar\frac{d}{dt}|\psi\rangle = H|\psi\rangle$ where $\hbar$ is *Planck's constant*. We normally choose units such that $\hbar = 1$.
2. The Hamiltonian H is not the Hadamard matrix/gate H !
3. In quantum computing systems, the hermitian operators are obtained by applying lasers, magnetic fields, etc., appropriately oriented and turned on or off for various time intervals. Of course, this means the quantum system under consideration is not closed. However, by taking it as part of a larger system which is then itself treated as closed, one can usually find a *time dependent* hermitian operator acting on the original quantum system under consideration, with parameters that can be varied as required, and which when used in Schrödinger's gives a good approximation to what happens in the original system.

Solution of Schrödinger's equation for time independent Hamiltonian

In this case, in terms of an orthonormal basis of eigenvectors for H,

$$|\psi\rangle = |\psi(t)\rangle = \begin{bmatrix}\psi_1\\ \vdots \\ \psi_n\end{bmatrix}, \quad H = \begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}. \tag{2.17}$$

It follows that (2.16) is equivalent to the following simple system of independent ordinary differential equations:

$$\frac{d\psi_j}{dt} = -i\lambda_j\psi_j, \quad j = 1, \ldots, n. \tag{2.18}$$

This has the solution

$$\psi_j(t) = e^{-i\lambda_j t}\psi_j(0). \tag{2.19}$$

That is, the solution $|\psi(t)\rangle$ of Schrödinger's equation at time t, with initial state $|\psi(0)\rangle$, is

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle, \quad U(t) := e^{-itH} = \begin{bmatrix}e^{-it\lambda_1} & & \\ & \ddots & \\ & & e^{-it\lambda_n}\end{bmatrix}. \tag{2.20}$$

See also Section F.4.

Example

2.4 Combining Quantum Systems

Chapter 3

Quantum Circuits

Chapter 4

Composite Systems

Appendices

Appendix A

Visualising Higher Dimensions

Sometimes it is convenient to try and visualise $\mathbb{R}^n$ for $n \geq 4$ and $\mathbb{C}^n$ for $n \geq 2$.

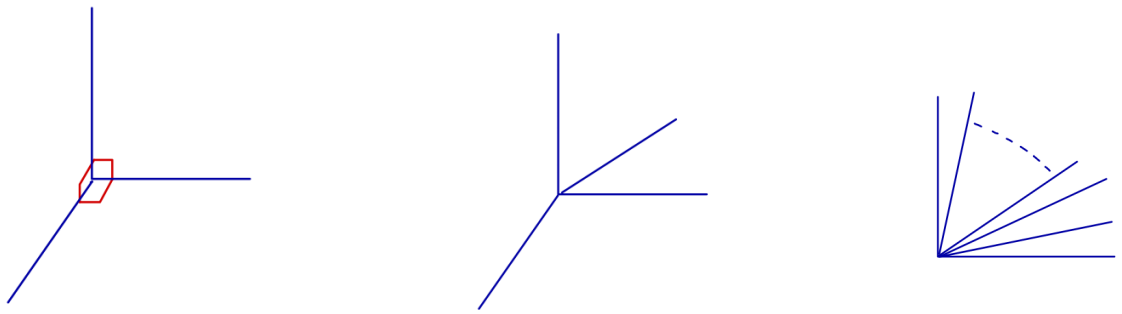

Figure A.1: 3 mutually orthogonal lines in $\mathbb{R}^3$, 4 in $\mathbb{R}^4$, n in $\mathbb{R}^n$.

Exercise: Explain to your 2-D friend *living in the 2-D plane of the paper*, why the 3 lines (segments) in the first diagram are projections into $\mathbb{R}^2$ of 3 mutually orthogonal lines in $\mathbb{R}^3$.

In a similar manner, explain to your 3-D friend in this room, why the 4 lines in the second diagram are actually the projections into $\mathbb{R}^3$ (and then projected onto $\mathbb{R}^2$, but forget that part) of 4 mutually orthogonal lines in $\mathbb{R}^4$.

Similarly, explain why the $n \geq 6$ lines in the third diagram are actually the projections into $\mathbb{R}^3$ of n mutually orthogonal lines in $\mathbb{R}^n$.

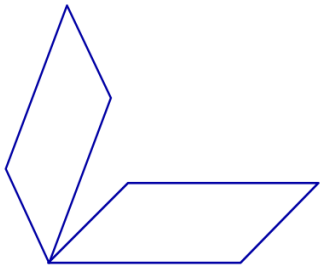

Figure A.2: 2 orthogonal 2-D real subspaces of $\mathbb{R}^4$, or 2 orthogonal 1-D complex lines in $\mathbb{C}^2$.

Exercise: Explain to someone why the above diagram represents (after "extending both rectangles to infinity in all directions") two 2-D real orthogonal subspaces of $\mathbb{R}^4$, whose only point of intersection is the origin.

Then explain why the diagram also represents (after again extending each rectangle) two orthogonal "complex lines"[1] (which of course are each of real dimension 2) in $\mathbb{C}^2$.

[1] A *complex line* in $\mathbb{C}^n$ consists of all vectors of the form αv for some fixed non-zero $v \in \mathbb{C}^n$ and all $\alpha \in \mathbb{C}$.

Appendix B

Hermitiann and Unitary Operators

I summarise the main results for linear operators $T : \mathbb{C}^n \to \mathbb{C}^n$ ($T : \mathbb{R}^n \to \mathbb{R}^n$) which are related to the inner product on $\mathbb{C}^n$ ($\mathbb{R}^n$).

Often I confuse the operator and the matrix of the operator with respect to some basis (the basis being understood from context).

If the proofs are not relatively straightforward I will indicate this.

I also discuss normal operators, but we will not use them in this set of notes.

B.1 Inner Product Space

Here we are interested in finite dimensional spaces, often $\mathbb{R}^n$ or $\mathbb{C}^n$.

An *inner product space* is a real or complex vector space together with an associated inner product $\langle\ ,\ \rangle$. The *inner product* is a map into $\mathbb{R}$ or $\mathbb{C}$ with the properties

- $\langle u, v\rangle = \overline{\langle v, u\rangle}$ (*symmetry*)
- $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$ (mathematics) or $\langle u, \alpha v\rangle = \alpha\langle u, v\rangle$ (physics, computer science) (*linearity* in one of the arguments and hence conjugate linearity in the other)
- $\langle u, u\rangle \geq 0$, and $= 0$ iff $u = \mathbf{0}$

The *standard inner products* on $\mathbb{R}^n$ and $\mathbb{C}^n$ are

- $\langle u, v\rangle = u_1v_1 + \cdots + u_nv_n$
- $\langle u, v\rangle = u_1\overline{v}_1 + \cdots + u_n\overline{v}_n$ (maths) or $\langle u, v\rangle = \overline{u}_1v_1 + \cdots + \overline{u}_nv_n$ (physics, comp sci)

The *norm* of u is $\|u\| = \langle u, u\rangle^{1/2}$, the *distance* between u and v is $\|u - v\|$. The distance is a metric.

The *Cauchy-Schwarz inequality* is $|\langle u, v\rangle| \leq \|u\|\,\|v\|$.

The *angle* between u and v is defined in the $\mathbb{C}^n$ case by $\theta \in [0, \pi/2]$ and $\cos\theta = \dfrac{|\langle u, v\rangle|}{\|u\|\,\|v\|}$. *Warning*: if $v = e^{i\phi}u$ then $\theta = 0$, $\theta \neq \phi$.

In the $\mathbb{R}^n$ case we define $\theta \in [0, \pi]$ by $\cos\theta = \dfrac{\langle u, v\rangle}{\|u\|\,\|v\|}$.

The inner product can be defined from its norm (so norm preserving maps are inner product preserving).

- real case: $\langle u, v\rangle = \frac{1}{4}\big(\|u + v\|^2 - \|u - v\|^2\big)$
- complex case: $\langle u, v\rangle = \frac{1}{4}\big(\|u + v\|^2 + i\|u + iv\|^2 - \|u - v\|^2 - i\|u - iv\|^2\big)$

Warning: Not every norm actually defines an inner product in this way. The necessary and sufficient condition is that the norm satisfies the *parallelogram identity*: $2\|u\|^2 + 2\|v\|^2 = \|u + v\|^2 + \|u - v\|^2$.

B.2 Operator Definitions in Terms of a "Good" Orthonormal Basis

B.2.1 Hermitian Operator

The operator T is *hermitian* ($\mathbb{C}^n$ case), *symmetric* ($\mathbb{R}^n$ case) and *self-adjoint* (either case) iff it corresponds to a (real) scaling in n orthogonal directions ($\mathbb{C}^n$ and $\mathbb{R}^n$ case). Note that in the $\mathbb{C}^n$ case, a "direction" corresponds to a 1-complex-dimensional subspace and hence to a 2-real-dimensional subspace. A scaling corresponds to multiplying by a real number in each of these n complex subspaces.

More precisely, T is *self-adjoint* iff there is an orthonormal basis w.r.t. which T is a real diagonal matrix.

$$T = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \tag{B.1}$$

(We allow negative λ_i.) Think of the unit ball at the origin being stretched into an ellipsoid.

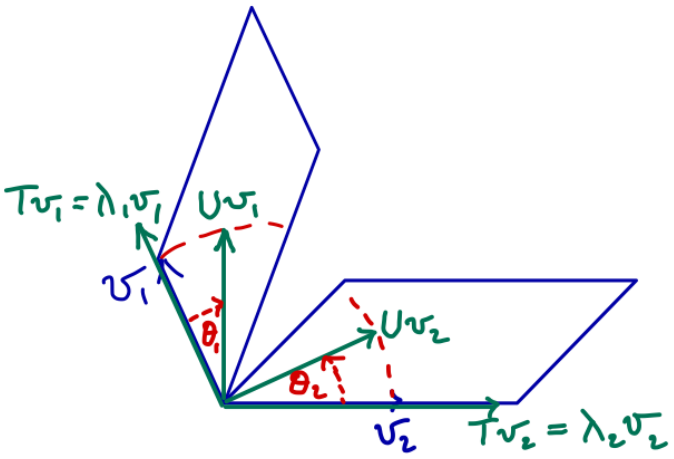

Figure B.1: Hermitian operator T and unitary operator U on $\mathbb{C}^2$, $v_1 \in \text{span}\,|0\rangle$, $v_2 \in \text{span}\,|1\rangle$.

Figure B.1 shows the action of $T = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ and $U = \begin{bmatrix} e^{i\theta_1} & 0 \\ 0 & e^{i\theta_2} \end{bmatrix}$ (see Section B.2.2) on vectors $v_1 \in \text{span}\,|0\rangle$ and $v_2 \in \text{span}\,|1\rangle$. In particular, Tv_1 is obtained by scaling v_1 with the real factor λ_1 and Uv_1 is obtained by rotating v_1 in the complex line $\text{span}\,|0\rangle$ (a 2 real-dimensional plane) anticlockwise[1] through the angle θ_1.

B.2.2 Unitary Operator

U is *unitary* ($\mathbb{C}$ case) iff it corresponds to a rotation (multiplication by $e^{i\theta_1}, \ldots, e^{i\theta_n}$) in n orthogonal directions $v_1, \ldots, v_n$.

More precisely, U is *unitary* iff there is an orthonormal basis w.r.t. which U is a diagonal matrix with all diagonal entries having absolute value 1.

$$U = \begin{bmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{bmatrix}$$

The corresponding notion for orthonormal matrices ($\mathbb{R}$ case) is not so clean. We say that O is orthonormal if there is an orthonormal basis such that O fixes some basis vectors, reflects other basis vectors, and rotates in orthogonal planes determined by pairs of the remaining basis vectors.

[1] Note that there is an orientation on $\text{span}\,|0\rangle$ given by complex multiplication.

More precisely, O is *orthonormal* iff there is an orthonormal basis with respect to which O has matrix

$$O = \begin{bmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & -1 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & -1 & & & & \\ & & & & & & \cos\theta_1 & -\sin\theta_1 & & \\ & & & & & & \sin\theta_1 & \cos\theta_1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & \cos\theta_k & -\sin\theta_k \\ & & & & & & & & \sin\theta_k & \cos\theta_k \end{bmatrix}$$

In particular, in $\mathbb{R}^3$ every orthonormal operator is a rotation around some vector followed perhaps by a reflection in the plane orthogonal to this vector!

Think of the unit ball at the origin being rotated into itself (perhaps there is also a reflection, thus reversing the orientation)

B.2.3 Normal Operator

T ($\mathbb{C}$ case) is *normal* iff it corresponds to multiplication by complex numbers in each of n orthogonal directions. That is, corresponds to a stretch followed by a rotation in each of the corresponding n real 2-dimensional subspaces.

More precisely, T is *normal* iff there is an orthonormal basis with respect to which T is diagonal (with complex, not necessarily real, entries). That is, T has matrix

$$T = \begin{bmatrix} \lambda_1 e^{i\theta_1} & & \\ & \ddots & \\ & & \lambda_n e^{i\theta_n} \end{bmatrix} \tag{B.2}$$

So hermitian and unitary operators are particular cases of normal operators.

B.3 Definitions in Terms of Adjoint Operator

B.3.1 Adjoint

The *adjoint* T^* of the operator T is defined by $\langle T^*u, v\rangle = \langle u, Tv\rangle$ for all $u, v \in \mathbb{C}^n$ ($\mathbb{R}^n$).

With respect to any orthonormal basis, the matrix of T^* is the conjugate (not relevant of course in the $\mathbb{R}$ case) of the transpose of T, i.e. $T^*_{ij} = \overline{T_{ji}} = T^*_{ji}$, where $\overline{T_{ji}} = T^*_{ji}$ is the complex conjugate of T_{ji}.

Note $(AB)^* = B^*A^*$.

B.3.2 Hermitian Operator

T is *hermitian* iff $T = T^*$.

The $\Longrightarrow$ case is easy but the converse requires work. It follows from the more general result for normal operators below. But in this special case it is easier to show the existence of an orthonormal basis of eigenvectors.

As noted in Section B.2.1, one often says T is *Hermitian* in the $\mathbb{C}$ case and *symmetric* in the $\mathbb{R}$ case.

B.3.3 Unitary Operator

U is *unitary* (O is *orthonormal*) iff $UU^* = U^*U = I$ ($OO^* = O^*O = I$).

As for hermitian/self-adjoint operators, $\implies$ is easy, the converse requires work, but follows from the more general result for normal operators.

For arbitrary matrices, being unitary is equivalent to the columns being orthonormal (norm one and inner products of different columns equalling zero) and also equivalent to the rows having the same property.

B.3.4 Normal Operator

T is normal iff $TT^* = T^*T$. ($\implies$ is clear from the definition. The converse requires work. The point is to show the existence of an orthonormal basis of eigenvectors. This is sketched in [NC10, p72].

B.4 Polar Decomposition

Every linear operator T ($\mathbb{C}$ case) can be written in the form $T = UH$ where U is unitary and H is Hermitian (i.e. self-adjoint) with nonnegative e-values. That is, real nonnegative stretching in orthogonal complex directions followed by rotations in another set of orthogonal complex directions.

If the rotations are in the *same* complex directions as the stretches, then the matrix is clearly normal.

In the $\mathbb{R}$ case, $T = OS$ where O is orthonormal and S is symmetric with nonnegative e-values. That is, T is again a nonnegative stretch in orthogonal directions and then apply orthogonal transformation.

Think of taking the unit ball sitting at the origin, stretching it to an ellipsoid, and then rotating.

Similarly ($\mathbb{C}$ case), $T = H'U$ where U is unitary and H' is self-adjoint with nonnegative e-values. Just take same U as before, and take $H' = UHU^*$.

Similarly for the $\mathbb{R}$ case.

B.5 Another Characterisation

Suppose $T : \mathbb{C}^n \to \mathbb{C}^n$. Then

1. T is *hermitian* iff $\langle Tu, u\rangle$ is real for all u.
2. T is *unitary* (orthonormal) iff $\|Tu\| = \|u\|$ for all u iff $\langle Tu, Tv\rangle = \langle u, v\rangle$ for all u, v.
3. T is *normal* iff $\|Tu\| = \|T^*u\|$ for all u.

For $\Longleftarrow$ in 1, suppose $\langle Tu, u\rangle$ is real for all u. Then

$$\langle Tu, u\rangle = \langle Tu, u\rangle^* = \langle u, T^*u\rangle^* = \langle T^*u, u\rangle.$$

So $\langle (T - T^*)u, u\rangle = 0$ for all u. But if $\langle Au, u\rangle = 0$ for all u then for all u, v

$$0 = \langle A(u + v), u + v\rangle = \langle Au, v\rangle + \langle Av, u\rangle.$$

Replacing v by iv and dividing through by i,

$$0 = \langle Au, v\rangle - \langle Av, u\rangle.$$

Adding, $\langle Au, v\rangle = 0$ for all u, v and so $A = 0$. So $T = T^*$.

(Note: If $A : \mathbb{R}^n \to \mathbb{R}^n$ then $\langle Au, u\rangle = 0$ for all u does not imply $A = 0$. Consider rotation by $\pi/2$ in $\mathbb{R}^2$.)

Appendix C

Determinant and Trace

The following applies to both real and complex spaces, provided "vol" is defined in the same way in both cases and is allowed to be complex if necessary.

C.1 Determinant

If A is a square $n \times n$ matrix then the *determinant* $\det A$ can be defined[1], for example, by

$$\det A = \sum_{\pi \in \Pi\{1,\ldots,n\}} (-1)^\pi A_{1\pi(1)} \cdot \ldots \cdot A_{n\pi(n)}, \tag{C.1}$$

where $\Pi\{1,\ldots,n\}$ is the set of permutations of $\{1,\ldots,n\}$ and $(-1)^\pi$ is $+1$ or -1 according as π is even or odd.

Important properties are

$$\det(I) = 1, \quad \det(AB) = \det(A)\,\det(B) \tag{C.2}$$

where I is the identity matrix and A, B are square matrices of the same size.

It follows that $\det(S^{-1}AS) = \det A$, and so $\det(A)$ *is independent of choice of basis.*

In particular, if A has eigenvalues $\lambda_1, \ldots, \lambda_n$ counted with multiplicities, then

$$\det(A) = \lambda_1 \cdot \ldots \cdot \lambda_n. \tag{C.3}$$

Moreover, if $A : V \to V$ where V is a finite dimensional vector space and A is a *linear operator*, we have shown there is a well defined notion of $\det(A)$ independent of any choice of basis.

Geometrically, and this is what footnote 1 says, $\det(A)$ is the volume change factor given by

$$\mathrm{vol}[Av_1, \ldots, Av_n] = \det(A) \times \mathrm{vol}[v_1, \ldots, v_n], \tag{C.4}$$

where $[w_1, \ldots, w_n]$ is the parallelopiped spanned by $w_1, \ldots, w_n$ and "vol" is the oriented volume.

C.2 Trace

The *trace* $\mathrm{tr}(A)$ of the square matrix A is the sum of its diagonal elements:

$$\mathrm{tr}(A) = \sum_i A_{ii}. \tag{C.5}$$

It is straightforward to check

$$\mathrm{tr}(I) = n, \quad \mathrm{tr}(AB) = \mathrm{tr}(BA). \tag{C.6}$$

[1] The most natural, coordinate-free and geometric way is to use exterior algebra and define $\det A$ for an operator $A : V \to V$ to be the unique constant such that $Av_1 \wedge \cdots \wedge Av_n = \det(A)\, v_1 \wedge \cdots \wedge v_n$ for any basis $\{v_1, \ldots, v_n\}$ of V.

Here A and B need *not* be square; A is $m \times n$ and B is $n \times m$ is the general situation. In general, assuming both A and B are square of the same size so that the following three traces are all defined,

$$\operatorname{tr}(AB) \neq \operatorname{tr}(A)\operatorname{tr}(B). \tag{C.7}$$

However, since $\operatorname{tr}(S^{-1}AS) = \operatorname{tr}(AS^{-1}S = \operatorname{tr}(A)$, $\operatorname{tr}(A)$ *is* independent of choice of basis.

Thus if $A : V \to V$ where V is a finite dimensional vector space and A is a linear operator, then there is a well defined notion of $\operatorname{tr}(A)$ independent of any choice of basis.

In particular, if A has eigenvalues $\lambda_1, \ldots, \lambda_n$ counted with multiplicities, then

$$\operatorname{tr}(A) = \lambda_1 + \ldots + \lambda_n. \tag{C.8}$$

Moreover, if $A : V \to V$ where V is a finite dimensional vector space and A is a *linear operator*, we have shown there is a well defined notion of $\operatorname{tr}(A)$ independent of any choice of basis.

For a *geometric interpretation* of $\operatorname{tr}(A)$ suppose A has e-values $\lambda_1, \ldots, \lambda_n$ (with multiplicities). Then the e-values of $I + tA$ are $1 + t\lambda_1, \ldots, 1 + t\lambda_n$, and so for small t,

$$\det(I + tA) = (1 + t\lambda_1) \cdot \ldots \cdot (1 + t\lambda_n) = 1 + t\operatorname{tr}(A) + O(t^2). \tag{C.9}$$

Hence

$$\left.\frac{d}{dt}\right|_{t=0} \det(I + tA) = \operatorname{tr}(A). \tag{C.10}$$

Noting (C.4), $\operatorname{tr}(A)$ is the rate of change of volume at $t = 0$ of a parallelopiped perturbed by tA.[2]

[2]By taking the matrix for A in the basis $\{v_1, \ldots, v_n\}$ as in (C.4), we see this rate of change of volume only depends on changes in each edge of the parallelopiped in the direction of that edge.

Appendix D

Dirac Bra-Ket Notation

This notation is particularly useful in quantum theory where we are working with vectors, dual vectors and hermitian operators. The terminology "bra-ket" is a play on words from the bracketed expression $\langle v|w\rangle$, in which $\langle v|$ is a bra vector and $|w\rangle$ is a ket vector, see below.

Let V be an inner product space which we usually take to be finite dimensional and usually $\mathbb{C}^n$ for some n.

D.1 Kets and Bras

An element of V is called a *ket vector*, written $|v\rangle$ for some appropriate symbol v, and pronounced "ket v". With respect to some orthonormal basis (understood from context, and often the standard basis in the case of $\mathbb{C}^n$) we have the column vector notation $|v\rangle = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$.

Examples in $\mathbb{C}^2$ are

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad |+\rangle := \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad |-\rangle := \frac{1}{\sqrt{2}}\big(|0\rangle - |1\rangle\big) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{D.1}$$

The dual vector in the dual space V^* which corresponds to $|v\rangle$ is written $\langle v|$ and pronounced "bra v". That is, interpreting $\langle v|$ as an operator acting on $|w\rangle \in V$ in the first expression below,

$$\langle v|(|w\rangle) := \big(|v\rangle, |w\rangle\big) = v_1^* w_1 + \cdots + v_n^* w_n = \begin{bmatrix} {v_1}^* \cdots {v_n}^* \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \langle v|\,|w\rangle = \langle v|w\rangle \tag{D.2}$$

where in the second expression $(\,,\,)$ is the inner product, the third expression is just the definition of the inner product, the fourth expression is rewriting the third as matrix multiplication, the fifth is also matrix multiplication where $\langle v|$ is interpreted as the row vector $\begin{bmatrix} {v_1}^* \cdots {v_n}^* \end{bmatrix}$, and the sixth expression is the standard abbreviation of the fifth in the Dirac notation.

In particular, $\langle v|w\rangle$ is the inner product of $|v\rangle$ with $|w\rangle$. Also, $\langle v| = |v\rangle^* = \begin{bmatrix} {v_1}^* \cdots {v_n}^* \end{bmatrix}$ as a row vector or $1 \times n$ matrix.

D.2 More on Inner Products

Let $A : \mathbb{C}^n \to \mathbb{C}^n$ be a linear operator. We often consider expressions of the form $\langle v|A|w\rangle$. In terms of matrix multiplication this is just (from left to right) the product of the row vector $\begin{bmatrix} {v_1}^* \cdots {v_n}^* \end{bmatrix}$ with the $n \times n$ matrix A and the column vector $|w\rangle = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$.

As an example of working with the Dirac notation, in the (associative) product of the three matrices $\langle v|A|w\rangle$ we can bracket the second and third terms:

$$\langle v|A|w\rangle = \langle v|\ (A\,|w\rangle) = |v\rangle^*(A\,|w\rangle), \tag{D.3}$$

so that $\langle v|A|w\rangle$ is the inner product of the ket vector $|v\rangle$ with the ket vector $A|w\rangle$. On the other hand, bracketing the first and second terms,

$$\langle v|A|w\rangle = (\langle v|\,A))\ |w\rangle = (A^*\,|v\rangle)^*\ |w\rangle, \tag{D.4}$$

so $\langle v|A|w\rangle$ is also the inner product of the ket vector $A^*|v\rangle$ with the ket vector $|w\rangle$.

If A is hermitian then $A = A^*$, so $\langle v|A|w\rangle$ is both the inner product of $|v\rangle$ with $A|w\rangle$ and the inner product of $A|v\rangle$ with $|w\rangle$.

D.3 Outer Products

Check that you can show all assertions in the following.

D.3.1 Definition and Matrix Version

For $|v\rangle, |w\rangle \in V$ define the *outer product*, a linear map $|v\rangle\langle w| : V \to V$, by

$$|v\rangle\langle w|\,(|\psi\rangle) = |v\rangle\,\langle w|\psi\rangle = \langle w|\psi\rangle\,|v\rangle\,, \quad \text{for } |\psi\rangle \in V. \tag{D.5}$$

So $|v\rangle\langle w|$ has range $|v\rangle$ and kernel equal to the orthogonal complement of $|w\rangle$.

The operator $|v\rangle\langle w|$ is called a *dyad*.

In terms of matrices, $|v\rangle\langle w|$ is given by matrix multiplication:

$$|v\rangle\langle w| = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \begin{bmatrix} w_1{}^* \cdots w_n{}^* \end{bmatrix} = \begin{bmatrix} v_1 w_1^* & \dots & v_1 w_n^* \\ \vdots & \ddots & \vdots \\ v_n w_1^* & \dots & v_n w_n^* \end{bmatrix}. \tag{D.6}$$

D.3.2 Orthogonal Projectons

If $|v\rangle$ is a unit vector then

$$|v\rangle\langle v| \tag{D.7}$$

is orthogonal projection onto (the space spanned by) $|v\rangle$.

If W is a subspace of V with orthonormal basis $\{|v_1\rangle, \dots, |v_k\rangle\}$ then

$$\sum_{i=1}^{k} |v_i\rangle\langle v_i| \tag{D.8}$$

is orthogonal projection onto W.

In particular, if $\{|v_1\rangle, \dots, |v_n\rangle\}$ is an orthonormal basis for V then the identity map I satisfies

$$I = \sum_{i=1}^{n} |v_i\rangle\langle v_i|\,. \tag{D.9}$$

If M has matrix elements M_{ij} w.r.t. the computational basis then

$$M = \sum_{i,j} M_{ij}\,|i\rangle\langle j|\,, \qquad M_{ij} = \langle i|M|j\rangle\,. \tag{D.10}$$

If T is hermitian with (necessarily real) eigenvalues λ_i and orthonormal eigenvectors $|v_i\rangle$ (forming a basis for V) then

$$T = \sum_i \lambda_i\,|v_i\rangle\langle v_i|\,. \tag{D.11}$$

D.4 Computing with Dirac Notation

In computing with the bra-ket notation one is usually dealing with expressions such as $\langle v|A|w\rangle$ or $|v\rangle\langle w|\psi\rangle$ etc., and often much more complicated, which can be treated as a product of matrices.

Consequence and points to note are:

- bracketing is allowed in any order consistent with the associative rule for multiplying matrices,
- expressions such as $\langle w|\psi\rangle$ are an abbreviation for the product $\langle w||\psi\rangle$,
- scalar quantities such as $\langle w|\psi\rangle$ can be moved through the expression,
- adjoints of a product are the product of the adjoints *multiplied in the reverse order*,
- adjoints of matrices are conjugate transposes, so adjoints of scalars are their complex conjugates, of hermitian matrices A are again A, of unitary matrices U are U^{-1}, of bras $\langle v|$ are kets $|v\rangle$, and of kets $|v\rangle$ are bras $\langle v|$.

There will be many examples in these notes.[1]

[1] See (2.10), ...

Appendix E

The Bloch Sphere

Connect back to Remark 4

This is a very useful way of representing qubits as points on the unit sphere in $\mathbb{R}^3$. Recall that a qubit is just an equivalence class of points on the unit sphere in $\mathbb{C}^2$, where two qubits are equivalent if they agree up to a global phase factor $e^{i\zeta}$ for some $\zeta \in \mathbb{R}$.

The unit sphere $S^3 \subset C^2$ has 3 real dimensions, and so after factoring out by the given equivalence relation is a 2 real dimensional surface.

The coordinate mapping Any qubit $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \in S^3$, after using polar coordinates for α and β and multiplying through by a (irrelevant) global phase factor, can be written in the form $|\psi\rangle = r_1\,|0\rangle + r_2 e^{i\phi}\,|1\rangle$ where $0 \le r_1,\ 0 \le r_2,\ r_1^2 + r_2^2 = 1,\ 0 \le \phi < 2\pi$. Hence $|\psi\rangle$ can be written in the form

$$|\psi\rangle = \cos\frac{\theta}{2} + \sin\frac{\theta}{2}\,e^{i\phi}, \quad 0 \le \theta \le \pi,\ 0 \le \phi < 2\pi. \tag{E.1}$$

This representation is unique unless $\theta = 0$ or $\theta = 1$.

By using $\{\theta, \phi\}$ as spherical coordinates we obtain a *unique* correspondence between qubits (remember that qubits correspond to equivalence classes of points on the unit sphere in $\mathbb{C}^2$) and points on the unit sphere in $\mathbb{R}^3$. See Figure E.1. (We need $\cos\frac{\theta}{2}$ with $0 \le \theta \le \pi$ rather than $\cos\theta$ with $0 \le \theta \le \pi/2$ for this.)

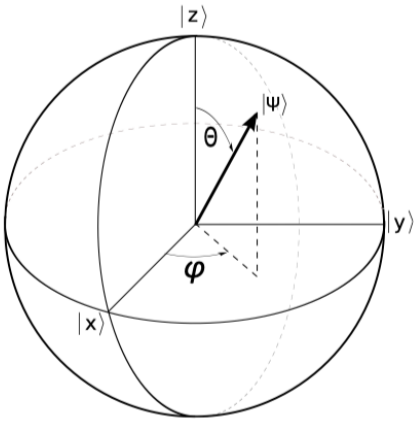

Figure E.1: Bloch Sphere. See (E.4) for notation. ([Vaz12, Notes Chapter 10, Fig. 10.1]).

The point on the unit sphere with spherical coordinates $\{\theta, \phi\}$ has Cartesian coordinates

$$(\sin\theta\,\cos\phi,\ \sin\theta\,\sin\phi,\ \cos\theta) = (\sin\theta\,e^{i\phi},\ \cos\theta), \tag{E.2}$$

where for the second expression we have identified the x–y plane with the complex plane. This is often convenient for calculations.

Antipodal points = orthonormal basis One point that can initially cause confusion is that the pair of *antipodal* points (θ, ϕ) and $(\pi - \theta, \pi + \phi)$ on the Bloch sphere correspond to an *orthonormal basis* in $\mathbb{C}^2$.[1]

Important pairs of antipodal points On the Bloch sphere the north pole corresponds to $|0\rangle$ (also written $|z\rangle$) since the spherical coordinates are $(0, \phi)$, and the south pole corresponds to $|1\rangle$ (also written $|-z\rangle$) since the spherical coordinates are (π, ϕ). The cartesian coordinates are $z = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^3$.

Similarly, with $x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $y = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and using the notation of (2.1),

$$\begin{aligned}
|+\rangle &= \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle = \cos\left(\frac{\pi}{4}\right)|0\rangle + \sin\left(\frac{\pi}{4}\right)|1\rangle = |x\rangle\,, & (\theta,\phi) &= \left(\frac{\pi}{2}, 0\right) \\
|-\rangle &= \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle = \cos\left(\frac{\pi}{4}\right)|0\rangle + \sin\left(\frac{\pi}{4}\right)e^{i\pi}|1\rangle = |-x\rangle\,, & (\theta,\phi) &= \left(\frac{\pi}{2}, \pi\right) \\
|i\rangle &= \frac{1}{\sqrt{2}}|0\rangle + i\frac{1}{\sqrt{2}}|1\rangle = \cos\left(\frac{\pi}{4}\right)|0\rangle + \sin\left(\frac{\pi}{4}\right)e^{i\pi/2}|1\rangle = |y\rangle\,, & (\theta,\phi) &= \left(\frac{\pi}{2}, \frac{\pi}{2}\right) \\
|-i\rangle &= \frac{1}{\sqrt{2}}|0\rangle - i\frac{1}{\sqrt{2}}|1\rangle = \cos\left(\frac{\pi}{4}\right)|0\rangle + \sin\left(\frac{\pi}{4}\right)e^{i3\pi/2}|1\rangle = |-y\rangle\,, & (\theta,\phi) &= \left(\frac{\pi}{2}, \frac{3\pi}{2}\right)
\end{aligned} \tag{E.3}$$

Summarising,

$$|z\rangle = |0\rangle\,,\ |-z\rangle = |1\rangle\,,\ |x\rangle = |+\rangle\,,\ |-x\rangle = |-\rangle\,,\ |y\rangle = |i\rangle\,,\ |-y\rangle = |-i\rangle\,. \tag{E.4}$$

[1] Since $\cos\frac{\pi-\theta}{2}|0\rangle + \sin\frac{\pi-\theta}{2}e^{i(\pi+\phi)}|1\rangle = \sin\frac{\theta}{2}|0\rangle - \cos\frac{\theta}{2}e^{i\phi}|1\rangle$ is clearly a unit vector orthogonal to $\cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}e^{i\phi}|1\rangle$.

Appendix F

Functions of Normal Operators

F.1 Examples of Interest

Let V be a finite dimensional complex inner product space. For us it will be a state space. You can think of $V = \mathbb{C}^n$.

Let $T : V \to V$ be a normal operator. Recall from Appendix B.2.3 that this means T has an orthonormal basis of eigenvectors. We are interested in the case T is hermitian or unitary, but the more general case is just as straightforward.

Let $f : \mathbb{C} \to \mathbb{C}$ be some function. Then can we make sense of $f(T)$ in some useful way?

We will be particularly concerned with

$$\cos(H), \quad \sin(H), \quad \exp(i\theta H), \tag{F.1}$$

where $\exp(u) = e^u$ is the exponential function and H is hermitian.

F.2 Power Series Definition

If f is the squaring function, i.e. $f(z) = z^2$, then we define $f(T) = T^2$. Similarly if $f(z) = z^k$ for some positive integer k we define $f(T) = T^k$. We can also define T^{-1} to be the inverse of T, provided T is invertible.

If $f(z) = a_0 + a_1 z + a_2 z^2 + \cdots$ converges for $|z| < R$ (i.e. f is analytic) then we can define $f(T) = a_0 + a_1 T + a_2 T^2 + \cdots$. This will converge if all entries in the standard matrix representation are $\leq R'$ for some R' depending on R. In particular, for the functions $\cos(z)$, $\sin z$ and e^{az} for any constant a, the series converges for all T. However, this is not the approach we take here. It is mentioned just for comparison with the literature.

F.3 Definition for Normal Operators

For normal matrices we have the following more natural and convenient approach. In particular, there are no differentiability requirements on f, and the definition of $f(T)$ depends only on the values of f for the eigenvalues of T.[1]

With respect to any orthonormal basis of eigenvectors we have

$$T = \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{bmatrix} \quad \Longrightarrow \quad f(T) := \begin{bmatrix} f(\alpha_1) & & \\ & \ddots & \\ & & f(\alpha_1) \end{bmatrix}. \tag{F.2}$$

Note this is not a coordinate dependent definition! We are using an orthonormal basis of eigenvectors for T. Any allowable change of basis vectors for T (e.g. reordering or phase changes) will have the same effect on $f(T)$, i.e. no change on the actual operator.

[1] This approach works for diagonalisable matrices in general. But it is not as stable under convergence of matrices as in the case of normal matrices where we are restricting to orthonormal bases. See [HJ91, §6.2.37 Theorem, p 433].

F.4 Matrices for the Schrödinger Equation

If H is hermitian as in (B.1) with matrix below for some orthonormal basis of eigenvectors $\{E_1, \ldots, E_n\}$ corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$, then $U(t) := e^{-itH}$ has the matrix shown for the same orthonormal basis of eigenvectors (now also for e^{-itH}).

$$H = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \qquad U(t) := e^{-itH} = \begin{bmatrix} e^{-it\lambda_1} & & \\ & \ddots & \\ & & e^{-it\lambda_n} \end{bmatrix}. \tag{F.3}$$

It follows immediately that $U(t)$ is unitary.

It is also clear from its matrix representation that for any unitary operator U there is a unique hermitian operator H such that $U = \exp(-iH)$.

If t is interpreted as time then U acts on the H eigenspace E_i for λ_i by rotation with velocity λ_i, and in the clockwise direction if $\lambda_i > 0$.

In Figure F.1 let $H = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ and $U = \begin{bmatrix} e^{-it\lambda_1} & 0 \\ 0 & e^{-it\lambda_2} \end{bmatrix} = \exp(-itH)$.

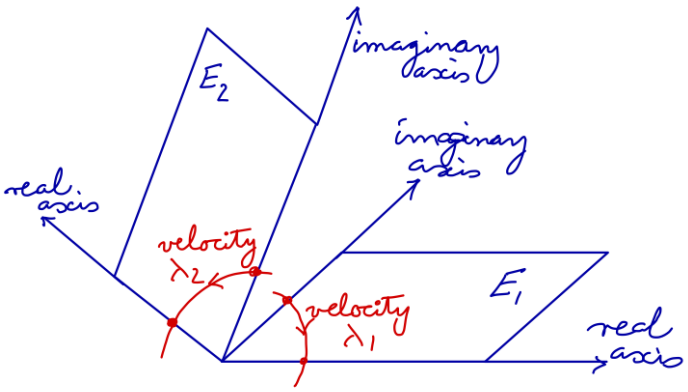

Figure F.1: Schematic representation for the 2 orthogonal eigenspaces $E_1, E_2 \subset \mathbb{C}^2$ and for the unitary operator $U = \exp(-itH)$ as it rotates vectors in these eigenspaces. The direction of rotation is shown in case $\lambda_1, \lambda_2 > 0$.

See also the Schrödinger Equation 2.16

Appendix G

The Pauli Matrices

Exercise: Verify all statements in the following.

The Pauli matrices are ubiquitous in the theory of quantum computation. Here they are with their standard notations.

$$\begin{aligned} \sigma_0 = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad & \sigma_1 = X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\ \sigma_2 = Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad & \sigma_3 = Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \end{aligned} \tag{G.1}$$

Usually σ_0 is omitted from the list. It should be clear from context when this is not the case.

They are *hermitian*. Clearly $\sigma_i = \sigma_i^*$.
They are *unitary*. It is easy to check that $\sigma_i \sigma_i^* = I$.
They are *trace free*, i.e. $\operatorname{tr}(\sigma_i) = 0$, and $\det(\sigma_i) = -1$. It follows that they each have eigenvalues ± 1.

They satisfy the following identities.[1]

$$\begin{aligned} \sigma_i^2 = I, \quad \sigma_i \sigma_j = -\sigma_j \sigma_i \text{ if } i \neq j, \\ \sigma_1 \sigma_2 = i\sigma_3, \quad \sigma_2 \sigma_3 = i\sigma_1, \quad \sigma_3 \sigma_1 = i\sigma_2. \end{aligned} \tag{G.2}$$

Using the notation of (2.1)

- σ_3 has eigenvectors $|0\rangle$ and $|1\rangle$ with eigenvalues $+1$ and -1 respectively,
- σ_1 has eigenvectors $|+\rangle$ and $|-\rangle$ with eigenvalues $+1$ and -1 respectively,
- σ_2 has eigenvectors $|i\rangle$ and $|-i\rangle$ with eigenvalues $+1$ and -1 respectively.

Every hermitian matrix H in $\mathbb{C}^2$ can be uniquely represented as follows with $(n_0, n_1, n_2, n_3) \in \mathbb{R}^4$:[2]

$$H = \begin{bmatrix} n_0 + n_3 & n_1 - in_2 \\ n_1 + in_2 & n_0 - n_3 \end{bmatrix} = n_0 I + n_1 \sigma_1 + n_2 \sigma_2 + n_3 \sigma_3. \tag{G.3}$$

The set of all such matrices is a real vector space $\mathcal{H}$ with basis vectors $\sigma_0, \sigma_1, \sigma_2, \sigma_3$.[3]

The trace free hermitian matrices form a subspace $\mathcal{H}_0 \subset \mathcal{H}$. Every $H_0 \in \mathcal{H}_0$ can be written uniquely in the form

$$H_0 = \begin{bmatrix} n_3 & n_1 - in_2 \\ n_1 + in_2 & -n_3 \end{bmatrix} = n_1 \sigma_1 + n_2 \sigma_2 + n_3 \sigma_3 \tag{G.4}$$

for some $(n_1, n_2, n_3) \in \mathbb{R}^3$. This gives a vector space isomorphism[4] between $\mathcal{H}_0$ and $\mathbb{R}^3$, which will be particularly useful in terms of the Bloch sphere map.

An important example is the Hadamard operator $H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Then $H \in \mathcal{H}_0$, $H^2 = I$, $H = \frac{1}{\sqrt{2}}(\sigma_1 + \sigma_2)$, $\det(H) = -1$, similar to the Pauli matrices themselves.

[1] Note the cyclic ordering in the second line. Clearly if follows $\sigma_1 \sigma_3 = -i\sigma_2, \sigma_2 \sigma_1 = -i\sigma_3, \sigma_3 \sigma_2 = -i\sigma_1$.

[2] The off diagonal entries are arbitrary complex conjugates, the main diagonal entries H_{11} and H_{22} are arbitrary reals — let $n_0 = \frac{1}{2}(H_{11} + H_{22})$, $n_3 = \frac{1}{2}(H_{11} - H_{22})$.

[3] In fact an inner product space with Hilbert Schmidt inner product $(A, B) := \frac{1}{2} \operatorname{tr} A^* B = \frac{1}{2} \operatorname{tr} AB$.

[4] In fact, an inner product space isomorphism.

Bibliography

[ADR82] Alain Aspect, Jean Dalibard, and Gérard Roger, *Experimental test of Bell's inequalities using time-varying analyzers*, Phys Rev Lett **49** (1982), 1804–1807, available at https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.49.1804.

[Bel64] John Bell, *On the Einstein Podolsky Rosen paradox*, Physics **1** (1964), 195–200, available at https://journals.aps.org/ppf/pdf/10.1103/PhysicsPhysiqueFizika.1.195. (Bell derives an experimentally testable inequality to distinguish between quantum predictions and Local Realism.)

[BS14] James Binney and David Skinner, *The Physics of Quantum Mechanics*, Oxford University Press, 2014.

[CHSH69] J. F. Clauser, A. Horne, A. Shimony, and R. A. Holt, *Proposed experiment to test local hidden-variable theories*, Phys. Rev. Lett. **23** (1969), 880–884, available at https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.23.880.

[dW19] Ronald de Wolf, *The Potential Impact of Quantum Computers on Society* (2019), https://arxiv.org/abs/1712.05380. (An Insightful Essay).

[Dir58] P.A.M. Dirac, *Principles of Quantu Mechanics*, fourth (revised) edition, Oxford University Press, 1958.

[EPR35] A Einstein, B Podolsky, and N Rosen, *Can quantum mechanical description of physical reality be considered complete?*, Phys. Rev. **47** (1935), 227, available at https://journals.aps.org/pr/pdf/10.1103/PhysRev.47.777.

[Fey] Richard Feynman, *The Character of Physical Law*, twelfth printing 1985, MIT Press. (1964 lectures given to Cornell students. Appear to be widely available online, see http://people.virginia.edu/~ecd3m/1110/Fall2014/The_Character_of_Physical_Law.pdf).

[FLS65] Richard P. Feynman, Robert B. Leighton, and Matthew Sands, *The Feynman Lectures on Physics, Volume III, Quantum Mechanics*, Caltech online edition http://www.feynmanlectures.caltech.edu, 1965.

[GHW09] Daniel Greenberger, Klaus Hentschel, and Friedel Weinert, *Compendium of Quantum Physics*, Springer, 2009.

[HJ91] Roger Horn and Charles R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.

[JS17] Des Johnston and Bernd Schroers, *Quantum Mechanics and Quantum Computing: an Introduction*, 2017, http://www.macs.hw.ac.uk/~des/qcnotesaims17.pdf. (A small update with a few corrections to Schroers' original notes, also in this bibliography .)

[MN19] Andy Matischak and Michael A. Nielsen, *Quantum computing for the very curious* (2019), https://quantum.country/qcvc. (A new and experimental online course by an excellent expositor. Covers much less and slower pace than these notes. Uses "spaced repetition learning" as in Anki cards. Google it!).

[NC10] Michael A. Nielsen and Isaac L. Chuang, *Quantum computation and quantum information*, Cambridge University Press, Cambridge, 2010. (Ten year Anniversary Edition of what is widely considered to be the gold standard in the field. Comprehensive. Very well written. No longer sold, but has been available free from the publisher's website and often available online for courses. For example, see the quantum computing course http://www.csis.pace.edu/ctappert/cs837-18spring/).

[Pre18] John Preskill, *Quantum Computing in the NISQ era and beyond*, Quantum **2** (2018), 79, available at https://quantum-journal.org/papers/q-2018-08-06-79/pdf/?.

[RP11] Eleanor Rieffel and Wolfgang Polak, *Quantum Computing, A Gentle Introduction*, MIT Press, Cambridge, Massachusetts, 2011. (A little slow at times, but some very interesting discussions.)

[Sch07a] Bernd Schroers, *Quantum Computing*, 2007/8, http://www.macs.hw.ac.uk/~bernd/F14ZD1/qcnotes.pdf. (Goes a fair distance, globally well structured, a quick intoduction to the field, follows Nielsen and Chuang. Locally not so good, some mistakes and overly complicated arguments. Some errors corrected in the slightly updated version used by Johnston).

[Sch07b] ———, *Quantum Computing: Problem Sheets*, 2007/8, http://www.macs.hw.ac.uk/~bernd/F14ZD1/. (Problems vary in difficulty. Many routine ones to check background algebra. Others to check the underlying concepts.)

[SW10] Benjamin Schumacher and Michael D. Westmoreland, *Quantum Processes, Systems, and Information*, Cambridge University Press, 2010. (A well written modern introduction to quantum processes at the advanced undergraduate level.)

[Vaz13] Umesh Vazirani, *Quantum Mechanics and Quantum Computation* (2013), https://courses.edx.org/courses/BerkeleyX/CS-191x/2013_August/course/. (A very well presented edX MOOC.)

[Vaz12] ———, *Chem/CS/Phys191:Qubits, Quantum Mechanics, and Computers* (2012), http://www-inst.eecs.berkeley.edu/~cs191/sp12/. (The Berkeley course on which the edX MOOC course was based. See the link to the course notes.)

[vN18] John von Neumann, *Mathematical Foundations of Quantum Mechanics*, new edition (Nicholas A. Wheeler, ed.), Princeton University Press, 2018.

[YCL+17] Juan Yin, Yuan Cao, Yu-Huai Li, et al., *Satellite-based entanglement distribution over 1200 kilometers*, Science **356** (2017), 1140–1144, available at http://science.sciencemag.org/content/356/6343/1140.full. Arxiv version https://arxiv.org/pdf/1707.01339.pdf.