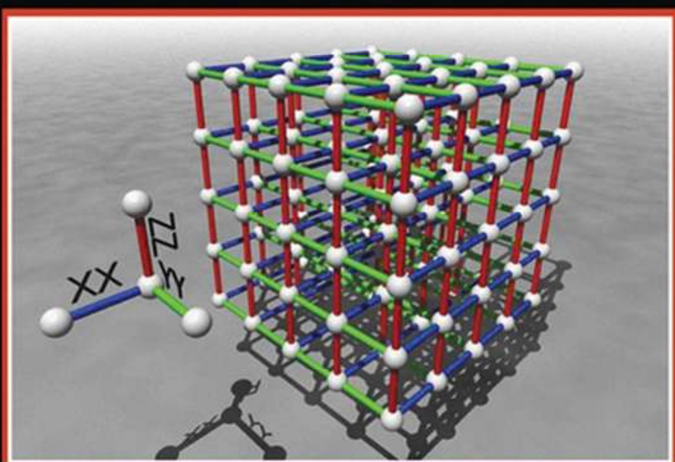
Quantum Computing Explained



DAVID McMAHON

QUANTUM COMPUTING EXPLAINED

The Wiley Bicentennial–Knowledge for Generations

Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation's journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it!

William J. Pesce
President and Chief Executive Officer

Peter Booth Wiley
Chairman of the Board

QUANTUM COMPUTING EXPLAINED

David McMahon

IEEE COMPUTER SOCIETY

BICENTENNIAL 1807 WILEY 2007 BICENTENNIAL

WILEY-INTERSCIENCE
A John Wiley & Sons, Inc., Publication

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Wiley Bicentennial Logo: Richard J. Pacifico

Library of Congress Cataloging-in-Publication Data:

McMahon, David (David M.)
Quantum computing expained / David McMahon.
p. cm.
Includes index.
ISBN 978-0-470-09699-4 (cloth)
1. Quantum computers. I. Title.
QA76.889.M42 2007
004.1–dc22

2007013725

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

CONTENTS

PREFACE

"In the twenty-first" century it is reasonable to expect that some of the most important developments in science and engineering will come about through interdisciplinary research. Already in the making is surely one of the most interesting and exciting development we are sure to see for a long time, *quantum computation*. A merger of computer science and physics, quantum computation came into being from two lines of thought. The first was the recognition that *information is physical*, which is an observation that simply states the obvious fact that information can't exist or be processed without a physical medium.

At the present time quantum computers are mostly theoretical constructs. However, it has been proved that in at least some cases quantum computation is much faster in principle than any done by classical computer. The most famous algorithm developed is Shor's factoring algorithm, which shows that a quantum computer, if one could be constructed, could quickly crack the codes currently used to secure the world's data. Quantum information processing systems can also do remarkable things not possible otherwise, such as teleporting the state of a particle from one place to another and providing unbreakable cryptography systems.

Our treatment is not rigorous nor is it complete for the following reason: this book is aimed primarily at two audiences, the first group being undergraduate physics, math, and computer science majors. In most cases these undergraduate students will find the standard presentations on quantum computation and information science a little hard to digest. This book aims to fill in the gap by providing undergraduate students with an easy to follow format that will help them grasp many of the fundamental concepts of quantum information science.

This book is also aimed at readers who are technically trained in other fields. This includes students and professionals who may be engineers, chemists, or biologists. These readers may not have the background in quantum physics or math that most people in the field of quantum computation have. This book aims to fill the gap here as well by offering a more "hand-holding" approach to the topic so that readers can learn the basics and a little bit on how to do calculations in quantum computation.

Finally, the book will be useful for graduate students in physics and computer science taking a quantum computation course who are looking for a calculationally oriented supplement to their main textbook and lecture notes.

The goal of this book is to open up and introduce quantum computation to these nonstandard audiences. As a result the level of the book is a bit lower than that found in the standard quantum computation books currently available. The presentation is informal, with the goal of introducing the concepts used in the field and then showing through explicit examples how to work with them. Some topics are left out entirely and many are not covered at the deep level that would be expected in a graduate level quantum computation textbook. An in-depth treatment of adiabatic quantum computation or cluster state computation is beyond this scope of this book. So this book could not be considered complete in any sense. However, it will give readers who are new to the field a substantial foundation that can be built upon to master quantum computation.

While an attempt was made to provide a broad overview of the field, the presentation is weighted more in the physics direction.

1

A BRIEF INTRODUCTION TO INFORMATION THEORY

In this chapter we will give some basic background that is useful in the study of quantum information theory. Our primary focus will be on learning how to quantify information. This will be done using a concept known as *entropy*, a quantity that can be said to be a measure of disorder in physics. Information is certainly the opposite of disorder, so we will see how entropy can be used to characterize the information content in a signal and how to determine how many bits we need to reliably transmit a signal. Later these ideas will be tied in with quantum information processing. In this chapter we will also briefly look at problems in computer science and see why we might find quantum computers useful. This chapter won't turn you into a computer engineer, we are simply going to give you the basic fundamentals.

CLASSICAL INFORMATION

Quantum computation is an entirely new way of information processing. For this reason traditional methods of computing and information processing you are familiar with are referred to as *classical information*. For those new to the subject, we begin with a simple and brief review of how information is stored and used in computers. The most basic piece of information is called a *bit*, and this basically represents a

Quantum Computing Explained, by David McMahon

yes–no answer to a question. To represent this mathematically, we use the fact that we're dealing with a two-state system and choose to represent information using base 2 or *binary* numbers. A binary number can be 0 or 1, and a bit can assume one or the other of these values. Physically we can implement a bit with an electrical circuit that is either at ground or zero volts (binary 0), or at say $+5$ volts (binary 1). The physical implementation of a computing system is not our concern in this book; we are only worried about the mathematics and logic of the system. As a first step in getting acquainted with the binary world we might want to learn how to count using base 2.

Before we do that, we need to know that the number of bits required to represent something can be determined in the following way: Suppose that some quantity can assume one of m different states. Then

$$2^n \geq m \tag{1.1}$$

for some n. The smallest n for which this holds tells us the number of bits we need to represent or encode that quantity.

To see how this works, suppose that we want to represent the numbers 0, 1, 2, 3 in binary. We have four items, and $2^2 = 4$. Therefore we need at least two bits to represent these numbers. The representation is shown in Table 1.1.

To represent the numbers 0 through 7, we have $2^3 = 8$, so we need three bits. The binary representation of the numbers 0 through 7 is shown in Table 1.2.

INFORMATION CONTENT IN A SIGNAL

Now that we know how to encode information, we can start thinking about how to quantify it. That is, given a message m, how much information is actually contained in that message?

A clue about how this quantification might be done can be found by looking at (1.1). Considering the case where we take the equal sign, let's take the base two logarithm of both sides. That is, we start with

$$m = 2^n$$

TABLE 1.1 Binary representation of the numbers 0–3

Decimal	Binary
0	00
1	01
2	10
3	11

TABLE 1.2 Binary representation of the numbers 0–7

Decimal	Binary
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111

Taking the base 2 log of both sides, we find that

$$\log_2 m = n \tag{1.2}$$

Equation (1.2) was proposed by Ralph Hartley in 1927. It was the first attempt at quantifying the amount of information in a message. What (1.2) tells us is that n bits can store m different messages. To make this more concrete, notice that

$$\log_2 8 = 3$$

That tells us that 3 bits can store 8 different messages. In Table 1.2 the eight messages we encoded were the numbers 0 through 7. However, the code could represent anything that had eight different possibilities.

You're probably familiar with different measurements of information storage capacity from your computer. The most basic word or unit of information is called a *byte*. A byte is a string of eight bits linked together. Now

$$\log_2 256 = 8$$

Therefore a byte can store 256 different messages. Measuring information in terms of logarithms also allows us to exploit the fact that logarithms are additive.

ENTROPY AND SHANNON'S INFORMATION THEORY

The Hartley method gives us a basic characterization of information content in a signal. But another scientist named Claude Shannon showed that we can take things a step further and get a more accurate estimation of the information content in a signal by thinking more carefully. The key step taken by Shannon was that he asked how *likely* is it that we are going to see a given piece of information? This is an important insight because it allows us to characterize how much information we actually *gain* from a signal.

If a message has a very high probability of occurrence, then we don't gain all that much new information when we come across it. On the other hand, if a message has a low probability of occurrence, when we are made of aware of it, we gain a significant amount of information. We can make this concrete with an example. A major earthquake occurred in the St. Louis area way back in 1812. Generally speaking, earthquakes in that area are relatively rare—after all, when you think of earthquakes, you think of California, not Missouri.

So most days people in Missouri aren't waiting around for an earthquake. Under typical conditions the probability of an earthquake occurring in Missouri is low, and the probability of an earthquake *not* occurring is high. If our message is that tomorrow there will *not* be an earthquake in Missouri, our message is a high probability message, and it conveys very little new information—for the last two hundred years day after day there hasn't been an earthquake. On the other hand, if the message is that tomorrow there will be an earthquake, this is dramatic news for Missouri residents. They gain *a lot* of information in this case.

Shannon quantified this by taking the base 2 logarithm of the probability of a given message occurring. That is, if we denote the information content of a message by I, and the probability of its occurrence by p, then

$$I = -\log_2 p \tag{1.3}$$

The negative sign ensures that the information content of a message is positive, and that the less probable a message, the higher is the information content. Let's suppose that the probability of an earthquake not happening tomorrow in St. Louis is 0.995. The information content of this fact is

$$I = -\log_2 0.995 = 0.0072$$

Now the probability that an earthquake does happen tomorrow is 0.005. The information content of this piece of information is

$$I' = -\log_2 0.005 = 7.6439$$

So let's summarize the use of logarithms to characterize the information content in a signal by saying:

- A message that is unlikely to occur has a low probability and therefore has a large information content.
- A message that is very likely to occur has a high probability and therefore has a small information content.

Next let's develop a more formal definition. Let X be a random variable characterized by a probability distribution $\vec{p}$, and suppose that it can assume one of the values $x_1, x_2, \ldots, x_n$ with probabilities $p_1, p_2, \ldots, p_n$. Probabilities satisfy $0 \leqslant p_i \leqslant 1$ and $\sum_i p_i = 1$.

The Shannon entropy of X is defined as

$$H(X) = -\sum_i p_i \log_2 p_i \tag{1.4}$$

If the probability of a given x_j is zero, we use $0 \log 0 = 0$. Notice that if we are saying that the logarithm of the probability of x gives the information content, we can also view the Shannon entropy function as a measure of the amount of uncertainty or randomnessin a signal.

We can look at this more concretely in terms of transmitted message signals as follows: Suppose that we have a signal that always transmits a "2," so that the signal is the string 22222222222. . . . What is the entropy in this case? The probability of obtaining a 2 is 1, so the entropy or disorder is

$$H = -\log_2 1 = 0$$

The Shannon entropy works as we expect—a signal that has all the same characters with no changes has no disorder and hence no entropy.

Now let's make a signal that's a bit more random. Suppose that the probability of obtaining a "1" is 0.5 and the probability of obtaining a "2" is 0.5, so the signal looks something like 11212221212122212121112. . . with approximately half the characters 1's and half 2's. What is the entropy in this case? It's

$$H = -\frac{1}{2}\log_2 \frac{1}{2} - \frac{1}{2}\log_2 \frac{1}{2} = \frac{1}{2} + \frac{1}{2} = 1$$

Suppose further that there are three equally likely possibilities. In that case we would have

$$H = -\frac{1}{3}\log_2 \frac{1}{3} - \frac{1}{3}\log_2 \frac{1}{3} = 0.528 + 0.528 + 0.528 = 1.585$$

In each case that we have examined here, the uncertainty in regard to what character we will see next in the message has increased each time—so the entropy also increases each time. In this view we can see that Shannon entropy measures the amount of uncertainty or randomness in the signal. That is:

- If we are certain what the message is, the Shannon entropy is zero.
- The more uncertain we are as to what comes next, the higher the Shannon entropy.

We can summarize Shannon entropy as

$$\text{Decrease uncertainty} \Rightarrow \text{Increase information}$$

$$\text{Increase uncertainty} \Rightarrow \text{Increase entropy}$$

Now suppose that we require l_i bits to represent each x_i in X. Then the *average bit rate* required to encode X is

$$R_X = \sum_{i=1}^{n} l_i p_i \tag{1.5}$$

The Shannon entropy is the lower bound of the average bit rate

$$H(X) \leqslant R_X \tag{1.6}$$

The worst-case scenario in which we have the least information is a distribution where the probability of each item is the same—meaning a uniform distribution. Again, suppose that it has n elements. The probability of finding each x_i if the distribution is uniform is $1/n$. So sequence X with n elements occurring with uniform probabilities $1/n$ has entropy $-\sum \frac{1}{n} \log_2 \frac{1}{n} = \sum \frac{1}{n} \log n = \log n$. This tells us that the Shannon entropy has the bounds

$$0 \leqslant H(X) \leqslant \log_2 n \tag{1.7}$$

The *relative entropy* of two variables X and Y characterized by probability distributions p and q is

$$H(X \| Y) = \sum p \log_2 \frac{p}{q} = -H(X) - \sum p \log_2 q \tag{1.8}$$

Suppose that we take a fixed value y_i from Y. From this we can get a conditional probability distribution $p(X|y_i)$ which are the probabilities of X given that we have y_i with certainty. Then

$$H(X|Y) = -\sum_j p(x_j|y_i) \log_2(p(x_j|y_i)) \tag{1.9}$$

This is known as the *conditional entropy*. The conditional entropy satisfies

$$H(X|Y) \leqslant H(X) \tag{1.10}$$

To obtain equality in (1.10), the variables X and Y must be independent. So

$$H(X, Y) = H(Y) + H(X|Y) \tag{1.11}$$

We are now in a position to define *mutual information* of the variables X and Y. In words, this is the difference between the entropy of X and the entropy of X

given knowledge of what value Y has assumed, that is,

$$I(X|Y) = H(X) - H(X|Y) \tag{1.12}$$

This can also be written as

$$I(X|Y) = H(X) + H(Y) - H(X, Y) \tag{1.13}$$

PROBABILITY BASICS

Before turning to quantum mechanics in the next chapter, it's a good idea to quickly mention the basics of probability. Probability is heavily used in quantum theory to predict the possible results of measurement.

We can start by saying that the probability p_i of an event x_i falls in the range

$$0 \leqslant p_i \leqslant 1 \tag{1.14}$$

The two extremes of this range are characterized as follows: The probability of an event that is *impossible* is 0. The probability of an event that is *certain to happen* is 1. All other probabilities fall within this range.

The probability of an event is simply the relative frequency of its occurrence. Suppose that there are n total events, the *jth* event occurs n_j times, and we have $\sum_{j=1}^{\infty} n_j = n$. Then the probability that the *jth* event occurs is

$$p_j = \frac{n_j}{n} \tag{1.15}$$

The sum of all the probabilities is 1, since

$$\sum_{j=1}^{\infty} p_j = \sum_{j=1}^{\infty} \frac{n_j}{n} = \frac{1}{n}\sum_{j=1}^{\infty} n_j = \frac{n}{n} = 1 \tag{1.16}$$

The average value of a distribution is referred to as the *expectation value* in quantum mechanics. This is given by

$$\langle j \rangle = \sum_{j=1}^{\infty} \frac{j n_j}{n} = \sum_{j=1}^{\infty} j p_j \tag{1.17}$$

The *variance* of a distribution is

$$\langle (\Delta j)^2 \rangle = \langle j^2 \rangle - \langle j \rangle^2 \tag{1.18}$$

Example 1.1

A group of students takes an exam. The number of students associated with each score is

Score	Students
95	1
85	3
77	7
71	10
56	3

What is the most probable test score? What is the expectation value or average score?

Solution

First we write down the total number of students

$$n = \sum n_j = 1 + 3 + 7 + 10 + 3 = 24$$

The probability of scoring 95 is

$$p_1 = \frac{n_1}{n} = \frac{1}{24} = 0.04$$

and the other probabilities are calculated similarly. The most probable score is 71 with probability

$$p_4 = \frac{n_4}{n} = \frac{10}{24} = 0.42$$

The expectation value is found using (1.17):

$$\langle j \rangle = \sum j\, p_j = 95(0.04) + 85(0.13) + 77(0.29) + 71(0.42) + 56(0.13) = 74.3$$

In the next chapter we will see how to quantum mechanics uses probability.

EXERCISES

1.1. *How many bits are necessary to represent the alphabet using a binary code if we only allow uppercase characters? How about if we allow both uppercase and lowercase characters?*

1.2. *Describe how you can create an OR gate using NOT gates and AND gates.*

1.3. *A kilobyte is 1024 bytes. How many messages can it store?*

1.4. *What is the entropy associated with the tossing of a fair coin?*

1.5. *Suppose that X consists of the characters A, B, C, D that occur in a signal with respective probabilities 0.1, 0.4, 0.25, and 0.25. What is the Shannon entropy?*

1.6. *A room full of people has incomes distributed in the following way:*

$$\begin{aligned} n(25.5) &= 3 \\ n(30) &= 5 \\ n(42) &= 7 \\ n(50) &= 3 \\ n(63) &= 1 \\ n(75) &= 2 \\ n(90) &= 1 \end{aligned}$$

What is the most probable income? What is the average income? What is the variance of this distribution?

2

QUBITS AND QUANTUM STATES

In this chapter we will expand on our discussion of the qubit and learn some basic facts and notation that are necessary when learning how to work with quantum states.

THE QUBIT

In the last chapter we saw that the basic unit of information processing in a modern-day computer is the bit, which can assume one of two states that we label 0 and 1. In an analogous manner, we can define a basic unit of information processing that can be used in quantum computation. This basic unit of information in quantum computing is called the *qubit*, which is short for *quantum bit*. While a qubit is going to look in some way superficially similar to a bit, we will see as we go along that it is fundamentally different and that its fundamental difference allows us to do information processing in new and interesting ways.

Like a bit, a qubit can also be in one of two states. In the case of a qubit, for reasons that for the moment will seem utterly obscure, we label these two states

Quantum Computing Explained, by David McMahon by $|0\rangle$ and $|1\rangle$. In quantum theory an object enclosed using the notation $|\ \rangle$ can be called a *state*, a *vector*, or a *ket*.

So how is a qubit any different than an ordinary bit? While a bit in an ordinary computer can be in the state 0 *or* in the state 1, a qubit is somewhat more general. A qubit can exist in the state $|0\rangle$ or the state $|1\rangle$, but it can also exist in what we call a *superposition* state. This is a state that is a linear combination of the states $|0\rangle$ and $|1\rangle$. If we label this state $|\psi\rangle$, a superposition state is written as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \tag{2.1}$$

Here α, β are complex numbers. That is, the numbers are of the form $z = x + iy$, where $i = \sqrt{-1}$.

While a qubit can exist in a superposition of the states $|0\rangle$ and $|1\rangle$, whenever we make a measurement we aren't going to find it like that. In fact, when a qubit is measured, it is only going to be found to be in the state $|0\rangle$ or the state $|1\rangle$. The laws of quantum mechanics tell us that the modulus squared of α, β in (2.1) gives us the probability of finding the qubit in state $|0\rangle$ or $|1\rangle$, respectively. In other words:

$|\alpha|^2$: Tells us the probability of finding $|\psi\rangle$ in state $|0\rangle$
$|\beta|^2$: Tells us the probability of finding $|\psi\rangle$ in state $|1\rangle$

The fact that probabilities must sum to one puts some constraints on what the multiplicative coefficients in (2.1) can be. Since the squares of these coefficients are related to the probability of obtaining a given measurement result, α and β are constrained by the requirement that

$$|\alpha|^2 + |\beta|^2 = 1 \tag{2.2}$$

Generally speaking, if an event has N possible outcomes and we label the probability of finding result i by p_i, the condition that the probabilities sum to one is written as

$$\sum_{i=1}^{N} p_i = p_1 + p_2 + \cdots + p_N = 1 \tag{2.3}$$

When this condition is satisfied for the squares of the coefficients of a qubit, we say that the qubit is *normalized*.

We can calculate the modulus of these numbers in the following way:

$$|\alpha|^2 = (\alpha)(\alpha^*)$$
$$|\beta|^2 = (\beta)(\beta^*)$$

where α^* is the complex conjugate of α and β^* is the complex conjugate of β. We recall that to form the complex conjugate of $z = x + iy$, we let $i \to -i$. Therefore the modulus of a complex number z is

$$|z|^2 = (x+iy)(x-iy) = x^2 + ixy - ixy + y^2$$
$$\Rightarrow |z| = \sqrt{x^2 + y^2}$$

Right now, it might seem as if we have traded in the reliability of a classical computer for a probabilistic guessing game. Later we will see that this isn't the case and that these strange propeies actually make a qubit an asset, rather than a liability, in information processing. For now let's look at a couple of examples to reinforce the basic ideas.

Example 2.1

For each of the following qubits, if a measurement is made, what is the probability that we find the qubit in state $|0\rangle$? What is the probability that we find the qubit in the state $|1\rangle$?

(a) $|\psi\rangle = \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}|1\rangle$

(b) $|\phi\rangle = \frac{i}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle$

(c) $|\chi\rangle = \frac{(1+i)}{\sqrt{3}}|0\rangle - \frac{i}{\sqrt{3}}|1\rangle$

Solution

To find the probability that each qubit is found in the state $|0\rangle$ or the state $|1\rangle$, we compute the modulus squared of the appropriate coefficient.

(a) In this case the probability of finding $|\psi\rangle$ in state $|0\rangle$ is

$$\left|\frac{1}{\sqrt{3}}\right|^2 = \frac{1}{3}$$

The probability that we find $|\psi\rangle$ in state $|1\rangle$ is

$$\left|\frac{\sqrt{2}}{\sqrt{3}}\right|^2 = \frac{2}{3}$$

When doing calculations in quantum mechanics or quantum computation, it is always a good idea to verify that the probabilities you find add to 1. We can label the probability that the system is in state $|0\rangle$ p_0 and the probability that the system is in state $|1\rangle$ p_1. In the context of this example $p_0 = 1/3$ and $p_1 = 2/3$, hence

$$\sum p_i = p_0 + p_1 = \frac{1}{3} + \frac{2}{3} = 1$$

(b) The next state has coefficients that are complex numbers. Remember we need to use the complex conjugate when calculating the modulus squared. We find that the probability of the system being in state $|0\rangle$ is

$$|\phi\rangle = \frac{i}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle, \Rightarrow p_0 = \left|\frac{i}{2}\right|^2 = \left(\frac{i}{2}\right)^* \left(\frac{i}{2}\right) = \left(\frac{-i}{2}\right)\left(\frac{i}{2}\right) = \frac{1}{4}$$

The probability that the system is in state $|1\rangle$ is

$$\left|\frac{\sqrt{3}}{2}\right|^2 = \frac{3}{4}$$

Again, we check that the probabilities sum to one:

$$\sum p_i = p_0 + p_1 = \frac{1}{4} + \frac{3}{4} = 1$$

(c) Finally, for the last state, the probability the system is in state $|0\rangle$ is

$$\left|\frac{1+i}{\sqrt{3}}\right|^2 = \left(\frac{1-i}{\sqrt{3}}\right)\left(\frac{1+i}{\sqrt{3}}\right) = \frac{1-i+i+1}{3} = \frac{2}{3}$$

The probability that the system is in state $|1\rangle$ is

$$\left|\frac{-i}{\sqrt{3}}\right|^2 = \left(\frac{-i}{\sqrt{3}}\right)^* \left(\frac{-i}{\sqrt{3}}\right) = \left(\frac{i}{\sqrt{3}}\right)\left(\frac{-i}{\sqrt{3}}\right) = \frac{1}{3}$$

Again, these probabilities sum to 1:

$$p_0 + p_1 = \frac{2}{3} + \frac{1}{3} = 1$$

VECTOR SPACES

While the basic unit of information in quantum computation is the qubit, the arena in which quantum computation takes place is a mathematical abstraction called a *vector space*. If you've taken elementary physics or engineering classes, then you know what a vector is. It turns out that quantum states behave mathematically in an analogous way to physical vectors—hence the term vector space. This type of space is one that shares with physical vectors the most basic properties that vectors have—for example, a length. In this section we are going to look at state vectors more generally and talk a little bit about the spaces they inhabit. We aren't going to be rigorous here, the purpose of this section is just to introduce some basic ideas and terminology. To avoid further confusion, let's just get down to the basic definitions. A vector space V is a nonempty set with elements u, v called *vectors* for which the following two operations are defined:

1. Vector addition: An operation that assigns the sum $w = u + v$, which is also an element of V; in other words, w is another vector belonging to the same space

2. Scalar multiplication: Defines multiplication of a vector by a number α such that the vector $\alpha u \in V$

In addition the following axioms hold for the vector space V:

Axiom 1: Associativity of addition. Given vectors u, v, and w,

$$(u + v) + w = u + (v + w)$$

Axiom 2: There is a vector belonging to V called the *zero vector* that satisfies

$$u + 0 = 0 + u = u$$

for any vector $u \in V$.

Axiom 3: For every $u \in V$ there exists an additive inverse of u such that

$$u + (-u) = (-u) + u = 0$$

Axiom 4: Addition of vectors is commutative:

$$u + v = v + u$$

There are other axioms that apply to vectors spaces, but these should suffice for most of our purposes.

One particular vector space that is important in quantum computation is the vector space $\mathbb{C}^n$, which is the vector space of "n-tuples" of complex numbers. When we say "n-tuple" this is just a fancy way of referring to an ordered collection of numbers. Following the notation we have introduced for qubits, we label the elements of $\mathbb{C}^n$ by $|a\rangle$, $|b\rangle$, $|c\rangle$. Then we write down an element of this vector space as an n-dimensional *column vector* or simply a list of numbers a_1, a_2, $\ldots$, a_n arranged in the following way:

$$|a\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \tag{2.4}$$

This type of notation can be used with qubits. Look back at the general notation used for qubit given in (2.1). We write this in a column vector format by putting the coefficient of $|0\rangle$ in the first row and the coefficient of $|1\rangle$ in the second row:

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

For a more concrete example, let's take a look back at the qubits we examined in Example 2.1. The first qubit was

$$|\psi\rangle = \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}|1\rangle$$

Hence the column vector representation of this qubit is

$$|\psi\rangle = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \sqrt{\frac{2}{3}} \end{pmatrix} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}$$

In general, the numbers a_i in (2.4), which we call the *components* of the vector, are complex—something we've already mentioned.

Multiplication of a vector by a scalar proceeds as

$$\alpha|a\rangle = \alpha \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \alpha a_1 \\ \alpha a_2 \\ \vdots \\ \alpha a_n \end{pmatrix} \tag{2.5}$$

It's easy to see that this produces another column vector with n complex numbers, so the result is another element in $\mathbb{C}^n$. So $\mathbb{C}^n$ is closed under scalar multiplication.

Vector addition is carried out component by component, producing a new, third vector:

$$|a\rangle + |b\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix} \tag{2.6}$$

This should all be pretty straightforward, but let's illustrate it with an example.

Example 2.2

We want to define the vectors

$$|u\rangle = \begin{pmatrix} -1 \\ 7i \\ 2 \end{pmatrix}, \quad |v\rangle = \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix}$$

and then compute $7|u\rangle + 2|v\rangle$.

Solution

First we compute $7|u\rangle$ by multiplying each of the components of the vector by 7:

$$7|u\rangle = 7\begin{pmatrix} -1 \\ 7i \\ 2 \end{pmatrix} = \begin{pmatrix} -7 \\ 49i \\ 14 \end{pmatrix}$$

Next we compute the scalar multiplication for the other vector:

$$2|v\rangle = 2\begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 8 \end{pmatrix}$$

Now we follow the vector addition rule, adding the vectors component by component:

$$7|u\rangle + 2|v\rangle = \begin{pmatrix} -7 \\ 49i \\ 14 \end{pmatrix} + \begin{pmatrix} 0 \\ 4 \\ 8 \end{pmatrix} = \begin{pmatrix} -7+0 \\ 49i+4 \\ 14+8 \end{pmatrix} = \begin{pmatrix} -7 \\ 4+49i \\ 22 \end{pmatrix}$$

LINEAR COMBINATIONS OF VECTORS

In the previous example we calculated an important quantity, a *linear combination* of vectors. Generally speaking, let α_i be a set of complex coefficients and $|v_i\rangle$ be a set of vectors. A linear combination of these vectors is given by

$$\alpha_1|v_1\rangle + \alpha_2|v_2\rangle + \cdots + \alpha_n|v_n\rangle = \sum_{i=1}^{n} \alpha_i|v_i\rangle \tag{2.7}$$

We've seen a linear combination before, when we considered a superposition state of a qubit.

Now, if a given set of vectors $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$ can be used to represent any vector $|u\rangle$ that belongs to the vector space V, we say that the set $\{|v_i\rangle\}$ *spans* the given vector space. For example, consider the three-dimensional vector space $\mathbb{C}^3$. We can write any three-dimensional column vector in the following way:

$$|u\rangle = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \beta \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \gamma \end{pmatrix} = \alpha\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \beta\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \gamma\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Therefore we see that the set of vectors defined as

$$|v_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |v_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |v_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Spans the space $\mathbb{C}^3$. We've already seen a spanning set for qubits when we considered the basic states a qubit can be found in, $\{|0\rangle, |1\rangle\}$. Recall from (2.1) that an arbitrary qubit can be written as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

Writing this in column vector notation, and then using the properties of vector addition and scalar multiplication, we have

$$|\psi\rangle = \begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}\alpha\\ 0\end{pmatrix} + \begin{pmatrix}0\\ \beta\end{pmatrix} = \alpha\begin{pmatrix}1\\ 0\end{pmatrix} + \beta\begin{pmatrix}0\\ 1\end{pmatrix}$$

This allows us to identify the vectors $|0\rangle$ and $|1\rangle$ as

$$|0\rangle = \begin{pmatrix}1\\ 0\end{pmatrix}, \quad |1\rangle = \begin{pmatrix}0\\ 1\end{pmatrix}$$

An important notion involving the linear combination of a set of vectors is that of *linear independence*. If

$$\alpha_1|v_1\rangle + \alpha_2|v_2\rangle + \cdots + \alpha_n|v_n\rangle = 0$$

and at least one of the $\alpha_i \neq 0$, we say the set $\{|v_i\rangle\}$ is linearly dependent. Another way of saying this is that if one vector of the set can be written as a linear combination of the other vectors in the set, then the set is *linearly dependent*.

Example 2.3

Show that the set

$$|a\rangle = \begin{pmatrix}1\\ 2\end{pmatrix}, \quad |b\rangle = \begin{pmatrix}-1\\ 1\end{pmatrix}, \quad |c\rangle = \begin{pmatrix}5\\ 4\end{pmatrix}$$

is linearly dependent; that is, one of the vectors can be written as a linear combination of the other two.

Solution

We can show that $|c\rangle$ can be expressed as a linear combination of the other two vectors. Let's start out by writing $|c\rangle$ as some arbitrary combination of the other two vectors:

$$\begin{pmatrix}5\\ 4\end{pmatrix} = \alpha\begin{pmatrix}1\\ 2\end{pmatrix} + \beta\begin{pmatrix}-1\\ 1\end{pmatrix}$$

Now we actually have two equations. These are

$$\alpha - \beta = 5, \quad \Rightarrow \beta = \alpha - 5$$
$$2\alpha + \beta = 4$$

Substitution of $\beta = \alpha - 5$ into the second equation gives

$$2\alpha + (\alpha - 5) = 3\alpha - 5 = 4, \Rightarrow \alpha = 3$$

So immediately we see that $\beta = -2$. Using these results, we see that the following relationship holds, demonstrating the linear dependence of this set of vectors:

$$3|a\rangle - 2|b\rangle + |c\rangle = 0$$

UNIQUENESS OF A SPANNING SET

A spanning set of vectors for a given space V is not unique. Once again, consider the complex vector space $\mathbb{C}^2$, consisting of column vectors with two elements. Earlier we saw that we can write any arbitrary column vector with two elements in the following way:

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

And therefore the set

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

spans $\mathbb{C}^2$, the vector space in which qubits live. Now consider the set:

$$|u_1\rangle = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad |u_2\rangle = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

This set also spans the space $\mathbb{C}^2$, and this means that we can also use these vectors to represent qubits. We can write any element of this space in terms of these vectors:

$$\begin{aligned} |\psi\rangle &= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\ &= \begin{pmatrix} \alpha \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \beta \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \alpha + \alpha \\ \alpha - \alpha \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \beta - \beta \\ \beta + \beta \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} \alpha \\ \alpha \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \alpha \\ -\alpha \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \beta \\ \beta \end{pmatrix} - \frac{1}{2}\begin{pmatrix} \beta \\ -\beta \end{pmatrix} \end{aligned}$$

Now we use the scalar multiplication rule to pull the constants α, β outside of the vectors:

$$\frac{1}{2}\begin{pmatrix}\alpha\\ \alpha\end{pmatrix}+\frac{1}{2}\begin{pmatrix}\alpha\\ -\alpha\end{pmatrix}+\frac{1}{2}\begin{pmatrix}\beta\\ \beta\end{pmatrix}-\frac{1}{2}\begin{pmatrix}\beta\\ -\beta\end{pmatrix}$$

$$=\frac{\alpha}{2}\begin{pmatrix}1\\ 1\end{pmatrix}+\frac{\alpha}{2}\begin{pmatrix}1\\ -1\end{pmatrix}+\frac{\beta}{2}\begin{pmatrix}1\\ 1\end{pmatrix}-\frac{\beta}{2}\begin{pmatrix}1\\ -1\end{pmatrix}$$

Rearranging a bit, we find that

$$\left(\frac{\alpha+\beta}{2}\right)\begin{pmatrix}1\\ 1\end{pmatrix}+\left(\frac{\alpha-\beta}{2}\right)\begin{pmatrix}1\\ -1\end{pmatrix}=\left(\frac{\alpha+\beta}{2}\right)|u_1\rangle+\left(\frac{\alpha-\beta}{2}\right)|u_2\rangle$$

BASIS AND DIMENSION

When a set of vectors is linearly independent and they span the space, the set is known as a *basis*. We can express any vector in the space V in terms of a linear expansion on a basis set. Moreover the *dimension* of a vector space V is equal to the number of elements in the basis set. We have already seen a basis set for $\mathbb{C}^2$, with the qubit states $|0\rangle$ and $|1\rangle$:

$$|0\rangle=\begin{pmatrix}1\\ 0\end{pmatrix},\quad |1\rangle=\begin{pmatrix}0\\ 1\end{pmatrix}$$

As we have shown, any vector in $\mathbb{C}^2$ can be written as a linear combination of these two vectors; hence they span the space. It is also pretty easy to show that these two vectors are linearly independent. Therefore they constitute a basis of $\mathbb{C}^2$.

A vector space can have many basis sets. We have already seen another basis set that can be used for $\mathbb{C}^2$ (and hence for qubits):

$$|u_1\rangle=\begin{pmatrix}1\\ 1\end{pmatrix},\quad |u_2\rangle=\begin{pmatrix}1\\ -1\end{pmatrix}$$

In term of basis vectors, quantum states in n dimensions are straightforward generalizations of qubits. A quantum state $|\psi\rangle$ can be written as a linear combination of a basis set $|v_i\rangle$ with complex coefficients of expansion c_i as

$$|\psi\rangle=\sum_{i=1}^{n}c_i|v_i\rangle=c_1|v_1\rangle+c_2|v_2\rangle+\cdots+c_n|v_n\rangle \tag{2.8}$$

The modulus squared of a given coefficient c_i gives the probability that measurement finds the system in the state $|v_i\rangle$.

INNER PRODUCTS

To compute the length of a vector, even if it's a length in an abstract sense, we need a way to find the *inner product*. This is a generalization of the dot product used with ordinary vectors in Euclidean space as you may already be familiar with. While the dot product takes two vectors and maps them into a real number, in our case the inner product will take two vectors from $\mathbb{C}^n$ and map them to a *complex* number. We write the inner product between two vectors $|u\rangle$, $|v\rangle$ with the notation $\langle u|v\rangle$. If the inner product between two vectors is zero,

$$\langle u|v\rangle = 0$$

We say that $|u\rangle$, $|v\rangle$ are *orthogonal* to one another. The inner product is a complex number. The conjugate of this complex number satisfies

$$\langle u|v\rangle^* = \langle v|u\rangle \tag{2.9}$$

We can use the inner product to define a norm (or length)—by computing the inner product of a vector with itself:

$$\|u\| = \sqrt{\langle u|u\rangle} \tag{2.10}$$

Notice that the norm is a real number, and hence can define a length. For any vector $|u\rangle$ we have

$$\langle u|u\rangle \geq 0 \tag{2.11}$$

with equality if and only if $|u\rangle = 0$. When considering the inner product of a vector with a superposition or linear combination of vectors, the following linear and *anti-linear* relations hold:

$$\begin{aligned} \langle u|\alpha v + \beta w\rangle &= \alpha\langle u|v\rangle + \beta\langle u|w\rangle \\ \langle \alpha u + \beta v|w\rangle &= \alpha^*\langle u|w\rangle + \beta^*\langle v|w\rangle \end{aligned} \tag{2.12}$$

To compute the inner product between two vectors, we must calculate the Hermitian conjugate of a vector

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

In quantum physics $\langle u|$ is sometimes called the *dual vector* or *bra* corresponding to $|u\rangle$.

If a ket is a column vector, the dual vector or bra is a row vector whose elements are the complex conjugates of the elements of the column vector. In other words, when working with column vectors, the Hermitian conjugate is computed in two steps:

1. Write the components of the vector as a row of numbers.
2. Take the complex conjugate of each element and arrange them in a row vector.

Generally, we have

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}^{\dagger} = \begin{pmatrix} a_1^* & a_2^* & \ldots & a_n^* \end{pmatrix} \tag{2.13}$$

With qubits, this is fairly easy. Let's go back to the state

$$|\phi\rangle = \frac{i}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle$$

The column vector representation of this state is

$$|\phi\rangle = \begin{pmatrix} \frac{i}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}$$

The bra or dual vector is found by computing the complex conjugate of each element and then arranging the result as a row vector. In this case

$$\langle\phi| = \begin{pmatrix} -\frac{i}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

Using the dual vector to find the inner product makes the calculation easy. The inner product $\langle a|b\rangle$ is calculated in the following way:

$$\langle a|b\rangle = \begin{pmatrix} a_1^* & a_2^* & \ldots & a_n^* \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1^* b_1 + a_2^* b_2 + \cdots + a_n^* b_n = \sum_{i=1}^{n} a_i^* b_i \tag{2.14}$$

Example 2.4

Two vectors in $\mathbb{C}^3$ are given by

$$|a\rangle = \begin{pmatrix} -2 \\ 4i \\ 1 \end{pmatrix}, \quad |b\rangle = \begin{pmatrix} 1 \\ 0 \\ i \end{pmatrix}$$

Find

(a) $\langle a|$, $\langle b|$

(b) $\langle a|b\rangle$, $\langle b|a\rangle$
(c) $|c\rangle = |a\rangle + 2|b\rangle$, $\langle c|a\rangle$

Solution

(a) We consider $|a\rangle$ first. Begin by rewriting the elements in a row vector format. This is known as the *transpose* operation:

$$(|a\rangle)^T = (-2 \quad 4i \quad 1)$$

The Hermitian conjugate is then found by computing the complex conjugate of each element:

$$\langle a| = (|a\rangle)^\dagger = \left(|a\rangle^T\right)^* = (-2 \quad -4i \quad 1)$$

A similar procedure for $|b\rangle$ yields

$$\langle b| = (1 \quad 0 \quad -i)$$

(b) The inner product, from (2.14), is given by

$$\langle a|b\rangle = (-2 \quad 4i \quad 1)\begin{pmatrix} 1 \\ 0 \\ i \end{pmatrix} = -2^*1 + 4i^*0 + 1^*i = -2 + i$$

Remember, the inner product on a complex vector space obeys $\langle a|b\rangle = \langle b|a\rangle^*$; therefore we should find $\langle b|a\rangle = -2 - i$. We verify this with an explicit calculation

$$\langle b|a\rangle = (1 \quad 0 \quad -i)\begin{pmatrix} -2 \\ 4i \\ 1 \end{pmatrix} = 1^* - 2 + 0^*4i + (-i)^*1 = -2 - i$$

(c) We apply the rules of vector addition and scalar multiplication to obtain

$$|c\rangle = |a\rangle + 2|b\rangle = \begin{pmatrix} -2 \\ 4i \\ 1 \end{pmatrix} + 2\begin{pmatrix} 1 \\ 0 \\ i \end{pmatrix} = \begin{pmatrix} -2 \\ 4i \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \\ 2i \end{pmatrix} = \begin{pmatrix} -2+2 \\ 4i+0 \\ 1+2i \end{pmatrix} = \begin{pmatrix} 0 \\ 4i \\ 1+2i \end{pmatrix}$$

Therefore the inner product is

$$\langle c|a\rangle = (0 \quad -4i \quad 1-2i)\begin{pmatrix} -2 \\ 4i \\ 1 \end{pmatrix} = 0^*(-2) + (-4i)(4i) + (1-2i)^*1$$
$$= 16 + 1 - 2i = 17 - 2i$$

As an exercise, verify that the form of $\langle c|$ is correct by computing the Hermitian conjugate of $|c\rangle$.

Example 2.5

Compute the norm of the following vectors:

$$|u\rangle = \begin{pmatrix} 2 \\ 4i \end{pmatrix}, \quad |v\rangle = \begin{pmatrix} -1 \\ 3i \\ i \end{pmatrix}$$

Solution

We start by computing the Hermitian conjugate of each vector. Remember, first take the transpose of the vector and write the list of numbers as a row; then compute the complex conjugate of each element:

$$(|u\rangle)^T = (2 \quad 4i), \quad \Rightarrow \langle u| = (|u\rangle)^\dagger = (|u\rangle^T)^* = (2 \quad -4i)$$
$$\langle v| = (-1 \quad 3i \quad i)^* = (-1 \quad -3i \quad -i)$$

We find that

$$\langle u|u\rangle = (2 \quad -4i)\begin{pmatrix} 2 \\ 4i \end{pmatrix} = 2 * 2 + (-4i)^* 4i = 4 + 16 = 20$$
$$\langle v|v\rangle = (-1 \quad -3i \quad -i)\begin{pmatrix} -1 \\ 3i \\ i \end{pmatrix} = -1^*(-1) + (-3i) * 3i + (-i) * i = 1 + 9 + 1 = 11$$

The norm is found by taking the square root of these quantities:

$$\|u\| = \sqrt{\langle u|u\rangle} = \sqrt{20}$$
$$\|v\| = \sqrt{\langle v|v\rangle} = \sqrt{11}$$

ORTHONORMALITY

When the norm of a vector is unity, we say that vector is *normalized*. That is, if

$$\langle a|a\rangle = 1$$

then we say that $|a\rangle$ is normalized. If a vector is not normalized, we can generate a normalized vector by computing the norm (which is just a number) and dividing the vector by it. For the vectors in the previous example, $|u\rangle$, $|v\rangle$ are not normalized, since we found that $\|u\| = \sqrt{20}$ and $\|v\| = \sqrt{11}$. But the vectors

$$|\tilde{u}\rangle = \frac{|u\rangle}{\|u\|} = \frac{1}{\sqrt{20}}|u\rangle$$

$$|\tilde{v}\rangle = \frac{|v\rangle}{\|v\|} = \frac{1}{\sqrt{11}}|v\rangle$$

are normalized. This is easy to see. Check the first case:

$$\langle\tilde{u}|\tilde{u}\rangle = \left(\frac{1}{\sqrt{20}}\langle u|\right)\left(\frac{1}{\sqrt{20}}|u\rangle\right) = \frac{1}{20}\langle u|u\rangle = \frac{20}{20} = 1$$

If each element of a set of vectors is normalized and the elements are orthogonal with respect to each other, we say the set is *orthonormal*. For example, consider the set $\{|0\rangle, |1\rangle\}$. Remember, we have the following definition:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Therefore, using (2.14), we have

$$\langle 0|0\rangle = (1 \quad 0)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1*1 + 0*0 = 1$$

$$\langle 0|1\rangle = (1 \quad 0)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1*0 + 0*1 = 0$$

$$\langle 1|0\rangle = (0 \quad 1)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0*1 + 1*0 = 0$$

$$\langle 1|1\rangle = (0 \quad 1)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0*0 + 1*1 = 1$$

By showing that $\langle 0|0\rangle = \langle 1|1\rangle = 1$, we showed the vectors are normalized, while showing that $\langle 0|1\rangle = \langle 1|0\rangle = 0$, we showed they were orthogonal. Hence the set is orthonormal. Earlier we saw that any qubit could also be written as a linear combination of the vectors:

$$|u_1\rangle = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad |u_2\rangle = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Is this set orthonormal? Well the vectors are orthogonal, since

$$\langle u_1|u_2\rangle = (1 \quad 1)\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1 - 1 = 0$$

But the vectors are not normalized. It's easy to see that $\|u_1\| = \|u_2\| = \sqrt{2}$. We can create an orthonormal set from these vectors by normalizing them. We denote these vectors by $|+\rangle$ and $|-\rangle$. Therefore another orthonormal basis set for $\mathbb{C}^2$ (and hence for qubits) is

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

GRAM-SCHMIDT ORTHOGONALIZATION

An orthonormal basis can be produced from an *arbitrary* basis by application of the *Gram-Schmidt orthogonalization* process. Let $\{|v_1\rangle, |v_2\rangle, \ldots\}|v_n\rangle$, be a basis for an inner product space V. The Gram-Schmidt process constructs an orthogonal basis $|w_i\rangle$ as follows:

$$\begin{aligned}
|w_1\rangle &= |v_1\rangle \\
|w_2\rangle &= |v_2\rangle - \frac{\langle w_1|v_2\rangle}{\langle w_1|w_1\rangle}|w_1\rangle \\
&\vdots \\
|w_n\rangle &= |v_n\rangle - \frac{\langle w_1|v_n\rangle}{\langle w_1|w_1\rangle}|w_1\rangle - \frac{\langle w_2|v_n\rangle}{\langle w_2|w_2\rangle}|w_2\rangle - \cdots - \frac{\langle w_{n-1}|v_n\rangle}{\langle w_{n-1}|w_{n-1}\rangle}|w_{n-1}\rangle
\end{aligned}$$

To form an orthonormal set using the Gram-Schmidt procedure, divide each vector by its norm. For example, the normalized vector we can use to construct $|w_2\rangle$ is

$$|w_2\rangle = \frac{|v_2\rangle - \langle w_1|v_2\rangle|w_1\rangle}{\||v_2\rangle - \langle w_1|v_2\rangle|w_1\rangle\|}$$

Many readers might find this a bit abstract, so let's illustrate with a concrete example.

Example 2.6

Use the Gram-Schmidt process to construct an orthonormal basis set from

$$|v_1\rangle = \begin{pmatrix}1\\2\\-1\end{pmatrix}, \quad |v_2\rangle = \begin{pmatrix}0\\1\\-1\end{pmatrix}, \quad |v_3\rangle = \begin{pmatrix}3\\-7\\1\end{pmatrix}$$

Solution

We use a tilde character to denote the unnormalized vectors. The first basis vector is

$$|\tilde{w}_1\rangle = |v_1\rangle$$

Now let's normalize this vector

$$\langle v_1|v_1\rangle = (1 \quad 2 \quad -1)\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = 1*1+2*2+(-1)*(-1) = 1+4+1 = 6$$

$$\Rightarrow |w_1\rangle = \frac{|\tilde{w}_1\rangle}{\sqrt{\langle v_1|v_1\rangle}} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

Looking back at the algorithm for the Gram-Schmidt process, we construct the second vector by using the formula

$$|\tilde{w}_2\rangle = |v_2\rangle - \frac{\langle \tilde{w}_1|v_2\rangle}{\langle \tilde{w}_1|\tilde{w}_1\rangle}|\tilde{w}_1\rangle$$

First we compute

$$\langle \tilde{w}_1|v_2\rangle = (1 \quad 2 \quad -1)\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = [1*0+2*1+(-1)*(-1)] = 3$$

We already normalized $|w_1\rangle$, and so

$$|\tilde{w}_2\rangle = |v_2\rangle - \frac{\langle \tilde{w}_1|v_2\rangle}{\langle \tilde{w}_1|\tilde{w}_1\rangle}|\tilde{w}_1\rangle = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} - \frac{3}{6}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix}$$

Now we normalize this vector

$$\langle \tilde{w}_2|\tilde{w}_2\rangle = \left(-\frac{1}{2} \quad 0 \quad -\frac{1}{2}\right)\begin{pmatrix} -\frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix} = \frac{1}{4}+0+\frac{1}{4} = \frac{1}{2}$$

So a second normalized vector is

$$|w_2\rangle = \frac{1}{\sqrt{\langle \tilde{w}_2|\tilde{w}_2\rangle}}|\tilde{w}_2\rangle = \sqrt{2}\begin{pmatrix} -\frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

We check to see that this orthogonal to $|w_1\rangle$:

$$\langle w_1|w_2\rangle = \frac{1}{\sqrt{6}}(1 \quad 2 \quad -1)\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{6}}\left[-\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\right] = 0$$

Finally, the third vector is found from

$$|\tilde{w}_3\rangle = |v_3\rangle - \frac{\langle\tilde{w}_1|v_n\rangle}{\langle\tilde{w}_1|\tilde{w}_1\rangle}|\tilde{w}_1\rangle - \frac{\langle\tilde{w}_2|v_3\rangle}{\langle\tilde{w}_2|\tilde{w}_2\rangle}|\tilde{w}_2\rangle$$

Now we have

$$\langle\tilde{w}_2|v_3\rangle = \begin{pmatrix} -\frac{1}{2} & 0 & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} 3 \\ -7 \\ 1 \end{pmatrix} = -\frac{3}{2} - \frac{1}{2} = -\frac{4}{2} = -2$$

Therefore

$$|\tilde{w}_3\rangle = |v_3\rangle - \frac{\langle\tilde{w}_1|v_3\rangle}{\langle\tilde{w}_1|\tilde{w}_1\rangle}|\tilde{w}_1\rangle - \frac{\langle\tilde{w}_2|v_3\rangle}{\langle\tilde{w}_2|\tilde{w}_2\rangle}|\tilde{w}_2\rangle = \begin{pmatrix} 3 \\ -7 \\ 1 \end{pmatrix} + \frac{12}{6}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \frac{2}{\left(\frac{1}{2}\right)}\begin{pmatrix} -\frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix}$$

$$= \begin{pmatrix} 3 \\ -7 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 4 \\ -2 \end{pmatrix} + \begin{pmatrix} -2 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \\ -3 \end{pmatrix}$$

Normalizing, we find that

$$\langle\tilde{w}_3|\tilde{w}_3\rangle = (3 \quad -3 \quad -3)\begin{pmatrix} 3 \\ -3 \\ -3 \end{pmatrix} = 9 + 9 + 9 = 27$$

Therefore

$$|w_3\rangle = \frac{1}{\sqrt{\langle\tilde{w}_3|\tilde{w}_3\rangle}}|\tilde{w}_3\rangle = \frac{1}{\sqrt{27}}\begin{pmatrix} 3 \\ -3 \\ -3 \end{pmatrix} = \frac{1}{3\sqrt{3}}\begin{pmatrix} 3 \\ -3 \\ -3 \end{pmatrix} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$$

BRA-KET FORMALISM

In quantum mechanics, once you have a state specified in terms of a basis it is not necessary to work directly with the components of each vector. An alternative way of doing calculations is to represent states in an orthonormal basis and do calculations using "bras" and "kets." We are already familiar with this formalism, remember a ket is just another label for the vector notation we have been using. Examples of kets are

$$|\psi\rangle, \quad |\phi\rangle, \quad |0\rangle$$

We saw above that "bras" or dual vectors are the Hermitian conjugates of the corresponding kets. Abstractly, the bras corresponding to the kets above are

$$\langle\psi|, \quad \langle\phi|, \quad \langle 0|$$

Let's demonstrate how to work in this formalism using an example.

Example 2.7

Suppose that $\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}$ is an orthonormal basis for a three-dimensional Hilbert space. A system is in the state

$$|\psi\rangle = \frac{1}{\sqrt{5}}|u_1\rangle - i\sqrt{\frac{7}{15}}|u_2\rangle + \frac{1}{\sqrt{3}}|u_3\rangle$$

(a) Is this state normalized?

(b) If a measurement is made, find the probability of finding the system in each of the states $\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}$

Solution

(a) We know that $\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}$ is an orthonormal basis. To do our calculations, we rely on the following:

$$\langle u_1|u_1\rangle = \langle u_2|u_2\rangle = \langle u_3|u_3\rangle = 1$$
$$\langle u_1|u_2\rangle = \langle u_1|u_3\rangle = \langle u_2|u_3\rangle = 0$$

To show the state is normalized, we must show that

$$\langle\psi|\psi\rangle = 1$$

First we calculate the Hermitian conjugate of the ket $|\psi\rangle$:

$$|\psi\rangle^\dagger = \langle\psi| = \left(\frac{1}{\sqrt{5}}|u_1\rangle\right)^\dagger - \left(i\sqrt{\frac{7}{15}}|u_2\rangle\right)^\dagger + \left(\frac{1}{\sqrt{3}}|u_3\rangle\right)^\dagger$$
$$= \frac{1}{\sqrt{5}}\langle u_1| - \left(-i\sqrt{\frac{7}{15}}\right)\langle u_2| + \frac{1}{\sqrt{3}}\langle u_3|$$

Next we compute the inner product

$$\langle\psi|\psi\rangle = \left(\frac{1}{\sqrt{5}}\langle u_1| + i\sqrt{\frac{7}{15}}\langle u_2| + \frac{1}{\sqrt{3}}\langle u_3|\right)\left(\frac{1}{\sqrt{5}}|u_1\rangle - i\sqrt{\frac{7}{15}}|u_2\rangle + \frac{1}{\sqrt{3}}|u_3\rangle\right)$$

$$= \frac{1}{5}\langle u_1|u_1\rangle + \left(\frac{1}{\sqrt{5}}\right)\left(-i\sqrt{\frac{7}{15}}\right)\langle u_1|u_2\rangle + \left(\frac{1}{\sqrt{5}}\right)\left(\frac{1}{\sqrt{3}}\right)\langle u_1|u_3\rangle$$

$$+ \left(i\sqrt{\frac{7}{15}}\right)\left(\frac{1}{\sqrt{5}}\right)\langle u_2|u_1\rangle + \left(i\sqrt{\frac{7}{15}}\right)\left(-i\sqrt{\frac{7}{15}}\right)\langle u_2|u_2\rangle$$

$$+ \left(i\sqrt{\frac{7}{15}}\right)\left(\frac{1}{\sqrt{3}}\right)\langle u_2|u_3\rangle + \left(\frac{1}{\sqrt{3}}\right)\left(\frac{1}{\sqrt{5}}\right)\langle u_3|u_1\rangle$$

$$+ \left(\frac{1}{\sqrt{3}}\right)\left(-i\sqrt{\frac{7}{15}}\right)\langle u_3|u_2\rangle + \left(\frac{1}{3}\right)\langle u_3|u_3\rangle$$

Because of the orthonormality relations the only terms that survive are

$$\langle u_1|u_1\rangle = \langle u_2|u_2\rangle = \langle u_3|u_3\rangle = 1$$

and so

$$\langle\psi|\psi\rangle = \frac{1}{5}\langle u_1|u_1\rangle + \left(i\sqrt{\frac{7}{15}}\right)\left(-i\sqrt{\frac{7}{15}}\right)\langle u_2|u_2\rangle + \frac{1}{3}\langle u_3|u_3\rangle$$

$$= \frac{1}{5} + \frac{7}{15} + \frac{1}{3} = \frac{3}{15} + \frac{7}{15} + \frac{5}{15} = \frac{15}{15} = 1$$

Therefore the state is normalized. Another easy way to verify this is to check that the probabilities sum to one.

(b) The probability that the system is found to be in state $|u_1\rangle$ upon measurement is

$$|\langle u_1|\psi\rangle|^2$$

So we obtain

$$\langle u_1|\psi\rangle = \langle u_1|\left(\frac{1}{\sqrt{5}}|u_1\rangle + i\sqrt{\frac{7}{15}}|u_2\rangle + \frac{1}{\sqrt{3}}|u_3\rangle\right)$$

$$= \frac{1}{\sqrt{5}}\langle u_1|u_1\rangle + i\sqrt{\frac{7}{15}}\langle u_1|u_2\rangle + \frac{1}{\sqrt{3}}\langle u_1|u_3\rangle$$

$$= \frac{1}{\sqrt{5}} * 1 + i\sqrt{\frac{7}{15}} * 0 + \frac{1}{\sqrt{3}} * 0 = \frac{1}{\sqrt{5}}$$

The probability is found by calculating the modulus squared of this term

$$p_1 = |\langle u_1|\psi\rangle|^2 = \left|\frac{1}{\sqrt{5}}\right|^2 = \frac{1}{5}$$

For $|u_2\rangle$ we obtain

$$\begin{aligned}\langle u_2|\psi\rangle &= \langle u_2|\left(\frac{1}{\sqrt{5}}|u_1\rangle + i\sqrt{\frac{7}{15}}|u_2\rangle + \frac{1}{\sqrt{3}}|u_3\rangle\right)\\ &= \frac{1}{\sqrt{5}}\langle u_2|u_1\rangle + i\sqrt{\frac{7}{15}}\langle u_2|u_2\rangle + \frac{1}{\sqrt{3}}\langle u_2|u_3\rangle\\ &= \frac{1}{\sqrt{5}} * 0 + i\sqrt{\frac{7}{15}} * 1 + \frac{1}{\sqrt{3}} * 0 = i\sqrt{\frac{7}{15}}\end{aligned}$$

The probability is

$$p_2 = |\langle u_2|\psi\rangle|^2 = \left|i\sqrt{\frac{7}{15}}\right|^2 = \left(-i\sqrt{\frac{7}{15}}\right)\left(i\sqrt{\frac{7}{15}}\right) = \frac{7}{15}$$

Finally, for $|u_3\rangle$ we find that

$$\begin{aligned}\langle u_3|\psi\rangle &= \langle u_2|\left(\frac{1}{\sqrt{5}}|u_1\rangle + i\sqrt{\frac{7}{15}}|u_2\rangle + \frac{1}{\sqrt{3}}|u_3\rangle\right)\\ &= \frac{1}{\sqrt{5}}\langle u_3|u_1\rangle + i\sqrt{\frac{7}{15}}\langle u_3|u_2\rangle + \frac{1}{\sqrt{3}}\langle u_3|u_3\rangle\\ &= \frac{1}{\sqrt{5}} * 0 + i\sqrt{\frac{7}{15}} * 0 + \frac{1}{\sqrt{3}} * 1 = \frac{1}{\sqrt{3}}\end{aligned}$$

Therefore the probability the system is found to be in state $|u_3\rangle$ is

$$p_3 = |\langle u_3|\psi\rangle|^2 = \left|\frac{1}{\sqrt{3}}\right|^2 = \frac{1}{3}$$

THE CAUCHY-SCHWARTZ AND TRIANGLE INEQUALITIES

Two important identities are the Cauchy-Schwarz inequality

$$|\langle\psi|\phi\rangle|^2 \leq \langle\psi|\psi\rangle\langle\phi|\phi\rangle$$

and the triangle inequality

$$\sqrt{\langle\psi+\phi|\psi+\phi\rangle} \leq \sqrt{\langle\psi|\psi\rangle} + \sqrt{\langle\phi|\phi\rangle}$$

Example 2.8

Suppose

$$|\psi\rangle = 3|0\rangle - 2i|1\rangle, \quad |\phi\rangle = |0\rangle + 5|1\rangle$$

(a) Show that these states obey the Cauchy-Schwarz and triangle inequalities.
(b) Normalize the states.

Solution

(a) First we compute $\langle\psi|\psi\rangle$, $\langle\phi|\phi\rangle$:

$$\begin{aligned}
\langle\psi| &= (3)^*\langle 0| + (-2i)^*\langle 1| = 3\langle 0| + 2i\langle 1| \\
\Rightarrow \langle\psi|\psi\rangle &= (3\langle 0| + 2i\langle 1|)(3|0\rangle - 2i|1\rangle) \\
&= 9\langle 0|0\rangle + (2i)(-2i)\langle 1|1\rangle \\
&= 9 + 4 = 13
\end{aligned}$$

For $|\phi\rangle$ we obtain

$$\begin{aligned}
\langle\phi| &= \langle 0| + 5\langle 1| \\
\Rightarrow \langle\phi|\phi\rangle &= (\langle 0| + 5\langle 1|)(|0\rangle + 5|1\rangle) \\
&= \langle 0|0\rangle + 25\langle 1|1\rangle \\
&= 1 + 25 = 26
\end{aligned}$$

The inner product $\langle\psi|\phi\rangle$ is given by

$$\begin{aligned}
\langle\psi|\phi\rangle &= (3\langle 0| + 2i\langle 1|)(|0\rangle + 5|1\rangle) \\
&= 3\langle 0|0\rangle + (2i)(5)\langle 1|1\rangle \\
&= 3 + 10i \\
&\Rightarrow |\langle\psi|\phi\rangle|^2 = \langle\psi|\phi\rangle\langle\psi|\phi\rangle^* = (3 + 10i)(3 - 10i) = 9 + 100 = 109
\end{aligned}$$

Now we have

$$\langle\psi|\psi\rangle\langle\phi|\phi\rangle = (13)(26) = 338 > 109$$

Therefore the Cauchy-Schwarz inequality is satisfied. Next we check the triangle inequality:

$$\sqrt{\langle\psi|\psi\rangle} + \sqrt{\langle\phi|\phi\rangle} = \sqrt{13} + \sqrt{26} \cong 8.7$$

For the left side we have

$$|\psi + \phi\rangle = 3|0\rangle - 2i|1\rangle + |0\rangle + 5|1\rangle = (3 + 1)|0\rangle + (5 - 2i)|1\rangle = 4|0\rangle + (5 - 2i)|1\rangle$$

The bra corresponding to this ket is

$$\langle\psi+\phi| = 4\langle 0| + (5-2i)^*\langle 1| = 4\langle 0| + (5+2i)\langle 1|$$

Then the norm of the state is found to be

$$\begin{aligned}\langle\psi+\phi|\psi+\phi\rangle &= [4\langle 0| + (5+2i)\langle 1|][4|0\rangle + (5-2i)|1\rangle]\\ &= 16\langle 0|0\rangle + (5+2i)(5-2i)\langle 1|1\rangle\\ &= 16+10+4 = 30\end{aligned}$$

With this result in hand we find that

$$\sqrt{\langle\psi+\phi|\psi+\phi\rangle} = \sqrt{30} \simeq 5.5 < 8.7$$

Therefore the triangle inequality is satisfied.

(b) The inner products of these states are

$$\langle\psi|\psi\rangle = 13, \quad \Rightarrow \||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle} = \sqrt{13}$$

$$\langle\phi|\phi\rangle = 26, \quad \Rightarrow \||\phi\rangle\| = \sqrt{\langle\phi|\phi\rangle} = \sqrt{26}$$

So we obtain the normalized states

$$|\tilde{\psi}\rangle = \frac{|\psi\rangle}{\sqrt{\langle\psi|\psi\rangle}} = \frac{1}{\sqrt{13}}(3|0\rangle - 2i|1\rangle) = \frac{3}{\sqrt{13}}|0\rangle - \frac{2i}{\sqrt{13}}|1\rangle$$

$$|\tilde{\phi}\rangle = \frac{|\phi\rangle}{\sqrt{\langle\phi|\phi\rangle}} = \frac{1}{\sqrt{26}}(|0\rangle + 5|1\rangle) = \frac{1}{\sqrt{26}}|0\rangle + \frac{5}{\sqrt{26}}|1\rangle$$

Example 2.9

The qubit trine is defined by the states

$$\begin{aligned}|\psi_0\rangle &= |0\rangle\\ |\psi_1\rangle &= -\frac{1}{2}|0\rangle - \frac{\sqrt{3}}{2}|1\rangle\\ |\psi_2\rangle &= -\frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle\end{aligned}$$

Find states $|\overline{\psi}_0\rangle$, $|\overline{\psi}_1\rangle$, $|\overline{\psi}_2\rangle$ that are normalized, superposition states of $\{|0\rangle, |1\rangle\}$, and orthogonal to $|\psi_0\rangle$,$|\psi_1\rangle$,$|\psi_2\rangle$, respectively.

Solution

We write the states as superpositions with unknown coefficients

$$|\overline{\psi}_0\rangle = \alpha_0|0\rangle + \beta_0|1\rangle$$

$$|\overline{\psi}_1\rangle = \alpha_1|0\rangle + \beta_1|1\rangle$$

$$|\overline{\psi}_2\rangle = \alpha_2|0\rangle + \beta_2|1\rangle$$

For the first of these states, the requirement that it be orthogonal to $|\psi_0\rangle$ allows us to write

$$\langle\overline{\psi}_0|\psi_0\rangle = (\alpha_0^*\langle 0| + \beta_0^*\langle 1|)(|0\rangle) = \alpha_0^*\langle 0|0\rangle = \alpha_0^*$$

The requirement that $\langle\overline{\psi}_0|\psi_0\rangle = 0$ tells us that $\alpha_0 = 0$. So we have reduced the state $|\overline{\psi}_0\rangle$ to

$$|\overline{\psi}_0\rangle = \beta_0|1\rangle$$

The states are normalized, and we obtain

$$\begin{aligned}1 &= \langle\overline{\psi}_0|\overline{\psi}_0\rangle = \beta_0^*\langle 1|(\beta_0|1\rangle) = |\beta_0|^2\langle 1|1\rangle = |\beta_0|^2 \\ &\Rightarrow \beta_0 = 1\end{aligned}$$

The orthogonality requirement for $|\overline{\psi}_1\rangle$ gives us

$$\begin{aligned}\langle\overline{\psi}_1|\psi_1\rangle &= (\alpha_1^*\langle 0| + \beta_1^*\langle 1|)\left(-\frac{1}{2}|0\rangle - \frac{\sqrt{3}}{2}|1\rangle\right) \\ &= -\frac{\alpha_1^*}{2}\langle 0|0\rangle - \beta_1^*\frac{\sqrt{3}}{2}\langle 1|1\rangle \\ &= -\frac{\alpha_1^*}{2} - \beta_1^*\frac{\sqrt{3}}{2}\end{aligned}$$

This term must be zero, so

$$\alpha_1^* = -\sqrt{3}\beta_1^*$$

The requirement that the state be normalized gives

$$\begin{aligned}1 = \langle\overline{\psi}_1|\overline{\psi}_1\rangle &= (\alpha_1^*\langle 0| + \beta_1^*\langle 1|)(\alpha_1|0\rangle + \beta_1|1\rangle) \\ &= |\alpha_1|^2\langle 0|0\rangle + |\beta_1|^2\langle 1|1\rangle \\ &= |\alpha_1|^2 + |\beta_1|^2 \\ \alpha_1^* &= -\sqrt{3}\beta_1^*, \Rightarrow \alpha_1 = -\sqrt{3}\beta_1 \\ 1 &= 3|\beta_1|^2 + |\beta_1|^2 = 4|\beta_1|^2\end{aligned}$$

We take

$$\beta_1 = -\frac{1}{2}$$

So the state is

$$|\overline{\psi}_1\rangle = \frac{\sqrt{3}}{2}|0\rangle - \frac{1}{2}|1\rangle$$

For the final state the orthogonality requirement gives us

$$\langle\overline{\psi}_2|\psi_2\rangle = (\alpha_2^*\langle 0| + \beta_2^*\langle 1|)\left(-\frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle\right)$$

$$= -\frac{\alpha_2^*}{2}\langle 0|0\rangle + \beta_2^*\frac{\sqrt{3}}{2}\langle 1|1\rangle$$

$$= -\frac{\alpha_2^*}{2} + \beta_2^*\frac{\sqrt{3}}{2}$$

Again, if these states are orthogonal, this term must be equal to zero

$$-\frac{\alpha_2^*}{2} + \frac{\sqrt{3}}{2}\beta_2^* = 0$$

$$\Rightarrow \alpha_2^* = \sqrt{3}\beta_2^*$$

From normalization we find that

$$1 = \langle\overline{\psi}_2|\overline{\psi}_2\rangle = (\alpha_2^*\langle 0| + \beta_2^*\langle 1|)(\alpha_2|0\rangle + \beta_2|1\rangle)$$

$$= |\alpha_2|^2\langle 0|0\rangle + |\beta_2|^2\langle 1|1\rangle$$

$$= |\alpha_2|^2 + |\beta_2|^2$$

$$\alpha_2^* = \sqrt{3}\beta_2^*, \quad \Rightarrow \alpha_2 = \sqrt{3}\beta_2$$

$$1 = 3|\beta_2|^2 + |\beta_2|^2 = 4|\beta_2|^2$$

Because $\beta_2 = \frac{1}{2}$, the state is

$$|\overline{\psi}_2\rangle = \frac{\sqrt{3}}{2}|0\rangle + \frac{1}{2}|1\rangle$$

SUMMARY

In this chapter we introduced the notion of a quantum state, with particular attention paid to the two-state system known as a qubit. An arbitrary qubit can be written as a superposition of the basis states $|0\rangle$ and $|1\rangle$ as $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, where $|\alpha|^2$ gives the probability of finding $|\psi\rangle$ in the state $|0\rangle$ and $|\beta|^2$ gives the probability of finding the system in the state $|1\rangle$. We then learned what a spanning set was and when a set of vectors is linearly independent and constitutes a basis. In the next chapter we will see how to work with qubits and other quantum states by considering the notion of operators and measurement.

EXERCISES

2.1. *A quantum system is in the state*

$$\frac{(1-i)}{\sqrt{3}}|0\rangle + \frac{1}{\sqrt{3}}|1\rangle$$

If a measurement is made, what is the probability the system is in state $|0\rangle$ or in state $|1\rangle$?

2.2. *Two quantum states are given by*

$$|a\rangle = \begin{pmatrix} -4i \\ 2 \end{pmatrix}, \quad |b\rangle = \begin{pmatrix} 1 \\ -1+i \end{pmatrix}$$

(A) Find $|a+b\rangle$.
(B) Calculate $3|a\rangle - 2|b\rangle$.
(C) Normalize $|a\rangle$, $|b\rangle$.

2.3. *Another basis for $\mathbb{C}^2$is*

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

Invert this relation to express $\{|0\rangle, |1\rangle\}$ in terms of $\{|+\rangle, |-\rangle\}$.

2.4. *A quantum system is in the state*

$$|\psi\rangle = \frac{3i|0\rangle + 4|1\rangle}{5}$$

(A) Is the state normalized?
(B) Express the state in the $|+\rangle$, $|-\rangle$ basis.

2.5. *Use the Gram-Schmidt process to find an orthonormal basis for a subspace of the four-dimensional space $\mathbb{R}^4$spanned by*

$$|u_1\rangle = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad |u_2\rangle = \begin{pmatrix} 1 \\ 2 \\ 4 \\ 5 \end{pmatrix}, \quad |u_3\rangle = \begin{pmatrix} 1 \\ -3 \\ -4 \\ -2 \end{pmatrix}$$

2.6. *Photon horizontal and vertical polarization states are written as* $|h\rangle$ *and* $|v\rangle$*, respectively. Suppose*

$$
\begin{aligned}
|\psi_1\rangle &= \frac{1}{2}|h\rangle + \frac{\sqrt{3}}{2}|v\rangle \\
|\psi_2\rangle &= \frac{1}{2}|h\rangle - \frac{\sqrt{3}}{2}|v\rangle \\
|\psi_3\rangle &= |h\rangle
\end{aligned}
$$

Find

$$
|\langle\psi_1|\psi_2\rangle|^2, \quad |\langle\psi_1|\psi_3\rangle|^2, \quad |\langle\psi_3|\psi_2\rangle|^2
$$