Introduction to Quantum Computing
Lecture Notes

Amir Karamlou, Francisca Vasconcelos, and Megan Yamoah

Massachusetts Institute of Technology
6.s089 IAP 2019
## Contents

1 Introduction 3

2 Introduction to Quantum Mechanics 3
   2.1 Linear Algebra Review 4
   2.2 Dirac Notation 5
   2.3 Linear Algebra in QM 6
      2.3.1 Inner Product 6
      2.3.2 Operators 7
      2.3.3 Tensor Product 8
   2.4 Postulates of QM 9
      2.4.1 First Postulate 9
      2.4.2 Second Postulate 10
      2.4.3 Third Postulate 10
      2.4.4 Fourth Postulate 11
      2.4.5 Fifth Postulate 12
      2.4.6 Sixth Postulate 12
   2.5 Uncertainty Principle 13
   2.6 Local Hidden Variables 15
   2.7 The Bloch Sphere 17
1 Introduction

This course, taught during the MIT January term over 11 classes of two hours each, aims to serve as a general introduction to the field of quantum computing. It is tailored to the motivated underclass undergraduate and assumes no background in either quantum mechanics or classical computing theory, though it is built on many of the simpler concepts in linear algebra. As such, the course begins with a general introduction to quantum mechanics, followed by a study of the formalism used in the quantum computing field. Then, we will cover fundamental quantum algorithms and their applications before diving into more complex topics such as quantum machine learning and the operational methods of currently researched quantum computing architectures.

These notes cover all topics presented in the lectures and cite applicable references in the field for the interested reader. By no means are they a fully comprehensive introduction to quantum computing, but they should, at minimum, lay a foundation for further investigation into the topic.

2 Introduction to Quantum Mechanics

Quantum mechanics is, in short, the theory which describes the workings of the entire universe but which only manifests itself in certain regimes where “classical” effects do not take over (namely at small lengths, low temperatures, low energies, high pressures, etc.). It presents various un-intuitive ideas as demonstrated by the Heisenberg Uncertainty Principle and Schrödinger’s cat. Yet, it describes with perfect accuracy numerous phenomena we experimentally observe at the microscopic level. Add to this the fact that we have yet to fully prove quantum mechanics, and one seriously doubts our current knowledge of the universe.

Clearly, quantum mechanics is not a trivial subject. But, we will cover the bare minimum of what one needs to know in order to understanding quantum computing theory and algorithms. As such, our treatment will gloss over some details and mathematical background.

We progress by covering some basic linear algebra before introducing the prevailing notation and formalism used in the field. It is important to note that the formalism of quantum mechanics is built on that of linear (and, more fundamentally, Lie) algebra(s). After the mathematics introduction, we will then discuss the six postulates of quantum mechanics in order to gain a more thorough understanding of quantum mechanics before moving on to topics more specific to quantum computing.
2.1 Linear Algebra Review

Before we dive into basic quantum mechanics, let’s review some important concepts and definitions from linear algebra which will also help us develop a common language with which to proceed.

First stop: vectors. Vectors are simply matrices with one of the dimensions set to one. We thus have two types of vectors: column vectors and row vectors. For a column vector \( v \) in \( \mathbb{R}^n \), we have:

\[
    v = \begin{pmatrix}
        v_1 \\
        v_2 \\
        \vdots \\
        v_n
    \end{pmatrix}
\]

where \( v_i \in \mathbb{R} \). A row vector \( w \) is then:

\[
    w = (w_1 \ w_2 \ \ldots \ w_n)
\]

with \( w_i \in \mathbb{R} \).

Quantum mechanical states exist in complex spaces. So, using the above examples, one would actually have vectors in \( \mathbb{C}^n \) such that \( v_i, w_i \in \mathbb{C} \). Specifically, for quantum computing, our quantum bits, or qubits, are two-state systems, so one operates in the smaller \( \mathbb{C}^2 \).

Next, we look at transposes and conjugates. Transposing simply makes each row in a matrix a column and each column a row. In 2 dimensions, we have the transpose of a matrix \( A \) where

\[
    A = \begin{pmatrix}
        a_{11} & a_{12} \\
        a_{21} & a_{22}
    \end{pmatrix}
\]

equal to

\[
    A^T = \begin{pmatrix}
        a_{11} & a_{21} \\
        a_{12} & a_{22}
    \end{pmatrix}
\]

If we then conjugate the matrix, we take the complex conjugate of each value in the matrix or vector. Namely, we’d now have:

\[
    A^\dagger = (A^T)^* = \begin{pmatrix}
        a_{11}^* & a_{12}^* \\
        a_{21}^* & a_{22}^*
    \end{pmatrix}
\]
where we have $A^*$ equal to the complex conjugation of each entry in $A$ and thus $A^\dagger$ defined as the conjugate transpose or Hermitian adjoint of the matrix $A$. When $A$ acts on a vector $v$, the conjugate transpose reverses the order of operations, i.e. $(Av)^\dagger = v^\dagger A^\dagger$, due to the transpose identity for multiplication.

Lastly, we consider eigenvalues. For a matrix $B \in \mathbb{C}^n$, we define eigenvectors $v_i$ and corresponding eigenvalues $\lambda_i$ such that

$$Bv_i = \lambda_i v_i.$$

While we refer to dedicated texts on linear algebra to discuss how eigenvectors and eigenvalues are calculated, we note that eigen-parameters are significant for quantum mechanics and thus quantum computing since they represent the results of a measurement operation. More details are discussed in Sections 2.3.2 and 2.4.3.

2.2 Dirac Notation

It is now useful to introduce a notation formalism used in quantum mechanics and, especially, in quantum computing: Dirac “bra-ket” notation. While it may at first look odd, it is in reality a simple translation of the “wordier” vector notations used in linear algebra.

Namely, we introduce two definitions: the “bra” and the “ket”. Together, when we take an inner product (Section 2.3.1), they form a “bra-ket” or bracket.\(^1\)

The ket, written as $|\psi\rangle$, where $\psi$ is an arbitrary variable used to label the ket, represents a column vector of arbitrary length. Again, in quantum computing, since we only concern ourselves with individual systems of two states, this becomes a 2-by-1 column vector. The bra, $\langle \psi |$ then represents the conjugate transpose of the corresponding bra. So, for our purposes, the bra is a 1-by-2 row vector. More concretely, for a ket $|\psi\rangle$ such that $|\psi\rangle = (\alpha \beta)$, the bra $\langle \psi | = (\alpha^* \beta^*)$.

For a qubit, we’ll define the ground and excited states with $|0\rangle$ and $|1\rangle$, respectively, as

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

which are the real basis vectors of $\mathbb{C}^2$. The full quantum state of one qubit will then be a linear, complex superposition of these two basis vectors which are, by definition,\(^1\)

\(^1\)Physicists are really quite banal when coming to naming anything. So much so that we are often mistaken to be clever.
normalized (magnitude one) and orthogonal (inner product zero). An arbitrary qubit state
\( |\phi\rangle = \alpha |0\rangle + \beta |1\rangle \) for \( \alpha, \beta \in \mathbb{C} \). The state has corresponding bra \( \langle \phi | = \alpha^* \langle 0 | + \beta^* \langle 1 | \). We also impose the restriction \( |\alpha|^2 + |\beta|^2 = 1 \) to normalize the state for reasons to be discussed
in Section 2.4.3.

2.3 Linear Algebra in QM

2.3.1 Inner Product

When applying linear algebra to quantum mechanics there are a few concepts to operations
and concepts to highlight. First is that of the inner product, which is a generalization
of the dot product. For vectors like kets an bras, it is in fact just the dot product. But, it is
useful to take a closer look using Dirac notation.

Dirac notation actually makes the inner product quite straight-forward. Let’s look at a
state \( |\psi\rangle \) with complex entries. To take the inner product of \( |\psi\rangle \) with itself in order to find
the magnitude of \( |\psi\rangle \), we simply multiply \( |\psi\rangle \) with it’s conjugate transpose, which is exactly
\( \langle \psi \rangle \). So, the inner product of \( |\psi\rangle \) with itself is \( \langle \psi | \psi \rangle \). For a normalized state, namely for
a state where we impose the condition that the square magnitudes of its coefficients on
basis states sum to one: \( \sum_{i=1}^{n} |\alpha_i|^2 = 1 \), the inner product will also be one. To see this, we
decompose the inner product of \( |\psi\rangle \) with itself.

Taking \( |\psi\rangle \) to be in \( \mathbb{C}^n \), we can represent \( |\psi\rangle \) as a complex superposition of the basis
vectors of \( \mathbb{C}^n \):

\[
|\psi\rangle = \sum_{i=1}^{n} \alpha_i |e_i\rangle
\]

for \( \alpha \in \mathbb{C} \). Now taking the inner product against itself, we find

\[
\langle \psi | \psi \rangle = \sum_{i,j=1}^{n} \alpha_j^* \alpha_i \langle e_j | e_i \rangle
\]

\[
= \sum_{i,j=1}^{n} \alpha_j^* \alpha_i \delta_{i,j}
\]

\[
= \sum_{i=1}^{n} \alpha_i^* \alpha_i
\]

\[
= \sum_{i=1}^{n} |\alpha_i|^2 = 1
\]

What about the inner product of two different states, say, \( |\psi\rangle \) and \( |\phi\rangle \)? Well, since the
inner product is again defined such that we take the complex transpose of one of the states,
we have, for the inner product of $|\psi\rangle$ and $|\phi\rangle$, $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle$. Defining $|\phi\rangle = \sum_{i=1}^{n} \beta_{i} |e_{i}\rangle$, we’d have $\langle\phi|\psi\rangle = \sum_{i=1}^{n} \beta_{i}^{*} \alpha_{i}$. For normalized states, this value ranges from 0 to 1 (Can you show it?) and represents the “overlap” between the two states.

The inner product is a very important concept in quantum mechanics and quantum computing, so it is worth getting comfortable with it using Dirac notation.

2.3.2 Operators

Next, we consider quantum operators. Simply put, operators act on state vectors and are represented by matrices in the corresponding n-dimensional complex space. Some are defined in infinite dimensional space and must be truncated for finite dimensions. These operators become the classical analog of “gates” in a quantum algorithm.

Quantum operators are represented as a variable (often capital) with a hat, eg. $\hat{H}$, $\hat{x}$, or $\hat{a}$, and, in technical terms, are defined as mapping one vector space $V$ to another $W$. For a quantum operator $\hat{A}$, $\hat{A} : V \rightarrow W$. They have (but are not limited to) the following properties:

1. linearity
   
   - For quantum operator $\hat{A}$ in $V \in \mathbb{C}^{n}$, $\hat{A} \sum_{i} |e_{i}\rangle = \sum_{i} \hat{A} |e_{i}\rangle$.

2. composites
   
   - For quantum operators $\hat{A}$, $\hat{B}$ in $V \in \mathbb{C}^{n}$, $(\hat{A}\hat{B})|\psi\rangle = \hat{A}(\hat{B}|\psi\rangle)$

3. commutation
   
   - “Order matters”
   - For quantum operators $\hat{A}$, $\hat{B}$ in $V \in \mathbb{C}^{n}$, $\hat{A}\hat{B}|\psi\rangle \neq \hat{B}\hat{A}|\psi\rangle$
   
   - In fact, we have the idea of the commutator: The commutator of $\hat{A}$ and $\hat{B}$ is $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ and their anti-commutator is $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$.
   - We say that two operators commute if their commutator is zero.

We have special names for operators with distinct properties. Namely, we call an operator with the property $\hat{A}^{\dagger} = \hat{A}$ hermitian and an operator with the property $\hat{A}^{\dagger}\hat{A} = \hat{A}\hat{A}^{\dagger} = \mathbb{I}$ unitary. Unitary operations are “reversible” since action on a state with $\hat{A}^{\dagger}$ will undo the action of $\hat{A}$. Both hermitian and unitary operators are significant for quantum mechanics. We will discuss this in more detail in Section 2.4.

Some important operators in quantum computing are
identity: $\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

the Pauli matrices:

$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$

$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

and the Hadamard: $\hat{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

We will discuss all of these operators and their function in more detail in future lectures.

### 2.3.3 Tensor Product

The tensor product is an important tool for thinking of multiple quantum systems. A single qubit, which, again, is a two-state system, will exist in a complex vector space, called a Hilbert space $\mathcal{H}_A$. If we introduce another qubit into our quantum computer, it will exist in its own Hilbert space $\mathcal{H}_B$. How then, do we talk about these two qubit together?

Well, with a tensor product. Our full quantum computer system is now the tensor product of the two Hilbert spaces, namely $\mathcal{H}_A \otimes \mathcal{H}_B$. The state of each qubit will exist only in its relevant Hilbert space and an operator in $\mathcal{H}_A$ will only act on a state in $\mathcal{H}_A$ and vice versa. If qubit A is in state $\ket{\psi}$ and qubit B is in state $\ket{\phi}$, then our two-qubit computer is in state $\ket{\psi}_A \otimes \ket{\phi}_B$. Of course, this is generalizable to infinite numbers of qubits, or tensored Hilbert spaces.

The tensor product carries a few properties:

1. For $c \in \mathbb{C}$, $c (\ket{\psi}_A \otimes \ket{\phi}_B) = c \ket{\psi}_A \otimes \ket{\phi}_B = \ket{\psi}_A \otimes c \ket{\phi}_B$

2. Distribution: $(\ket{\psi_1}_A + \ket{\psi_2}_A) \otimes \ket{\phi}_B = \ket{\psi_1}_A \otimes \ket{\phi}_B + \ket{\psi_2}_A \otimes \ket{\phi}_B$

3. Operators act in their own spaces: $\hat{A}_A \otimes \hat{B}_B (\ket{\psi}_A \otimes \ket{\phi}_B) = \hat{A}_A \ket{\psi}_A \otimes \hat{B}_B \ket{\phi}_B$

We also make a notation simplification:

$\ket{\psi}_A \otimes \ket{\phi}_B = \ket{\psi} \ket{\phi} = \ket{\psi, \phi} = \ket{\psi\phi}$

The tensor product acts as following on vectors and matrices:

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \begin{pmatrix} \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} \alpha \gamma \\ \alpha \delta \\ \beta \gamma \\ \beta \delta \end{pmatrix}$$
For an $n \times m$ matrix $A$ and $p \times q$ matrix $B$:

$$A \otimes B = \begin{pmatrix}
A_{11}B & A_{12}B & \cdots & A_{1n}B \\
A_{21}B & A_{22}B & \cdots & A_{2n}B \\
\vdots & \vdots & \ddots & \vdots \\
A_{n1}B & A_{n2}B & \cdots & A_{nn}B
\end{pmatrix}$$

We call two-qubit states that can be written as the tensor product of two states product states and those that cannot entangled states. For example, $|\psi\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |10\rangle - |01\rangle - |11\rangle)$ is a product state while $|\Psi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$ is not.

### 2.4 Postulates of QM

To gain a more technical and intuitive understanding of quantum mechanics, we discuss the postulates of the theory. There are six postulates which dictate the basis of quantum mechanical theory. We model our discussion around MIT Professor Jaffe’s notes from 8.05 as taught in 1996 (basic quantum mechanics has not really changed since then). This section will provide many of the important consequences involving quantum operators and measurement that are significant for quantum computing.

#### 2.4.1 First Postulate

A quantum state is represented by a ket, some $|\psi\rangle$ in the state space, called a wavefunction. It completely specifies the quantum state.

The space of states is a vector space which means that the superposition of two states is again a state of the system. This allows us to define arbitrary states $|\psi\rangle = \sum_i a_i |e_i\rangle$ for basis vectors $|e_i\rangle$ of the $n$-dimensional state space. This describes phenomena like the double-slit diffraction of electrons.

As a Hilbert space, the space of states has with it a defined inner product. Namely,

$$\langle \psi, \phi \rangle \equiv \langle \psi | \phi \rangle = \int dx \psi^*(x) \phi(x) \quad (1)$$

We first define the inner product as an operation acting on two states in the ket space. Then the second step introduces the “bra space” of the same dimension as the ket space, and defines the inner product as an operation involving one element of each space. Finally, we reduce to the familiar continuous inner product which represents the overlap of $|\psi\rangle$ and $|\phi\rangle$. This notation allows us to see clearly that

$$\langle \psi | \phi \rangle^* = \langle \phi | \psi \rangle. \quad (2)$$
2.4.2 Second Postulate

Classical observables are introduced into quantum mechanics using operators. Specifically, every observable (measurable property) of a physical system is described by an operator that acts state kets.

Classical observables (such as position $x$ or momentum $p$) are made “quantum” by transforming them into quantum operators. One simply adds the operator hat ($\hat{}$) on top of the observable’s variable.

Notation-wise, an operator acts on a state from the left, i.e. $\hat{A}|\psi\rangle$. In general, an operator acting on a state will change the state. For every operator, there are, however, states which are not changed by the action of the operator, giving, for example,

$$\hat{A}|\psi_i\rangle = a_i|\psi_i\rangle$$  \hspace{1cm} (3)

These states are eigenstates to the operator and the value $a_i$ is the eigenvalue which represents the observable’s value. An operator acting on a state which is not one of its eigenstates, will produce a probabilistic outcome based on the states representation in the basis of eigenstates for the operator. In this case, the operator will also change the quantum state of the system.

2.4.3 Third Postulate

The result of a measurement of an observable with an operator $\hat{A}$ will only ever be an eigenvalue of $\hat{A}$.

This postulates leads to the “quantum” or “quantized” nature of quantum mechanics. Some operators, like the position $\hat{x}$ or momentum $\hat{p}$ operators, have a set of continuous eigenvalues, so this proposition is not far removed from what we are used to from classical mechanics. However, for operators with discrete eigenvalues such as the Hamiltonian (ie. energy) operator for particle in a bound well, this proposition creates an unintuitive result.

We note an important requirement here. First, measured values, in other words, the eigenvalues, must be real since these are real, classically measured observables. This means that operators corresponding to an observable must be Hermitian (remember Section 2.3.2). Hermitian operators also have eigenstates which form a complete basis which means that any state in the space can be decomposed into a linear superposition of eigenstates of the operator. These operators can be re-written as $\hat{A} = \sum_i \lambda_i |a_i\rangle \langle a_i|$ for eigenvalues $\lambda_i$ and corresponding eigenvectors $|a_i\rangle$. 

2.4.4 Fourth Postulate

When a measurement of an observable with operator $\hat{A}$ is made on a generic state $|\psi\rangle$, the probability of obtaining a certain eigenvalue $a_i$ is given by the square of the inner product of $|\psi\rangle$ with the corresponding eigenstate $|\langle a_i | \psi \rangle|^2$.

Professor Jaffe gives a very good treatment of this postulate, so we abbreviate them here.

We assume that the states are normalized which is in general possible. States are usually normalized to unity,

$$\langle \psi | \psi \rangle = 1$$
$$\langle a_j | a_k \rangle = \delta_{jk} \quad (4)$$

The complex number, $\langle a_i | \psi \rangle$ is known as the "probability amplitude", to measure $a_i$ as the value for the observable $\hat{A}$ in the state $|\psi\rangle$.

The postulate suggests the following algebraic exercise. First, any state can be expanded as a superposition of $\hat{A}$-eigenstates (see Postulate 3),

$$|\psi\rangle = \sum_n c_n |a_n\rangle \quad (5)$$

Next use the orthonormality of the $\hat{A}$-eigenstates to find an expression for the expansion coefficients $c_n$, 

$$\langle a_j | \psi \rangle = \sum_n c_n \langle a_j | a_n \rangle = c_j \quad (6)$$

So,

$$|\psi\rangle = \sum_n \langle a_n | \psi \rangle |a_n\rangle \quad (8)$$

where we have the complex number $\langle a_n | \psi \rangle$ and the state $|a_n\rangle$. The component of $|\psi\rangle$ along the direction of the nth eigenstate of $\hat{A}$ is given by $\langle a_n | \psi \rangle$. The measurement operation yields the result $a_n$ with a probability proportional to the square of this component, $|\langle a_n | \psi \rangle|^2$.

The probability of obtaining some result at all is unity. For states normalized to unity,

$$|\langle \psi | \psi \rangle|^2 = \left| \sum_m \sum_n c_m^* c_n \langle a_m | a_n \rangle \right|^2. \quad (9)$$
Using $|\langle \psi | \psi \rangle|^2 = 1$ and $\langle a_m | a_n \rangle = \delta_{mn}$, we get
\[
\sum_n |c_n|^2 = 1.
\] (10)

According to the usual rules of probability, we can compute the "expectation value" of the observable $A$. If the probability to observe $a_n$ is $|c_n|^2$, then the expected value (denoted $\langle A \rangle$) is
\[
\langle A \rangle = a_n |c_n|^2
\] (11)
as long as each eigenstate has a unique eigenvalue (the operator $\hat{A}$ has no degenerate eigenvalues).

2.4.5 Fifth Postulate

Immediately after the measurement of an observable $A$ with a value $a_n$, the state of the system is the normalized eigenstate $|a_n\rangle$.

This postulate gives way to the idea of the "collapse of the wavepacket". While measuring a state $|\psi\rangle$ for the observable $A$ gives the value $a_n$ with probability $|\langle a_n | \psi \rangle|$, results for remeasuring immediately after the first measurement are not statistically distributed. Namely, after measurement, the state $|\psi\rangle$ is found in the renormalized state $|a_n\rangle$ since the outcome is always $a_n$.

Wavefunction collapse is one of the more unintuitive concepts in quantum mechanics, but is supported by experimental results with repeated measurements.

2.4.6 Sixth Postulate

Quantum states generally change (evolve) with time. This time evolution preserves the normalization of the state. The time evolution of the state of a quantum system is described by $|\psi(t)\rangle = \hat{U}(t,t_0) |\psi(t_0)\rangle$, for some unitary operator $\hat{U}$.

A quantum state must evolve by a unitary operator since the norm of the state must be preserved. While individual probabilities of a observing certain eigenvalues of an operator may change, the total sum ($\sum_n |c_n|^2$ from before) remains the same. The probability for measuring some value of any operator remains unity.
As such, for a state evolving with time $|\psi(t)\rangle$, we maintain $|\langle \psi(t) | \psi(t) \rangle|^2 = 1$ for all $t$. So,

$$
|\langle \psi(t) | \psi(t) \rangle|^2 = 1 
= |\langle \psi(t_0) | U(t, t_0)^\dagger U(t, t_0) | \psi(t_0) \rangle|^2 
= |\langle \psi(t_0) | \psi(t_0) \rangle|^2 = 1,
$$

which is consistent with our earlier requirement for a unitary operator. This also means that time evolution is a reversible process. We reverse some time evolution given by $U(t, t_0)$ we simply act with $U(t_0, t) = U(t, t_0)^\dagger$.

This postulate actually leads to the Schrödinger equation which requires that a state $|\psi\rangle$ evolves under

$$
i\hbar \frac{d}{dt} |\psi(t)\rangle = \mathcal{H} |\psi(t)\rangle
$$

for a Hermitian operator $\mathcal{H}$ which is a system-dependent operator called the Hamiltonian.

### 2.5 Uncertainty Principle

The postulates of quantum mechanics highlight the probabilistic nature of the theory. This in turn gives rise to some interesting (and sometime unintuitive or even concerning) phenomena. One of these is the uncertainty principle which states that there is a limit to the precision with which one is able to determine the values of conjugate observables (such as position and momentum and energy and time).

Stepping back to classical probability theory, we can define an expectation value $\langle A \rangle$, variance $(\Delta A)^2$, and standard deviation $\Delta A$ such that for possible values $a_i$ each with probability $p_i$,

$$
\langle A \rangle = \sum_i a_i p_i,
$$

$$
(\Delta A)^2 = \langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2,
$$

and

$$
\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}.
$$

Using the fourth postulate, we can then apply this to quantum mechanics and relate the product of the standard deviations of two operators to their commutator. More explicitly, we
use $p_i = |\langle \psi | \phi_i \rangle|^2$ to find $\Delta A \Delta B$ in terms of $[A, B]$. The standard deviation $\Delta A$ measures the uncertainty in our knowledge of the value of the observable $A$, so we will be placing a limit on our this uncertainty.

First, in quantum mechanics, the expectation value of an observable simplifies to

$$
\langle A \rangle = \sum_i a_i p_i
$$

$$
= \sum_i a_i \langle \psi | \phi_i \rangle \langle \phi_i | \psi \rangle
$$

$$
= \langle \psi | \left( \sum_i a_i | \phi_i \rangle \langle \phi_i | \right) | \psi \rangle
$$

$$
= \langle \psi | A | \psi \rangle.
$$

To derive the uncertainty principle, let

$$
\langle \psi | AB | \psi \rangle = re^{i\phi}
$$

where $re^{i\phi}$ represents some arbitrary complex number. Using $\langle \psi | AB | \psi \rangle^\dagger = \langle \psi | BA | \psi \rangle$ for Hermitian operators $\hat{A}$ and $\hat{B}$, one finds

$$
\langle \psi | AB - BA | \psi \rangle = re^{i\phi} - re^{-i\phi} = 2ir \sin \phi,
$$

$$
\langle \psi | AB + BA | \psi \rangle = re^{i\phi} + re^{-i\phi} = 2r \cos \phi,
$$

and

$$
|\langle \psi | AB + BA | \psi \rangle|^2 + |\langle \psi | AB - BA | \psi \rangle|^2 = 4r^2.
$$

Thus follows

$$
|\langle \psi | AB - BA | \psi \rangle|^2 \leq 4 |\langle \psi | AB | \psi \rangle|^2
$$

since both arguments on the left side of Eq. 24 are necessarily greater than or equal to zero. One further reduces using the inner product inequality from linear algebra to find

$$
\frac{|\langle \psi | [A, B] | \psi \rangle|^2}{4} \leq \langle \psi | A^2 | \psi \rangle \langle \psi | B^2 | \psi \rangle.
$$

Now, define $A = C - \langle C \rangle$ and $B = D - \langle D \rangle$ where $[A, B] = [C, D]$, $\langle \psi | A^2 | \psi \rangle = (\Delta C)^2$, and $\langle \psi | B^2 | \psi \rangle = (\Delta D)^2$. So,

$$
\left| \frac{\langle \psi | [C, D] | \psi \rangle}{2} \right|^2 \leq (\Delta C)^2 (\Delta D)^2,
$$
which brings us to the Heisenberg Uncertainty principle:

\[
\frac{\langle \psi | [C, D] | \psi \rangle}{2} \leq \Delta C \Delta D. \tag{28}
\]

This statement gains significance when we considering the position and momentum conjugate variables. Since \([\hat{x}, \hat{p}] = i\hbar\), \(\Delta \hat{x} \Delta \hat{p} \geq \frac{\hbar}{2}\), so one can only measure together the position and momentum variables of a system to a minimum of \(\frac{\hbar}{2}\). Increasing knowledge about the position of a system necessarily decreases the knowledge an observer can have about the momentum of a system (and vice versa). This is true for all non-commuting observables.

### 2.6 Local Hidden Variables

Einstein was rather uncomfortable with the idea that quantum mechanics was a fundamentally probabilistic theory\(^2\). He say quantum mechanics as akin to thermodynamics which describes macroscopic effects using derivations from microscopic or “hidden” variables. We illustrate this idea with an example using a system of three (two-state) quantum spins (like an atom or a qubit, for example). This system has a total Hilbert space

\[
\mathcal{H} = \mathcal{H}_{\text{spin}}^{\otimes 3} = \mathcal{H}_A^{\text{spin}} \otimes \mathcal{H}_B^{\text{spin}} \otimes \mathcal{H}_C^{\text{spin}}. \tag{29}
\]

For a given state \(|\psi\rangle\), one cannot predict the outcome of measuring each spin along a single axis. However, one could assume that each spin has a secret (secret to the observer) setting: either \(\uparrow = +1\) or \(\downarrow = -1\) for each of \(x\), \(y\), and \(z\) for each spin. We label these \(S_x^1\) for an \(\hat{x}\) measurement on the first qubit, and so on. As such, the system has its own truth table, so to speak, which determines the outcome of any measurement on the system (say, \(\hat{x}_1 \hat{y}_2 \hat{y}_3\) or \(\hat{z}_1 \hat{x}_2 \hat{y}_3\)).

In 1993, physicists Greenberger, Horne, and Zeilinger attempted to show the invalidity of local hidden variables. They wrote down a state, known as the GHZ state:

\[
|\Psi\rangle_{\text{GHZ}} = \frac{1}{\sqrt{2}} (|000\rangle - |111\rangle) \tag{30}
\]

for a system of three spins. We will measure the following operators on the GHZ state: \(\hat{x}_1 \hat{x}_2 \hat{x}_3\), \(\hat{x}_1 \hat{y}_2 \hat{y}_3\), \(\hat{y}_1 \hat{x}_2 \hat{y}_3\), \(\hat{y}_1 \hat{y}_2 \hat{x}_3\).

\(^2\)as shown by his famous statement “God does not play dice.”
The expectation value of the first observable $x_1 x_2 x_3$ on the GHZ state is
\[
\langle x_1 x_2 x_3 \rangle_{GHZ} = \langle \Psi |_{GHZ} \sigma_x^1 \sigma_x^2 \sigma_x^3 | \Psi \rangle_{GHZ}
\]
\[
= \frac{1}{\sqrt{2}} \langle \Psi |_{GHZ} \sigma_x^1 \sigma_x^2 \sigma_x^3 (|000\rangle - |111\rangle)
\]
\[
= \frac{1}{\sqrt{2}} \langle \Psi |_{GHZ} (|111\rangle - |000\rangle)
\]
\[
= -\frac{1}{\sqrt{2}} \langle \Psi |_{GHZ} (|000\rangle - |111\rangle)
\]
\[
= -\langle \Psi | \Psi \rangle_{GHZ}
\]
\[
= -1.
\] (31)

For $x_1 y_2 y_3$, the expectation value is
\[
\langle \hat{x} \hat{y} \hat{y} \rangle_{GHZ} = \frac{1}{\sqrt{2}} \langle \Psi |_{GHZ} \hat{x} \hat{y} \hat{y} \rangle_{GHZ}
\]
\[
= \frac{1}{\sqrt{2}} \langle \Psi |_{GHZ} (|1\rangle_A (i |1\rangle_B) (i |1\rangle_C) - |0\rangle_A (-i |0\rangle_B) (-i |0\rangle_C))
\]
\[
= -\frac{1}{\sqrt{2}} \langle \Psi |_{GHZ} (-|111\rangle + |000\rangle)
\]
\[
= \langle \Psi | \Psi \rangle_{GHZ}
\]
\[
= 1
\] (32)

as is the expectation values for $y_1 x_2 y_3$ and $y_1 y_2 x_3$.

We have, then
\[
\langle \hat{x} \hat{x} \hat{x} \rangle \langle \hat{x} \hat{y} \hat{y} \rangle \langle \hat{y} \hat{x} \hat{x} \rangle = (-1)(+1)(+1)(+1)
\]
\[
= -1.
\] (33)

However, if we calculate this same value using the hidden variables approach, we find
\[
(S_x^1 S_x^2 S_x^3) (S_x^1 S_y^2 S_y^3) (S_y^1 S_x^2 S_y^3) (S_y^1 S_y^2 S_x^3) = (S_x^1)^2 (S_y^1)^2 (S_x^2)^2 (S_y^2)^2 (S_x^3)^2 (S_y^3)^2
\]
\[
= (\pm 1)^2
\]
\[
= +1,
\] (35)

which is clearly a contradiction. So, if quantum mechanics is correct, as research and experiment suggest, our world really is probabilistic after all\(^3\).

\(^3\)Sorry, Mr. Einstein.
2.7 The Bloch Sphere

The Bloch sphere is a useful tool for visualizing qubit states (or any quantum state in $\mathcal{H}^2$. A pure quantum state is represented by a vector on the surface of a unit sphere. The state $|0\rangle$ exists at the top of the sphere (the north pole) and the state $|1\rangle$ exists at the bottom (the south pole). Using spherical coordinates, arbitrary qubit states are defined on the sphere with $|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle$ with polar angle $\theta$ and azimuthal angle $\phi$. Any state can be written $|\psi\rangle = e^{i\gamma} \left( \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle \right)$ where we factor out the phase associated with the $|0\rangle$ basis state. However, in quantum mechanics, the arbitrary phase $e^{i\gamma}$ has no observable effects on the state. So, any qubit state is described by a point on the unit sphere.

Using this formula, we can define the states which match the positive and negative $x$, $y$, and $z$-axes as

\begin{align*}
+z &= |0\rangle \\
-z &= |1\rangle \\
+\bar{x} &= |+\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \\
-\bar{x} &= |\bar{-}\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \\
+\bar{y} &= |+i\rangle = \frac{1}{\sqrt{2}} (|0\rangle + i |1\rangle) \\
-\bar{y} &= |\bar{-}i\rangle = \frac{1}{\sqrt{2}} (|0\rangle - i |1\rangle)
\end{align*}

which one can prove to oneself using appropriate values for $\theta$ and $\phi$.

Let’s consider the action of the Pauli matrices on states on the Bloch sphere. We will find that they, when exponentiated, generate rotations. Specifically, a rotation along some axis $\hat{n}$ is represented by

\begin{align*}
R_{\hat{n}}(\theta) &= e^{-i\frac{\theta}{2} \hat{n} \cdot \vec{\sigma}} \\
&= \cos \frac{\theta}{2} I - i \sin \frac{\theta}{2} \hat{n} \cdot \vec{\sigma}
\end{align*}

where $\vec{\sigma}$ is the “vector” of Pauli matrices. In the second line, we use the fact that $\sigma^2_i = 1$ for the Paulis. In the following, we show that the above truly is a rotation on the Bloch sphere.

To do so, it is useful to consider the density operator formed by the outer product of
the qubit state:

\[ \rho = |\psi\rangle \langle \psi| = |\psi\rangle \otimes |\psi\rangle \]

\[ = \left( \begin{array}{cc}
\cos^2 \frac{\theta}{2} & e^{-i\phi} \cos^2 \frac{\theta}{2} \sin \frac{\theta}{2} \\
\cos \frac{\theta}{2} \sin \frac{\theta}{2} & \sin^2 \frac{\theta}{2}
\end{array} \right) \]

\[ = \frac{1}{2} \left( \begin{array}{cc}
1 + \cos \theta & \cos \phi \sin \theta - i \sin \phi \sin \theta \\
\cos \phi \sin \theta + i \sin \phi \sin \theta & 1 - \cos \theta
\end{array} \right) \]  

\[ = \frac{1}{2} (I + X \cos \phi \sin \theta + Y \sin \phi \sin \theta + Z \cos \theta) \]

\[ = \frac{1}{2} (I + \vec{r}_\rho \cdot \vec{\sigma}) \]  

where we have \( I, X, Y, Z \) as the identity and Pauli matrices and \( \vec{r}_\rho = (r_x, r_y, r_z) = (\cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta) \) as a unit vector in spherical coordinates.

The action of an operator on a quantum state must obey unitary evolution, preserving the norm of the state. The action of a unitary operator on a state can we used to find the action on the density operator \( \rho \). Since a unitary operator acting on a state \( |\psi\rangle \) is given by \( |\psi\rangle \rightarrow \hat{U} |\psi\rangle \), the action of the operator on \( \rho \) is

\[ \rho = |\psi\rangle \langle \psi| \rightarrow \hat{U} |\psi\rangle \langle \psi| \hat{U}^\dagger. \]  

Our rotation matrix is unitary since the Pauli matrices are Hermitian. Let’s observe the action of a z-axis rotation on our density matrix. For a rotation around the z-axis, \( R_z(\theta) = \cos^2 \frac{\theta}{2} I - i \sin \frac{\theta}{2} \sigma_z \). Since \( \rho \) is a sum of the Pauli matrices, we can evaluate \( R_z(\theta) \) on each individually:

\[ \rho' = R_z(\theta) \rho R_z(\theta)^\dagger \]

\[ = \frac{1}{2} \left( I + r_x R_z(\theta) X R_z(\theta)^\dagger + r_y R_z(\theta) Y R_z(\theta)^\dagger + r_z R_z(\theta) Z R_z(\theta)^\dagger \right). \]

For the first term, the we use the unitary property of the rotation matrix.

Expanding the \( X \) term, we have

\[ R_z(\theta) X R_z(\theta)^\dagger = \left( \cos \frac{\theta}{2} I - i \sin \frac{\theta}{2} \sigma_z \right) X \left( \cos \frac{\theta}{2} I + i \sin \frac{\theta}{2} \sigma_z \right) \]

\[ = \cos^2 \frac{\theta}{2} X + i \sin \frac{\theta}{2} \cos \frac{\theta}{2} \sigma_z X - i \sin \frac{\theta}{2} \sigma_z X \cos \frac{\theta}{2} \sigma_z X \]

\[ = \cos^2 \frac{\theta}{2} X + \sin \frac{\theta}{2} \cos \frac{\theta}{2} Y + \sin \frac{\theta}{2} \cos \frac{\theta}{2} Y - \sin^2 \frac{\theta}{2} X \]

\[ = \left( \cos \frac{\theta}{2} - \sin^2 \frac{\theta}{2} \right) X + 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} Y \]

\[ = \cos \theta X + \sin \theta Y. \]
Similarly, we can show

\begin{align}
R_z(\theta)YR_z(\theta)^\dagger &= \cos \theta Y - \sin \theta X \\
R_z(\theta)ZR_z(\theta)^\dagger &= Z.
\end{align}

Now, rewriting \( \rho' \),

\begin{align}
\rho' &= \frac{1}{2} \left( I + r_x R_z(\theta) X R_z(\theta)^\dagger + r_y R_z(\theta) Y R_z(\theta)^\dagger + r_z R_z(\theta) Z R_z(\theta)^\dagger \right) \\
&= \frac{1}{2} (I + r_x \cos \theta X + \sin \theta Y + r_y \cos \theta Y - \sin \theta X + r_z Z) \\
&= \frac{1}{2} (I + (r_x \cos \theta - r_y \sin \theta) X + (r_x \sin \theta + r_y \cos \theta) Y + r_z Z) \\
&= \frac{1}{2} \left( I + r'_x X + r'_y Y + r'_z Z \right)
\end{align}

where we have the transformed positions

\begin{align}
r'_x &= r_x \cos \theta - r_y \sin \theta \\
r'_y &= r_x \sin \theta + r_y \cos \theta \\
r'_z &= r_z
\end{align}

which match exactly what we expect from a rotation around the z-axis. We can then construct the \( R_z(\theta) \) rotation matrix

\[
\begin{pmatrix}
\cos \theta & -\sin \theta & 0 \\
\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{pmatrix}.
\]

We can do something similar to construct the matrix for a rotation around an arbitrary axis by making composite rotations with \( R_x, R_y \), and \( R_z \).

\[\text{Give it shot!}\]