Theory of Quantum Information and Quantum Computing

Maso*

October 18, 2024

Contents

Lecture 1

mer 02 ott 2024 16:30

Topics. A review of quantum mechanics. The first part is about qubits and the theory behind them: entanglement, Bell's inequalities. In this part one defines the operation of the quantum computer, define quantum gates; one defines algorithms and looks for the ones that are faster than classical ones; Shor's algorithm and others. Quantum error correction (not done: fault tolerance). The second part is about the physical realization of quantum computers: ion traps and superconducting qubits. A qubit is just a two-level quantum system: for ion traps the two levels are two well-separated excitation modes of an ion. The superconducting qubits are also called artificial atoms: they are made from superconducting circuits, there is no resistance and the circuit behaves as an harmonic oscillator.

Knowledge required. Quantum mechanics from the bachelor.

References. Main reference is the Nielsen–Chuang, "Quantum computation and quantum information". It is a compendium, not really technical and not written for physicists. For quantum computation and quantum algorithms there is Mermin, "Quantum computer science". Some online lectures are: Aaronson at Austin University for educated general audiences, Preskill at Caltech (fault tolerance), IBM webpage.

For the second part of the course, see professor's notes and refs online.

Currently there are several quantum computers. The first quantum computer was based on quantum annealing which is good for optimization problems, but this is not treated in the course. The course studies the superconducting qubits at IBM and Google. Right now, the order of qubits is about a thousand and not much can be done. The majority of those qubits

*https://github.com/M-a-s-o/notes

are allocated for error correction. The problem with qubits is transfer the classical information to the two levels in an isolated way. It is difficult to keep stable an excitation and it is even more difficult to maintain a superposition with a given relative phase. With time, the phases get washed out by the interactions with the environment. Keeping coherence of a two-level quantum system is difficult.

Exam. The exam is just an oral evaluation. Can be done whenever with enough notice. If one wishes, one may give a presentation about a topic not done in the course; the questions are more general; the presentation must be agreed upon.

1 Quantum mechanics

See Nielsen §§1.2, 1.3, 2.1, 2.2.

Notation. The state of the two-level system are called

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

One considers these two states as carriers of information: the quantum analog of the classic bit. The main difference between classical and quantum computation is that quantum state may exist in superposition

$$\alpha\,|0\rangle + \beta\,|1\rangle \ , \quad \alpha, \beta \in \mathbb{C}$$

Quantum mechanics is inherently probabilistic. If the system is in the state $|0\rangle$, then, by performing a measurement to see in which state the system is, one finds $|0\rangle$. One may state with certainty that the system is in the state $|0\rangle$. Same for $|1\rangle$. However, if the state is in a superposition, it is not obvious in which state the system is. One may find either state with probability

$$P(|0\rangle) = |\alpha|^2 \, , \quad P(|1\rangle) = |\beta|^2$$

If the two above are probabilities, then it must hold

$$|\alpha|^2 + |\beta|^2 = 1$$

From Quantum Mechanics, one knows that the global phase is irrelevant. [r] therefore there are two real numbers that characterizes the state. The quantum states carry a lot of information, the information related to two real numbers. Even if the two numbers are real which in theory may encode everything, the amount of information one may extract is much lower. One extracts the state, not the number. The two numbers may be recovered by performing many measurements on the system. This is true with a large number of initial states, but for a quantum computer there is a single quantum object. When performing a measurement, the wave function collapses and instantly becomes the state just observed. For a single measurement, the outcome must be either state and one may not do another measurement to extract the probabilities: it is impossible to extract the information.

The collapse of the wave function is a very discussed topic in physics. The collapse may be used in quantum computers, one exploits it to extract some information in a clever way.

1.1 Qubits

Definition 1.1 (I, State)**.** In quantum mechanics, states are rays in a Hilbert space.

Let $|\psi\rangle \in \mathcal{H}$ be a vector in the Hilbert space $\mathcal{H}$. A ray is the equivalence class given by

$$|\psi\rangle \sim \mathrm{e}^{\mathrm{i}\alpha}\,|\psi\rangle$$

with $|\psi\rangle$ having unit length.

Remark 1.2. Notice that not all states in the Hilbert space is a good quantum state. One needs to normalize it.

Dirac notation. The scalar product of two states is a complex number

$$\langle\phi|\psi\rangle \in \mathbb{C}$$

The length is defined as

$$\||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle}$$

The Hilbert space is a vector space, therefore given two vectors, every linear combination of them is also a vector in the Hilbert space.

The bra $\langle\phi|$ is a covector, a functional that acts on vectors

$$\langle\phi| : \mathcal{H} \to \mathbb{C}\,, \quad |\psi\rangle \mapsto \langle\phi|\psi\rangle$$

For a single qubit, one needs a Hilbert space of dimension 2. All vector spaces of dimension 2 are isomorphic to $\mathcal{H} \simeq \mathbb{C}^2$. One identifies

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

The generic vector is a pair of complex numbers linear combination of the two previous ones

$$\begin{bmatrix}z_1\\z_2\end{bmatrix} = z_1\begin{bmatrix}1\\0\end{bmatrix} + z_2\begin{bmatrix}0\\1\end{bmatrix} = z_1\,|0\rangle + z_2\,|1\rangle$$

The scalar product between two arbitrary vectors

$$|\phi\rangle = \begin{bmatrix}w_1\\w_2\end{bmatrix}\,, \quad |\psi\rangle = \begin{bmatrix}z_1\\z_2\end{bmatrix}$$

is the typical Hermitian scalar product

$$\langle\phi|\psi\rangle = w_1^* z_1 + w_2^* z_2 = \begin{bmatrix}w_1^* & w_2^*\end{bmatrix}\begin{bmatrix}z_1\\z_2\end{bmatrix}$$

One identifies

$$\langle\phi| = \begin{bmatrix}w_1^* & w_2^*\end{bmatrix} = (|\phi\rangle)^\dagger$$

The two vectors $|0\rangle$ and $|1\rangle$ have been chosen to be an orthonormal basis of the Hilbert space

$$\langle 0|0\rangle = \langle 1|1\rangle = 1\,, \quad \langle 0|1\rangle = 0$$

Form of a generic state. The most general state of a qubit is a generic ray in the Hilbert space. Therefore, one requires that

$$|z_1|^2 + |z_2|^2 = 1$$

There is a convenient parametrization of the parameters

$$|\psi\rangle = \begin{bmatrix}z_1\\z_2\end{bmatrix} = \cos\frac{\theta}{2}\mathrm{e}^{\mathrm{i}\phi_1}\,|0\rangle + \sin\frac{\theta}{2}\mathrm{e}^{\mathrm{i}\phi_2}\,|1\rangle$$

Applying phase invariance, one may eliminate a phase

$$|\psi\rangle = \begin{bmatrix}z_1\\z_2\end{bmatrix} = \sin\frac{\theta}{2}\,|0\rangle + \cos\frac{\theta}{2}\mathrm{e}^{\mathrm{i}\phi}\,|1\rangle$$

This is not the only way, but it is the typical one.

As noted before, one needs two real numbers to describe the content of the qubit.

Bloch sphere. Consider unit vector in $\mathbb{R}^3$

$$\mathbf{n} = \begin{bmatrix}\sin\theta\cos\varphi\\\sin\theta\sin\varphi\\\cos\theta\end{bmatrix}\,, \quad \theta \in [0,\pi]\,, \quad \varphi \in [0,2\pi]$$

One covers all possible points on the surface of a sphere. The angle θ is the colatitude and φ is the longitude. There is a parallel between the state of a qubit and a point on this Bloch sphere. The state $|0\rangle$ corresponds to $\theta = 0$, the north pole; while $|1\rangle$ corresponds to $\theta = \pi$. The mapping between the Hilbert space and the Bloch sphere is not an isometry due to the half angle present.

1.2 Observables and measurements

Definition 1.3 (II, Observables)**.** The observables correspond to self-adjoint operators

$$A^\dagger = A$$

where the adjoint is the transpose conjugate $A^\dagger = (A^\top)^* = (A^*)^\top$.

Theorem 1.4 (Spectral)**.** Let A be a self-adjoint operator on a Hilbert space. It exists an orthonormal basis of eigenvectors of such operator

$$A\left|n\right\rangle = a_n \left|n\right\rangle \ , \quad \left\langle n|m\right\rangle = \delta_{nm}$$

Therefore, any vector may be written in terms of the basis

$$\left|\psi\right\rangle = \sum_n \alpha_n \left|n\right\rangle \ , \quad \alpha \in \mathbb{C}$$

Projection operator. A self-adjoint operator of dimension d has d eigenvalues (not necessarily different). If the eigenvalues are not all different, then there is a degeneracy and one may group the corresponding eigenvectors. For each block, one may define the projection operator

$$P_{a_p}\left|\psi\right\rangle = \alpha_{p_1}\left|p_1\right\rangle + \cdots + \alpha_{p_n}\left|p_n\right\rangle$$

Example 1.5. Consider just one eigenvalue α_n that is not degenerate. The projection of the state $\left|\psi\right\rangle$ is done along the eigenvector $\left|n\right\rangle$ to give

$$\left|n\right\rangle\left\langle n|\psi\right\rangle = P\left|\psi\right\rangle$$

The second factor is a complex number. One may formally write that

$$P_{a_n} = \left|n\right\rangle\!\left\langle n\right|$$

Qubit. [r] In $\mathbb{C}^2$, therefore a qubit, one may parametrize a self-adjoint matrix as

$$A = \begin{bmatrix} a+b & c-\mathrm{i}d \\ c+\mathrm{i}d & a-d \end{bmatrix} = aI + b\sigma_3 + c\sigma_1 + d\sigma_2$$

where σ_i are the Pauli matrices

$$X = \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \ , \quad Y = \sigma_2 = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix} \ , \quad Z = \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The two states $\left|0\right\rangle$ and $\left|1\right\rangle$ are two eigenvectors of σ_3:

$$\sigma_3\left|0\right\rangle = \left|0\right\rangle \ , \quad \sigma_3\left|1\right\rangle = -\left|1\right\rangle$$

[r] If σ_3 is an observable, there must be an eigenbasis. One may check that the eigenbasis of σ_1 is given by

$$\left|+\right\rangle = \frac{\left|0\right\rangle + \left|1\right\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \ , \quad \left|-\right\rangle = \frac{\left|0\right\rangle - \left|1\right\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

The eigenbasis of σ_2 is given by

$$\left|+\mathrm{i}\right\rangle = \frac{\left|0\right\rangle + \mathrm{i}\left|1\right\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix} \ , \quad \left|-\mathrm{i}\right\rangle = \frac{\left|0\right\rangle - \mathrm{i}\left|1\right\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}$$

On the Bloch sphere, the eigenvectors of σ_i lie on the axis i. For a generic point on the sphere, one may define the spin in a generic direction

$$\boldsymbol{\sigma}\cdot\mathbf{n} = \sigma_1 n_1 + \sigma_2 n_2 + \sigma_3 n_3$$

the generic state $\left|\psi\right\rangle$ is an eigenvector

$$\boldsymbol{\sigma}\cdot\mathbf{n}\left|\psi\right\rangle = \left|\psi\right\rangle$$

Lecture 2

mer 09 ott 2024 16:30

Definition 1.6 (III, Measurement)**.** The result of a measurement of an observable A on the ket state $|\psi\rangle$ is its eigenvalues a_n with probability

$$P(a_n) = \|P_{a_n} |\psi\rangle\|^2$$

After the measurement, the state instantly becomes

$$|\psi\rangle \to \frac{P_{a_n} |\psi\rangle}{\|P_{a_n} |\psi\rangle\|}$$

and any further measurement will give a_n with probability 1. This postulate is also called Born rule.

In physics, the observable A (like the spin, the energy, etc.) is the fundamental object. In quantum computing, the observable A often just selects a basis $|n\rangle$ and the operator can be forgotten. One that says that one "measures" in the basis $|n\rangle$. One also uses the labels "n" to identify the state, instead of the physical quantity a_n measured. So, $|0\rangle$and $|1\rangle$ are eigenstates of spin along the z direction

$$S_z = \frac{\hbar}{2}\sigma_3$$

with eigenvalues $\pm\hbar/2$, but one refers to them as "0" and "1".

The simplest case is when all eigenvalues are different. Then one has N distinct outcomes of a measurement and a basis $|n\rangle$ in $\mathbb{C}^N$ with

$$|\psi\rangle = \sum_{n=1}^{N} \alpha_n |n\rangle = \alpha_1 |1\rangle + \cdots + \alpha_N |N\rangle = P_{a_1} |\psi\rangle + \cdots + P_{a_N} |\psi\rangle$$

The probability is just

$$P(a_n) = \|P_{a_n} |\psi\rangle\|^2 = \|\alpha_n |n\rangle\|^2 = |\alpha_n|^2$$

The state collapses to the "normalized" state

$$P_{a_n} |\psi\rangle = \alpha_n |n\rangle \ , \quad |\psi\rangle \to \frac{\alpha_n}{|\alpha_n|} |n\rangle \sim |n\rangle$$

Notice that the normalization may give a phase, but one is interested in rays as quantum states. The probability α_n disappears because all the states $|\psi\rangle$ must be of unit length.

For a qubit, one may measure in several bases, for example $(|0\rangle, |1\rangle)$ and $(|+\rangle, |-\rangle)$. Physically these correspond to measurements of S_z and S_x. Given a state $|\psi\rangle$, one may expand it

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle = \gamma |+\rangle + \delta |-\rangle$$

Then measure in the computational basis $(|0\rangle, |1\rangle)$

$$P(0) = |\alpha|^2 , \quad P(1) = |\beta|^2$$

corresponding respectively to spin up and spin down. One may instead measure in the basis $|\pm\rangle$

$$P(+) = |\gamma|^2 , \quad P(-) = |\delta|^2$$

Notice that a state $|+\rangle$ has probability 1 of measuring + (positive spin along the x direction), but equal probability of having either 0 or 1 (spin up or down along the z direction). This is because σ_1 and σ_3 do not commute. However, if one measures in the basis $(|0\rangle, |1\rangle)$ and finds $|0\rangle$, then the state $|+\rangle$ becomes $|0\rangle$ and any further measurements of σ_3 gives 0, but now the spin along the x direction is undetermined.

There is an equivalent formulation of the third postulate using non-orthogonal bases of operators (POVM formulation) that is useful to study non-isolated systems (i.e. open systems). See Nielson–Chuang and Preskill. For the moment the textbook's Born rule (also called "projective measurements") suffices.

1.3 Time evolution

Definition 1.7 (IV, Time evolution)**.** States evolve according to the Schrödinger equation

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

where the Hamiltonian $H(t)$ is self-adjoint.

The solution to the Schrödinger equation

$$|\psi(t)\rangle = U(t) |\psi(0)\rangle$$

can be written in terms of a time-evolution operator. This operator is immediate to find when the Hamiltonian H is time-independent

$$U(t) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}Ht}$$

The important point is that $U(t)$ is unitary

$$U^\dagger(t)U(t) = U(t)U^\dagger(t) = I$$

This also implies that time evolution is invertible. Unitary operators preserve scalar products and total probability

$$\langle U\phi|U\psi\rangle = \langle\phi| U^\dagger U |\psi\rangle = \langle\phi|\psi\rangle$$

Therefore the probability

$$\langle\psi(t)|\psi(t)\rangle = \langle\psi(0)|\psi(0)\rangle = 1$$

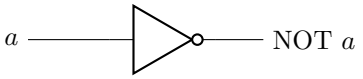
is conserved. The fact that $U(t)$ is unitary follows from the Schrödinger equation

$$\mathrm{d}_t \langle\psi(t)|\psi(t)\rangle = \frac{\mathrm{i}}{\hbar} \langle\psi(t)| H^\dagger |\psi(t)\rangle - \frac{\mathrm{i}}{\hbar} \langle\psi(t)| H |\psi(t)\rangle = 0$$

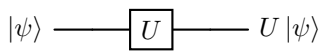
this holds since the Schrödinger equation also implies

$$-\mathrm{i}\hbar\, \mathrm{d}_t \langle\psi(t)| = \langle\psi(t)| H^\dagger\,, \quad H^\dagger = H$$

The only thing that may happen to a classical bit in its time evolution is to be flipped. In computer science one introduces logic gates. The standard 1-bit gate is the NOT gate



For one qubit the situation is more interesting. Using a "circuit model" to denote evolution (in quantum computing this is a qubit passing through a gate)



The matrix U can be any unitary matrix, if operator H is general enough. There are many possibilities. The Pauli matrices are not only Hermitian, but also unitary

$$\sigma_i^\dagger = \sigma_i\,, \quad \sigma_i^2 = I \implies \sigma_i^\dagger\sigma_i = I$$

So if one is clever enough, one may construct the gates (i.e. the physical systems) that perform the unitary operations

$$I\,, \quad X = \sigma_1\,, \quad Y = \sigma_2\,, \quad Z = \sigma_3$$

Notice that

$$\sigma_1\sigma_3 = -\mathrm{i}\sigma_2 \iff XZ = -\mathrm{i}Y$$

and $-\mathrm{i}Y$ is also unitary. Therefore one may use

$$I\,, \quad X\,, \quad Z\,, \quad XZ$$

which are denoted as U above. Since the first Pauli matrix is

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

it acts as a quantum NOT gate

$$X|0\rangle = |1\rangle \,, \quad X|1\rangle = |0\rangle$$

$$|0\rangle \text{ —}\boxed{X}\text{— } |1\rangle \,, \qquad |1\rangle \text{ —}\boxed{X}\text{— } |0\rangle$$

while Z flips signs

$$Z|0\rangle = |0\rangle \,, \quad Z|1\rangle = -|1\rangle \,, \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$|0\rangle \text{ —}\boxed{Z}\text{— } |0\rangle \,, \qquad |1\rangle \text{ —}\boxed{Z}\text{— } -|1\rangle$$

Notice that flipping a sign is equivalent to adding a relative phase of $\mathrm{e}^{\mathrm{i}\pi}$

$$a|0\rangle + b|1\rangle \text{ —}\boxed{Z}\text{— } a|0\rangle - b|1\rangle$$

There are also other interesting unitary matrices. In quantum computing, an important example is the Hadamard gate

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \,, \qquad \text{———}\boxed{H}\text{———}$$

The gate switches the bases $(|0\rangle, |1\rangle)$ and $(|+\rangle, |-\rangle)$

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

Other examples are

$$S = \sqrt{Z} = \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{i} \end{bmatrix} \,, \quad T = \sqrt{S} = \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\frac{\pi}{4}} \end{bmatrix}$$

There is a continuum of two-by-two unitary matrices. The generic matrix H is

$$H = aI + b_i \sigma_i \,, \quad a, b_i \in \mathbb{R}$$

and the associated unitary matrix is

$$U = \mathrm{e}^{-\mathrm{i}Ht} = \mathrm{e}^{-\mathrm{i}at}\mathrm{e}^{-\mathrm{i}\mathbf{b}\cdot\boldsymbol{\sigma}t} = \mathrm{e}^{-\mathrm{i}at}\mathrm{e}^{-\mathrm{i}b\mathbf{n}\cdot\boldsymbol{\sigma}t} \,, \quad \hbar = 1 \,, \quad b = \|\mathbf{b}\|$$

So U depends on four real parameters. In particular, the matrix

$$\boxed{R_{\mathbf{n}}(\lambda) = \mathrm{e}^{-\mathrm{i}\frac{\lambda}{2}(\mathbf{n}\cdot\boldsymbol{\sigma})}} = \cos\frac{\lambda}{2} - \mathrm{i}(\mathbf{n}\cdot\boldsymbol{\sigma})\sin\frac{\lambda}{2}$$

is unitary. In quantum mechanics this is a rotation about the direction $\mathbf{n}$. The unitary matrix is then

$$U = \mathrm{e}^{-\mathrm{i}\alpha} R_{\mathbf{n}}(2bt)$$

where $\alpha \in \mathbb{R}$ is just a number. The product of matrices of the form $e^{-\mathrm{i}\alpha}R$ generate all unitary operators. For example, one may show that any unitary matrix U can be written as

$$U = \mathrm{e}^{-\mathrm{i}\alpha} \begin{bmatrix} \mathrm{e}^{-\mathrm{i}\frac{\beta}{2}} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\frac{\beta}{2}} \end{bmatrix} \begin{bmatrix} \cos\frac{\gamma}{2} & -\sin\frac{\gamma}{2} \\ \sin\frac{\gamma}{2} & \cos\frac{\gamma}{2} \end{bmatrix} \begin{bmatrix} \mathrm{e}^{-\mathrm{i}\frac{\delta}{2}} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\frac{\delta}{2}} \end{bmatrix} = \mathrm{e}^{-\mathrm{i}\alpha} R_z(\beta) R_x(\gamma) R_z(\delta)$$

in terms of four real parameters $\alpha, \beta, \gamma, \delta \in \mathbb{R}$.

One may wonder if quantum computing is richer than classical computing.

Exercise 1.8. Prove

$$\mathrm{e}^{-\mathrm{i}\frac{\lambda}{2}(\mathbf{n}\cdot\boldsymbol{\sigma})} = \cos\frac{\lambda}{2} - \mathrm{i}(\mathbf{n}\cdot\boldsymbol{\sigma})\sin\frac{\lambda}{2}$$

Exercise 1.9. Prove that $R_{\mathbf{n}}(x)$ acts as a rotation on a state on the Bloch sphere. See Nielsen, exercise 4.6.

1.4 Multipartite systems

One would like to study multiple qubits.

Definition 1.10 (V, Multipartite systems)**.** If a quantum system is composed by two subsystems A and B, then

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$$

Consider two qubits. Each can be in the state 0 or 1. There are four possible choices of the total system

$$00\,, \quad 01\,, \quad 10\,, \quad 11$$

There must exist four states

$$|00\rangle\,, \quad |01\rangle\,, \quad |10\rangle\,, \quad |11\rangle$$

and by the rules of quantum mechanics, all possible linear combinations too

$$|\psi\rangle = \alpha_{00}\,|00\rangle + \alpha_{01}\,|01\rangle + \alpha_{10}\,|10\rangle + \alpha_{11}\,|11\rangle\,, \quad \sum_x |\alpha_x|^2 = 1$$

If one measures both qubits, one may get the state $x = 00, 01, 10, 11$ with probability $|\alpha_x|^2$. One may be interested only in the first qubit, regardless of the other. The result is can be 0 or 1 and the probability is given by the Born rule. The state can be rewritten as

$$|\psi\rangle = (\alpha_{00}\,|00\rangle + \alpha_{01}\,|01\rangle) + (\alpha_{10}\,|10\rangle + \alpha_{11}\,|11\rangle) = P_0\,|\psi\rangle + P_1\,|\psi\rangle$$

The probability of getting 0 in the first qubit is

$$P = \|P_0\,|\psi\rangle\|^2 = |\alpha_{00}|^2 + |\alpha_{01}|^2$$

One sees that one has to to sum the squared values of all the coefficients containing 0 in the first qubit. After the measurement, the state collapses

$$|\psi\rangle \to \frac{P_0\,|\psi\rangle}{\|P_0\,|\psi\rangle\|} = \frac{\alpha_{00}\,|00\rangle + \alpha_{01}\,|01\rangle}{\sqrt{|\alpha_{00}|^2 + |\alpha_{01}|^2}}$$

where the denominator is needed to normalize the state.

More generally, given $\mathcal{H}_A$ and $\mathcal{H}_B$ with orthogonal bases $|n\rangle$ and $|m\rangle$, the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$ is the formal linear combination of vectors

$$\sum_{n,m} \alpha_{nm}\,|nm\rangle\,, \quad \alpha_{nm} \in \mathbb{C}$$

More precisely, the product $\mathcal{H}_A \otimes \mathcal{H}_B$ is a Hilbert space of dimension $\dim \mathcal{H}_A \cdot \dim \mathcal{H}_B$ with an operation on vectors (the tensor product)

$$\otimes : \mathcal{H}_A \times \mathcal{H}_B \to \mathcal{H}_A \otimes \mathcal{H}_B\,, \quad (|n\rangle\,, |m\rangle) \mapsto |n\rangle \otimes |m\rangle = |nm\rangle$$

extended by linearity to all other vectors

$$\left[\sum_n \alpha_n\,|n\rangle\right] \otimes \left[\sum_m \beta_m\,|m\rangle\right] = \sum_{n,m} \alpha_n \beta_m\,|nm\rangle$$

Vectors in the tensor product space can be written as

$$|\psi\rangle_A \otimes |\phi\rangle_B$$

and therefore are called **separable**. They are a tiny fraction of all the vectors in the product space. For them it holds

$$\alpha_{nm} = \alpha_n \beta_m$$

and only few matrices are products of vectors (they must have rank one). Vectors that are not separable are called **entangled** and this notion fundamental in quantum computing.

The separable states are a subset with measure zero of all possible states in the product space.

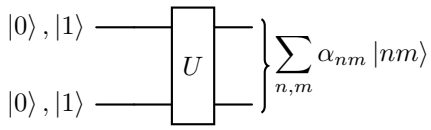
Example 1.11. The following state is separable

$$\frac{|00\rangle + |01\rangle}{\sqrt{2}} = \frac{|0\rangle \otimes |0\rangle + |0\rangle \otimes |1\rangle}{\sqrt{2}} = |0\rangle \otimes \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

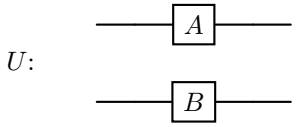
while the following states are entangled

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}}, \quad \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

A system of two qubits is depicted as



It is equivalent to a four dimensional vector. The unitary matrices U acting on it are 4×4. Some of the matrices can be realized by acting on each qubit independently



In other words, there exist 2×2 matrices A and B such that $U = A \otimes B$, meaning that it acts as

$$U[|\psi\rangle \otimes |\phi\rangle] = A\,|\psi\rangle \otimes B\,|\phi\rangle$$

The action of the matrix U on a generic vector is uniquely fixed by linearity. One may want to switch from a 2×2 description of the single qubits to a 4×4 vector description. Each single qubit can be put in correspondence with a vector in $\mathbb{C}^2$

$$\alpha_0\,|0\rangle + \alpha_1\,|1\rangle \to \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} \in \mathbb{C}^2$$

Similarly, for two qubits one may use a vector in $\mathbb{C}^4$

$$\alpha_{00}\,|00\rangle + \alpha_{01}\,|01\rangle + \alpha_{10}\,|10\rangle + \alpha_{11}\,|11\rangle \to \begin{bmatrix} \alpha_{00} \\ \alpha_{01} \\ \alpha_{10} \\ \alpha_{11} \end{bmatrix} \in \mathbb{C}^4$$

The order of the components must be chosen from the beginning and always be the same. To go from one description to the other one may use the Kronecker product. Consider two generic matrices A and B, then their Kronecker product is

$$A \otimes B = \begin{bmatrix} A_{11}B & A_{12}B & \cdots & A_{1q}B \\ A_{21}B & A_{22}B & \cdots & A_{2q}B \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1}B & A_{p2}B & \cdots & A_{pq}B \end{bmatrix}$$

For vectors this is

$$\begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} \otimes \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \alpha_0 \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \\ \alpha_1 \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \alpha_0\beta_0 \\ \alpha_0\beta_1 \\ \alpha_1\beta_0 \\ \alpha_1\beta_1 \end{bmatrix}$$

which is just

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

One may use σ_1 and σ_3 as the operators and the 4×4 matrix is

$$\sigma_1 \otimes \sigma_3 = X \otimes Z = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

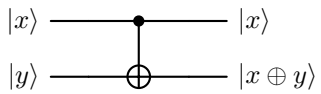
As for vectors, few 4×4 unitary matrices U are of the form $A \otimes B$. The other are actually more important for quantum computing: matrices that act on multiple qubits and not on individual ones independently. In classical computing there are several classical gates (e.g. AND, OR, XOR, etc) that cannot be factorized. The most important quantum gate is the controlled NOT or CNOT, a quantum generalization of the XOR. On a basis, it acts as

$$|00\rangle \to |00\rangle \ , \quad |01\rangle \to |01\rangle \ , \quad |10\rangle \to |11\rangle \ , \quad |11\rangle \to |10\rangle$$

It does nothing if the first qubit is 0 and flips the second if the first is 1. The first qubit is called **control** qubit and the second is called **target** qubit. As a matrix it is

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

Letting x be the control and y the target, it is denoted as

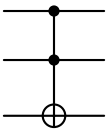


where $x, y = 0, 1$ and $\oplus$ denotes the XOR or, equivalently, the sum mod 2. In fact, if $x = 0$ nothing happens, but if $x = 1$ then

$$y \to y \oplus 1 = \begin{cases} 1 \, , & y = 0 \\ 2 \mod 2 = 0 \, , & y = 1 \end{cases}$$

Therefore, the gate CNOT is indeed a XOR.

A difference between quantum and classical gates is that quantum gates are always reversible: if one can apply U, one can also apply U^{-1} since both are unitary. Classical gates are not reversible in general, but one can prove that all classical circuits can be replaced by reversible classical circuits using the Toffoli gate, a two-bit XOR



Another difference is the following. In classical computing there are a finite number of gates, while in quantum computing one may choose any unitary matrix. For two-qubit systems, the most general unitary matrix has 16 real parameters.

Lecture 3

lun 14 ott 2024 16:30

1.5 No-cloning

In particle physics, one works with beams that contain many particles in the same initial state. The output of an experiment is related to the probability of the state

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

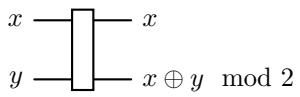
With many copies of the initial state $|\psi\rangle$, one may perform may measurements and extract the values of $|\alpha|^2$ and $|\beta|^2$.

In quantum computing, one typically has only one qubit in some state. In such situation, one cannot know much about the coefficients α and β. An experiment gives either $|0\rangle$ or $|1\rangle$ randomly and, after the experiment, if the result is 0, then the state becomes

$$|\psi\rangle \to |0\rangle$$

and all the information about α and β is lost. It is often said that quantum computing is superior to classical computing because the state $|\psi\rangle$ contains infinite information (i.e. the coefficients are understood as long classical bits), but the problem is that one cannot extract such information. The art of quantum computing is learning something about the coefficient without disturbing too much the system: sometimes this information is greater than the classical counterpart, but only sometimes. One learns how in the following.

This problem would be solved if states could be duplicated. In classical computing this is possible. One needs to use a "classical" CNOT gate



that can easily be realized. Then one starts with the bit x that is to be duplicated and the bit y initialized to zero. The output is two copies of x: the bit has been cloned.

In quantum mechanics, one can do something similar using a CNOT gate. Starting with $|x\rangle = |0\rangle, |1\rangle$ and $|y\rangle = 0$, the CNOT gate acts as its classical counterpart. In other words, starting with $|x\rangle = |0\rangle$, nothing happens to the target qubit initialized to $|y\rangle = |0\rangle$

$$|00\rangle \to |00\rangle$$

One has two copies of the state $|0\rangle$. Starting with $|x\rangle = |1\rangle$, the target qubit is flipped

$$|10\rangle \to |11\rangle$$

so one has two copies of the state $|1\rangle$. Therefore

$$|\psi\rangle \otimes |0\rangle \to |\psi\rangle \otimes |\psi\rangle$$

and the state $|\psi\rangle$ has been cloned. However, this is only true because one chooses the state $|\psi\rangle$ to be an element of an orthogonal basis. One may study the case for a linear combination

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

Starting with

$$[\alpha |0\rangle + \beta |1\rangle] \otimes |0\rangle = \alpha |00\rangle + \beta |10\rangle$$

the CNOT gate acts as

$$|\psi\rangle \otimes |0\rangle \to \alpha |00\rangle + \beta |11\rangle$$

which is different from

$$|\psi\rangle \otimes |\psi\rangle = \alpha^2 |00\rangle + \alpha\beta |01\rangle + \beta\alpha |10\rangle + \beta^2 |11\rangle$$

In the above result there are no states $|01\rangle$ and $|10\rangle$: the two coincide if and only if $\alpha\beta = 0$ which implies $\alpha = 0$ or $\beta = 0$, equivalently the states can be $|\psi\rangle = |0\rangle$ or $|\psi\rangle = |1\rangle$ only. This is a general rule: one may only clone orthogonal states.

Theorem 1.12 (No-cloning)**.** Given a normalized state $|\phi\rangle$, there is no unitary operator that acts as

$$|\psi\rangle \otimes |\phi\rangle \to U(|\psi\rangle \otimes |\phi\rangle) = |\psi\rangle \otimes |\psi\rangle$$

for all normalized states $|\psi\rangle$.

Proof. By contradiction, suppose the thesis to be true. Then, for two arbitrary states, one has

$$U(|\psi_1\rangle \otimes |\phi\rangle) = |\psi_1\rangle \otimes |\psi_1\rangle \ , \quad U(|\psi_2\rangle \otimes |\phi\rangle) = |\psi_2\rangle \otimes |\psi_2\rangle$$

The operator U is unitary and so it preservers the scalar product

$$(\langle\psi_2| \otimes \langle\psi_2|)(|\psi_1\rangle \otimes |\psi_1\rangle) = (\langle\psi_2| \otimes \langle\phi|)(|\psi_1\rangle \otimes |\phi\rangle)$$
$$(\langle\psi_2|\psi_1\rangle)^2 = \langle\psi_2|\psi_1\rangle \langle\phi|\phi\rangle$$

Since the state $|\phi\rangle$ is normalized, one obtains

$$\langle\psi_2|\psi_1\rangle (\langle\psi_2|\psi_1\rangle - 1) = 0$$

By the zero-product property of the complex numbers, it must be that either $\langle\psi_2|\psi_1\rangle = 0$ or $\langle\psi_2|\psi_1\rangle = 1$. In general, the Cauchy–Schwarz inequality holds

$$|\langle\psi_2|\psi_1\rangle| \leq \||\psi_1\rangle\|\||\psi_2\rangle\| = 1$$

It is saturated when

$$|\psi_1\rangle = \mathrm{e}^{\mathrm{i}\alpha} |\psi_2\rangle$$

So the two states $|\psi_j\rangle$ are orthogonal or are the same. However, this cannot be the case for two arbitrary states: this is the contradiction. Therefore, a single unitary operator U cannot clone a general quantum state. □

In quantum computing, one is only able to clone the states belonging to a basis, but this is enough.

2 Entanglement

See Nielsen, §§1.3.6, 1.3.7, 2.3, 2.6, Mermin, §6.2.

2.1 Bell states

Most of the difference between classical probability theory and quantum mechanics comes from entanglement.

Recall that a state in the product space $\mathcal{H}_A \otimes \mathcal{H}_B$ is entangled if it cannot be written as a product

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$$

A simple example of a two-qubit entangled state is

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

In the interpretation where $|0\rangle = |\uparrow\rangle$ and $|1\rangle = |\downarrow\rangle$, this is the singlet state of two spin-half particles, the one with angular momentum zero: something very common in atomic physics and not too difficult to realize.

The entangled state $|\psi\rangle$ behaves weirdly and is the source of many paradoxes. First of all is the Einstein–Podolsky–Rosen (EPR) paradox. When measuring the first qubit, one obtains 0 or 1 with equal probability. The same happens when measuring instead the second qubit. However, if one measures the first qubit, and obtains 0 for example, then the state collapses to

$$|\psi\rangle \to |01\rangle$$

and if one measures the second qubit, one obtains 1 with certainty. The qubits are (anti-)correlated. If the two qubits are separated by a great distance, then the measurement of one of the two instantly affects the other. This violates the principle of locality (in particular Einstein realism or local realism) which states that an object is influenced directly only by its immediate surroundings. However, causality is still retained and no information can be sent faster than the speed of light. From the point of view of quantum mechanics, the observer of the

first qubit and the one of the second qubit just get randomly 0 or 1. The result of one observer is the opposite of the one of the other, but each observer does not know that and this information has to be transmitted in another sub-luminal method. Therefore, information cannot travel between them using the qubits only.

One may even define a basis (the Bell, or EPR, basis) of entangled states for two qubits

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

This is a basis, so one may perform measurements in such basis, according to the Born rule. Every vector can be written as

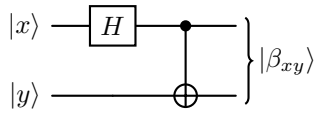
$$|\psi\rangle = \sum_{n,m=0}^{1} \alpha_{nm} |\beta_{nm}\rangle$$

and the probability is given by

$$P(\beta_{nm}) = |\alpha_{nm}|^2$$

This is called a measurement in the Bell basis, to distinguish it from a measurement in the standard (called computational) basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$.

The Bell states can be created easily with a unitary transformation from the computational basis. In the circuit language, one uses the Hadamard gate and a CNOT gate



The various inputs and outputs are

$$|00\rangle \xrightarrow{H} \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes |0\rangle \xrightarrow{\text{CNOT}} \frac{|00\rangle + |11\rangle}{\sqrt{2}} = |\beta_{00}\rangle$$
$$|01\rangle \xrightarrow{H} \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes |1\rangle \xrightarrow{\text{CNOT}} \frac{|01\rangle + |10\rangle}{\sqrt{2}} = |\beta_{01}\rangle$$
$$|10\rangle \xrightarrow{H} \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes |0\rangle \xrightarrow{\text{CNOT}} \frac{|00\rangle - |11\rangle}{\sqrt{2}} = |\beta_{10}\rangle$$
$$|11\rangle \xrightarrow{H} \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes |1\rangle \xrightarrow{\text{CNOT}} \frac{|01\rangle - |10\rangle}{\sqrt{2}} = |\beta_{11}\rangle$$

recalling that the Hadamard gates acts as

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

and the CNOT gate flips the second qubit if and only if the first qubit is $|1\rangle$.

2.2 Applications

There are many applications of entanglement in physics.

Superdense coding. If Observer A wants to send two classical bits of information to Observer B situated far away, then observer A has to necessarily send both bits. However, the same classical information can be encoded into two entangled qubits and only one of them has to be sent. Consider the entangled qubits

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} = |\beta_{00}\rangle$$

Suppose that the second qubit is in possession of observer B without modification. Observer A may perform an operation on their entangled qubit and then send that qubit to observer B who may perform a measurement to determine which Bell state the entangled pair is in. Observer A

can do four operations, each one corresponding to a different Bell state. To send $|\beta_{00}\rangle$, one may do nothing. To send $|\beta_{01}\rangle$, one applies X

$$01: \quad (X \otimes I)\,|\psi\rangle = \frac{|10\rangle + |01\rangle}{\sqrt{2}} = |\beta_{01}\rangle$$

To send $|\beta_{10}\rangle$, one applies Z to the first qubit

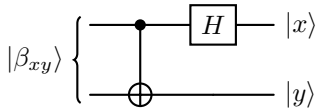
$$10: \quad (Z \otimes I)\,|\psi\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}} = |\beta_{10}\rangle$$

To send $|\beta_{11}\rangle$, one applies ZX

$$11: \quad (ZX \otimes I)\,|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}} = |\beta_{11}\rangle$$

Observer A may encode two classical bits $xy \to |\beta_{xy}\rangle$. Observer B receives the first qubit and may perform a measurement of the entangled pair in the Bell basis to read the value of xy.

Remark 2.1. This is the kind of operation that can be done with a circuit: observer A's operation is unitary and invertible. Then observer B can undo it and read the output with a measurement in the standard basis



since

$$(\mathrm{CNOT} \cdot H)^{-1} = H^{-1}\mathrm{CNOT}^{-1} = H \cdot \mathrm{CNOT}$$

Teleportation. Teleportation is the art of reconstructing a qubit faraway. Suppose observer A has a qubit $|\psi\rangle$. They do not know its state, but they want to deliver it to observer B using only classical channels. This is possible if both observers share an entangled pair of qubits in addition.

The problem looks difficult. Observer A cannot use quantum channels so they cannot physically send the qubit $|\psi\rangle$. They also cannot measure the qubit to read the coefficients

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

because the result is randomly 0 or 1 and no information is kept about α and β that can be sent classically. Also the state cannot be cloned. Even if it could be, and as such used to learn about α and β, these coefficients are continuous variables: they are infinitely long classical bits. It would take forever to send them over a classical channel with arbitrary precision.

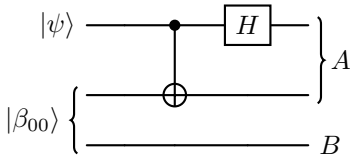
However, by sharing an entangled pair

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

the first qubit is observer A's and the second is observer B's. So observer A's state is

$$|\mathrm{ObA}\rangle = (\alpha\,|0\rangle + \beta\,|1\rangle) \otimes \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{\alpha}{\sqrt{2}}\,|0\rangle \otimes (|00\rangle + |11\rangle) + \frac{\beta}{\sqrt{2}}\,|1\rangle \otimes (|00\rangle + |11\rangle)$$

The first two qubits belong to A and the third to B. Observer A sends the qubits through a CNOT and then the first qubit through an Hadamard gate



to have

$$
\begin{aligned}
|\mathrm{ObA}\rangle &\overset{\text{CNOT}}{\to} \frac{1}{\sqrt{2}}\Big[\alpha\,|0\rangle\,(|00\rangle + |11\rangle) + \beta\,|1\rangle\,(|10\rangle + |01\rangle)\Big] \\
&\overset{H}{\to} \frac{1}{2}\Big[\alpha(|0\rangle + |1\rangle)(|00\rangle + |11\rangle) + \beta(|0\rangle - |1\rangle)(|10\rangle + |01\rangle)\Big] \\
&= \frac{1}{2}\Big[\,|00\rangle\,(\alpha\,|0\rangle + \beta\,|1\rangle) + |01\rangle\,(\alpha\,|1\rangle + \beta\,|0\rangle) + |10\rangle\,(\alpha\,|0\rangle - \beta\,|1\rangle) + |11\rangle\,(\alpha\,|1\rangle - \beta\,|0\rangle)\Big]
\end{aligned}
$$

where one has reorganized the terms dividing them between observer A's and B's (the parentheses). Now observer A measures their two qubits in the computational basis. According to this result, the qubit in the possession of B collapses to

Observer A	Observer B	Transformation
00	$\alpha\,\|0\rangle + \beta\,\|1\rangle$	I
01	$\alpha\,\|1\rangle + \beta\,\|0\rangle$	X
10	$\alpha\,\|0\rangle - \beta\,\|1\rangle$	Z
11	$\alpha\,\|1\rangle - \beta\,\|0\rangle$	ZX

Depending on what observer B gets after A measures, B can apply the corresponding transformation to obtain the qubit $|\psi\rangle$ observer A had. In particular, if A finds 00, then B has the original $|\psi\rangle$. If A gets 01, B has $\alpha\,|1\rangle + \beta\,|0\rangle$ which is not the original qubit $|\psi\rangle$, but can be recovered by applying X. Similarly, if A get 10 or 11, then B can apply Z or ZX to recover the qubit $|\psi\rangle$.

The task is fulfilled. The observer A measures xy and, via classical channels, tells B what the message xy is. With xy, B knows what operation to perform on their own qubit to recover $|\psi\rangle$. Only classical information has been transmitted.

Remark 2.2. Notice that there is no teleportation faster than the speed of light. Observer A needs to communicate classically (i.e. with sub-luminal methods) before B may reconstruct the qubit $|\psi\rangle$.

Remark 2.3. Notice also that the qubit $|\psi\rangle$ is not cloned. From the point of view of A, the qubit $|\psi\rangle$ is not cloned, but destroyed, since it needs to be measured.

[r] The EPR paradox lead people to construct other theories of quantum mechanics, especially deterministic. [r] It is possible to disprove the existence of a deterministic theory of quantum mechanics. The combination of realism and locality that need to be satisfied in a deterministic theory imposes restriction on the probabilities. For realism, one is talking about the existence of properties before observation. There are a set inequalities that can distinguish between classical and quantum probabilities.

Lecture 4

mar 16 ott 2024 16:30

2.3 Bell inequalities

Classical and quantum probabilities are physically different. Classical bits can be correlated, but classical correlation can be experimentally distinguished from quantum correlation.

Einstein and others were convinced that the behaviour of quantum mechanics is due to the ignorane of a more fundamental deterministic theory. Observables should assume values independently of measurement (realism) and the physics happening in a place can only affect the nearby physics (locality). Many hidden-variables theories were proposed where the result of a measurement of an observable A is determined by two values (a, λ): the result a and an experimentally inaccessible variable λ. To know the full theory, one needs to measure (a, λ). Due to ignorance, one may only give a probability $P(a, \lambda)$ of obtaining the result a. This is a description with a classical probability. It simply parametrizes ignorance.

For just one qubit or spin, it is straightforward to provide such a hidden-variable theory, a classical probability distribution covering all weirdness of quantum mechanics. For a single qubit

in a reference state, e.g. $|0\rangle$ without loss of generality, one may measure the general observable $\boldsymbol{\sigma}\cdot\mathbf{n}$, with $\|\mathbf{n}\|=1$. The point on the Bloch sphere corresponding to the vector $\mathbf{n}$

$$|\mathbf{n}\rangle = \cos\frac{\theta}{2}\,|0\rangle + \mathrm{e}^{\mathrm{i}\phi}\sin\frac{\theta}{2}\,|1\rangle$$

is an eigenstate of the spin along $\mathbf{n}$

$$(\boldsymbol{\sigma}\cdot\mathbf{n})\,|\mathbf{n}\rangle = |\mathbf{n}\rangle$$

and that the state

$$|-\mathbf{n}\rangle = |\theta\to\pi-\theta,\,\phi\to\phi+\pi\rangle = \sin\frac{\theta}{2}\,|0\rangle - \mathrm{e}^{\mathrm{i}\phi}\cos\frac{\theta}{2}\,|1\rangle$$

is also an eigenstate

$$(\boldsymbol{\sigma}\cdot\mathbf{n})\,|-\mathbf{n}\rangle = -\,|-\mathbf{n}\rangle$$

The spin

$$\mathbf{S} = \frac{\hbar}{2}\boldsymbol{\sigma}$$

is always $\pm\hbar/2$ along any axis. But then the quantum mechanical probability distribution for $\boldsymbol{\sigma}\cdot\mathbf{n}$ starting from the state $|0\rangle$ is

$$P(\boldsymbol{\sigma}\cdot\mathbf{n}=1) = |\langle\mathbf{n}|0\rangle|^2 = \cos^2\frac{\theta}{2}\,,\quad P(\boldsymbol{\sigma}\cdot\mathbf{n}=-1) = |\langle-\mathbf{n}|0\rangle|^2 = \sin^2\frac{\theta}{2}$$

A classical distribution that gives the same result is based on a variable $a=\pm1$ and a hidden random variable λ with uniform distribution in $[0,1]$. From physics one knows that, when one measures the spin along the direction $\mathbf{n}$ for a very complicated deterministic theory, one ignores the result of λ [r]? to have

$$\begin{cases} |\uparrow\rangle\,, & 0\le\lambda\le\cos^2\frac{\theta}{2} \\ |\downarrow\rangle\,, & \cos^2\frac{\theta}{2}<\lambda\le 1 \end{cases}$$

Therefore, one obtains the same result as the quantum theory.

Where classical and quantum theories differ is with entanglement. Bell showed this with his inequalities. Many others have been written afterwards. This course discusses the Clauser–Horne–Shimony–Holt (CHSH) inequality which is the one used in experiments.

Observers A and B are given a correlated pair of bits. Observer A can measure two observables a and a' each with possible result ±1. Observer B can measure b and b' with possible outcomes ±1. The two observers measure simultaneously and choose randomly whether to measure the primed or non-primed observable. The two cannot communicate. In a hidden variable theory, there would be a probability distribution $P(a,a',b,b')$ and each variable existing in a given state ±1 even before the measurement. The probability distribution is only due to ignorance. Consider

$$C = (a+a')b + (a-a')b'$$

Since all variables exist with values ±1, the term $a+a'$ is either 0 or ±2 while $a-a'$ is the opposite. In any case $C=\pm2$. The mean value of C is

$$\begin{aligned} |\langle C\rangle| &= \Big|\sum_{aa'bb'=\pm1} P(a,a',b,b')[(a+a')b+(a-a')b]\Big| \\ &\le \sum_{aa'bb'=\pm1} P(a,a',b,b')|(a+a')b+(a-a')b| = \langle|C|\rangle \\ &= 2\sum_{aa'bb'=\pm1} P(a,a',b,b') = 2 \end{aligned}$$

This is the CHSH inequality

$$\boxed{|\langle C\rangle| \le \langle|C|\rangle = 2}$$

valid for all classical correlations.

If the two observers share a quantum correlation, an entangled qubit, for example

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

with the first qubit given to A and the second to B, they can measure the spin along chosen directions

$$a = \boldsymbol{\sigma}\cdot\mathbf{a}\,, \quad a' = \boldsymbol{\sigma}\cdot\mathbf{a}'\,, \quad b = \boldsymbol{\sigma}\cdot\mathbf{b}\,, \quad b' = \boldsymbol{\sigma}\cdot\mathbf{b}'$$

where $\mathbf{a}$, $\mathbf{a}'$, $\mathbf{b}$ and $\mathbf{b}'$ are unit vectors. It is straightforward to show that the quantum mechanical mean value is

$$\langle\psi|\,(\boldsymbol{\sigma}\cdot\mathbf{c})\otimes(\boldsymbol{\sigma}\cdot\mathbf{d})\,|\psi\rangle = -\mathbf{c}\cdot\mathbf{d} = -\cos\theta_{cd}$$

where θ_{cd} is the angle between the two unit vectors $\mathbf{c}$ and $\mathbf{d}$. If the two observers choose their axis to be separated by 45° going counter-clockwise in the sequence a', b, a, b', then

$$\begin{aligned}\langle C\rangle &= \langle\psi|\,(ab + a'b + ab' - a'b')\,|\psi\rangle = -\sum \cos(\text{relative angle})\\ &= -\frac{1}{\sqrt{2}}\,[1+1+1-(-1)] = -\frac{4}{\sqrt{2}} = -2\sqrt{2}\end{aligned}$$

This violates the CHSH inequality

$$|\langle C\rangle| = 2\sqrt{2} > 2$$

One can show through the Tsirelson's bound that $2\sqrt{2}$ is the maximum possible violation. See Nielsen, problem 2.3 and Preskill, §4.3.

Experiments have been done that support the violation. Physics is quantum. The most famous experiment (Aspect '82[1][2][3]) still had loopholes:

1. locality, a signal could have gone from A to B, the two must be separated by a space-like distance;
2. detection, the experiment used photons, the detectors may not have been perfect and failed to detect photons, perhaps the failure is also described by hidden variable theory.

Experiments that evaded one of the two loopholes above were made and in 2015 an experiment evading both was done. There remains the possibility that hidden-variable theories involve also observers which are not able to perform independent measurements. Experiments were proposed and done where the observers decide what to measure by looking at distant independent fluctuations of brightness of well-separated stars, but the conspiracy is on the size of the universe.

Exercise 2.4. Show that

$$\langle\psi|\,(\boldsymbol{\sigma}\cdot\mathbf{c})(\boldsymbol{\sigma}\cdot\mathbf{d})\,|\psi\rangle = -\cos\theta$$

Solution. Notice that $|\psi\rangle$ is a singlet state and has angular momentum zero

$$(\boldsymbol{\sigma}\otimes I + I\otimes\boldsymbol{\sigma})\,|\psi\rangle = 0$$

Then

$$\begin{aligned}\langle\psi|\,(\boldsymbol{\sigma}\cdot\mathbf{c})\otimes(\boldsymbol{\sigma}\cdot\mathbf{d})\,|\psi\rangle &= -\,\langle\psi|\,(\boldsymbol{\sigma}\cdot\mathbf{c})(\boldsymbol{\sigma}\cdot\mathbf{d})\otimes I\,|\psi\rangle\\ &= -\frac{1}{2}\Big[(\langle 01| - \langle 10|)(\boldsymbol{\sigma}\cdot\mathbf{c})(\boldsymbol{\sigma}\cdot\mathbf{d})\otimes I(|01\rangle - |10\rangle)\Big]\\ &= -\frac{1}{2}\Big[\,\langle 0|\,(\boldsymbol{\sigma}\cdot\mathbf{c})(\boldsymbol{\sigma}\cdot\mathbf{d})\,|0\rangle + \langle 1|\,(\boldsymbol{\sigma}\cdot\mathbf{c})(\boldsymbol{\sigma}\cdot\mathbf{d})\,|1\rangle\,\Big]\\ &= -\frac{1}{2}\,\mathrm{Tr}[(\boldsymbol{\sigma}\cdot\mathbf{c})(\boldsymbol{\sigma}\cdot\mathbf{d})]\\ &= -\frac{1}{2}\cdot 2\mathbf{c}\cdot\mathbf{d} = -\mathbf{c}\cdot\mathbf{d}\end{aligned}$$

where one uses

$$\sigma_i\sigma_j c_i d_j = c_i d_i\,, \quad \sigma_i\sigma_j = \delta_{ij}I + \mathrm{i}\varepsilon_{ijk}\sigma_k\,, \quad \mathrm{Tr}\,\sigma_i = 0$$

[1] A. Aspect, P. Grangier, G. Roger. "Experimental Tests of Realistic Local Theories via Bell's Theorem", in Phys. Rev. Lett., vol. 47, pp. 460–463, 1981.

[2] A. Aspect, P. Grangier, G. Roger. "Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities", in Phys. Rev. Lett., vol. 49, pp. 91–94, 1982.

[3] A. Aspect, J. Dalibard, G. Roger. "Experimental Test of Bell's Inequalities Using Time-Varying Analyzers", in Phys. Rev. Lett., vol. 49, pp. 1804–1807, 1982

2.4 Quantum encryption

Even before the first algorithm devised to break RSA encryption (Shor's) was written, it was pointed out that quantum mechanics can give a secure basis for the exchange of secret message. Sending quantum states (e.g. polarized photons) along optical fibers in a protected way is not so difficult.

Classical cryptography. The two observers A and B can have an unbreakable code if they have identical sequences of random bits, S. Observer A takes the message, encodes it in a sequence of bits M and creates the bitwise sum modulo two (i.e. XOR) $S \oplus M$ which is a random sequence of bits with no information anymore (this is valid if one does not know S). This sum is sent to observer B who recovers the message by adding S in the same manner

$$S \oplus (S \oplus M) = (S \oplus S) \oplus M = 0 \oplus M = M$$

The sequence S is a one-time codepad (OTP). It can only be used once. In fact, if an eavesdropper E intercepts the two messages, M_1 and M_2, encoded both with S

$$S \oplus M_1\,, \quad S \oplus M_2$$

by adding them bitwise

$$(S \oplus M_1) \oplus (S \oplus M_2) = M_1 \oplus M_2$$

can recover the bitwise sum of the original messages. With a bit of guesswork and letter frequency analysis, E can recover part of the information.

Therefore, one needs to share a new random code S every time a message is sent with the risk of being intercepted.

Protocol BB84. Quantum mechanics may help solve the problem of creating a new random code. The quantum version, developed by Bennett and Brassard in 1984 (BB84), is the following. Observer A sends randomly two types of qubits: type C qubits $|0\rangle$ or $|1\rangle$, or type H qubits $|+\rangle$ or $|-\rangle$. The type and the state are both randomly chosen. Observer B receives the qubits, identifies them by the sequence in which they arrive and makes the choice of randomly measuring the result in the computational or Hadamard basis (measuring in the Hadamard basis is the same as applying the Hadamard gate and them measuring in the computational basis).

After B's measurements, A tells them, over an insecure channel, the type of each qubit, but not the state of each. For a given qubit, if B chooses the same basis for the measurement, then they know that the result of the measurement is the same that A wanted to send. If the measurement is in a different basis, B gets a random result, uncorrelated with A's. Observer B tells A over an insecure channel, which qubits were measured in the same basis and they both discard the others. What remains is a random series of 0s and 1s that can be used as a codepad.

Notice that if B does not immediately measure the qubits and instead considers waiting for A to tell them the bases used, then both of them do not waste qubits. However, storing qubits is complicated and it is better to measure them on arrival.

The best an eavesdropper E can do is make measurements while in transit, but they do not know which basis to use. The eavesdropper either gets lucky with the correct result or randomly gets a state disturbing the initial qubit. Observer B can recognize this discrepancy during the comparison with A. Though this is possible in half of the cases due to randomness. In fact if A sends $|0\rangle$, E may measure with $|\pm\rangle$ and obtain $|+\rangle$, but the B measures again in $|0, 1\rangle$ and may agree or disagree with A.

The two observers can know if the eavesdropper is listening by further comparing a fraction of the results. The previously agreeing results would now disagree a fourth of the time (half due to E and half due to B). By comparison, A and B would know if someone is monitoring the exchange. If the disagreement is small, the two observers could also decide to proceed by discarding the different qubits. They could also choose which fraction to compare (since it has to be discarded) to be reasonably sure.

Quantum non-demolition. One wonders if the eavesdropper E may use non-demolishing measurements: measure the state and restore it to a previous state. This is not possible for reasons similar to the no-cloning theorem. Let the four possibles states be

$$|\phi_\mu\rangle = (|0\rangle, |1\rangle, |+\rangle, |-\rangle)$$

and $|\Phi\rangle$ be a state belonging to E. The eavesdropper E would like to perform a unitary operation

$$U(|\phi_\mu\rangle \otimes |\Phi\rangle) = |\phi_\mu\rangle \otimes |\Phi_\mu\rangle$$

with four different $|\Phi_\mu\rangle$ that can be distinguished. If E measures in the base $|\Phi_\mu\rangle$, they would be able to read the value of μ corresponding to the state sent by A and then pass on the untouched state $|\phi_\mu\rangle$. However, the operator U preserves scalar products, so

$$\begin{aligned}(\langle\phi_\mu| \otimes \langle\Phi|)(|\phi_\nu\rangle \otimes |\Phi\rangle) &= (\langle\phi_\mu| \otimes \langle\Phi_\mu|)(|\phi_\nu\rangle \otimes |\Phi_\nu\rangle)\\ \langle\phi_\mu|\phi_\nu\rangle \langle\Phi|\Phi\rangle &= \langle\phi_\mu|\phi_\nu\rangle \langle\Phi_\mu|\Phi_\nu\rangle\\ \langle\phi_\mu|\phi_\nu\rangle &= \langle\phi_\mu|\phi_\nu\rangle \langle\Phi_\mu|\Phi_\nu\rangle\end{aligned}$$

Since $\langle\phi_\mu|\phi_\nu\rangle \neq 0$ for $\mu\nu = 02, 03, 12, 13$ then

$$\langle\Phi_\mu|\Phi_\nu\rangle = 1$$

but the product of normalized states can be 1 only when they are the same state so

$$|\Phi_0\rangle = |\Phi_1\rangle = |\Phi_2\rangle = |\Phi_3\rangle$$

and the eavesdropper fails.

Lecture 5

lun 21 ott 2024 16:30

3 Quantum computing

See Nielsen, §§1.4.2, 1.4.3, 1.4.4, Mermin, §§2.1, 2.2, 2.4, 3.4, 3.5, 3.7, 3.10, 4.1, 4.2.

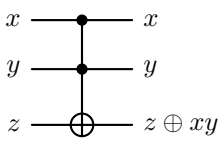
3.1 Quantum and classical circuits

One has already seen examples and differences between classical and quantum circuits.

Reversibility. One difference is reversibility. Quantum circuits are reversible because gates are unitary

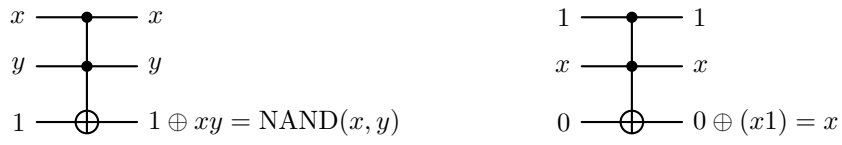
If U is unitary, so is its inverse U^{-1}. In general, classical circuits are not reversible. The NOT gate is reversible, but others are not, for example the AND and OR gates

However, all classical computations can be written in a reversible way using multiple-bit gates. For example, the Toffoli gate is



where the product xy is the AND operation. This gate flips the third bit if and only if the first two are both 1. It is reversible because it is just a permutation of (xyz) which is invertible.

All classical computations can be done with NAND and FANOUT (and enough bits). The Toffoli gate can simulate them both



Continuity. Another difference is that quantum computing gates are continuous. For example, consider one-qubit gates. The unitary operator U is a 2×2 matrix $UU^\dagger = U^\dagger U = I$. This type of matrices from the group U(2). All unitary matrices U can be written in terms of a Hermitian matrix

$$U = \mathrm{e}^{-\mathrm{i}H}$$

in fact

$$U^\dagger U = \mathrm{e}^{\mathrm{i}H^\dagger}\mathrm{e}^{-\mathrm{i}H} = \mathrm{e}^{\mathrm{i}H}\mathrm{e}^{-\mathrm{i}H} = I$$

The matrix H is of the form

$$H = x_0 I + \sum_{i=1}^{3} x_i \sigma_i\,, \quad x_j \in \mathbb{R}$$

This is the most general 2×2 Hermitian matrix. There are four real continuous parameters. Writing $\mathbf{x} = \|\mathbf{x}\|\mathbf{n}$, with $\mathbf{n}$ a unit vector, and defining $\lambda \equiv 2\|\mathbf{x}\|$, $\alpha = -x_0$ one finds

$$H = -\alpha I + \frac{\lambda}{2}\mathbf{n}\cdot\boldsymbol{\sigma}$$

So that the most general unitary operator is

$$U = \mathrm{e}^{-\mathrm{i}H} = \mathrm{e}^{\mathrm{i}\alpha} R_{\mathbf{n}}(\lambda)$$

where one recalls

$$R_{\mathbf{n}}(\lambda) = \exp\left[-\mathrm{i}\frac{\lambda}{2}\mathbf{n}\cdot\boldsymbol{\sigma}\right] = I\cos\frac{\lambda}{2} - \mathrm{i}(\mathbf{n}\cdot\boldsymbol{\sigma})\sin\frac{\lambda}{2}$$

This matrix $R_{\mathbf{n}}(\lambda)$ is interesting: in quantum mechanics it implements rotations of an angle λ around the direction $\mathbf{n}$ on the space of a spin-half particle. For example, it acts by rotating a state represented by a point on the Bloch sphere (See Nielsen, exercise 4.6). This is not proven in this course, but one may check the explicit form of $R_{\mathbf{n}}(\lambda)$. Notice that

$$(\mathbf{n}\cdot\boldsymbol{\sigma})^2 = \sigma_i\sigma_j n_i n_j = n_i n_j \frac{\sigma_i\sigma_j + \sigma_j\sigma_i}{2} = n_i n_j \delta_{ij} I = \|n\|^2 I = I$$

therefore, if $A^2 = I$ it follows

$$\begin{aligned}
\mathrm{e}^{-\mathrm{i}At} &= \sum_{n=0}^{\infty} \frac{(-\mathrm{i}At)^n}{n!} = \sum_{n=0}^{\infty} \frac{(-\mathrm{i}At)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(-\mathrm{i}At)^{2n+1}}{(2n+1)!} \\
&= I\sum_{n=0}^{\infty}(-1)^n \frac{t^{2n}}{(2n)!} - \mathrm{i}A\sum_{n=0}^{\infty}(-1)^n \frac{t^{2n+1}}{(2n+1)!} \\
&= I\cos t - \mathrm{i}A\sin t
\end{aligned}$$

then let $t = \lambda/2$ and $A = \mathbf{n}\cdot\boldsymbol{\sigma}$. Notice that rearranging the terms is justified because the series is absolutely convergent otherwise the Riemann series theorem holds.

The general unitary matrix U depends on an overall phase α and a unit vector $\mathbf{n}$: four real parameters. Since all rotation can be obtained by combining cleverly three elementary ones (z, then x, then z) one may write

$$U = \mathrm{e}^{\mathrm{i}\alpha}\begin{bmatrix} \mathrm{e}^{-\mathrm{i}\frac{\beta}{2}} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\frac{\beta}{2}} \end{bmatrix}\begin{bmatrix} \cos\frac{\gamma}{2} & -\sin\frac{\gamma}{2} \\ \sin\frac{\gamma}{2} & \cos\frac{\gamma}{2} \end{bmatrix}\begin{bmatrix} \mathrm{e}^{-\mathrm{i}\frac{\delta}{2}} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\frac{\delta}{2}} \end{bmatrix} = \mathrm{e}^{\mathrm{i}\alpha} R_z(\beta) R_x(\gamma) R_z(\delta)$$

with four real parameters α, β, γ and δ. Multiple qubits are messier: there are 2^n states and the unitary matrices $U \in \mathrm{U}(2^n)$ contain 2^{2n} real parameters.

Exercise 3.1. Count the number of parameters in the most general Hermitian matrix.

3.2 Universality

One can show that all classical computations can be done using only NAND and FANOUT (plus lost of ancilla bits) or just using Toffoli gates. The Toffoli gate forms a universal set of gates.

One would like to find the quantum analogue. It is important to know because unitary matrices are continuous matrices. It could be difficult to experimentally control 2^{2n} parameters of a $\mathrm{U}(2^n)$ n-qubit gate. It is also difficult to work with more than 2 qubits simultaneously.

The reuslt (see Nielsen, §4 for proof) is that Hadamard, phase (S), $\pi/8$ (T) and CNOT gates are a universal set of gates. Recall that $T^4 = S^2 = Z = \sigma_3$. The T gate is also called $\pi/8$ because

$$T = \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\frac{\pi}{4}} \end{bmatrix} = \mathrm{e}^{\mathrm{i}\frac{\pi}{8}} \begin{bmatrix} \mathrm{e}^{-\mathrm{i}\frac{\pi}{8}} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\frac{\pi}{8}} \end{bmatrix}$$

it introduces the phases $\pm\pi/8$. The system is redundant since $T^2 = S$, but it is customary to keep both S and T because they are important for fault-tolerant computations.

In the proof, one notices that the generic unitary operator on n qubits $U \in \mathrm{U}(2^n)$ can be written as a product of 2^{2n} 2-qubit unitary operators. Then any 4×4 unitary matrix U can be written as products of 1-qubit operators and CNOTs. For 4×4 operator and 1-qubit operator one should understand an operator U of dimension $2^n \times 2^n$ that acts only on 2 and 1 qubits respectively. The qubits do not need to be adjacent to each other, because they can be swapped with little extra cost (polynomial in n). In total, one needs an order of $O(n^2 2^{2n})$ single (i.e. 1-qubit) and CNOT gates. This is extremely inefficient, some gates are difficult to construct. It is part of the art in quantum computing to create algorithms that run efficiently.

Single qubit gates are still continuous and infinite in number. It is impossible to get an arbitrary continuous object with just discrete ones. The best one can do is approximate it, allowing for small errors. One would like to quantify such errors. One can measure the difference between two unitary matrices U and V with

$$\text{error} = \max_{|\psi\rangle} \|(U - V)\,|\psi\rangle\| \equiv E(U, V)$$

Then one can show that, given $U \in \mathrm{U}(2)$, there exists a string of Hs and Ts such that the error is arbitrarily small

$$E(U, \cdots H^{n_1} T^{n_2} H^{n_3} T^{n_4} \cdots) < \varepsilon$$

One may wonder how efficiently this can be done. The Solovay–Kitaev theorem (see Nielse, appendix 3) shows that one needs $O(\ln^c(1/\varepsilon))$ gates H and T with $c \approx 2$. So a circuit with m CNOT and single gates can be replaced with a circuit with $O(m \ln^c(m/\varepsilon))$ elements of the discrete set to accuracy ε. It is only a logarithmic loss and one can do all 2-qubits operations.