Quantum Computation and Quantum Information by Michael A. Nielsen and Isaac L. Chuang

Chris Doble

May 2024

Contents

Part I
Fundamental concepts

1 Introduction and overview

1.2 Quantum bits

- The special states $|0\rangle$ and $|1\rangle$ form an orthonormal basis and are known as **computational basis states**.
- A quantum bit (**qubit**) is a linear combination of the computational basis states
$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$$
where α and β are complex numbers.
- When we measure a qubit we either get $|0\rangle$ with probability $|\alpha|^2$ or $|1\rangle$ with probability $|\beta|^2$. Thus, $|\alpha|^2 + |\beta|^2 = 1$ and a qubit can be thought of as a unit vector in a two-dimensional complex vector space.
- If a qubit is in the state
$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$
there's a 50/50 chance of measuring $|0\rangle$ or $|1\rangle$.

- If we let
$$\alpha = e^{i\gamma} \cos \frac{\theta}{2}$$
and
$$\beta = e^{i\gamma} e^{i\varphi} \sin \frac{\theta}{2}$$
then
$$\begin{aligned} |\alpha|^2 + |\beta|^2 &= \alpha^* \alpha + \beta^* \beta \\ &= \cos^2 \frac{\theta}{2} + \sin^2 \frac{\theta}{2} \\ &= 1 \end{aligned}$$
so the qubit is still normalised and it can be written
$$|\psi\rangle = e^{i\gamma} \left(\cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle \right).$$
It turns out that $e^{i\gamma}$ has no observable effects and we can effectively write
$$|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle .$$
This defines a point on a three-dimensional sphere known as the **Bloch sphere** where θ and φ take on their usual roles in a spherical coordinate system.

- Before measurement a qubit is in a linear combination of $|0\rangle$ and $|1\rangle$ but when measured you get one or the other and the state of the system changes to match the measured result.

1.2.1 Multiple Bits

- A two qubit system has four computational basis state $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$ so the general expression for the state of such a system is
$$|\psi\rangle = \alpha_{00} |00\rangle + \alpha_{01} |01\rangle + \alpha_{10} |10\rangle + \alpha_{11} |11\rangle .$$

- If you were to measure the first qubit, you would get $|0\rangle$ with probability $|\alpha_{00}|^2 + |\alpha_{01}|^2$ and the system would be left in the state
$$|\psi\rangle = \frac{\alpha_{00} |00\rangle + \alpha_{01} |01\rangle}{\sqrt{|\alpha_{00}|^2 + |\alpha_{01}|^2}},$$
i.e. it is renormalised such that the normalisation condition still holds.

1.3 Quantum Computation

1.3.1 Single Qubit Gates

- The quantum NOT changes $|0\rangle$ to $|1\rangle$ and $|1\rangle$ to $|0\rangle$. It acts linearly on superpositions of those states, i.e. it turns $\alpha|0\rangle + \beta|1\rangle$ into $\beta|0\rangle + \alpha|1\rangle$. If a quantum state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is written in vector notation as

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

  then the quantum NOT gate can be expressed in matrix form as

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

- In order to preserve the normalisation condition, matrix representations of quantum gates must be unitary, i.e. $M^\dagger M = I$ where I is the identity matrix.

- An arbitrary unitary 2x2 matrix can be decomposed into a finite set of other 2x2 matrices. This means an arbitrary single qubit gate can be generated by a finite set of other gates.

1.3.2 Multiple Qubit Gates

- The controlled-NOT or CNOT gate is a multi-qubit gate that has two input qubits known as the control qubit and the target qubit. If the control qubit is set to $|0\rangle$ the target qubit is left alone, but if it's set to $|1\rangle$ the target qubit is flipped. Another way of writing this is $|A, B\rangle \rightarrow |A, A \oplus B\rangle$ where $\oplus$ is modulo-two addition. Yet another way of writing this is in matrix form

$$U_{CN} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

  where the first column describes what happens to the $|00\rangle$ basis state, etc.

- Other classical gates like NAND or XOR can't be represented as quantum gates as they're irreversibile. For example, given the output $A \oplus B$ from an XOR gate it's not possible to determine what the inputs A and B were.

- Any multiple qubit logic gate may be composed from CNOT and single qubit gates.

1.3.3 Measurements in bases other than the computational basis

- Given any basis states $|a\rangle$ and $|b\rangle$ it is possible to express an arbitrary state as a linear combination $\alpha|a\rangle + \beta|b\rangle$ of those states. For example, if $|a\rangle = |+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $|b\rangle = |-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$ then

$$|\phi\rangle = \alpha|0\rangle + \beta|1\rangle = \alpha\frac{|+\rangle + |-\rangle}{\sqrt{2}} + \beta\frac{|+\rangle - |-\rangle}{\sqrt{2}} = \frac{\alpha+\beta}{\sqrt{2}}|+\rangle + \frac{\alpha-\beta}{\sqrt{2}}|-\rangle\,.$$

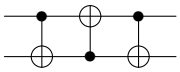
- If $|a\rangle$ and $|b\rangle$ are orthonormal, it's also possible to perform a measurement with respect to that basis. In the example above you would measure $|+\rangle$ with probability $|\alpha+\beta|^2/2$ and $|-\rangle$ with probability $|\alpha-\beta|^2/2$.

1.3.4 Quantum circuits

- Applying a CNOT gate three times swaps the state of two qubits:

$$\begin{aligned}|a,b\rangle &\to |a, a\oplus b\rangle \\ &\to |a\oplus(a\oplus b), a\oplus b\rangle = |b, a\oplus b\rangle \\ &\to |b, (a\oplus b)\oplus b\rangle = |b,a\rangle\,.\end{aligned}$$

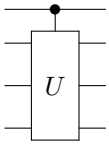
  This can be represented in a quantum circuit diagram



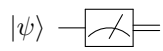
  which is equivalent to

- Quantum circuits don't allow loops (they are acyclic), they don't allow multiple wires to be joined into one, and they don't allow one wire to be split into multiple wires.

- If U is a unitary matrix operating on n qubits then U can be regarded as a quantum gate. Then we can define a controlled-U gate which takes a single control qubit and n target qubits. If the control qubit is set to 0 then nothing happens to the target qubits. If it's set to 1 then the U gate is applied to the target qubits.



- Measurement is represented as



  where the double lines represent a classical bit.

1.3.5 Qubit copying circuit?

- A CNOT gate can be used to copy a classical bit. As input it takes a bit in state x and a "scratchpad" bit set to zero. The output is a bit in state x and a bit in state 0 if $x = 0$ and 1 if $x = 1$ — both bits are in state x.
- However, a CNOT gate cannot be used to copy quantum information. As input it takes a qubit in the state $|\psi\rangle = a|0\rangle + b|1\rangle$ and a "scratchpad" qubit set to $|0\rangle$. The input state may be written as

$$[a|0\rangle + b|1\rangle]|0\rangle = a|00\rangle + b|10\rangle.$$

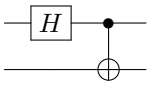
The gate negates the second bit if the first bit is 1 so the output is $a|00\rangle + b|11\rangle$. The desired output is

$$|\psi\rangle|\psi\rangle = a^2|00\rangle + ab|01\rangle + ab|10\rangle + b^2|11\rangle.$$

Unless $a = 0$ or $b = 0$, i.e. it's a classical bit, the qubit hasn't been copied. This is known as the **no-cloning theorem**.

1.3.6 Example: Bell states

- The Hadamard gate acts on a single qubit, mapping $|0\rangle$ to $(|0\rangle + |1\rangle)/\sqrt{2}$ and $|1\rangle$ to $(|0\rangle - |1\rangle)/\sqrt{2}$.
- This circuit



applies a Hadamard gate to the first qubit which then acts as the control input to the CNOT. The output states

$$|\beta_{00}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}},$$
$$|\beta_{01}\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}},$$
$$|\beta_{10}\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \text{ and}$$
$$|\beta_{11}\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

are known as the **Bell states**, **EPR states**, or **EPR pairs**.

1.3.7 Example: quantum teleportation

- If Alice and Bob prepare a Bell pair, Alice takes the first qubit, and Bob takes the second qubit, it's possible to teleport an arbitrary qubit from Alice to Bob at a later time by communicating only classical information.

- This doesn't enable faster-than-light communication because Alice still needs to communicate classical information to Bob which is limited by the speed of light. Bob then uses this information to conditionally apply further gates to his qubit to get it into the teleported state.

1.4 Quantum algorithms

1.4.1 Classical computations on a quantum computer

- Classical circuits may contain irreversible operations, e.g. NAND gates, but quantum circuits must be reversible. In order to simulate classical circuits on a quantum computer we must find a way to make these irreversible operations reversible.

- Any classical circuit can be replaced by an equivalent classical circuit containing only reversible elements by making use of the **Toffoli gate**. This gate has three input bits and three output bits. Two bits are control bits and are unaffected by the gate. The third bit is a target bit that is flipped if both of the control bits are 1 and is otherwise left alone. Applying the Toffoli gate twice has the effect $(a, b, c) \to (a, b, c \oplus ab) \to (a, b, c)$ and thus the gate is reversible.

- The Toffoli gate can also be used to simulate NAND gates and to do FANOUT. With this it is possible to make any classical circuit reversible.

- The equivalent quantum Toffoli gate works as expected, e.g. it changes $|110\rangle$ to $|111\rangle$. This ensures quantum computers are capable of performing any classical computation.

- If the classical computer is non-deterministic, i.e. it can generate random numbers, this can be simulated on a quantum computer by preparing a qubit in the state $|0\rangle$, sending it through a Hadamard gate to produce $(|0\rangle + |1\rangle)/\sqrt{2}$, and then measuring the state. The result will be $|0\rangle$ or $|1\rangle$ with 50/50 probability.

1.4.2 Quantum parallelism

- Quantum parallelism allows quantum computers to evaluate a function $f(x)$ over many different values of x simultaneously. For example, if you have a gate that accepts two input qubits $|x, y\rangle$ and produces two output qubits $|x, y \oplus f(x)\rangle$ you can start with $|0, 0\rangle$, send the first bit through a Hadamard gate to produce $(|0\rangle + |1\rangle)\,|0\rangle\,/\sqrt{2}$, then send both bits through the $f(x)$ gate to produce $(|0, f(0)\rangle + |1, f(1)\rangle)/\sqrt{2}$.

- The above process can be generalised to functions on an arbitrary number of bits by using a general operation known as the **Hadamard transform**

which is n Hadamard gates operating in parallel on n qubits. This is denoted $H^{\otimes n}$, e.g.

$$H^{\otimes 2} = \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle).$$

- However, quantum parallelism isn't particularly useful in itself because upon measurement you only get a single value $|x, f(x)\rangle$.

1.4.3 Deutsch's algorithm

- Deutsch's algorithm can determine $f(0) \oplus f(1)$ — a global property of $f(x)$ — using only one evaluation of f. A classical computer would require at least two.

- This demonstrates a property of quantum computers where in the state $|0\rangle |f(0)\rangle + |1\rangle |f(1)\rangle$ it's possible to make the two alternatives inferere with each other resulting in some global property of the function f.

1.4.4 The Deutsch-Jozsa algorithm

- The Deutsch-Jozsa algorithm can be used to determine if a function f is constant (always returns the same value) or balanced (returns 0 for half of its domain and 1 for the other half) with a single evaluation of f. The same problem on a classical computer requires $2^n/2 + 1$ evaluations of f.

1.4.5 Quantum algorithms summarized

- There are three classes of quantum algorithms which provide an advantage over known classical algorithms:
  - algorithms based on quantum versions of the Fourier transform,
  - quantum search algorithms, and
  - quantum simulation where a quantum computer is used to simulate a quantum system.

- Classically the fast Fourier transform takes roughly $N \log(N) = n2^n$ steps to transform $N = 2^n$ numbers. On a quantum computer this can be accomplished in about $\log^2(N) = n^2$ steps which is an exponential saving.

- However, the Fourier transform is being performed on the amplitudes of the quantum state, not on the input vector itself. This means we need clever ways to extract this information.

- Classical search algorithms are typically $\mathcal{O}(n)$ whereas the quantum equivalents are $\mathcal{O}(\sqrt{n})$.

- **Computational complexity theory** is the subject of classifying the difficuly of various computational problems. The most basic idea is a **complexity class** which can be though of as a collection of computational problems that all share some common feature with respect to the computational resources needed to solve those problems.
- One complexity class is **P** — the class of problems that can be solved quickly on a classical computer.
- Another is **NP** — the class of problems whose solutions can be verified quickly on a classical computer.
- **P** is a subset of **NP**, but it's not known if the're equal.
- **NP**-complete problems are an important subclass of **NP** problems. They're important because: there are many problems that are known to be **NP**-complete, and any given **NP**-complete problem is known to be "at least as hard" as all other **NP** problems, i.e. an algorithm to solve a specific **NP**-complete problem can be modified to solve any other **NP** problem. If $\mathbf{P} \neq \mathbf{NP}$ then it will follow that no **NP**-complete problem can be efficiently solved on a classical computer.

1.5 Experimental quantum information processing

1.5.1 The Stern-Gerlach experiment

- The Stern-Gerlach device sends an atom through a magnetic field and measures its deflection. The atom is electrically neutral so it is not deflected by the Lorentz force. The electron is in the $n = 1$ level so there one possible angular momentum value. You would expect there to be a single measured value but there are two. This is due to the spin of the electron.
- If you chain Stern-Gerlach devices, the first measuring in the $\hat{z}$ direction and the $|+z\rangle$ output passing into another device measuring in the $\hat{x}$ direction you see a 50/50 split in the output. If you pass the $|+x\rangle$ output into another device measuring in the $\hat{z}$ direction you again see a 50/50 split.

1.6 Quantum information

1.6.1 Quantum information theory: example problems

- Shannon's noiseless channel coding theorem quantifies how many bits are required to store information being emitted by a source of information. A classical information source is described by a set of probabilities p_j, $j = 1, 2, \ldots, d$. Each use of the source results in the "letter" j being emitted, chosen at random with probability p_j, independently for each use of the source. If letters that occur more frequently can be represented using fewer bits, the information can be compressed. Shannon's noiseless channel

coding theorem quantifies exactly how well such a compression scheme can work.

- Shannon's noisy channel coding theorem quantifies the amount of information that can be transmitted through a noisy channel. The idea is to encode the information using error-correcting codes so that any noise introduced by the channel can be corrected at the other end. For example, a single bit of information may be encoded across two bits, in which case the channel has a capacity of half a bit.

2 Introduction to quantum mechanics

2.1 Linear algebra

2.1.1 Bases and linear independence

- A **spanning set** for a vector space is a set of vectors $|v_1\rangle, \ldots, |v_n\rangle$ such that any vector $|v\rangle$ in the vector space can be written as a linear combination $|v\rangle = \sum_i a_i |v\rangle_i$ of vectors in that set. We say that the vectors $|v_1\rangle, \ldots, |v_n\rangle$ **span** the vector space.

- A set of non-zero vectors $|v_1\rangle, \ldots, |v_n\rangle$ are **linearly dependent** if there exists a set of complex numbers $a_1, \ldots, a_n$ with $a_i \neq 0$ for at least one value of i such that

$$a_1 |v_1\rangle + \ldots + a_n |v_n\rangle = 0.$$

  A set of vectors is **linearly independent** if it's not linearly dependent.

- Any two sets of linearly independent vectors which span a vector space V contain the same number of elements. We call such a set a **basis** for V.

2.1.2 Linear operators and matrices

- A **linear operator** between vector spaces V and W is defined to be any function $A : V \rightarrow W$ which is linear in its inputs

$$A\left(\sum_i a_i |v_i\rangle\right) = \sum_i a_i A(|v_i\rangle).$$

- We usually write $A|v\rangle$ to denote $A(|v\rangle)$.

- Linear operators can be expressed as matrices. Suppose $A : V \rightarrow W$ is a linear operator between the vector spaces V and W, $|v_1\rangle, \ldots, |v_m\rangle$ is a basis for V, and $|w_1\rangle, \ldots, |w_n\rangle$ is a basis for W. Then for each j in the range $1, \ldots, m$ there exist complex numbers A_{1j} throught A_{nj} such that

$$A|v_j\rangle = \sum_i A_{ij} |w_i\rangle,$$

i.e. the result of applying A to each basis vector $|v_j\rangle$ can be expressed as a linear combination of the basis vectors $|w_i\rangle$ and the coefficients A_{ij} form the matrix representation of A.

2.1.3 The Pauli matrices

- The Pauli matrices are defined as

$$\begin{aligned}
\sigma_0 &= I \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
\sigma_1 &= \sigma_x \\
&= X \\
&= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
\sigma_2 &= \sigma_y \\
&= Y \\
&= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\
\sigma_3 &= \sigma_z \\
&= Z \\
&= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}
\end{aligned}$$

2.1.4 Inner products

- An **inner product** is a function which takes as input two vectors $|v\rangle$ and $|w\rangle$ from a vector space and produces a complex number as output. It can be denoted $(|v\rangle, |w\rangle)$, but this isn't the standard quantum mechanical notation. The standard notation is $\langle v|w\rangle$ where $\langle v|$ is the dual vector to the vector $|v\rangle$. The dual vector is a linear operator from the inner product space V to the complex numbers $\mathbb{C}$, defined by

$$\langle v|\,(|w\rangle) = \langle v|w\rangle = (|v\rangle, |w\rangle).$$

- A function $(\cdot, \cdot)$ from $V \times V$ to $\mathbb{C}$ is an inner product if
  - it is linear in the second argument

$$\left(|v\rangle, \sum_i \lambda_i |w_i\rangle\right) = \sum_i \lambda_i (|v\rangle, |w_i\rangle),$$

  - $(|v\rangle, |w\rangle) = (|w\rangle, |v\rangle)^*$, and
  - $(|v\rangle, |v\rangle) \geq 0$ with equality if and only if $|v\rangle = 0$.

- A vector space equipped with an inner product is called an **inner product space**.
- In finite dimensional complex vector spaces, a Hilbert space is the same as an inner product space.
- Vectors $|w\rangle$ and $|v\rangle$ are **orthogonal** if their inner product is zero.
- The **norm** of a vector is defined as
$$|| \, |v\rangle \, || = \sqrt{\langle v, v \rangle}.$$
- A **unit vector** is a vector $|v\rangle$ such that $|| \, |v\rangle \, || = 1$. We also say such a vector is **normalised**.
- A vector can be normalised by dividing by its norm.
- A set $|i\rangle$ of vectors with index i is **orthonormal** if each vector is a unit vector and distinct vectors in the set are orthogonal, i.e. $\langle i|j\rangle = \delta_{ij}$.
- If $|v\rangle$ is a vector in an inner product space V and $|w\rangle$ is a vector in an inner product space W, then $|w\rangle \langle v|$ is the linear operator from V to W whose action is defined by
$$(|w\rangle \langle v|)(|v'\rangle) = |w\rangle \langle v|v'\rangle = \langle v|v'\rangle \, |w\rangle \, .$$
- Let $|i\rangle$ be any orthonormal basis for the vector space V, so an arbitrary vector $|v\rangle$ can be written $|v\rangle = \sum_i v_i \, |i\rangle$ for some set of complex numbers v_i. Therefore
$$\left(\sum_i |i\rangle \langle i| \right) |v\rangle = \sum_i |i\rangle \langle i|v\rangle = \sum_i v_i \, |i\rangle = |v\rangle$$
and thus
$$\sum_i |i\rangle \langle i| = I.$$
This is known as the **completeness relation**.
- Suppose $A : V \to W$ is a linear operator, $|v_i\rangle$ is an orthonormal basis for V, and $|w_j\rangle$ is an orthonormal basis for W. We can use the completeness relation twice to obtain
$$\begin{aligned} A &= I_W A I_V \\ &= \sum_{ij} |w_j\rangle \langle w_j|A|v_i\rangle \langle v_i| \\ &= \sum_{ij} \langle w_k|A|v_i\rangle \, |w_j\rangle \langle v_i| \, . \end{aligned}$$
This is known as the **outer product representation** of A. From this we can see the matrix representation of A has $\langle w_k|A|v_i\rangle$ in the ith column and jth row.

2.1.5 Eigenvector and eigenvalues

- An **eigenvector** of a linear operator A on a vector space is a non-zero vector $|v\rangle$ such that $A|v\rangle = v|v\rangle$ where v is a complex number known as the **eigenvalue** of A corresponding to $|v\rangle$.
- The **characteristic function** of an operator is defined to be $c(\lambda) = \det|A - \lambda I|$ and the solutions to $c(\lambda) = 0$ are the eigenvalues of the operator A.
- The **eigenspace** corresponding to en eigenvalue v is the set of vectors which have eigenvalue v.
- A **diagonal representation** of an operator A on a vector space V is a representation $A = \sum_i \lambda_i |i\rangle\langle i|$ where the vectors $|i\rangle$ form an orthonormal set of eigenvectors for A with corresponding eigenvalues λ_i. In other words, the basis vectors are chosen to be orthonormal eigenvectors. The matrix representation in this basis has the eigenvalues along the diagonal and zeroes everywhere else.
- When an eigenspace is more than one-dimensional (i.e. there are multiple eigenvectors associated with a particular eigenvalue) we say that it is **degenerate**.

2.1.6 Adjoints and Hermitian operators

- If A is any linear operator on a Hilbert space V, then there exists a unique linear operator $A^\dagger$ on V such that for all vectors $|v\rangle, |w\rangle \in V$,

  $$(|v\rangle, A|w\rangle) = (A^\dagger|v\rangle, |w\rangle).$$

  This operator is known as the **adjoint** or **Hermitian conjugate** of the operator A. From definition it can be seen that $(AB)^\dagger = B^\dagger A^\dagger$, $|v\rangle^\dagger = \langle v|$, and $(A|v\rangle)^\dagger = \langle v|A^\dagger$.
- In a matrix representation of an operator A, the action of the Hermitian conjugation operation is to take the conjugate transpose of its matrix $A^\dagger = (A^*)^T$.
- An operator that is its own adjoint is known as **Hermitian** or **self-adjoint**.
- Suppose W is a k-dimensional vector subspace of the d-dimensional vector space V. Using the Gram-Schmidt procedure it's possible to construct an orthonormal basis $|1\rangle, \ldots, |d\rangle$ for V such that $|1\rangle, \ldots, |k\rangle$ is an orthonormal basis for W. For example, $\hat{x}, \hat{y}, \hat{z}$ is an orthonormal basis for $\mathbb{R}^3$ where $\hat{x}, \hat{y}$ is an orthonormal basis for the subspace $\mathbb{R}^2$.

  By definition,

  $$P = \sum_{i=1}^{k} |i\rangle\langle i|$$

is the **projector** onto the subspace W — it "picks out" only those components of the vector present in W. $|v\rangle\langle v|$ is Hermitian for any vector $|v\rangle$

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

and thus P is also Hermitian $P^\dagger = P$.

The **orthogonal complement** of P is the operator $Q = I - P$. This is a projector onto the vector space spanned by $|k+1\rangle, \ldots, |d\rangle$. This can be seen as

$$Q|v\rangle = (I-P)|v\rangle = I|v\rangle - P|v\rangle,$$

i.e. subtract from $|v\rangle$ the components present in W leaving only those outside of W.

- An operator A is said to be **normal** if $AA^\dagger = A^\dagger A$. An operator that is Hermitian is guaranteed to be normal.
- The **spectral decomposition theorem** states that any normal operator M on a vector space V is diagonal with respect to some orthonormal basis for V and, conversely, any diagonalisable operator is normal.
- A matrix U is said to be unitary if $U^\dagger U = I$. Similarly an operator U is unitary if $U^\dagger U = I$. A unitary operator also satisfies $UU^\dagger = I$ and therefore U is normal and has a spectral decomposition.
- Unitary operators preserve inner products between vectors.
- A **positive operator** A is defined to be an operator such that for any vector $|v\rangle$, $(|v\rangle, A|v\rangle)$ is a real, non-negative number. If $(|v\rangle, A|v\rangle)$ is strictly greater than zero for all $|v\rangle \neq 0$ then we say A is **positive definite**.

2.1.7 Tensor Products

- Suppose V and W are vector spaces of dimension m and n respectively. Then $V \otimes W$ (read "V tensor W") is an mn dimensional vector space. The elements of $V \otimes W$ are linear combinations of "tensor products" $|v\rangle \otimes |w\rangle$ of elements $|v\rangle$ of V and $|w\rangle$ of W. In particular, if $|i\rangle$ and $|j\rangle$ are orthonormal bases for the spaces V and W then $|i\rangle \otimes |j\rangle$ is a basis for $V \otimes W$. We often use the abbreviated notations $|v\rangle|w\rangle$, $|v, w\rangle$, and $|vw\rangle$ for $|v\rangle \otimes |w\rangle$.
- The tensor product satisfies the following properties:
  1. For an arbitrary scalar z and elements $|v\rangle$ of V and $|w\rangle$ of W,
  $$z(|v\rangle \otimes |w\rangle) = (z|v\rangle) \otimes |w\rangle = |v\rangle \otimes (z|w\rangle).$$
  2. For arbitrary $|v_1\rangle$ and $|v_2\rangle$ in V and $|w\rangle$ in W,
  $$(|v_1\rangle + |v_2\rangle) \otimes |w\rangle = |v_1\rangle \otimes |w\rangle + |v_2\rangle \otimes |w\rangle.$$

3. For arbitrary $|v\rangle$ in V and $|w_1\rangle$ and $|w_2\rangle$ in W,

$$|v\rangle \otimes (|w_1\rangle + |w_2\rangle) = |v\rangle \otimes |w_1\rangle + |v\rangle \otimes |w_2\rangle .$$

- Suppose $|v\rangle$ and $|w\rangle$ are vectors in V and W, and A and B are linear operators on V and W, respectively. Then we can define a linear operator $A \otimes B$ on $V \otimes W$ by the equation

$$(A \otimes B)(|v\rangle \otimes |w\rangle) = A|v\rangle \otimes B|w\rangle .$$

- An arbitrary linear operator C mapping $V \otimes W$ to $V' \otimes W'$ can be represented as a linear combination of tensor products of operators mapping V to V' and W to W',

$$C = \sum_i c_i A_i \otimes B_i.$$

- The inner product on $V \otimes W$ is defined as

$$\left(\sum_i a_i |v_i\rangle \otimes |w_i\rangle , \sum_j b_j |v'_j\rangle \otimes |w'_j\rangle \right) = \sum_{ij} a_i^* b_j \langle v_i|v'_j\rangle \langle w_i|w'_j\rangle .$$

- Suppose A is an m by n matrix and B is a p by q matrix. **Kronecker product** representation of the tensor product is defined as

$$A \otimes B = \begin{bmatrix} A_{11}B & A_{12}B & \cdots & A_{1n}B \\ A_{21}B & A_{22}B & \cdots & A_{bn}B \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1}B & A_{m2}B & \cdots & A_{mn}B \end{bmatrix}$$

where the terms $A_{ij}B$ represent a p by q submatrices.

For example, the tensor product of the Pauli matrices X and Y is

$$X \otimes Y = \begin{bmatrix} 0Y & 1Y \\ 1Y & 0Y \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}.$$

- The notation $|\psi\rangle^{\otimes k}$ means $|\psi\rangle$ tensored with itself k times.

2.1.8 Operator functions

- Generally speaking, given a function $f : \mathbb{C} \to \mathbb{C}$, it is possible to define a corresponding matrix function on normal matrices (or some subclass such as the Hermitian matrices). Let

$$A = \sum_i \lambda_i |i\rangle \langle i|$$

be the spectral decomposition for a normal operator A. Then, keeping in mind that $|i\rangle\langle i|$ is a matrix then

$$f(A) = \sum_i f(i)\,|i\rangle\langle i|\,.$$

For example,

$$e^{\theta Z} = \begin{bmatrix} e^{\theta} & 0 \\ 0 & e^{-\theta} \end{bmatrix}.$$

- The **trace** of a matrix A is the sum of its diagonal elements

$$\operatorname{tr} A = \sum_i A_{ii}.$$

- The trace is cyclic $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ and linear $\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$, $\operatorname{tr}(zA) = z \operatorname{tr} A$. From the cyclic property it can be seen that the trace is invariant under the unitary similarity transformation $A \to UAU^\dagger$ as $\operatorname{tr}(UAU^\dagger) = \operatorname{tr}(U^\dagger UA) = \operatorname{tr} A$.

- The trace of an operator A is defined as the trace of any matrix representation of A under any orthonormal basis

$$\operatorname{tr} A = \sum_i \langle i|A|i\rangle\,.$$

- To evaluate $\operatorname{tr}(A\,|\psi\rangle\langle\psi|)$ we can use the Gram-Schmidt procedure to express $|\psi\rangle$ in an orthonormal basis $|i\rangle$ that includes the normalised $|\psi\rangle$ as an element

$$\begin{aligned} \operatorname{tr}(A\,|\psi\rangle\langle\psi|) &= \sum_i \langle i|(A\,|\psi\rangle\langle\psi|)|i\rangle \\ &= \sum_i \langle i|A|\psi\rangle\langle\psi|i\rangle \\ &= \langle\psi|A|\psi\rangle\,. \end{aligned}$$

2.1.9 The commutator and anti-commutator

- The **commutator** of two operators A and B is defined to be

$$[A, B] = AB - BA.$$

If $[A, B] = 0$ then we say A commutes with B.

- The **anti-commutator** of two operators A and B is defined to be

$$\{A, B\} = AB + BA.$$

If $\{A, B\} = 0$ we say A anti-commutes with B.

- The **simultaneous diagonalisation theorem** states that, if A and B are Hermitian operators, then $[A, B] = 0$ if and only if there exists an orthonormal basis such that both A and B are diagonal with respect to that basis. That is, they share a common set of eigenvectors in that basis.

2.1.10 The polar and singular value decompositions

- If A is a linear operator on a vector space V then there exists unitary U and positive operators J and K such that

$$A = UJ = KU$$

  where the unique positive operators J and K satisfying these equations are defined by $J = \sqrt{A^\dagger A}$ and $K = \sqrt{AA^\dagger}$. Moveover, if A is invertible then U is unique.

- If A is a square matrix then there exist unitary matrices U and V, and a diagonal matrix D with non-negative entries such that

$$A = UDV.$$

  The diagonal elements of D are called the singular values of A.

2.2 The postulates of quantum mechanics

2.2.1 State space

- Associated to any isolated physical system is a complex vector space with inner product (i.e. a Hilbert space) known as the **state space** of the system. The system is completely described by its **state vector**, which is a unit vector in the system's state space.

- The simplest quantum mechanical system is the qubit. It has a two-dimensional state space where $|0\rangle$ and $|1\rangle$ forn an orthonormal basis for the state space.

- An arbitrary state vector in this state space can be written

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle\,.$$

  The requirement that $|\psi\rangle$ be a unit vector means that

$$|a|^2 + |b|^2 = 1.$$

2.2.2 Evolution

- The evolution of a closed quantum system is described by a **unitary transformation**. That is, the state $|\psi\rangle$ of the system at time t_1 is related to the state $|\psi'\rangle$ of the system at time t_2 by a unitary operator U which depends only on the times t_1 and t_2,

$$|\psi'\rangle = U\,|\psi\rangle\,.$$

  As a reminder, a unitary operator is one where

$$U^\dagger U = UU^\dagger = I.$$

- The time evolution of the state of a closed quantum system is described by the **Schrödinger equation**,

$$i\hbar \frac{d\left|\psi\right\rangle}{dt} = H\left|\psi\right\rangle$$

where $\hbar$ is **Planck's constant**, H is a Hermitian operator known as the **Hamiltonian** of the system.

- Because the Hamiltonian is a Hermitian operator it has a spectral decomposition

$$H = \sum_E E\left|E\right\rangle\left\langle E\right|$$

with eigenvalues E and normalised eigenvectors $\left|E\right\rangle$. The states $\left|E\right\rangle$ are known as the **energy eigenstates** and E is the energy of the state $\left|E\right\rangle$.

2.2.3 Quantum measurement

- Quantum measurements are described by a collection $\{M_n\}$ of **measurement operators**. These are operators acting on the state space of the system being measured. The index m refers to the measurement outcomes that may occur during the experiment. If the state of the quantum system is $\left|\psi\right\rangle$ immediately before the measurement then the probability that result m occurs is given by

$$p(m) = \left\langle\psi\right| M_m^\dagger M_m \left|\psi\right\rangle,$$

and the state of the system after measurement is

$$\frac{M_m\left|\psi\right\rangle}{\sqrt{\left\langle\psi\right| M_m^\dagger M_m \left|\psi\right\rangle}}.$$

- For example, to measure a qubit in the computational basis there are two measurement operators, $M_0 = \left|0\right\rangle\left\langle 0\right|$ and $M_1 = \left|1\right\rangle\left\langle 1\right|$. Each measurement operator is Hermitian $M_0^2 = M_0$, $M_1^2 = M_1$. If the state is $\left|\psi\right\rangle = a\left|0\right\rangle + b\left|1\right\rangle$ then the probability of measuring 0 is

$$p(0) = \left\langle\psi\right| M_0^\dagger M_0 \left|0\right\rangle = \left\langle\psi\right| M_0 \left|\psi\right\rangle = \left\langle\psi|0\right\rangle\left\langle 0|\psi\right\rangle = |a|^2.$$

The state after the measurement is

$$\frac{M_0\left|\psi\right\rangle}{|a|} = \frac{a}{|a|}\left|0\right\rangle,$$

and because $a/|a|$ is unitary it has no practical impact and the result is effectively $\left|0\right\rangle$.

2.2.4 Distinguishing Quantum States

- If Alice chooses a state $|\psi_i\rangle$ $(1 \leq i \leq n)$ from a fixed set of states and gives it to Bob, it's only possible for Bob to reliably distinguish the state if $|\psi_i\rangle$ are orthonormal.

2.2.5 Projective Measurements

- A **projective measurement** is described by an observable, M, a Hermitian operator on the state space of the system being observed. The observable has a spectral decomposition

$$M = \sum_m m P_m$$

  where P_m is the projector onto the eigenspace of M with eigenvalue m. The possible outcomes of the measurement correspond to the eigenvalues m of the observable. Upon measuring the state $|\psi\rangle$, the probability of getting result m is given by

$$p(m) = \langle\psi|P_m|\psi\rangle\,.$$

  Given that m occurred, the state of the quantum system immediately after the measurement is

$$\frac{P_m\,|\psi\rangle}{\sqrt{p(m)}}.$$

- In other words, on measuring the state $|\psi\rangle$ you'll get one of the eigenvalues m of M with probability

$$p(m) = \langle\psi|P_m|\psi\rangle$$

  and the state will collapse to the eigenvector associated with m.

- The average (or expected) value of a measurement is

$$\begin{aligned}\mathbf{E}(M) &= \sum_m m\,p(m) \\ &= \sum_m m\,\langle\psi|P_m|\psi\rangle \\ &= \langle\psi|\left(\sum_m m P_m\right)|\psi\rangle \\ &= \langle\psi|M|\psi\rangle\,.\end{aligned}$$

  This is often written $\langle M\rangle$.

- The standard deviation of a measurement is

$$\Delta(M) = \sqrt{\langle M^2\rangle - \langle M\rangle^2}.$$

2.2.6 POVM measurements

- The general measurement postulate gives us ways to both determine the probability of measuring a particular value $p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$, and to determine the state of the system after measuring that value

$$|\psi'\rangle = \frac{M_m\,|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}.$$

  However, in some cases we don't care about the post-measurement state in which case we can define

$$E_m = M_m^\dagger M_m$$

  and the probabilities are given by $p(m) = \langle\psi|E_m|\psi\rangle$. The set of operators E_m are called the **POVM elements** and they are sufficient to determine the probabilities of the different measurement outcomes. The complete set $\{E_m\}$ is known as **a POVM**.

- Projective measurements assume that performing the same measurement multiple times leaves the system unchanged which isn't always the case, e.g. if performing the measurement destroys the system. For this reason the general measurement postulate is often more useful.

2.2.7 Phase

- The term "phase" can have several meanings in quantum mechanics.

- One is **global phase factor** — a unitary complex factor applied to an entire quantum state. For example, applying a measurement operator M to the quantum state $e^{i\theta}\,|\psi\rangle$ where θ is a real number gives $\langle\psi|e^{-i\theta}M^\dagger M e^{i\theta}|\psi\rangle = \langle\psi|M^\dagger M|\psi\rangle$, i.e. the factor has no effect on measurement.

- Another is **relative phase factor** — complex factors applied to different basis vectors in a quantum state. For example, in the states

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

  the amplitudes of the $|0\rangle$ basis vectors are both $1/\sqrt{2}$ while the amplitudes of the $|1\rangle$ basis vectors are $1/\sqrt{2}$ and $-1/\sqrt{2}$. The magnitude of the $|1\rangle$ amplitudes are the same but they differ in sign. More generally, we say that two amplitudes, a and b, differ by a relative phase if there is a real θ such that $a = e^{i\theta}b$. Two states are said to differ by a relative phase in some basis if each of the amplitudes in that basis is related by such a phase factor.

- States that differ by relative phase factors in some basis have observable differences in measurement statistics, i.e. we can distinguish them.

2.2.8 Composite Systems

- The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if we have systems numbered 1 through n, and system number i is prepared in the state $|\psi_i\rangle$, then the joint state of the total system is $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle$.
- Subscript notation is sometimes used to denote states and operators on different systems within a composite system. For example, in a system containing three qubits, X_2 is the Pauli σ_x operator acting on the second qubit.