



uantum Foundations
and
Computation

Seminar Notes

John Hutchinson

April 30, 2019 version

PRELIMINARY & INCOMPLETE

Introduction

Quantum Foundations 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. 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]

Switching from Newton’s mechanics to Einstein’s, for a mathematician, must to some extent be like switching from a good old provincial dialect to the latest Parisian slang. In contrast, switching to quantum mechanics, I imagine, is like switching from French to Chinese. Alexandre Grothendieck, *Récoltes et Semailles* (1986) (tr. from French)

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

Structure of the Notes

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

¹ \mathbb{C}^n is the vector space of n -tuples $a = (a_1, \dots, a_n)$ with inner product $a \cdot b = a_1^* b_1 + \dots + a_n^* b_n$ (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.

provided in Appendix [B](#). Other mathematics is developed in the appendices and referred to as needed. This way you will get to very interesting quantum material very quickly!

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.

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.

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

1	Overview	6
1.1	Qubits	6
1.1.1	Bits and Qubits	6
1.1.2	State Space, Physical Reality and Measuring Qubits	7
1.1.3	Superposition and Ensembles	8
1.1.4	The private world of qubits	8
1.2	Young's Two-Slit Experiment	10
1.2.1	Experimental Set-Up and Outcomes	10
1.2.2	The State Space	11
1.3	Entangled Pairs of Qubits	14
1.3.1	Quantum states for pairs of qubits	14
1.3.2	Separable qubits and entangled qubits	14
1.3.3	The Bell state	14
1.4	The EPR Paradox	16
1.4.1	What does it mean for quantum theory to be complete?	16
1.4.2	The EPR Experiment	16
1.4.3	Quantum theory <i>is</i> complete	17
2	Quantum Postulates for Closed Systems	18
2.1	State Space	19
2.1.1	State Space Postulate	19
2.1.2	Three Physical Examples of Qubits	20
2.2	Measurement	22
2.2.1	Measurement Postulate (Orthonormal Basis Version)	22
2.2.2	Measuring Qubits	23
2.2.3	Distinguishing Relative Phase	23
2.2.4	Hermitian Operators and Measurements	24
2.2.5	Measurement Postulate (Hermitian Operator Version)	24
2.2.6	Pauli Matrices and Measurement	26
2.3	Qubits and Spin 1/2 Particles	28
2.3.1	Spin-Orientation	28
2.3.2	Stern-Gerlach Experimental Set-Up	28
2.3.3	Multiple Particles with the Same Spin-Orientation	29
2.3.4	Spin-Orientation and Qubit Correspondence: Bloch Sphere Map	29
2.3.5	Spin-Orientation and Qubit Correspondence: Physical Justification	30
2.4	Evolution	31
2.4.1	Evolution Postulate (Discrete Version)	31
2.4.2	Evolution Postulate (Continuous Version)	31
2.5	Composite System	34
2.5.1	Introduction	34
2.5.2	Composite System Postulate	34
2.6	EPR, Bell, CHSH	37
2.6.1	Bell's Inequality	37

3	Quantum Circuits	39
3.1	Quantum Gates	39
3.2	No Cloning Theorem	39
3.3	Teleportation of Qubits	39
3.4	Superdense Coding	39
3.5	Deutsch-Jozsa Algorithm	39
4	Composite Systems	40
4.1	Ensembles and Density Operators	40
4.2	The Bloch Ball	40
4.3	Subsystems and the Reduced Density Operator	40
4.4	Schmidt Decomposition and Purification	40
4.5	Bell's Inequality, CHSH Inequality, GHZ Inequality	40
A	Visualising Higher Dimensions	41
B	Hermitian and Unitary Operators	43
B.1	Inner Product Space	43
B.1.1	Inner Product	43
B.1.2	Norm	44
B.1.3	Norms and Inner Products	44
B.1.4	Angle	44
B.2	Operator Definitions in Terms of a “Good” Orthonormal Basis	45
B.2.1	Hermitian Operator	45
B.2.2	Unitary Operator	45
B.2.3	Normal Operator	46
B.3	Definitions in Terms of Adjoint Operator	46
B.3.1	Adjoint	46
B.3.2	Hermitian Operator	47
B.3.3	Unitary Operator	47
B.3.4	Normal Operator	47
B.4	Polar Decomposition	47
B.5	Another Characterisation	47
C	Determinant and Trace	49
C.1	Determinant	49
C.2	Trace	50
D	Dirac Bra-Ket Notation	51
D.1	Kets and Bras	51
D.2	More on Inner Products	52
D.3	Outer Products	52
D.3.1	Definition and Matrix Representation	52
D.3.2	Orthogonal Projections	52
D.3.3	Operator as a Sum of Dyads	53
D.4	Computing with Dirac Notation	53
E	The Pauli Matrices	54
E.1	Definitions, Basic properties, Eigenvalues and Eigenvectors	54
E.2	As Basis Vectors	54
E.3	Pauli Vector	55
E.3.1	Definitions and Examples	55
E.3.2	Properties	55

F	The Bloch Sphere	57
F.1	Coordinate Approach	57
F.2	Pairs of antipodal points	58
F.2.1	Orthonormal Bases & Antipodal Points	58
F.2.2	Important Examples	58
F.3	Projection Mapping Approach	59
F.4	Angle Doubling	61
G	Functions of Normal Operators	62
G.1	Examples of Interest	62
G.2	Power Series Definition	62
G.3	Definition for Normal Operators	62
G.4	Matrices for the Schrödinger Equation	63
H	Tensor Products	64
H.1	Informal Presentation	64
H.1.1	Bilinear Requirements	64
H.1.2	Basis and Dimension	65
H.1.3	Entangled and Product Elements	65
H.1.4	Other Bilinear Products	65
H.1.5	Inner Product	65
H.1.6	Examples	65
H.2	Basis Definition of Tensor Products	66
	Bibliography	66

Chapter 1

Overview

Contents

1.1 Qubits	6
1.1.1 Bits and Qubits	6
1.1.2 State Space, Physical Reality and Measuring Qubits	7
1.1.3 Superposition and Ensembles	8
1.1.4 The private world of qubits	8
1.2 Young’s Two-Slit Experiment	10
1.2.1 Experimental Set-Up and Outcomes	10
1.2.2 The State Space	11
1.3 Entangled Pairs of Qubits	14
1.3.1 Quantum states for pairs of qubits	14
1.3.2 Separable qubits and entangled qubits	14
1.3.3 The Bell state	14
1.4 The EPR Paradox	16
1.4.1 What does it mean for quantum theory to be complete?	16
1.4.2 The EPR Experiment	16
1.4.3 Quantum theory <i>is</i> complete	17

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.

¹“Bit” is a contraction of “binary information digit”.

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

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

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 coefficients 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

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

²“Qubit” — pronounced “q(ueue)bit” — is a contraction of “quantum bit”.

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

⁴This 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 $\frac{|v\rangle}{\|v\|}$.

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

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

- (d) More generally, suppose the measurement corresponds to the orthonormal basis $\{|u_1\rangle, |u_2\rangle\}$. After this measurement the state will change/collapse⁷ 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 $\mathcal{E} = \{(|0\rangle, p), (|1\rangle, q)\}$, 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 .

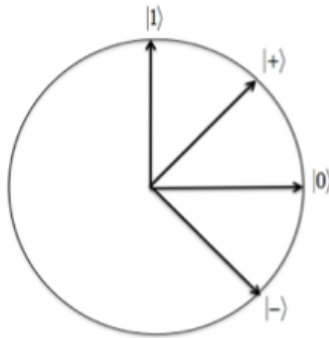


Figure 1.1: A section of \mathbb{C}^2 . (from [Vaz12])

The qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is not the same as the ensemble $\mathcal{E} = \{(|0\rangle, |\alpha|^2), (|1\rangle, |\beta|^2)\}$. For example, let $\alpha = \beta = 1/\sqrt{2}$. Define⁸

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

See Figure 1.2. Note $|\psi\rangle = |+\rangle$. Also see (2.2).

Measurements⁹ of $|\psi\rangle$ and \mathcal{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 \mathcal{E} measurement in the $\{|+\rangle, |-\rangle\}$ basis gives $|+\rangle$ and $|-\rangle$ each with probability $1/2$. *Exercise*

In the case of the ensemble \mathcal{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

⁷I 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.

⁸These are important qubits.

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

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

¹⁰If $\alpha = r_1 e^{i\theta_1}$, $\beta = r_2 e^{i\theta_2}$ 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 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.

1.2.1 Experimental Set-Up and Outcomes

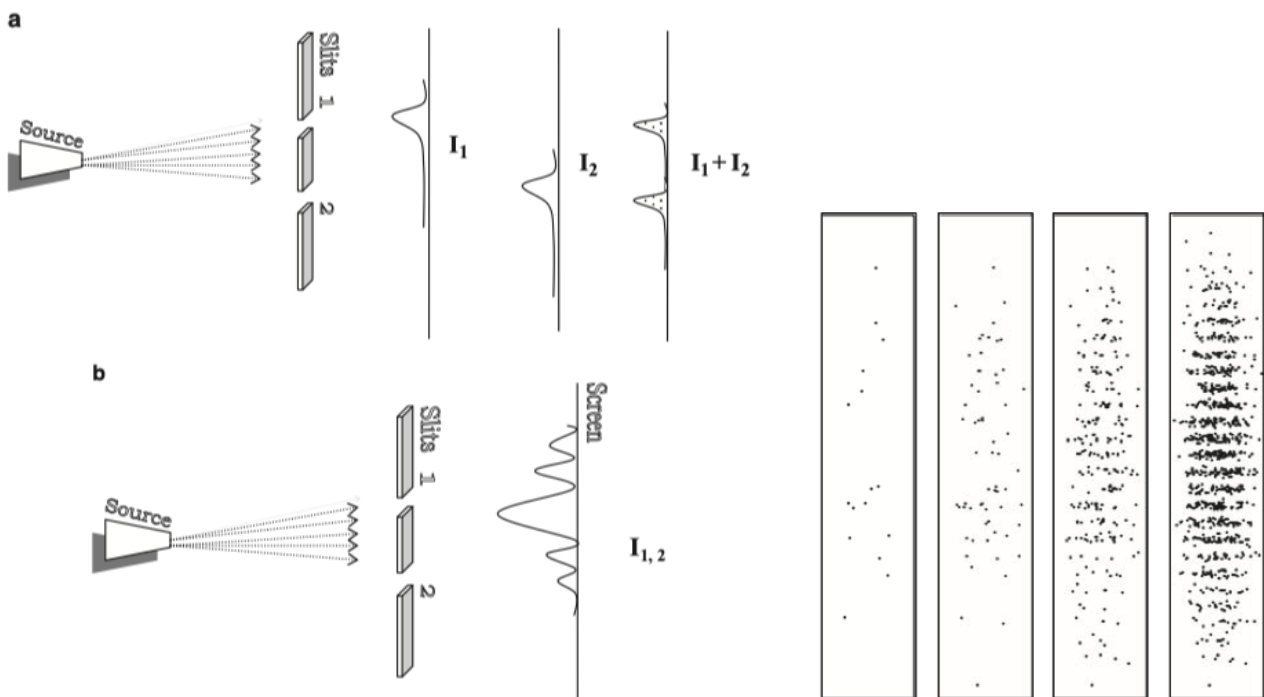


Figure 1.2: Double-slit experiment with photons/electrons or bullets. I_1 is the result *if* slit 2 is covered, I_2 is the result *if* slit 1 is covered. $I_1 + I_2$ is the result if both slits open and bullets are used, $I_{1,2}$ if photons/electrons are used ([GHW09, p 177]). The diagram on the right 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 obtained if a steady water wave pattern is sent towards the panel, as in Figure 1.3. 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!

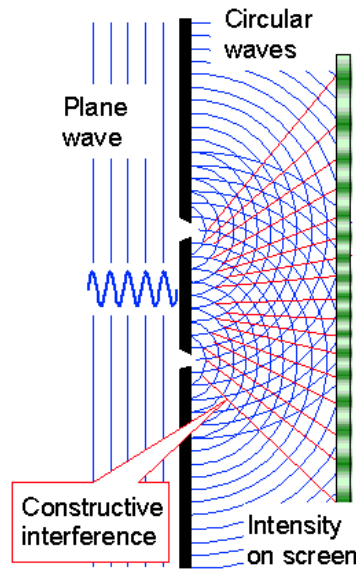


Figure 1.3: Water waves. (from https://www.tf.uni-kiel.de/matwis/amat/iss/kap_4/illustr/s4_2_2.html)

1.2.2 The State Space

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

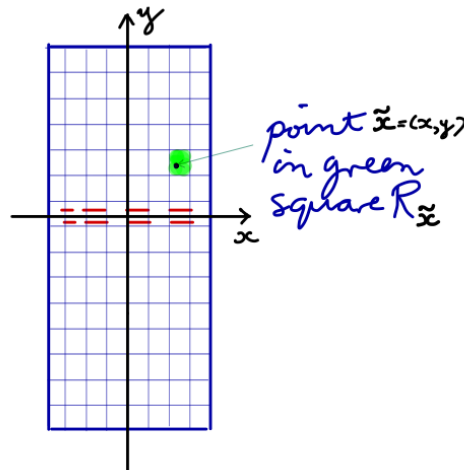


Figure 1.4: Screen as in Figure 1.2. (The dotted red lines are just to indicate that they are directly behind the two slits on the front panel.)

The possible outcomes of a particle striking the screen are given by the N possible values of \mathbf{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

¹¹So $N = 8 \times 15 = 120$ here!

an orthonormal basis consisting of vectors $|\mathbf{x}\rangle$. 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

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

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

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

The probability $p_j(\mathbf{x})$ that such a particle will arrive in $R_{\mathbf{x}}$ can be measured by repeated such experiments. It is given by $p_j(\mathbf{x}) = |\alpha_j(\mathbf{x})|^2$ according to the analogue of the discussion in Section 1.1.2, and thus $|\alpha_j(\mathbf{x})|$ can be estimated in this manner. The phase of the particle just prior to arriving at $R_{\mathbf{x}}$ is denoted by $\phi_j(\mathbf{x})$. Writing $\mathbf{x} = (x, y)$ in cartesian coordinates, where $y = 0$ is the vertical height corresponding to midway between the slits, $p_j(\mathbf{x} = (x, y))$ 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 roughly 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.

Suppose 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 superposition¹²

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

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

$$\frac{1}{\sqrt{2}} (\alpha_1(\mathbf{x}) + \alpha_2(\mathbf{x})) = \frac{1}{\sqrt{2}} (r_1 e^{i\phi_1} + r_2 e^{i\phi_2}), \quad (1.6)$$

¹²This 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 $\|\psi\| = 1$. 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,

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

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 \mathbf{x} is

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

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

It is not difficult to show¹³ that, for $s, y \ll L$, the distance between stripes in Figure 1.2 is approximately $\lambda L/s$ and

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

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.*

¹³The 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 $\phi_1(\mathbf{x}) - \phi_2(\mathbf{x}) \approx 2\pi \frac{sy}{\lambda L}$.

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 .

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

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

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

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 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

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

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*

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

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.

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

¹⁵Not 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.

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 probability¹⁶ $1/2$, 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

¹⁶By “probability” we just mean what occurs over many experiments in a frequency analysis.

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

1.4 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.

1.4.2 The EPR Experiment

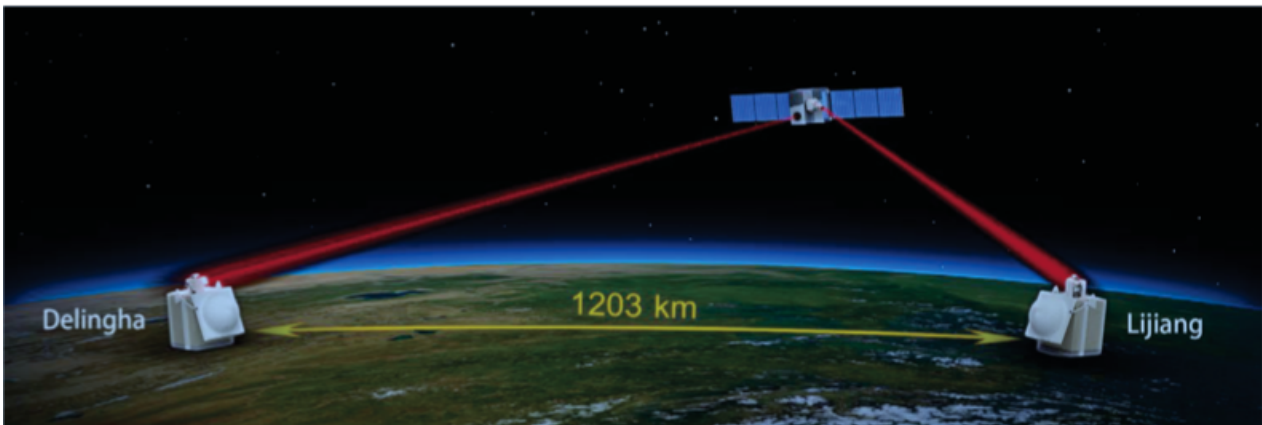


Figure 1.5: Experimental set-up of satellite-based entanglement distribution. (Arxiv version of [YCL⁺17])



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

¹⁸In 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

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.

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 than the speed of light in these cases! We will return to this later in these notes.

manner.

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

²⁰For 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

Contents

2.1 State Space	19
2.1.1 State Space Postulate	19
2.1.2 Three Physical Examples of Qubits	20
2.2 Measurement	22
2.2.1 Measurement Postulate (Orthonormal Basis Version)	22
2.2.2 Measuring Qubits	23
2.2.3 Distinguishing Relative Phase	23
2.2.4 Hermitian Operators and Measurements	24
2.2.5 Measurement Postulate (Hermitian Operator Version)	24
2.2.6 Pauli Matrices and Measurement	26
2.3 Qubits and Spin 1/2 Particles	28
2.3.1 Spin-Orientation	28
2.3.2 Stern-Gerlach Experimental Set-Up	28
2.3.3 Multiple Particles with the Same Spin-Orientation	29
2.3.4 Spin-Orientation and Qubit Correspondence: Bloch Sphere Map	29
2.3.5 Spin-Orientation and Qubit Correspondence: Physical Justification	30
2.4 Evolution	31
2.4.1 Evolution Postulate (Discrete Version)	31
2.4.2 Evolution Postulate (Continuous Version)	31
2.5 Composite System	34
2.5.1 Introduction	34
2.5.2 Composite System Postulate	34
2.6 EPR, Bell, CHSH	37
2.6.1 Bell's Inequality	37

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 quantum 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 Appendices F.1 and F.2 on the Bloch sphere as they are used in the discussion of qubit measurement in Section 2.2 and provide 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 separable¹ 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 are² the basis vectors

$$|0\rangle := \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle := \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad (2.1)$$

and

$$\begin{aligned} |+\rangle &:= \frac{1}{\sqrt{2}} |0\rangle + \frac{1}{\sqrt{2}} |1\rangle, & |-\rangle &:= \frac{1}{\sqrt{2}} |0\rangle - \frac{1}{\sqrt{2}} |1\rangle, \\ |i\rangle &:= \frac{1}{\sqrt{2}} |0\rangle + i \frac{1}{\sqrt{2}} |1\rangle, & |-i\rangle &:= \frac{1}{\sqrt{2}} |0\rangle - i \frac{1}{\sqrt{2}} |1\rangle. \end{aligned} \quad (2.2)$$

Each of these three pairs is an orthonormal basis for \mathbb{C}^2 . *Exercise.*

A 3-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

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle + \gamma|2\rangle \in \mathbb{C}^3, \quad \text{where } |\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1, \quad \alpha, \beta, \gamma \in \mathbb{C}. \quad (2.3)$$

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

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”).³
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.⁴
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 2.2.⁵ We call $e^{i\theta}$ a *global phase factor*. The two vectors $|\psi\rangle$ and

¹finite dimensional spaces are always complete and separable

²“:=” means “by definition is equal to”.

³In 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 .

⁴And 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.

⁵Thus the state space is really a complex projective space, but this is usually not the most convenient/intuitive way to think about it.

$e^{i\theta} |\psi\rangle$ 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.
6. We will see in Section 2.2 that mutually exclusive outcomes of a measurement correspond to orthonormal vectors in state space. Mutually exclusive *and* exhaustive outcomes correspond to an orthonormal *basis*.

2.1.2 Three Physical Examples of Qubits

Please look at Appendices F.1 and F.2 for the Bloch Sphere representation of qubits.

In particular, every qubit $|\psi\rangle \in \mathbb{C}^2$ can be written in the form

$$|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} e^{i\phi} |1\rangle, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \phi < 2\pi, \quad (2.4)$$

and this corresponds to a unique point on the *Bloch sphere*, the unit sphere $S^2 \subset \mathbb{R}^3$ with spherical coordinates (θ, ϕ) .

Every two-level quantum system corresponds to 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.⁶

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 the case of spin 1/2 in Section 2.3.

Energy levels of an electron in an atom

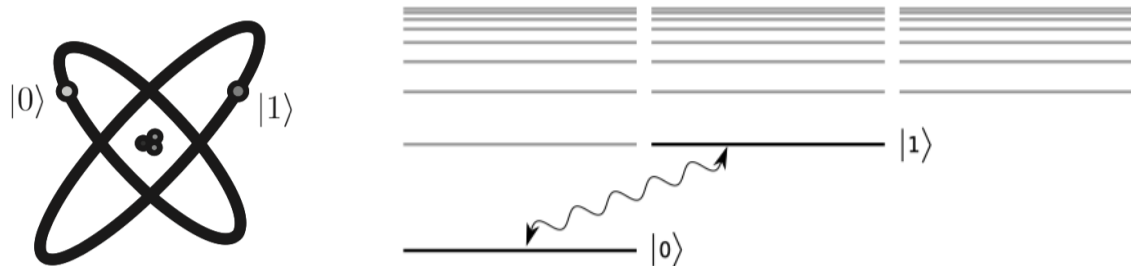


Figure 2.1: Electron energy levels of an electron in an atom. ([NC10, p14 Fig 1.2], [Vaz12, Notes Ch 1])

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 $|1\rangle$.

⁶If 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.

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.

Polarised photon

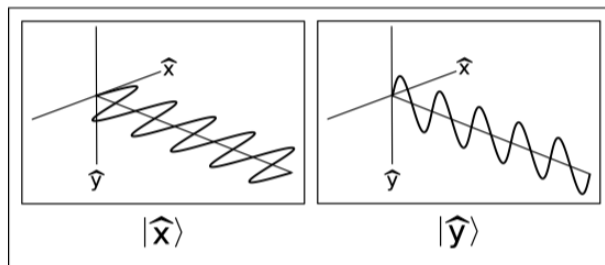


Figure 2.2: Horizontal and vertical polarisation of light. ([Vaz12, Lecture Notes Chapter 1])

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 phase shifter is a transparent slab which shifts the phase by an amount depending on the thickness and transparency of the slab.

Spin 1/2 particle

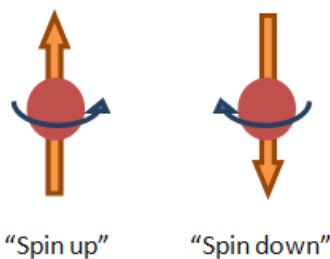


Figure 2.3: Schematic representation of spin 1/2 particle. Spin up = $|0\rangle = |\uparrow\rangle$, Spin down = $|1\rangle = |\downarrow\rangle$. (<http://www.sussex.ac.uk/physics/iqt/research/researchers/simulation.html>)

All protons, neutrons, electrons, neutrinos, quarks, and certain atoms such as hydrogen and silver, have a property called *spin* 1/2⁷ which, after measurement in any direction in \mathbb{R}^3 , takes one of two values, typically denoted by $1/2$ and $-1/2$,⁸ or by \uparrow (spin up, i.e. “right hand thumb rule in the direction of the measuring device”) and \downarrow (spin down, i.e. “right hand thumb rule in the opposite direction from the measuring device”). See Figure 2.3.

In this case each possible state of the qubit corresponds to a unit vector giving the “alignment” of the spin in \mathbb{R}^3 . We discuss this in some detail in Sections 2.3 and 2.3.5.

⁷This is 2-state spin, which is the most common, but n states for any larger integer n are allowed in principle.

⁸The 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.

2.2 Measurement

2.2.1 Measurement Postulate (Orthonormal Basis Version)

Recall that we are only considering finite dimensional state spaces.

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

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

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

Remarks

1. You might like to think of the outcome as a mark v_j on a screen, or a number that shows on a dial, there being only a finite number n of possibilities in the situations we are considering. We typically observe the outcome v_j rather than the state $|v_j\rangle$. But sometimes we will be loose in our language and refer to the outcome $|v_j\rangle$.
2. If the orthonormal basis $\{|v_1\rangle, \dots, |v_n\rangle\} \subset \mathbb{C}^n$ is the standard/computational basis $\{e_1, \dots, e_n\}$, which is usually written $\{|0\rangle, \dots, |n-1\rangle\}$, then we refer to *measurement in the computational basis* or *measurement in the standard basis*.
3. 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 $|\beta|^2$.
4. In the previous example, if the outcome was 0, then all subsequent measurements in the computational basis will give $|0\rangle$.⁹ *Why?*
5. 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 interpreted as elements of state space for spin 1/2 particles, the orthonormal basis $\{|0\rangle, |1\rangle\}$ corresponds to measurement of spin in the z -direction typically by means of a Stern-Gerlach apparatus as in Section 2.3, with outcomes spin 1/2 (i.e. spin \uparrow , i.e. new state $|0\rangle$) or spin -1/2 (i.e. spin \downarrow , i.e. new state $|1\rangle$).

Similarly, as we will see later, all measurements for spin 1/2 particles correspond to measuring spin in some direction in \mathbb{R}^3 , and the two mutually exhaustive and exclusive outcomes for each such measurement correspond to an orthonormal basis in \mathbb{C}^2 .¹⁰

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

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.

⁹To be precise, this assumes the particle was not disturbed by the measurement process. Typically this would mean a 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 $|0\rangle$.

¹⁰If the measurement direction has spherical coordinates (θ, ϕ) in \mathbb{R}^3 , then the orthonormal basis in \mathbb{C}^2 is

$$\left\{ \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} e^{i\phi} |1\rangle, \quad \cos \frac{\pi - \theta}{2} |0\rangle + \sin \frac{\pi - \theta}{2} e^{i(\phi + \pi)} |1\rangle = \sin \frac{\theta}{2} |0\rangle - \cos \frac{\theta}{2} e^{i\phi} |1\rangle \right\},$$

corresponding to the antipodal points in \mathbb{R}^3 with spherical coordinates (θ, ϕ) and $(\pi - \theta, \phi + \pi)$.

6. Conversely to the previous remark, mutually exclusive outcomes of a measurement correspond to orthonormal vectors in state space. Mutually exclusive *and* exhaustive outcomes correspond to an orthonormal *basis*.
7. Since $e^{i\theta} |\psi\rangle = \sum_i (e^{i\theta} \alpha_i) |v_i\rangle$ it follows from Postulate 2 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 .

2.2.2 Measuring Qubits

We say a measurement is with respect to (or is measured by) 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 so 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 Figure 2.4)

By the preliminary version of Postulate 2 this is the same as measurement w.r.t.¹¹ the (ordered) orthonormal basis $\{|v\rangle, |v\rangle^\perp\}$. Thus w.l.o.g.¹² we can consider measurements given by a single ket. Of course, this is only true in \mathbb{C}^2 .

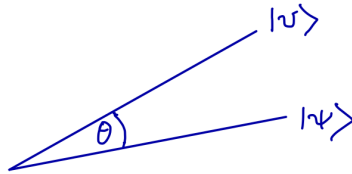


Figure 2.4: A section of \mathbb{C}^2 . $|\psi\rangle$ is measured w.r.t. $|v\rangle$ — the outcome is $|v\rangle$ with probability $\cos^2(\theta)$.

2.2.3 Distinguishing Relative Phase

We noted in Section 2.2.1 Remark 7 that measurements do not distinguish between global phase factors. However, they *can* distinguish relative phase factors such as $e^{i\phi}$ in

$$|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} e^{i\phi} |1\rangle. \quad (2.5)$$

To see this, suppose we have a large number of similarly prepared copies of $|\psi\rangle$. This may be done in the case of spin by passing a large number of spin 1/2 particles through a Stern-Gerlach apparatus pointing in the \mathbb{R}^3 direction with spherical coordinates (θ, ϕ) and blocking those particles which come out of the apparatus with spin down. Of course, if we set up the Stern-Gerlach apparatus ourselves we know (θ, ϕ) and hence the state $|\psi\rangle$. But suppose someone else set up the apparatus and sent the identically prepared qubits on to us. Suppose we are told all the particles are in the same state, but we are not told what is the state.

Step 1. If we measure such a qubit in the computational basis $\{|0\rangle, |1\rangle\}$, the probability that after measurement the qubit will be in the state $|0\rangle$ is $\cos^2 \frac{\theta}{2}$. So we do a frequency count of what happens to a sufficiently large number of such similarly prepared qubits $|\psi\rangle$. This will give a good approximation to $\cos^2 \frac{\theta}{2}$ and hence to θ .¹³

¹¹“w.r.t.” is an abbreviation for “with respect to”

¹²“w.l.o.g.” is an abbreviation for “without loss of generality”.

¹³Remember that $0 \leq \theta \leq \pi$ and so in particular $\cos \frac{\theta}{2}$ is ≥ 0 .

Step 2 Such measurements in the computational basis gave no information about ϕ , only about θ . But suppose we also measure a large number of other copies of $|\psi\rangle$ in the $\{|+\rangle, |-\rangle\}$ basis.¹⁴ From (2.2)

$$|0\rangle = \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle, \quad |1\rangle = \frac{1}{\sqrt{2}}|+\rangle - \frac{1}{\sqrt{2}}|-\rangle. \quad (2.6)$$

So from (2.5) and (2.6)

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(\cos \frac{\theta}{2} + \sin \frac{\theta}{2} e^{i\phi} \right) |+\rangle + \frac{1}{\sqrt{2}} \left(\cos \frac{\theta}{2} - \sin \frac{\theta}{2} e^{i\phi} \right) |-\rangle =: \alpha' |+\rangle + \beta' |-\rangle. \quad (2.7)$$

Therefore

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

So given information about θ from Step 1, we can estimate the relative phase ϕ by measuring a large number of similarly produced qubits $|\psi\rangle$ in the $\{|+\rangle, |-\rangle\}$ basis.

2.2.4 Hermitian Operators and Measurements

Every orthonormal basis $\{v_1, \dots, v_n\}$ in an n -dimensional state space V determines the eigenvectors of an hermitian operator $H : V \rightarrow V$, but not the eigenvalues. The eigenvalues are essentially labels, and do not normally have a canonical physical interpretation in our setting.¹⁵

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

$$V = E_1 \oplus \dots \oplus E_k, \quad H = \sum_i \lambda_i P_i, \quad I = \sum_i P_i, \quad (2.8)$$

where $I : V \rightarrow V$ is the identity operator and $V = E_1 \oplus \dots \oplus E_k$ just says the E_i are mutually orthogonal and every $v \in V$ has a unique decomposition $v = v_1 + \dots + v_k$ with $v_i \in E_i$. See Figure 2.7.

The other important properties of the P_i are

$$P_i = P_i^*, \quad P_i^2 = P_i, \quad P_i P_j = O \quad \text{if } i \neq j, \quad (2.9)$$

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 .

2.2.5 Measurement Postulate (Hermitian Operator Version)

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

Postulate 2. *Suppose $H : V \rightarrow V$ is hermitian with decomposition as in (2.8). 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 .*

¹⁴Not copies we have already measured in the computational basis, of course!! Those qubits will be in the $|0\rangle$ or $|1\rangle$ state, and so of no use for our purpose.

¹⁵However, 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 3. In the finite dimensional setting the eigenvalues could give the energy, after normalisation, of the corresponding eigenstate.

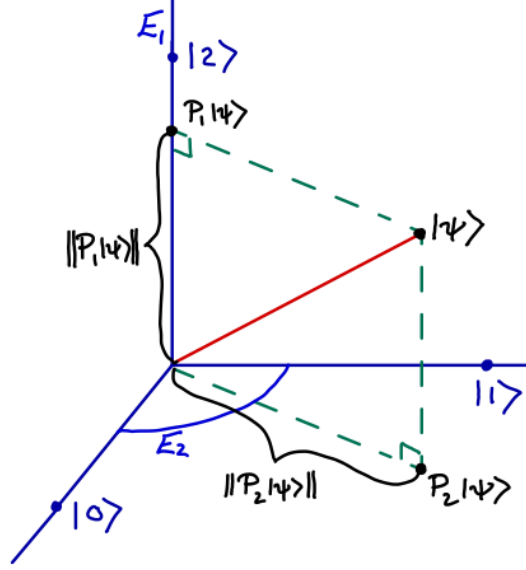


Figure 2.5: Schematic representation for (2.8) of $V = \mathbb{C}^3 = E_1 \oplus E_2$,
 $|\psi\rangle = P_1 |\psi\rangle + P_2 |\psi\rangle$, $\|\psi\rangle\|^2 = \|P_1 |\psi\rangle\|^2 + \|P_2 |\psi\rangle\|^2$.

Remarks

1. It is *NOT* normally 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

$$\begin{aligned} |\psi\rangle &= P_1 |\psi\rangle + \dots + P_k |\psi\rangle && \text{by (2.8)} \\ &= \|P_1 |\psi\rangle\| \frac{P_1 |\psi\rangle}{\|P_1 |\psi\rangle\|} + \dots + \|P_k |\psi\rangle\| \frac{P_k |\psi\rangle}{\|P_k |\psi\rangle\|}, \end{aligned} \quad (2.10)$$

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

$$P_i^* = P_i, \quad P_i^2 = P_i. \quad (2.11)$$

Hence¹⁶

$$p_{|\psi\rangle}(\lambda_i) := \|P_i |\psi\rangle\|^2 = (P_i |\psi\rangle)^* P_i |\psi\rangle = \langle \psi | P_i^* P_i | \psi \rangle = \langle \psi | P_i | \psi \rangle. \quad (2.12)$$

Summarising, the important point is that

$$p_{|\psi\rangle}(\lambda_i) := \|P_i |\psi\rangle\|^2 = \langle \psi | P_i | \psi \rangle. \quad (2.13)$$

5. Note $\sum_i p_{|\psi\rangle}(\lambda_i) = 1$. We can see this two ways:

- Either, from (2.13) and (2.10), $\sum_i p_{|\psi\rangle}(\lambda_i) = \sum_i \|P_i |\psi\rangle\|^2 = \langle \psi | \psi \rangle = 1$,
- or, from (2.13) and (2.8), $\sum_i p_{|\psi\rangle}(\lambda_i) = \sum_i \langle \psi | P_i | \psi \rangle = \langle \psi | \sum_i P_i | \psi \rangle = \langle \psi | I | \psi \rangle = 1$.

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.

¹⁶See Appendix D.4 for the third “=”.

Distinction between the Orthonormal Basis and Hermitian Operator Approaches

The essential aspect of both approaches is the set of eigenspaces, the eigenvalues essentially just being labels.

In the orthonormal basis approach the eigenspaces are all one dimensional. In the hermitian operator approach we allow higher dimensional eigenspaces.

In addition, it is often technically more convenient to deal with an hermitian operator rather than an orthonormal basis.

2.2.6 Pauli Matrices and Measurement

Each of the Pauli matrices (see Appendix E) gives an important example of an hermitian operator on qubits. Denoting eigenspaces by E_j and the corresponding eigenvectors and projection operators by λ_j and P_j , we have the following.

Pauli σ_3

If in (2.8) $H = \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = P_1 - P_2$, then

$$E_1 = \text{span} |0\rangle, P_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \lambda_1 = 1; \quad E_2 = \text{span} |1\rangle, P_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \lambda_2 = -1.$$

The physical realisation of this for spin is measurement of spin in the $\Phi(|0\rangle) = Z$ direction in \mathbb{R}^3 . See Sections 2.3 and 2.3.5, and (F.6).

Pauli σ_1

If in (2.8) $H = \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ then only the algebra is a little more complicated. In this case (see (2.2))

$$E_1 = \text{span} |+\rangle, P_1 = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \lambda_1 = 1; \quad E_2 = \text{span} |-\rangle, P_2 = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \lambda_2 = -1.$$

See Appendix E for the eigenvectors and values of σ_1 . The matrices P_1 and P_2 are the operators $|+\rangle\langle+|$ and $|-\rangle\langle-|$ (see Appendix D.3.2) and may be calculated from this. See Appendix D.3.1:

$$P_1 = |+\rangle\langle+| = \frac{1}{2} (|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

$$P_2 = |-\rangle\langle-| = \frac{1}{2} (|0\rangle\langle 0| - |0\rangle\langle 1| - |1\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

So indeed $\sigma_1 = P_1 - P_2$.

The physical realisation of this for spin is measurement of spin in the $\Phi(|+\rangle) = X$ direction in \mathbb{R}^3 . See Section 2.3 and (F.6).

Pauli σ_2

Similarly, if $H = \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, then

$$E_1 = \text{span} |i\rangle, P_1 = \frac{1}{2} \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}, \lambda_1 = 1; \quad E_2 = \text{span} |-i\rangle, P_2 = \frac{1}{2} \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}, \lambda_2 = -1.$$

See Appendix E for the eigenvectors and values of σ_2 . The projection operators are

$$P_1 = |i\rangle\langle i| = \frac{1}{2} (|0\rangle\langle 0| - i|0\rangle\langle 1| + i|1\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2} \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix},$$

$$P_2 = |-i\rangle\langle -i| = \frac{1}{2} (|0\rangle\langle 0| + i|0\rangle\langle 1| - i|1\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2} \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}.$$

So $\sigma_2 = P_1 - P_2$.

The physical realisation of this for spin is measurement of spin in the $\Phi(|i\rangle) = Y$ direction in \mathbb{R}^3 . See Section 2.3 and (F.6).

2.3 Qubits and Spin 1/2 Particles

For spin 1/2 particles there is a particularly nice physical interpretation of qubit states and of the measurement process.

Initially we will speak of a spin 1/2 particle which *corresponds* to a qubit $|\psi\rangle$, and of the outcome *probability* of various measurements. Later we make these correspondences and probabilities “concrete” on the basis of experimental results and certain frequency counts.

A key point will be that one can produce multiple spin 1/2 particles, all spin-oriented in the same direction, see Sections 2.3.1 and 2.3.3.

2.3.1 Spin-Orientation

In the following a particle (or a qubit) is referred to as having “spin-orientation in the direction \hat{n} ”, or just “spin-orientation \hat{n} ”, where \hat{n} is a unit vector in \mathbb{R}^3 . This is also referred to as spin up, spin \uparrow or spin $+1/2$, in the direction \hat{n} . Spin orientation in the direction $-\hat{n}$, as well as being referred to as spin \uparrow or spin $1/2$ in the direction $-\hat{n}$, may also be referred to as spin down, spin \downarrow or spin $-1/2$ in the direction \hat{n} .

It is important to realise that a particle with “spin-orientation in the direction \hat{n} ” does *not* possess a definite property of “spin-orientation in the \hat{n} -direction” which can be discovered by means of some appropriate measurement.

To the extent such a particle with “spin-orientation \hat{n} ” has a “definite” property it is of the following *probabilistic form*. (See Section 2.3.2 for Stern-Gerlach apparatus and Section 2.3.3 for preparing a large sample of particles with the same spin-orientation.)

Experimental Observation Suppose spin 1/2 particles with the same spin-orientation \hat{n} are measured by a Stern-Gerlach apparatus oriented in direction \hat{k} , and $\eta \in [0, \pi]$ is the angle between \hat{n} and \hat{k} . Then after measurement the spin-orientation is \hat{k} with probability $\cos^2 \frac{\eta}{2}$ and is $-\hat{k}$ with probability $1 - \cos^2 \frac{\eta}{2}$. See Figure 2.6.

These probabilities have been observed repeatedly and extremely accurately.

Note that if $\hat{n} = \hat{k}$ ($\hat{n} = -\hat{k}$) then the first (second) probability is 1 and so in these cases, and only these cases, the outcome is definite.

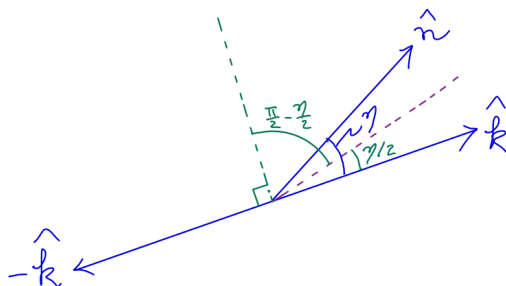


Figure 2.6: Particles with spin-orientation \hat{n} are measured by a Stern-Gerlach apparatus oriented in direction \hat{k} , with angle η between \hat{n} and \hat{k} . After measurement the spin-orientation is \hat{k} with probability $\cos^2 \frac{\eta}{2}$ and $-\hat{k}$ with probability $1 - \cos^2 \frac{\eta}{2}$.

2.3.2 Stern-Gerlach Experimental Set-Up

A Stern-Gerlach apparatus generates a non homogeneous magnetic field in some direction in \mathbb{R}^3 , as in Figure 2.7 where the field is in the vertical or z -direction.

When a spin 1/2 particle is passed horizontally through this magnetic field the particle is either deflected up or down. More accurately, after passing through the field the state of the particle is

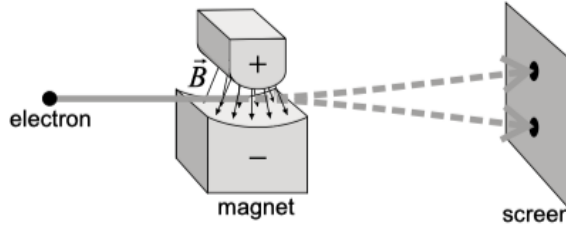


Figure 2.7: Spin 1/2 particle measured by a Stern-Gerlach apparatus. The apparatus has magnetic field, from $-$ to $+$, oriented in the z (vertical) direction, or more generally in an arbitrary \hat{k} direction. ([Lvo18, p203 Fig 4.7]).

in a superposition of spin up \uparrow and spin down \downarrow in the z direction in \mathbb{R}^3 , with the spin amplitudes unchanged from before passing through the field. But there is now correlation with spatial position in a larger Hilbert space — spin up is correlated with being vertically higher and spin down being vertically lower.

When the particle arrives at the screen it has to, in effect, “make a decision” whether to be spin up in the z -direction, i.e. have spin-orientation z , or to be spin down in the z -direction, i.e. have spin-orientation $-z$, and this will be indicated by which of the two possible spots on the screen register as being “hit”.

More generally, replacing z by a unit vector \hat{k} , suppose the Stern-Gerlach apparatus has magnetic field oriented in the \hat{k} direction. Suppose the particle is passed through the magnetic field in a direction orthogonal to \hat{k} . Then the particle after measurement will have spin-orientation in either the \hat{k} or $-\hat{k}$ direction.

The probability for each of the two possible outcomes $\pm\hat{k}$, assuming the initial spin-orientation \hat{n} is known, is given by the Experimental Observation in Section 2.3.1. Note that the two possible outcomes, as opposed to their probabilities, are *independent* of the particles initial spin-orientation direction. (Except that only one outcome is possible if $\hat{n} = \pm\hat{k}$.)

2.3.3 Multiple Particles with the Same Spin-Orientation

From Section 2.3.2, by passing spin 1/2 particles through a Stern-Gerlach apparatus oriented in the direction \hat{k} , we can prepare particles spin-oriented in either the \hat{k} direction or $-\hat{k}$ direction.

However, once the particle hits the detection screen it is destroyed. But if we remove the upper half of the screen in Figure 2.7 then only those particles which “decide” at the time of measurement to have spin-orientation $-\hat{k}$ are destroyed, and the upper “path” gives a stream of particles with spin-orientation \hat{k} .¹⁷

If these particles with spin-orientation \hat{k} are passed through another Stern-Gerlach apparatus oriented in the direction \hat{k} ¹⁸ (or the direction $-\hat{k}$) they will remain with spin-orientation \hat{k} .

Moreover, if these particles are passed through yet another Stern-Gerlach apparatus oriented in the \hat{k}' direction then they will be observed to have spin-orientation $\pm\hat{k}'$ with probabilities in accordance with the Experimental Observation in Section 2.3.1.

2.3.4 Spin-Orientation and Qubit Correspondence: Bloch Sphere Map

In Appendix F we discuss from a *mathematical* perspective the Bloch sphere one-one correspondence Φ between qubits $|\psi\rangle \in \mathbb{C}^2$ and unit vectors $\hat{n} \in \mathbb{R}^3$. This correspondence can be given directly via spherical coordinates as in (F.1)–(F.3) or via the projection map $|\psi\rangle\langle\psi|$ as in (F.7) and (F.11).

¹⁷A critical point concerning quantum theory is that quantum particles do not “make a decision” re physical space until required to do so. Thus the state of the qubit is a superposition of up and down, albeit with different spatial correlations in each case, until measurement time arrives.

¹⁸You may object that after the previous (slight) deflection (in the direction of \hat{k}) the path of the particle is no longer orthogonal to the direction \hat{k} . However, it is possible to apply an intermediate magnetic field to correct for this.

For our purposes we use the projection map correspondence. From (F.7)

$$\Phi|\psi\rangle = (n_1, n_2, n_3) \quad \text{if} \quad |\psi\rangle\langle\psi| = \frac{1}{2}(I + n_1\sigma_1 + n_2\sigma_2 + n_3\sigma_3) =: \frac{1}{2}(I + \hat{n} \cdot \vec{\sigma}), \quad (2.14)$$

where $\hat{n} = (n_1, n_2, n_3) \in S^2 \subset \mathbb{R}^3$ is uniquely determined by $|\psi\rangle$, and the qubit $|\psi\rangle$ is uniquely determined by \hat{n} . Recall also that the correspondence between qubits and projection maps $|\psi\rangle\langle\psi|$ is one-one, see the first paragraph of Appendix F.3.

So after we fix an orthonormal frame in space to correspond to the unit vectors $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1) \in \mathbb{R}^3$ we also regard the Bloch Sphere map Φ as a correspondence between qubits in \mathbb{C}^2 and directions in space.

2.3.5 Spin-Orientation and Qubit Correspondence: Physical Justification

So finally we come to the key idea.

All spin measurements are essentially of the Stern-Gerlach type. By the Experimental Observation in Section 2.3.1, for a fixed measurement direction, the observable consequences (which as we have emphasised are probabilistic) are precisely determined by the measurement direction and the spin-orientation of the particle being measured.

For this reason we want to identify qubit states with spin-orientations, i.e. with unit vectors in \mathbb{R}^3 , or equivalently with directions in physical space after choosing an orthonormal frame. One candidate for such an identification is the Bloch Sphere map. The intention here is to justify on physical/experimental grounds that this must be the correct identification.

Consider two unit directions in physical space, \hat{n} and \hat{k} , and let $\eta \in [0, \pi]$ be the angle between them.

Let $|\hat{n}\rangle$ and $|\hat{k}\rangle$ be the qubits in \mathbb{C}^2 such that $\Phi|\hat{n}\rangle = \hat{n}$ and $\Phi|\hat{k}\rangle = \hat{k}$. Then *the angle between $|\hat{n}\rangle$ and $|\hat{k}\rangle$ is $\eta/2$* from the angle doubling proposition in Appendix F.4.

From Section 2.3.1 recall:

Experimental Observation. If spin 1/2 particles are prepared with spin-orientation \hat{n} and are measured by a Stern-Gerlach apparatus oriented in the direction \hat{k} , then the spin-orientation after measurement is \hat{k} with probability $\cos^2 \frac{\eta}{2}$ and is $-\hat{k}$ with probability $1 - \cos^2 \frac{\eta}{2}$. See Figure 2.6.

From the version of the measurement postulate at the end of Section 2.2.1 under the heading “Measuring Qubits” we have, since the angle between $|n\rangle$ and $|k\rangle$ is $\eta/2$:

Measurement Postulate Prediction. If $|\hat{n}\rangle$ is measured by (i.e. is measured w.r.t.) $|\hat{k}\rangle$, then the state after measurement is $|\hat{k}\rangle$ with probability $\cos^2 \frac{\eta}{2}$ and is $-\hat{k}\rangle$ with probability $1 - \cos^2 \frac{\eta}{2}$.

Conclusion: Through the Bloch sphere mapping $|\hat{n}\rangle \leftrightarrow \hat{n} = \Phi|\hat{n}\rangle$, all experimental observations (for probabilities via frequency counts) agree with the measurement postulate prediction for the corresponding qubits. Since the predicted results of measurements are essentially all we can know about state space, the Bloch sphere map is the correct and unique identification between state space and spin orientations in physical space.

2.4 Evolution

2.4.1 Evolution Postulate (Discrete Version)

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

$$U(\alpha|v\rangle + \beta|w\rangle) = \alpha U|v\rangle + \beta U|w\rangle, \quad \|U|v\rangle\| = 1. \quad (2.15)$$

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

Examples

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

Here are three examples of unitary matrices, each acting on the state space $V = \mathbb{C}^2$. It is easy to see they are unitary, for example the two columns in each case are of unit length and orthogonal to one another.

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

$$\text{---} \boxed{X} \text{---} \quad \text{or} \quad \alpha|0\rangle + \beta|1\rangle \text{---} \boxed{X} \text{---} \beta|0\rangle + \alpha|1\rangle, \quad (2.16)$$

with the input and output indicated in the second case.

As for all quantum circuits, this should be interpreted as change in time, moving from left to right.

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

$$\text{---} \boxed{Z} \text{---} \quad \text{or} \quad \alpha|0\rangle + \beta|1\rangle \text{---} \boxed{Z} \text{---} \alpha|0\rangle - \beta|1\rangle. \quad (2.17)$$

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

$$\text{---} \boxed{H} \text{---} \quad . \quad \text{Note } |0\rangle \text{---} \boxed{H} \text{---} |+\rangle, \quad |1\rangle \text{---} \boxed{H} \text{---} |-\rangle. \quad (2.18)$$

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 also Appendix E equation (E.4).

2.4.2 Evolution Postulate (Continuous Version)

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*

$$\frac{d}{dt} |\psi(t)\rangle = -iH |\psi(t)\rangle, \quad (2.19)$$

where $|\psi(t)\rangle$ is the state of the system at time t . H is a time-independent hermitian operator known as the Hamiltonian of the system.

Remarks

1. In (2.23) we see that the solution of Schrödinger's equation for a given $|\psi(0)\rangle$ is $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ where $U(t) = \exp(-itH)$. Moreover, $U(t)$ is unitary for all t . This is consistent with the discrete time version in Section 2.4.1.
2. It is also clear from the matrix representation (G.3) that *every* unitary matrix U can be written in the form $U = \exp(-iH)$ for some Hamiltonian H . So there is a one-one correspondence between the discrete and continuous time descriptions. In understanding the theory of quantum computation, as opposed to its experimental set-up, one is usually concerned with the discrete time formulation.
3. 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 $\hbar = 1$.
4. The Hamiltonian H is *not* the Hadamard matrix/gate H !
5. In quantum computing systems, the hermitian operators in Schrödinger's equation are obtained by applying lasers, magnetic fields, etc., appropriately oriented and turned on or off for various time intervals, typically picoseconds¹⁹. 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 equation gives a good approximation to what happens in the original system.

Solution of Schrödinger's equation for *time-independent* Hamiltonian

Since the Hamiltonian H in (2.19) is time-independent, in terms of an orthonormal basis of eigenvectors for H we can write

$$|\psi\rangle = |\psi(t)\rangle = \begin{bmatrix} \psi_1(t) \\ \vdots \\ \psi_n(t) \end{bmatrix}, \quad H = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}. \quad (2.20)$$

It follows that (2.19) is equivalent to the following simple system of independent ordinary differential equations:

$$\frac{d}{dt} \psi_j(t) = -i\lambda_j \psi_j(t), \quad j = 1, \dots, n. \quad (2.21)$$

This has the solution

$$\psi_j(t) = e^{-i\lambda_j t} \psi_j(0). \quad (2.22)$$

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

$$|\psi(t)\rangle = U(t) |\psi(0)\rangle, \quad \text{where } U(t) := e^{-itH} = \begin{bmatrix} e^{-i\lambda_1 t} & & \\ & \ddots & \\ & & e^{-i\lambda_n t} \end{bmatrix}. \quad (2.23)$$

Note that $U(t)$ is indeed unitary. See Section G.4.

Example

As a prototypical example consider the Hamiltonian

$$H = -\frac{\omega}{2} \sigma_3 = \begin{bmatrix} -\frac{\omega}{2} & 0 \\ 0 & \frac{\omega}{2} \end{bmatrix}. \quad (2.24)$$

Here σ_3 is the Pauli matrix discussed briefly in Section 2.4.1 Example 2. The factor ω is the angular velocity in (2.26) with the factor $-1/2$ chosen in (2.24) accordingly.

¹⁹A picosecond is 10^{-12} of a second.

Let the arbitrary initial state, as in (2.4), be

$$|\psi(0)\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + \sin\left(\frac{\theta}{2}\right) e^{i\phi} |1\rangle. \quad (2.25)$$

Since H is already in diagonal form, from (2.23)

$$\begin{aligned} |\psi(t)\rangle &= e^{-itH} |\psi(0)\rangle = \begin{bmatrix} e^{i\frac{\omega}{2}t} & 0 \\ 0 & e^{-i\frac{\omega}{2}t} \end{bmatrix} \begin{bmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} e^{i\phi} \end{bmatrix} \\ &= \begin{bmatrix} \cos\frac{\theta}{2} e^{i\frac{\omega}{2}t} \\ \sin\frac{\theta}{2} e^{i(\phi-\frac{\omega}{2}t)} \end{bmatrix} \sim \begin{bmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} e^{i(\phi-\omega t)} \end{bmatrix} \quad (\text{removing the (time dependent) global phase factor } e^{i\frac{\omega}{2}t}) \\ &= \cos\left(\frac{\theta}{2}\right) |0\rangle + \sin\left(\frac{\theta}{2}\right) e^{i(\phi-\omega t)} |1\rangle. \end{aligned} \quad (2.26)$$

Thus in \mathbb{C}^2 the initial state $|\psi(0)\rangle$ is fixed in the $|0\rangle$ complex line and rotated with angular velocity ω clockwise in the $|1\rangle$ complex line. Compare with Figure G.1.

In terms of the Bloch sphere representation the initial state $|\psi(0)\rangle$ is at latitude θ radians down from the north pole and is rotated anticlockwise around the north pole with angular velocity ω in the westerly direction.

2.5 Composite System

2.5.1 Introduction

Consider two classical bits, which we refer to as the first and second bit. Suppose they have interacted some time in the recent past in a manner we may not know.

Their possible values are 00, 01, 10 and 11. For example, 10 is the case in which the first bit has value 1 and the second has value 0.

Guided by the State Space Postulate in Section 2.1.1, and Remarks 5 and 6 in Section 2.2.1, we might hope to model the quantum situation for a *pair of qubits* by considering \mathbb{C}^4 with basis vectors

$$|00\rangle := \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad |01\rangle := \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad |10\rangle := \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad |11\rangle := \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad (2.27)$$

An arbitrary quantum qubit pair would be given by

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle = \begin{bmatrix} \alpha_{00} \\ \alpha_{01} \\ \alpha_{10} \\ \alpha_{11} \end{bmatrix}, \quad \alpha_{jk} \in \mathbb{C}, \quad \sum |\alpha_{jk}|^2 = 1. \quad (2.28)$$

Moreover, $|\psi\rangle$ and $e^{i\phi}|\psi\rangle$ should be physically indistinguishable and so represent the same qubit.

Indeed, this is correct. The consequences and predictions are very counterintuitive, but have been borne out in innumerable experiments.

2.5.2 Composite System Postulate

Postulate 4. *Suppose we have N physical systems with state spaces V_1, \dots, V_N respectively. Then the state space V of the composite physical system is given by the tensor product $V = V_1 \otimes \dots \otimes V_N$. If the i th system is in the state $|\psi_i\rangle$ then the state of the composite system is $|\psi_1\rangle \otimes \dots \otimes |\psi_N\rangle$.*

Tensor Products

The vector space $V = V_1 \otimes V_2$ is the *tensor product* of V_1 and V_2 . The state $|\psi_1\rangle \otimes |\psi_2\rangle$ is the *tensor product* of $|\psi_1\rangle \in V_1$ and $|\psi_2\rangle \in V_2$.

See Appendix H. But here we will need to know little about tensors beyond the notation and the following discussion.²⁰

Just think of $|\psi_1\rangle \otimes |\psi_2\rangle$ as the state of two qubits where the first is in the state $|\psi_1\rangle \in V_1$ and the second is in the state $|\psi_2\rangle \in V_2$.

N -Qubit Space

In quantum computing we will mostly be concerned with N -qubit space

$$\mathbb{C}^{2^N} = \overset{1}{\mathbb{C}^2} \otimes \dots \otimes \overset{N}{\mathbb{C}^2}. \quad (2.29)$$

Note how the dimension 2^N grows exponentially in N .

The canonical basis for $\overset{1}{\mathbb{C}^2} \otimes \dots \otimes \overset{N}{\mathbb{C}^2}$ is in one-one correspondence with the set of all 2^N possible classical N -bits,

$$(0, \dots, 0, 0, 0), \quad (0, \dots, 0, 0, 1), \quad (0, \dots, 0, 1, 0), \quad \dots, \quad (1, \dots, 1, 1, 1), \quad (2.30)$$

²⁰Unfortunately, many courses on general relativity present tensors in an unnecessarily unpleasant and unintuitive manner.

here ordered lexicographically. We write the corresponding 2^N basis vectors as

$$|0\dots 000\rangle, |0\dots 001\rangle, |0\dots 010\rangle, \dots, |1\dots 111\rangle. \quad (2.31)$$

Thus for a single qubit we have canonical basis vectors $\{|0\rangle, |1\rangle\} \subset \mathbb{C}^2$. For a pair of qubits the canonical basis vectors are

$$|00\rangle, |01\rangle, |10\rangle, |11\rangle \subset \mathbb{C}^4. \quad (2.32)$$

For a triplet of qubits the canonical basis vectors in \mathbb{C}^8 are

$$|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle, |101\rangle, |110\rangle, |111\rangle.$$

2-Qubit Space

The most general state for a pair of qubits is a superposition of the 4 basis states (2.32):

$$\alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle \quad \sum_{j,k} |\alpha_{jk}|^2 = 1. \quad (2.33)$$

If the first qubit is prepared in the state $\alpha_0|0\rangle + \alpha_1|1\rangle$ and the second in the state $\beta_0|0\rangle + \beta_1|1\rangle$ then the state of the pair of qubits is represented by

$$(\alpha_0|0\rangle + \alpha_1|1\rangle)(\beta_0|0\rangle + \beta_1|1\rangle) = \alpha_0\beta_0|00\rangle + \alpha_0\beta_1|01\rangle + \alpha_1\beta_0|10\rangle + \alpha_1\beta_1|11\rangle.$$

This is a *product* state and is not very interesting, since we can act on and measure the two qubits independently and nothing particularly new will occur.

Entangled Qubits

It is also possible to prepare a pair of qubits in the Bell state

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \in \mathbb{C}^4. \quad (2.34)$$

This is an *entangled* state, not a product state. See Appendix H.1.3.

It is routine to entangle a pair of qubits in this and other ways on a quantum computer. Moreover, after entanglement the two qubits may be sent to different locations and remain entangled. In the EPR experiment discussed in Section 1.4 entangled photons were separated by over 1200km and remained entangled. In principle, particles could be entangled and remain entangled after one of them is sent off to a different galaxy.

Measuring a Pair of Qubits

By way of example, we consider the Bell pair in (2.34).

According to Quantum postulates, it is possible to measure just the first qubit and consider instantaneous effect on the entangled pair as a whole system, even though the individual qubits may be (in principle) spatially separated by a large distance.

For example, suppose we measure the first qubit in the $\{|0\rangle, |1\rangle\}$ frame, which corresponds in the spin case to measuring the spin (“up” or “down”) in the Z direction in \mathbb{R}^3 .

Measurement on the first qubit is given by σ_3 , on the second by I and on the pair by $\sigma \otimes I$. We discuss this in a little detail in Section ***. But here the effect is straightforward.

Applying measurement σ_3 to a single qubit $\alpha|0\rangle + \beta|1\rangle$ gives new state $|0\rangle$ (outcome 0, or up) with probability $|\alpha|^2$ and new state $|1\rangle$ with probability $|\beta|^2$. That is, the qubit is “projected” onto span $|0\rangle$ with probability $|\alpha|^2$ and onto span $|1\rangle$ with probability $|\beta|^2$.

Applying $\sigma \otimes I$ to $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \in \mathbb{C}^4 = \mathbb{C}^2 \otimes \mathbb{C}^2$ gives $|00\rangle$ with probability 1/2 and $|11\rangle$ with probability 1/2.

If the outcome of the first measurement is 0 and new state is hence $|0\rangle$ then the state of the second qubit is also immediately $|0\rangle$, even if the qubits are very distant from each other. Similarly if the

outcome for the first qubit is 1 and new state is hence $|1\rangle$, then the state of the second qubit is is also immediately $|1\rangle$

This was very puzzling to Einstein, “spooky action at a distance”, even though technically it did not contradict relativity or casuality, as we will discuss later.

A nice property of $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$

Let $U = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}$ be any *real* rotation in \mathbb{C}^2 . ($\alpha, \beta \in \mathbb{R}$, $\alpha^2 + \beta^2 = 1$.) Then $U^{-1} = U^* = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$.

Let $|v\rangle = U|0\rangle$ and $|v^\perp\rangle = U|1\rangle$ be any orthonormal basis obtained by a real rotation of the standard basis.

Then

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|vv\rangle + |v^\perp v^\perp\rangle). \quad (2.35)$$

Proof. $|v\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|v^\perp\rangle = -\beta|0\rangle + \alpha|1\rangle$. Hence

$$\begin{aligned} \frac{1}{\sqrt{2}}(|vv\rangle + |v^\perp v^\perp\rangle) &= \frac{1}{\sqrt{2}}\left((\alpha|0\rangle + \beta|1\rangle)(\alpha|0\rangle + \beta|1\rangle) + (-\beta|0\rangle + \alpha|1\rangle)(-\beta|0\rangle + \alpha|1\rangle)\right) \\ &= \frac{1}{\sqrt{2}}\left((\alpha^2 + \beta^2)|00\rangle + (\alpha^2 + \beta^2)|11\rangle\right) = |\Phi^+\rangle. \end{aligned}$$

□

This means that if we measure the first qubit of $|\Phi^+\rangle$ in *any* frame $\{|v\rangle, |v^\perp\rangle\}$ obtained by a real rotation of the standard frame, we obtain $|vv\rangle$ or $|v^\perp v^\perp\rangle$ with equal probability 1/2. This is surprising! Knowledge of one qubit in any rotated frame completely determines the other qubit.

Missing Information

If we consider a coin toss, it may seem random that it tuns up heads or tails. But *in principle*, if we had complete information, we could predict the outcome.

Einstein looked for more information that *in principle* would predict the outcome of quantum experiments.

Note that in the case of $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ this additional information (“hidden variables”) could be such that at the time of the entanglement for say 1000 pairs, the additional information for half of the pairs was that both members of the pair would give ”spin up” in the Z direction, and for the other half of the pairs both would give spin down. This does not really solve the problem if we measure in other directions, but may point in the correct direction.

If one includes information for all possible outcomes, one contradicts the inbuilt probability notions in QM. For example. it a qubit will definitely give spin up in the Z direction then it will in X direction give spin up exactly half the time and spin down exactly 1/2 of the time.

Bohr and Einstein discussed/argued this for many decades, beginning 1927 or so, with most physicists agreeing with Bohr, not concerning themselves with the foundations, and dismissing continued discussion as mere pilosophical rumination. Einstein agreed that much had been achieved, but felt the foundations were still lacking.

In 1935 Einstein sharpened his arguments in the Einstein, Podolsky, Rosen paper. I will give two versions

Then Bell came along. Essentially the difficulty was to explain the high degree of correlations in entangled particles. In 1964 he wrote a paper coming up with an experiment to decide between QM explanations and any potential explanation via classical means (including probabilistic ones).

Beginning in 1972 these experiments were conducted with increasing efficiency and closing all conceivable “loopholes”. We discussed the latest in Chapter 1. QM wins convincingly.

I will discuss the two versions of EPR, Bell’s inequalities and the subsequent physical experiments.

2.6 EPR, Bell, CHSH

2.6.1 Bell's Inequality

A pair of qubits is entangled as $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, the first is sent to Alice and the second to Bob, with a long distance between them. See Figures 1.5 and 1.6. The experiment is repeated many times, say 10,000.

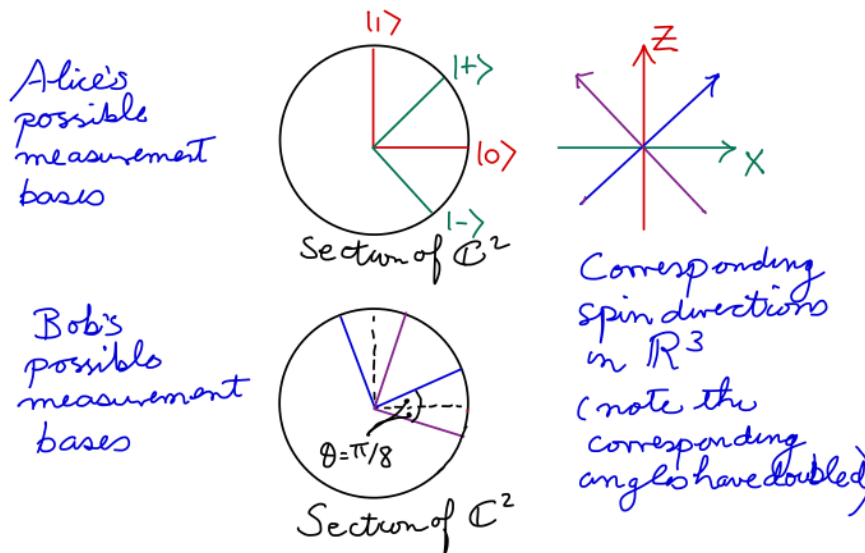


Figure 2.8: The Experiment

Now see Figure 2.8

Alice has a choice of two measurement bases, $\{|0\rangle, |1\rangle\}$ or $\{|+\rangle, |-\rangle\}$. In the case of spin particles this corresponds to measuring the particles in the Z or X directions. Note the second basis is obtained by rotating the first (standard basis) by $\pi/4$. The corresponding oriented directions for measuring spin differ by $\pi/2$ (double the angle).

Bob has a choice of two measurement bases, the $\pi/8$ basis obtained by rotating the standard basis by $\pi/8$ and the $-\pi/8$ basis obtained by rotating the standard basis by $-\pi/8$. The corresponding directions for spin measurement are “half way” between the oriented X and Z directions in the first case, and half way between the oriented X and $-Z$ directions in the second case.

The measurement bases/directions are chosen randomly and independently with probability $1/2$ in each case for Alice and probability $1/2$ in each case for Bob.

Each time the experiment is conducted Alice notes the choice of measurement and the outcome, and Bob does the same. The experiments are done in such a way that Bob cannot know the choice made by Alice, nor can Alice know the choice made by Bob.

We know that if Bob and Alice had chosen the same measurement frame / direction of measurement (which is not actually possible in this experiment) they would both have the same result (same state and same outcome spin up, or spin down).

In the situation here, there are four possibilities for the combined choices of Alice and Bob, with frame noted first and corresponding direction in \mathbb{R}^3 second.

$A = \{ 0\rangle, 1\rangle\}$ Z red	$B = \{ \pi/8\rangle, 5\pi/8\rangle\}$ blue
$A = \{ 0\rangle, 1\rangle\}$ Z red	$B = \{ -\pi/8\rangle, 3\pi/8\rangle\}$ purple
$A = \{ +\rangle, -\rangle\}$ X green	$B = \{ \pi/8\rangle, 5\pi/8\rangle\}$ blue
$A = \{ +\rangle, -\rangle\}$ X green	$B = \{ -\pi/8\rangle, 3\pi/8\rangle\}$ purple

Note that in the first 3 cases the two frames differ only by $\pi/8$ (or $\pi/4$ for the directions of spin measurement). In the last case the difference is $3\pi/8$ (or $3\pi/4$ for directions of spin measurement)

In the first 3 cases, if Alice measures first then the outcome is 0 or 1 (or whatever labels we use) and the new new state is either $|0\rangle$ or $|1\rangle$. Since Bob's frame is just rotated by $\pm\pi/8$ from Alice's frame, with probability

$$\cos^2 \frac{\pi}{8} = \frac{1}{2} \left(1 + \cos \frac{\pi}{4} \right) = \frac{2 + \sqrt{2}}{4} \approx .86.$$

the outcome will be the same 0 or 1. But in the fourth case, the outcomes will be the same with probability $\cos^2(3\pi/8) = \sin^2(\pi/2 - 3\pi/8) = 1 - \cos^2(\pi/8)$, and hence differ with probability $\cos^2(\pi/8)$.

So with probability $\approx .86$ we have correlation in the first 3 cases and anti-correlation in the fourth case.

The same applies if Bob measures first.

This can also be obtained by applying the hermitian operators $\sigma_3 \otimes \frac{1}{\sqrt{2}}(\sigma_3 + \sigma_1)$ in the first case, and similarly for the others.

What is the best we can expect from a classical information? Even with hidden classical information and hidden classical probabilities, the best we can expect for these correlations is a probability of .75.

Remember that Bob cannot know what experiment Alice chooses, and conversely.

If the "hidden information" is always output 0 then there will be agreement in 3/4 of the cases, the exception being those where the fourth case of experiment choice arose. (One could also have "hidden information" that says always output 1.)

Chapter 3

Quantum Circuits

Contents

3.1 Quantum Gates	39
3.2 No Cloning Theorem	39
3.3 Teleportation of Qubits	39
3.4 Superdense Coding	39
3.5 Deutsch-Jozsa Algorithm	39

- 3.1 Quantum Gates
- 3.2 No Cloning Theorem
- 3.3 Teleportation of Qubits
- 3.4 Superdense Coding
- 3.5 Deutsch-Jozsa Algorithm

Chapter 4

Composite Systems

Contents

4.1	Ensembles and Density Operators	40
4.2	The Bloch Ball	40
4.3	Subsystems and the Reduced Density Operator	40
4.4	Schmidt Decomposition and Purification	40
4.5	Bell's Inequality, CHSH Inequality, GHZ Inequality	40

4.1 Ensembles and Density Operators

4.2 The Bloch Ball

4.3 Subsystems and the Reduced Density Operator

4.4 Schmidt Decomposition and Purification

4.5 Bell's Inequality, CHSH Inequality, GHZ Inequality

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 $n \geq 2$.

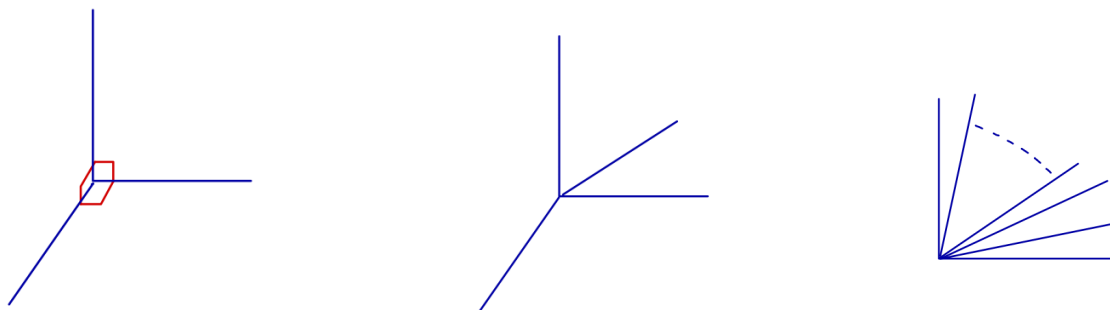


Figure A.1: 3 mutually orthogonal lines in \mathbb{R}^3 , 4 in \mathbb{R}^4 , n in \mathbb{R}^n .

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 .

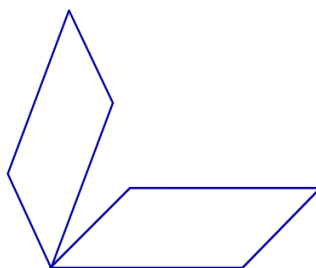


Figure A.2: 2 orthogonal 2-D real subspaces of \mathbb{R}^4 , or 2 orthogonal 1-D complex lines in \mathbb{C}^2 .

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 .

¹A *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

Hermitian and Unitary Operators

Contents

B.1 Inner Product Space	43
B.1.1 Inner Product	43
B.1.2 Norm	44
B.1.3 Norms and Inner Products	44
B.1.4 Angle	44
B.2 Operator Definitions in Terms of a “Good” Orthonormal Basis	45
B.2.1 Hermitian Operator	45
B.2.2 Unitary Operator	45
B.2.3 Normal Operator	46
B.3 Definitions in Terms of Adjoint Operator	46
B.3.1 Adjoint	46
B.3.2 Hermitian Operator	47
B.3.3 Unitary Operator	47
B.3.4 Normal Operator	47
B.4 Polar Decomposition	47
B.5 Another Characterisation	47

I summarise the main results for linear operators $T : \mathbb{C}^n \rightarrow \mathbb{C}^n$ ($T : \mathbb{R}^n \rightarrow \mathbb{R}^n$) 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 real or complex finite dimensional spaces V , often \mathbb{R}^n or \mathbb{C}^n .

B.1.1 Inner Product

An *inner product space* is a real or complex vector space together with an associated inner product $\langle \cdot, \cdot \rangle$. 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

- $\langle u, v \rangle = u_1 v_1 + \cdots + u_n v_n$
- $\langle u, v \rangle = u_1 \bar{v}_1 + \cdots + u_n \bar{v}_n$ (maths) or $\langle u, v \rangle = \bar{u}_1 v_1 + \cdots + \bar{u}_n v_n$ (physics, comp sci)

B.1.2 Norm

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 $|\langle u, v \rangle| \leq \|u\| \|v\|$.

B.1.3 Norms and Inner Products

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

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

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.1.4 Angle

The *angle* θ between u and v in a *real* inner product space is defined by $\cos \theta = \frac{\langle u, v \rangle}{\|u\| \|v\|}$ with $\theta \in [0, \pi]$, and hence by $\cos \theta = \langle u, v \rangle$ if u and v are unit vectors.

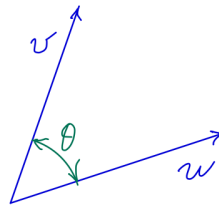


Figure B.1: $u, v \in \mathbb{R}^n$ are unit vectors, $\langle u, v \rangle = \theta \in [0, \pi]$ is the angle between them.

For *complex* inner product spaces the *angle* between u and v is defined by $\cos \theta = \frac{|\langle u, v \rangle|}{\|u\| \|v\|}$ with $\theta \in [0, \pi/2]$, and in particular by $\cos \theta = |\langle u, v \rangle|$ if u and v are unit vectors.

The definition of θ is invariant under replacing u by αu and v by βv for any $\alpha, \beta \in \mathbb{C}$. In particular, the definition is invariant under “phase changes”, i.e. multiplication by $e^{i\eta}$ and $e^{i\eta'}$ for u and v respectively. So the angle is really an angle between the one-dimensional complex (and hence real two-dimensional) subspaces $\text{Span } u$ and $\text{Span } v$.¹

¹For a connection with angles in real vector spaces, consider the natural correspondence between \mathbb{C}^n and \mathbb{R}^{2n} given by

$$\mathbb{C}^n \ni u = (x_1 + iy_1, \dots, x_n + iy_n) \longleftrightarrow (x_1, y_1, \dots, x_n, y_n) =: \tilde{u} \in \mathbb{R}^{2n}. \quad (\text{B.1})$$

Then it is easy to check that

$$\text{Re}\langle u, v \rangle = \langle \tilde{u}, \tilde{v} \rangle_{\mathbb{R}} =: \langle u, v \rangle_{\mathbb{R}}, \quad (\text{B.2})$$

where $\text{Re}\langle u, v \rangle$ is the real part of $\langle u, v \rangle$, the vectors $\tilde{u}, \tilde{v} \in \mathbb{R}^{2n}$ correspond to $u, v \in \mathbb{C}^n$ respectively, and $\langle u, v \rangle_{\mathbb{R}} := \langle \tilde{u}, \tilde{v} \rangle_{\mathbb{R}}$ is the \mathbb{R}^{2n} (real) inner product.

Proposition. *Let u and v be unit vectors in \mathbb{C}^n and let $\cos \theta = |\langle u, v \rangle|$ where $\theta \in [0, \pi]$.*

Then θ is the least angle $\zeta \in [0, \pi/2]$ such that $\cos \zeta = \langle u', v' \rangle_{\mathbb{R}}$ for unit vectors $u' \in \text{Span } u$ and $v' \in \text{Span } v$.

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.

$$T = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \quad (\text{B.4})$$

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

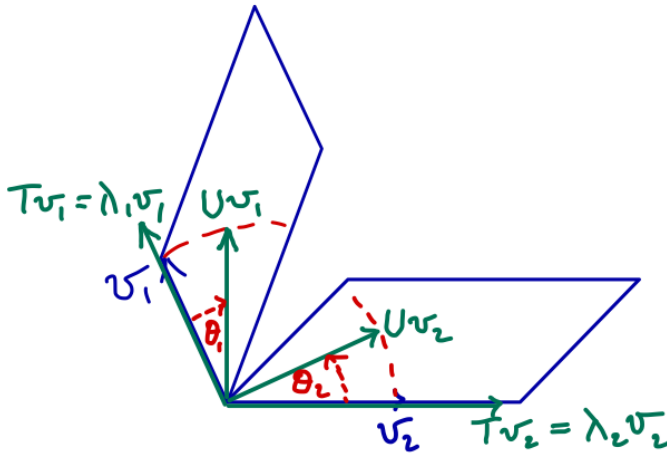


Figure B.2: Hermitian operator T and unitary operator U on \mathbb{C}^2 , $v_1 \in \text{span } |0\rangle$, $v_2 \in \text{span } |1\rangle$.

Figure B.2 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) anticlockwise² through the angle θ_1 .

B.2.2 Unitary Operator

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

Proof. Just as $|z| = \sup_{\eta} \text{Re}(e^{i\eta} z)$ for $z \in \mathbb{C}$, in the same manner

$$|\langle u, v \rangle| = \sup_{\eta, \eta'} \text{Re} \langle e^{i\eta} u, e^{i\eta'} v \rangle = \max_{\eta, \eta'} \text{Re} \langle e^{i\eta} u, e^{i\eta'} v \rangle = \max_{\eta, \eta'} \langle e^{i\eta} u, e^{i\eta'} v \rangle_{\mathbb{R}}. \quad (\text{B.3})$$

The first max is achieved, since $\text{Re} \langle e^{i\eta} u, e^{i\eta'} v \rangle$ is continuous and (η, η') can be restricted to the compact set $[0, 2\pi] \times [0, 2\pi]$. The last equality is from (B.2).

Assuming the max is achieved at η and η' , let $u' = e^{i\eta} u$ and $v' = e^{i\eta'} v$, and define $\theta \in [0, \pi]$ by $\cos \theta = \langle u', v' \rangle_{\mathbb{R}}$. Note $\langle u', v' \rangle_{\mathbb{R}} \geq 0$ (why?) and so in fact $\theta \in [0, \pi/2]$. \square

²Note that there is an orientation on $\text{span } |0\rangle$ given by complex multiplication.

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.

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

$$T = \begin{bmatrix} \lambda_1 e^{i\theta_1} & & \\ & \ddots & \\ & & \lambda_n e^{i\theta_n} \end{bmatrix} \tag{B.5}$$

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 \implies 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 $T : \mathbb{C}^n \rightarrow \mathbb{C}^n$. 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 \Leftarrow 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 $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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

Contents

C.1 Determinant	49
C.2 Trace	50

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 defined¹, for example, by

$$\det A = \sum_{\pi \in \Pi\{1, \dots, n\}} (-1)^\pi A_{1\pi(1)} \cdots A_{n\pi(n)}, \quad (\text{C.1})$$

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

Important properties are

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

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 $\lambda_1, \dots, \lambda_n$ counted with multiplicities, then

$$\det(A) = \lambda_1 \cdots \lambda_n. \quad (\text{C.3})$$

Moreover, if $A : V \rightarrow V$ 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

$$\text{vol}[Av_1, \dots, Av_n] = \det(A) \times \text{vol}[v_1, \dots, v_n], \quad (\text{C.4})$$

where $[w_1, \dots, w_n]$ is the parallelpiped spanned by w_1, \dots, w_n and “vol” is the oriented volume.

¹The most natural, coordinate-free and geometric way is to use exterior algebra and define $\det A$ for an operator $A : V \rightarrow V$ 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 $\{v_1, \dots, v_n\}$ of V .

C.2 Trace

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

$$\text{tr}(A) = \sum_i A_{ii}. \quad (\text{C.5})$$

It is straightforward to check

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

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,

$$\text{tr}(AB) \neq \text{tr}(A)\text{tr}(B). \quad (\text{C.7})$$

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

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

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

$$\text{tr}(A) = \lambda_1 + \dots + \lambda_n. \quad (\text{C.8})$$

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

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

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

Hence

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

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

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

Appendix D

Dirac Bra-Ket Notation

Contents

D.1 Kets and Bras	51
D.2 More on Inner Products	52
D.3 Outer Products	52
D.3.1 Definition and Matrix Representation	52
D.3.2 Orthogonal Projections	52
D.3.3 Operator as a Sum of Dyads	53
D.4 Computing with Dirac Notation	53

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 $|v\rangle = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$, $|w\rangle = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$, etc.

Examples in \mathbb{C}^2 are

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

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,

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

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 $[v_1^* \cdots v_n^*]$, 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, $\langle v| = |v\rangle^* = [v_1^* \cdots v_n^*]$ as a row vector or $1 \times n$ matrix.

D.2 More on Inner Products

Let $A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ 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 $[v_1^* \cdots v_n^*]$ 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:

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

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,

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

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 Representation

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

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

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,

$$|v\rangle\langle w| = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} [w_1^* \cdots w_n^*] = \begin{bmatrix} v_1 w_1^* & \cdots & v_1 w_n^* \\ \vdots & \ddots & \vdots \\ v_n w_1^* & \cdots & v_n w_n^* \end{bmatrix}. \quad (\text{D.6})$$

In particular, if e_1, \dots, e_n is an orthonormal basis for V then $|e_i\rangle\langle e_j|$ has matrix whose ij entry is 1 and every other entry is 0.

D.3.2 Orthogonal Projections

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

$$|v\rangle\langle v| \quad (\text{D.7})$$

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

$$\sum_{i=1}^k |v_i\rangle\langle v_i| \quad (\text{D.8})$$

is orthogonal projection onto W .¹

Orthogonal projections are hermitian, since there exists an orthonormal base of eigenvectors and the eigenvalues are either 0 or 1.

¹ $\sum_{i=1}^k |v_i\rangle\langle v_i|$ is independent of choice of orthonormal basis for W . To see this, note we can write $w \in V$ uniquely in the form $w_1 + w_2$ where $w_1 \in W$ and $w_2 \in W^\perp$. Now check $(\sum_{i=1}^k |v_i\rangle\langle v_i|)w = w_1$ by expressing w_1 and w_2 in terms of an orthonormal basis for V which extends $\{|v_1\rangle, \dots, |v_k\rangle\}$.

D.3.3 Operator as a Sum of Dyads

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

$$I = \sum_{i=1}^n |v_i\rangle\langle v_i|. \quad (\text{D.9})$$

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

$$M = \sum_{i,j} M_{ij} |i\rangle\langle j|, \quad M_{ij} = \langle i|M|j\rangle. \quad (\text{D.10})$$

If H is hermitian with (necessarily real) eigenvalues λ_j and corresponding orthonormal eigenvectors $|v_j\rangle$ (forming a basis for V) then

$$H = \sum_j \lambda_j |v_j\rangle\langle v_j|. \quad (\text{D.11})$$

Similarly, if U is unitary with eigenvalues $e^{i\theta_j}$ and corresponding orthonormal eigenvectors $|v_j\rangle$ (forming a basis for V) then

$$U = \sum_j e^{i\theta_j} |v_j\rangle\langle v_j|. \quad (\text{D.12})$$

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 expressions, 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.²

²See (2.12), (D.3), (D.4), (F.9), (F.12), ...

Appendix E

The Pauli Matrices

Exercise: Verify all statements in the following.

E.1 Definitions, Basic properties, Eigenvalues and Eigenvectors

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

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

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. $\text{tr}(\sigma_i) = 0$, and $\det(\sigma_i) = -1$. It follows that they each have eigenvalues ± 1 .

They satisfy the following identities.¹

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

Using the notation of (2.2)

- 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.

E.2 As Basis Vectors

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

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

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

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

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

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

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

for some $(n_1, n_2, n_3) \in \mathbb{R}^3$. This gives a vector space isomorphism⁴ between \mathcal{H}_0 and \mathbb{R}^3 , which will be particularly useful in terms of the Bloch sphere map.

E.3 Pauli Vector

E.3.1 Definitions and Examples

Although one can interpret $\bar{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ as a vector of Pauli matrices called the *Pauli vector*, we just need the expression $\hat{n} \cdot \bar{\sigma}$ as an abbreviation:

$$\hat{n} \cdot \bar{\sigma} := n_1\sigma_1 + n_2\sigma_2 + n_3\sigma_3 = \begin{bmatrix} n_3 & n_1 - in_2 \\ n_1 + in_2 & -n_3 \end{bmatrix}, \quad \text{where } \hat{n} = (n_1, n_2, n_3) \in \mathbb{R}^3. \quad (\text{E.5})$$

Unless otherwise clear from context, we assume

$$\hat{n} = (n_1, n_2, n_3) \text{ is a unit vector.}$$

Examples are $\hat{n} = (1, 0, 0)$ so $\hat{n} \cdot \bar{\sigma} = \sigma_1$; similarly for σ_2 and σ_3 ; and $\hat{n} = \frac{1}{\sqrt{2}}(1, 0, 1)$ which gives the Hadamard operator $H = \frac{1}{\sqrt{2}}(\sigma_1 + \sigma_3) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

E.3.2 Properties

Proposition.

$$(\hat{n} \cdot \bar{\sigma})^2 = I, \quad \text{tr } \hat{n} \cdot \bar{\sigma} = 0, \quad (\text{E.6})$$

$$\det \hat{n} \cdot \bar{\sigma} = -1, \quad \text{the eigenvalues of } \hat{n} \cdot \bar{\sigma} \text{ are } \pm 1. \quad (\text{E.7})$$

Proof. Using (E.2) it follows $(\hat{n} \cdot \bar{\sigma})^2 = I$. Since $\text{tr } \sigma_i = 0$ for each i , $\text{tr } \hat{n} \cdot \bar{\sigma} = 0$.

Since $\text{tr } \hat{n} \cdot \bar{\sigma} = 0$, the eigenvalues of $\hat{n} \cdot \bar{\sigma}$ are $-\lambda$ and λ for some $\lambda > 0$. (The eigenvalues are real since $\hat{n} \cdot \bar{\sigma}$ is hermitian.) Hence $\det \hat{n} \cdot \bar{\sigma} = -\lambda^2 < 0$.

But $1 = \det I = \det(\hat{n} \cdot \bar{\sigma})^2 = (\det \hat{n} \cdot \bar{\sigma})^2$. Hence $\det \hat{n} \cdot \bar{\sigma} = -1$ (not $+1$) and the eigenvalues are ± 1 . \square

In the following we find the eigenvectors of $\hat{n} \cdot \bar{\sigma}$ (*without doing any matrix calculations!*).

Proposition. *The unit eigenvectors of $\hat{n} \cdot \bar{\sigma}$ are, for $n_3 \neq \pm 1$,*

$$\begin{aligned} & \frac{1}{\sqrt{2(1+n_3)}} \begin{bmatrix} 1+n_3 \\ n_1+in_2 \end{bmatrix} \text{ with eigenvalue } +1, \\ & \frac{1}{\sqrt{2(1-n_3)}} \begin{bmatrix} 1-n_3 \\ -(n_1+in_2) \end{bmatrix} \text{ with eigenvalue } -1. \end{aligned} \quad (\text{E.8})$$

If $n_3 = \pm 1$ then $\hat{n} \cdot \bar{\sigma} = \pm \sigma_3$; the eigenvectors are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ with eigenvalue $+1$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ with eigenvalue -1 .

Proof. The case $\hat{n} \cdot \bar{\sigma} = \pm \sigma_3$ is immediate and already noted in Section E.1. So assume $n_3 \neq \pm 1$.

Since $\hat{n} \cdot \bar{\sigma}$ has eigenvalues $\{+1, -1\}$ respectively,

$$\frac{1}{2}(I + \hat{n} \cdot \bar{\sigma}) = \frac{1}{2} \begin{bmatrix} 1+n_3 & n_1-in_2 \\ n_1+in_2 & 1-n_3 \end{bmatrix} \quad (\text{E.9})$$

has eigenvalues $\{1, 0\}$ respectively with the *same* respective eigenvectors (*why?*).

But this matrix is then an orthogonal projection operator, so either column spans the range, and so either column is an eigenvector with eigenvalue 1. We choose the first column just because the “leading” entry is real and positive.

⁴In fact, an inner product space isomorphism.

The squared norm of $\begin{bmatrix} 1 + n_3 \\ n_1 + in_2 \end{bmatrix}$ is $(1 + n_3)^2 + n_1^2 + n_2^2 = 2(1 + n_3)$, and this then gives the first eigenvector in (E.8).

Similarly, by considering $\frac{1}{2}(I - \hat{n} \cdot \bar{\sigma})$ we get the second eigenvector. □

Appendix F

The Bloch Sphere

Contents

F.1 Coordinate Approach	57
F.2 Pairs of antipodal points	58
F.2.1 Orthonormal Bases & Antipodal Points	58
F.2.2 Important Examples	58
F.3 Projection Mapping Approach	59
F.4 Angle Doubling	61

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

In fact, this map $S^3 \rightarrow S^2$, where the inverse of any point $\mathbf{x} \in S^2$ is a great circle $S^1 \subset S^3$, is the famous Hopf fibration. See, for example, [MD01, §§1,2].

F.1 Coordinate Approach

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 $|\psi\rangle = r_1|0\rangle + r_2e^{i\phi}|1\rangle$ where $0 \leq r_1, 0 \leq r_2, r_1^2 + r_2^2 = 1, 0 \leq \phi < 2\pi$. Hence $|\psi\rangle$ can be written in the form

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

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

By using (θ, ϕ) 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 F.1. (We need $\cos \frac{\theta}{2}$ with $0 \leq \theta \leq \pi$ rather than $\cos \theta$ with $0 \leq \theta \leq \pi/2$ for this.)

If $|\psi\rangle$ is a qubit then the corresponding point on the Bloch sphere is denoted by

$$\Phi(|\psi\rangle). \tag{F.2}$$

In cartesian coordinates,

$$\Phi(|\psi\rangle) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) = (\sin \theta e^{i\phi}, \cos \theta), \tag{F.3}$$

where on the right side we have identified the x - y plane with the complex plane. This is often convenient for calculations.

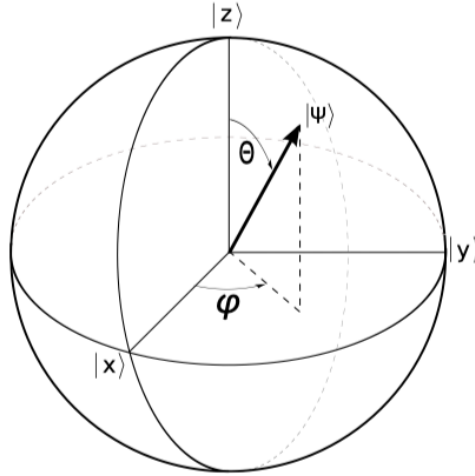


Figure F.1: Bloch Sphere. See (F.6) for notation. ([Vaz12, Notes Chapter 10, Fig. 10.1]).

F.2 Pairs of antipodal points

F.2.1 Orthonormal Bases & Antipodal Points

An arbitrary *orthonormal basis*¹

$$\begin{aligned} |\psi\rangle &= \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} e^{i\phi} |1\rangle, \\ |\psi'\rangle &= \sin \frac{\theta}{2} |0\rangle - \cos \frac{\theta}{2} e^{i\phi} |1\rangle = \cos \frac{\pi - \theta}{2} |0\rangle + \sin \frac{\pi - \theta}{2} e^{i(\pi + \phi)} |1\rangle, \end{aligned} \quad (\text{F.4})$$

for \mathbb{C}^2 corresponds to a pair of *antipodal* points (θ, ϕ) and $(\pi - \theta, \pi + \phi)$ on the Bloch sphere, and conversely.

F.2.2 Important Examples

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) .

In cartesian coordinates, let

$$z = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad y = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \in \mathbb{R}^3. \quad (\text{F.5})$$

¹Why can every orthonormal basis be written this way, up to a global phase factor for each basis vector?

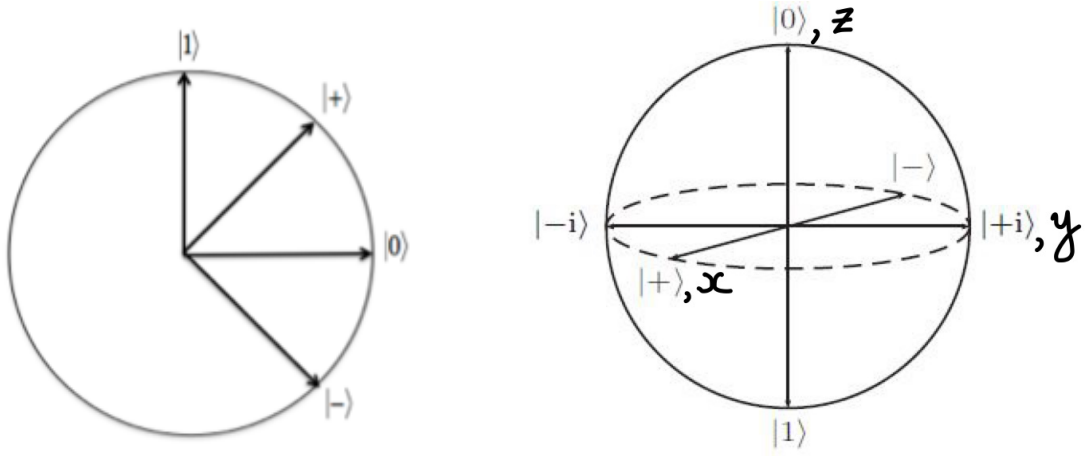


Figure F.2: A 2(real)-dim slice of \mathbb{C}^2 and the Bloch sphere $S^2 \subset \mathbb{R}^3$.

Using the notation of (2.2),

$$\begin{aligned}
|\psi\rangle &= \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} e^{i\phi} |1\rangle, & (\theta, \phi) \\
|0\rangle = |z\rangle &= \cos 0 |0\rangle + \sin 0 e^{i0} |1\rangle, & (\theta, \phi) = (0, 0) \\
|1\rangle = |-z\rangle &= \cos \left(\frac{\pi}{2}\right) |0\rangle + \sin \left(\frac{\pi}{2}\right) e^{i0} |1\rangle, & (\theta, \phi) = (\pi, 0) \\
|+\rangle = |x\rangle = \frac{1}{\sqrt{2}} |0\rangle + \frac{1}{\sqrt{2}} |1\rangle &= \cos \left(\frac{\pi}{4}\right) |0\rangle + \sin \left(\frac{\pi}{4}\right) e^{i0} |1\rangle = |x\rangle, & (\theta, \phi) = \left(\frac{\pi}{2}, 0\right) \\
|-\rangle = |-x\rangle = \frac{1}{\sqrt{2}} |0\rangle - \frac{1}{\sqrt{2}} |1\rangle &= \cos \left(\frac{\pi}{4}\right) |0\rangle + \sin \left(\frac{\pi}{4}\right) e^{i\pi} |1\rangle = |-x\rangle, & (\theta, \phi) = \left(\frac{\pi}{2}, \pi\right) \\
|i\rangle = |y\rangle = \frac{1}{\sqrt{2}} |0\rangle + i \frac{1}{\sqrt{2}} |1\rangle &= \cos \left(\frac{\pi}{4}\right) |0\rangle + \sin \left(\frac{\pi}{4}\right) e^{i\pi/2} |1\rangle = |y\rangle, & (\theta, \phi) = \left(\frac{\pi}{2}, \frac{\pi}{2}\right) \\
|-i\rangle = |-y\rangle = \frac{1}{\sqrt{2}} |0\rangle - i \frac{1}{\sqrt{2}} |1\rangle &= \cos \left(\frac{\pi}{4}\right) |0\rangle + \sin \left(\frac{\pi}{4}\right) e^{i3\pi/2} |1\rangle = |-y\rangle, & (\theta, \phi) = \left(\frac{\pi}{2}, \frac{3\pi}{2}\right)
\end{aligned} \tag{F.6}$$

F.3 Projection Mapping Approach

For each qubit $|\psi\rangle$ the mapping $|\psi\rangle\langle\psi|$ is orthogonal projection onto the one dimensional subspace of \mathbb{C}^2 spanned by $|\psi\rangle$ (see Section D.3.2). Moreover, this projection map depends only on the *equivalence class* for $|\psi\rangle$.²

For $|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} e^{i\phi} |1\rangle$, we defined the Bloch sphere representation $\Phi |\psi\rangle \in S^2 \subset \mathbb{R}^3$ to be (θ, ϕ) in spherical coordinates and hence $(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ in cartesian coordinates. See (F.3). However, a frequently more convenient approach is to write the hermitian matrix $|\psi\rangle\langle\psi|$ in terms of the hermitian basis matrices $I, \sigma_1, \sigma_2, \sigma_3$ and to read off the coefficients, as in the following proposition.

Proposition. *Let $|\psi\rangle$ be a qubit. Since the projection map $|\psi\rangle\langle\psi|$ is hermitian it has a unique expansion*

$$|\psi\rangle\langle\psi| = \frac{1}{2}(I + n_1\sigma_1 + n_2\sigma_2 + n_3\sigma_3) = \frac{1}{2} \begin{bmatrix} 1 + n_3 & n_1 - in_2 \\ n_1 + in_2 & 1 - n_3 \end{bmatrix}. \tag{F.7}$$

Then $\Phi |\psi\rangle = \hat{n} = (n_1, n_2, n_3)$ is the Bloch sphere representation of $|\psi\rangle$ in cartesian coordinates.

Proof. That the projection map $|\psi\rangle\langle\psi|$ is hermitian was noted in Section D.3.2, and the unique expansion of any hermitian matrix in terms of the Pauli matrices was noted in Appendix E equation (E.3).

²Let $|\psi'\rangle = e^{i\phi} |\psi\rangle$. Then $\langle\psi'| = e^{-i\phi} \langle\psi|$ and so $|\psi'\rangle\langle\psi'| = |\psi\rangle\langle\psi|$. See Section D.4.

Setting $|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} e^{-i\phi} |1\rangle$, using some basic trigonometric identities³ and (E.3)

$$\begin{aligned}
|\psi\rangle\langle\psi| &= \begin{bmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i\phi} \end{bmatrix} \begin{bmatrix} \cos \frac{\theta}{2} & \sin \frac{\theta}{2} e^{-i\phi} \end{bmatrix} \\
&= \begin{bmatrix} \cos^2 \frac{\theta}{2} & \cos \frac{\theta}{2} \sin \frac{\theta}{2} e^{-i\phi} \\ \cos \frac{\theta}{2} \sin \frac{\theta}{2} e^{i\phi} & \sin^2 \frac{\theta}{2} \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} 1 + \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & 1 - \cos \theta \end{bmatrix} \\
&= \frac{1}{2} (I + \sin \theta \cos \phi \sigma_1 + \sin \theta \sin \phi \sigma_2 + \cos \theta \sigma_3).
\end{aligned} \tag{F.8}$$

This gives (F.7) with $(n_1, n_2, n_3) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$, which from (F.3) is the Bloch sphere representation of $|\psi\rangle$ in cartesian coordinates. \square

Remark The fact the coefficient of I in F.7 is $\frac{1}{2}$ can be seen directly. First note that

$$\text{tr } |\psi\rangle\langle\psi| = \text{tr } \langle\psi|\psi\rangle = \langle\psi|\psi\rangle = 1. \tag{F.9}$$

The first equality is by the cyclic property of trace and the second equality because the trace of a scalar is the scalar. Since also

$$\text{tr } I = 2, \text{tr } \sigma_1 = \text{tr } \sigma_2 = \text{tr } \sigma_3 = 0,$$

the result follows.

Remark The fact $|\hat{n}| = 1$ can also be seen directly from (F.7). As in the proof of (F.13),

$$1 = (\langle\psi|\psi\rangle)^2 = \text{tr}(\langle\psi|\psi\rangle)^2 = \text{tr}(|\psi\rangle\langle\psi| |\psi\rangle\langle\psi|) = \frac{1}{2} (1 + |\hat{n}|^2).$$

Hence $|\hat{n}| = 1$.

Example Find $\Phi |\psi\rangle$ if $|\psi\rangle = \frac{1}{2} |0\rangle + \frac{\sqrt{3}}{2} i |1\rangle$.

Compute

$$\begin{aligned}
|\psi\rangle\langle\psi| &= \begin{bmatrix} \frac{1}{2} \\ \frac{\sqrt{3}i}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}i}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & -\frac{\sqrt{3}i}{4} \\ \frac{\sqrt{3}i}{4} & \frac{3}{4} \end{bmatrix} \\
&= \frac{1}{2} I + \begin{bmatrix} -\frac{1}{4} & -\frac{\sqrt{3}i}{4} \\ \frac{\sqrt{3}i}{4} & \frac{1}{4} \end{bmatrix} = \frac{1}{2} \left(I + \frac{\sqrt{3}}{2} \sigma_2 - \frac{1}{2} \sigma_3 \right)
\end{aligned}$$

Hence $\Phi |\psi\rangle = \left(0, \frac{\sqrt{3}}{2}, -\frac{1}{2}\right)$ in cartesian coordinates.

From the Bloch Sphere back to the Qubit Given a unit vector $(n_1, n_2, n_3) \in S^2 \subset \mathbb{R}^3$ we can readily find the coordinates of the ket $|\psi\rangle$ which is represented by (n_1, n_2, n_3) as follows.

From (F.7) and (E.1) we have

$$|\psi\rangle\langle\psi| = \frac{1}{2} \begin{bmatrix} 1 + n_3 & n_1 - in_2 \\ n_1 + in_2 & 1 - n_3 \end{bmatrix}. \tag{F.10}$$

We can read off the coordinates of $|\psi\rangle\langle\psi|$ directly as follows.

Since $|\psi\rangle\langle\psi|$ is a projection operator it has rank 1. Hence the two columns in its matrix must be linearly dependent and the image of $|\psi\rangle\langle\psi|$, which is $|\psi\rangle$, is spanned by either column:

$$\text{Span } |\psi\rangle = \text{Span} \begin{bmatrix} 1 + n_3 \\ n_1 + in_2 \end{bmatrix} = \text{Span} \begin{bmatrix} n_1 - in_2 \\ 1 - n_3 \end{bmatrix}.$$

³There is not much to remember, just $\sin 2\theta = 2 \sin \theta \cos \theta$, $\cos 2\theta = 2 \cos^2 \theta - 1$.

Using the first column vector since it gives a real positive coefficient for $|0\rangle$, and since its norm squared is $(1+n_3)^2 + |n_1+in_2|^2 = 2(1+n_3)$, up to an arbitrary phase factor and assuming $n_3 \neq -1$,

$$|\psi\rangle = \frac{\sqrt{1+n_3}}{\sqrt{2}} |0\rangle + \frac{n_1+in_2}{\sqrt{2(1+n_3)}} |1\rangle. \quad (\text{F.11})$$

If $n_3 = -1$ then $|\psi\rangle = |1\rangle$ up to a phase factor.

F.4 Angle Doubling

Let

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

as in (F.1). Since $|\langle 0|\psi\rangle| = \langle 0|\psi\rangle = \cos \frac{\theta}{2}$, the angle between $|0\rangle$ and $|\psi\rangle$ is $\frac{\theta}{2}$. See Appendix B.1.4.

On the other hand, from (F.3), $\Phi|\psi\rangle = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) =: \hat{k}$ and $\Phi|0\rangle = (0, 0, 1) := z$, so $\langle z, \hat{k} \rangle = \cos \theta$ and the angle between $\Phi|0\rangle$ and $\Phi|\psi\rangle$ is θ .

Thus there is angle doubling in passing from $S^3 \subset \mathbb{C}^2$ to $S^2 \subset \mathbb{R}^3$. We will see in the following proposition that this is always true. To prove this from the coordinate representation of qubits is very messy and not recommended.

The following proof uses a couple of important properties of trace, see Appendix C.2.

Proposition. *Suppose $\Phi|\psi\rangle = \hat{n}$ and $\Phi|\phi\rangle = \hat{k}$. Let $|\langle \psi|\phi\rangle| = \cos \zeta$ with $\zeta \in [0, \pi/2]$, and let $\langle \hat{n}, \hat{k} \rangle = \cos \theta$ with $\theta \in [0, \pi]$. Then $\zeta = \theta/2$, i.e. $\theta = 2\zeta$.*

Proof. First note that

$$\cos^2 \zeta = |\langle \psi|\phi\rangle|^2 = \langle \psi|\phi\rangle \langle \phi|\psi\rangle = \text{tr}(\langle \psi|\phi\rangle \langle \phi|\psi\rangle) = \text{tr}(|\psi\rangle\langle\psi| |\phi\rangle\langle\phi|). \quad (\text{F.12})$$

The third equality is just because the trace of a scalar is the scalar. The fourth is by (C.6), noting $\langle \psi|\phi\rangle = \langle \psi||\phi\rangle$ and $\langle \phi|\psi\rangle = \langle \phi||\psi\rangle$.

From (F.7), with $\hat{n} = (n_1, n_2, n_3)$ and $\hat{k} = (k_1, k_2, k_3)$,

$$\begin{aligned} \text{tr}(|\psi\rangle\langle\psi| |\phi\rangle\langle\phi|) &= \frac{1}{4} \text{tr} \begin{bmatrix} 1+n_3 & n_1-in_2 \\ n_1+in_2 & 1-n_3 \end{bmatrix} \begin{bmatrix} 1+k_3 & k_1-ik_2 \\ k_1+ik_2 & 1-k_3 \end{bmatrix} \\ &= \frac{1}{4} \text{tr} \begin{bmatrix} (1+n_3)(1+k_3) + (n_1-in_2)(k_1+ik_2) & \star \\ \star & (n_1+in_2)(k_1-ik_2) + (1-n_3)(1-k_3) \end{bmatrix} \\ &= (1+n_3+k_3+n_3k_3) + (n_1k_1+in_1k_2-in_2k_1+n_2k_2) \\ &\quad + (n_1k_1-in_1k_2+in_2k_1+n_2k_2) + (1-n_3-k_3+n_3k_3) \\ &= \frac{1}{2} (1 + \langle \hat{n}, \hat{k} \rangle) = \frac{1}{2} (1 + \cos \theta) = \cos^2 \frac{\theta}{2}. \end{aligned} \quad (\text{F.13})$$

The proposition now follows from (F.12) and (F.13). \square

Appendix G

Functions of Normal Operators

Contents

G.1 Examples of Interest	62
G.2 Power Series Definition	62
G.3 Definition for Normal Operators	62
G.4 Matrices for the Schrödinger Equation	63

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

We will be particularly concerned with

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

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

G.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 $f(z) = a_0 + a_1z + a_2z^2 + \dots$ converges for $|z| < R$ (i.e. f is analytic) then we can define $f(T) = a_0 + a_1T + a_2T^2 + \dots$. 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.

G.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

¹This 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].

With respect to any orthonormal basis of eigenvectors we have

$$T = \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{bmatrix} \implies f(T) := \begin{bmatrix} f(\alpha_1) & & \\ & \ddots & \\ & & f(\alpha_n) \end{bmatrix}. \quad (\text{G.2})$$

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.

G.4 Matrices for the Schrödinger Equation

If H is hermitian as in (B.4) with matrix below for some orthonormal basis of eigenvectors $\{E_1, \dots, E_n\}$ corresponding to eigenvalues $\lambda_1, \dots, \lambda_n$, then $U(t) := e^{-itH}$ has the matrix shown for the same orthonormal basis of eigenvectors (now also for e^{-itH}).

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

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

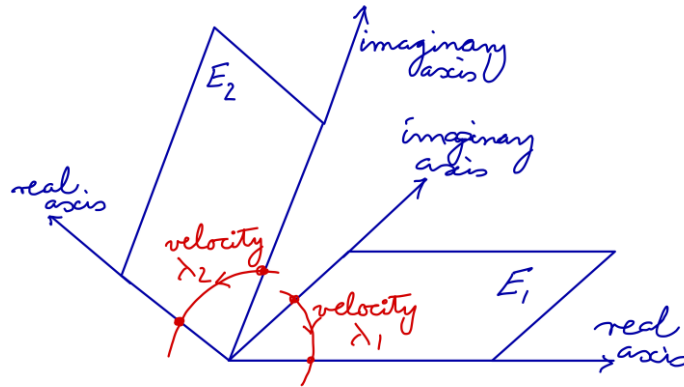


Figure G.1: Schematic representation for the 2 orthogonal eigenspaces $E_1, E_2 \subset \mathbb{C}^2$ and for the unitary operator $U = \exp(-itH)$ as it rotates vectors in these eigenspaces. The direction of rotation is shown in case $\lambda_1, \lambda_2 > 0$.

See also the Schrödinger Equation 2.19

Appendix H

Tensor Products

Contents

H.1 Informal Presentation	64
H.1.1 Bilinear Requirements	64
H.1.2 Basis and Dimension	65
H.1.3 Entangled and Product Elements	65
H.1.4 Other Bilinear Products	65
H.1.5 Inner Product	65
H.1.6 Examples	65
H.2 Basis Definition of Tensor Products	66

H.1 Informal Presentation

H.1.1 Bilinear Requirements

Consider two finite dimensional complex¹ inner product spaces V and W , with orthonormal bases e_1, \dots, e_m and $\epsilon_1, \dots, \epsilon_n$ respectively. If $v \in V$ and $w \in W$ let

$$v = v_1 e_1 + \dots + v_m e_m, \quad w = w_1 \epsilon_1 + \dots + w_n \epsilon_n. \tag{H.1}$$

We want to form a new “tensor product”² vector space $V \otimes W$ which consists of all linear combinations of elements of the form $v \otimes w$ where $v \in V$ and $w \in W$. We call $v \otimes w$ the “tensor product of v and w ” or just “ v tensor w ”. We similarly say the “tensor product of V and W ” or just “ V tensor W ” for the vector space $V \otimes W$.

We want $V \otimes W$ to “preserve” the linear vector space structure of V and of W . For example, we require

$$\begin{aligned} (2v + 3v') \otimes (w - 3w') &= 2(v \otimes (w - 3w')) + 3(v' \otimes (w - 3w')) \\ &= 2(v \otimes w) - 6(v \otimes w') + 3(v' \otimes w) - 9(v' \otimes w') \\ &= 2v \otimes w - 6v \otimes w' + 3v' \otimes w - 9v' \otimes w', \quad \text{etc.} \end{aligned}$$

In particular we unambiguously write $2v \otimes w$ since we will require $(2v) \otimes w = 2(v \otimes w)$.

More precisely, we allow repeated applications of the following rules:

$$\begin{aligned} (\alpha v + \alpha' v') \otimes w &= \alpha v \otimes w + \alpha' v' \otimes w \\ v \otimes (\beta w + \beta' w') &= \beta v \otimes w + \beta' v \otimes w' \end{aligned} \tag{H.2}$$

These rules say that the map from $V \times W$ into $V \otimes W$ given by $(v, w) \mapsto v \otimes w$ is linear in v and is linear in w . In other words, the map from $V \times W$ into $V \otimes W$ is *bilinear*.

¹Essentially everything we do also applies in the real case.

²Tensor products can be defined for any two vector spaces, not necessarily complex, or finite dimensional or with an inner product.

H.1.2 Basis and Dimension

When we calculate $v \otimes w$ in this way, with v and w as in (H.1), we obtain

$$\begin{aligned} v \otimes w &= (v_1 e_1 + \cdots + v_m e_m) \otimes (w_1 \epsilon_1 + \cdots + w_n \epsilon_n) \\ &= v_1 w_1 e_1 \otimes \epsilon_1 + v_1 w_2 e_1 \otimes \epsilon_2 + \cdots + v_1 w_n e_1 \otimes \epsilon_n \\ &\quad + v_2 w_1 e_2 \otimes \epsilon_1 + \cdots + v_2 w_n e_2 \otimes \epsilon_n + \cdots + v_m w_n e_m \otimes \epsilon_n. \end{aligned} \tag{H.3}$$

Thus we can write $v \otimes w$ as a linear combination of the elements $e_j \otimes \epsilon_k$. Moreover, using the rules (H.2), $v \otimes w + v' \otimes w'$ and $\alpha(v \otimes w)$ will also be a linear combination of the $e_j \otimes \epsilon_k$.

This implies $V \otimes W$ should be a vector space spanned by $\{e_j \otimes \epsilon_k : 1 \leq j \leq m, 1 \leq k \leq n\}$. Moreover, we will require that the $e_j \otimes \epsilon_k$ be linearly independent, so that they actually form a *basis* for $V \otimes W$ and $V \otimes W$ is mn -dimensional. In particular,

$$\sum_{j,k} \alpha_{jk} e_j \otimes \epsilon_k = \sum_{j,k} \beta_{jk} e_j \otimes \epsilon_k \quad \text{implies} \quad \alpha_{jk} = \beta_{jk} \text{ for all } j \text{ and } k. \tag{H.4}$$

H.1.3 Entangled and Product Elements

It is important to note that the map from $V \times W$ into $V \otimes W$ is *not* onto. Because $V \otimes W$ is a vector space we can add its members, for example $e_1 \otimes \epsilon_1 + e_2 \otimes \epsilon_2 \in V \otimes W$. But do there exist v and w such that $e_1 \otimes \epsilon_1 + e_2 \otimes \epsilon_2 = v \otimes w$? The answer is *NO*.

To see this let v and w be as in (H.1). Then from (H.3) and (H.4),

$$v_1 w_1 = 1, \quad v_1 w_2 = 0, \quad v_2 w_1 = 0, \quad v_2 w_2 = 1.$$

Since $v_1 w_2 = 0$, either $v_1 = 0$ or $w_2 = 0$. If $v_1 = 0$ this contradicts $v_1 w_1 = 1$ and if $w_2 = 0$ this contradicts $v_2 w_2 = 1$.

We say that $e_1 \otimes \epsilon_1 + e_2 \otimes \epsilon_2$ is *entangled*. On the other hand,

$$e_1 \otimes \epsilon_1 + e_1 \otimes \epsilon_2 + e_2 \otimes \epsilon_1 + e_2 \otimes \epsilon_2 = (e_1 + e_2) \otimes (\epsilon_1 + \epsilon_2),$$

We say $e_1 \otimes \epsilon_1 + e_1 \otimes \epsilon_2 + e_2 \otimes \epsilon_1 + e_2 \otimes \epsilon_2$ is a *product*.

In the quantum setting, we refer to *entangled* and *product* states.

H.1.4 Other Bilinear Products

The requirements in (H.2) are the *only* restrictions on forming the tensor product of vectors from V and W . Another way of expressing this is through the linear independence of the $e_j \otimes \epsilon_k$ as in (H.4).

To clarify this, consider (in the real case) the cross product map from $\mathbb{R}^3 \times \mathbb{R}^3$ into \mathbb{R}^3 . This *is* bilinear. But $e_1 \times e_1 = 0$ and $e_1 \times e_2 + e_2 \times e_1 = 0$, whereas $e_1 \otimes e_1 \neq 0$ and $e_1 \otimes e_2 + e_2 \otimes e_1 \neq 0$. So the cross product is not a tensor product. Moreover, the tensor product space $\mathbb{R}^3 \otimes \mathbb{R}^3$ is in fact 9-dimensional, and as we see later is unique up to an appropriate isomorphism.

H.1.5 Inner Product

H.1.6 Examples

The basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$ constructed from the canonical basis $\{|0\rangle, |1\rangle\}$ for \mathbb{C}^2 consists of the four elements

$$|0\rangle \otimes |0\rangle, \quad |0\rangle \otimes |1\rangle, \quad |1\rangle \otimes |0\rangle, \quad |1\rangle \otimes |1\rangle.$$

In standard abbreviated form these are written as

$$|0\rangle|0\rangle, \quad |0\rangle|1\rangle, \quad |1\rangle|0\rangle, \quad |1\rangle|1\rangle; \quad \text{or} \quad |00\rangle, \quad |01\rangle, \quad |10\rangle, \quad |11\rangle.$$

The basis for $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ consists of the 8 elements

$$|000\rangle, \quad |001\rangle, \quad |010\rangle, \quad |011\rangle, \quad |100\rangle, \quad |101\rangle, \quad |110\rangle, \quad |111\rangle.$$

H.2 Basis Definition of Tensor Products

Definition. With V and W as above, the *tensor product* $V \otimes W$ is the mn dimensional inner product space consisting of all formal linear combinations of the form $\sum_{j,k} \alpha_{jk} e_j \otimes \epsilon_k$, with vector addition, scalar multiplications and inner product defined componentwise. In particular,

$$\left(\sum_{j,k} \alpha_{jk} e_j \otimes \epsilon_k, \sum_{j,k} \beta_{jk} e_j \otimes \epsilon_k \right) := \sum_{j,k} \overline{\alpha_{jk}} \beta_{jk}$$

and so the $e_j \otimes \epsilon_k$ form an orthonormal basis of $V \otimes W$.

Define

$$\widehat{\otimes} : V \times W \rightarrow V \otimes W, \quad \widehat{\otimes}(v, w) = v \otimes w = \sum_{j,k} v_j w_k e_j \otimes \epsilon_k.$$

Note that $\widehat{\otimes}(\cdot, \cdot)$ is bilinear, i.e. linear in each slot. But $\widehat{\otimes}$ is typically *not* onto.

Remarks Although the definition of $V \otimes W$ involves the choice of bases e_1, \dots, e_m and $\epsilon_1, \dots, \epsilon_n$, we will soon see that the tensor product is a basis independent notion. It is also possible to give a basis free definition, but it is less “concrete” than the “pedestrian” approach above.

Note that here we are just identifying $V \otimes W$ with the inner product space \mathbb{C}^{mn} , modulo a choice of orthonormal basis vectors e_1, \dots, e_m and $\epsilon_1, \dots, \epsilon_n$ and a choice of identifying the $e_j \otimes \epsilon_k$ with the standard basis vectors in \mathbb{C}^{mn} .

Proposition. *If $B : V \times W \rightarrow E$ is a bilinear map into a vector space E then there is a unique linear map $L : V \otimes W \rightarrow E$ such that $B(v, w) = L(v \otimes w)$. Conversely, B can be recovered from L .*

$$\begin{array}{ccc} V \times W & \xrightarrow{\widehat{\otimes}} & V \otimes W \\ B \downarrow & \nearrow L & \\ E & & \end{array}$$

Proof. Define $L(e_j \otimes \epsilon_k) = B(e_j, \epsilon_k)$ and extend to $L : V \otimes W \rightarrow E$ by linearity. It follows

$$B(v, w) = \sum_{j,k} v_j w_k B(e_j, \epsilon_k) = \sum_{j,k} v_j w_k L(e_j \otimes \epsilon_k) = L\left(\sum_{j,k} v_j w_k e_j \otimes \epsilon_k\right) = L(v \otimes w).$$

The existence and uniqueness of L are now immediate, and clearly B can be recovered from L . \square

Remark The point is that any *bilinear* map $B : V \times W \rightarrow E$ corresponds to a unique *linear* map $L : V \otimes W \rightarrow E$. We say $(\widehat{\otimes}, V \otimes W)$ satisfies the *universality property* for V and W .

We can also think of $\widehat{\otimes} : V \times W \rightarrow V \otimes W$ as the “freest” bilinear map on $V \times W$ in that the only conditions imposed are the ones of bilinearity. In particular, the bilinear maps given by scalar product and vector product on $\mathbb{R}^3 \times \mathbb{R}^3$ are not free in this sense. For example, $e_1 \cdot e_2 = 0$ but $e_1 \otimes e_2 \neq 0$, $e_1 \times e_1 = 0$ but $e_1 \otimes e_1 \neq 0$.

This leads to the following definition and proposition, which show that

Definition. Let $C : V \times W \rightarrow Z$ be a bilinear map into the vector space Z . Then (C, Z) satisfies the *universality property* for V and W if for any vector space E and bilinear map $B : V \times W \rightarrow E$ there exists a unique linear map $L_B : Z \rightarrow E$ such that the following diagram commutes:

$$\begin{array}{ccc} V \times W & \xrightarrow{C} & Z \\ B \downarrow & \nearrow L_B & \\ E & & \end{array}$$

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 Quantum 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.)
- [Lvo18] A.I. Lvovsky, *Quantum Physics: An Introduction Based on Photons*, Springer, 2018. Well written. Emphasises the photon polarisation interpretation of qubits. Lots of problems with solutions available through the publishers web site.
- [MD01] Rémy Mosseri and Rosen Dandoloff, *Geometry of entangled states, Bloch spheres and Hopf fibrations*, Journal of Physics A: Mathematical and General **34** (2001), no. 47, 10243, available at <https://arxiv.org/pdf/quant-ph/0108137.pdf>.
- [MN19] Andy Matuschak 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 introduction 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>.