Theory of Quantum Information and Quantum Computing

Maso*

November 7, 2024

Contents

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

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 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

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

Applying phase invariance, one may eliminate a phase

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

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

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

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

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

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

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

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

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

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

$$P_{a_n} = |n\rangle\langle n|$$

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 $|0\rangle$ and $|1\rangle$ are two eigenvectors of σ_3:

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

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

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

The eigenbasis of σ_2 is given by

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

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 $|\psi\rangle$ is an eigenvector

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

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

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

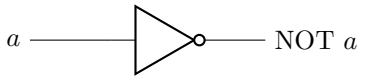
is conserved. The fact that $U(t)$ is unitary follows from the Schrödinger equation

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

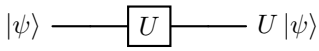
this holds since the Schrödinger equation also implies

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

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

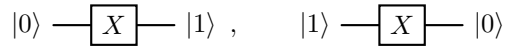
$$I\,, \quad X\,, \quad Z\,, \quad XZ$$

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

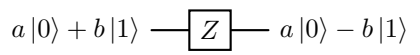
$$X|0\rangle = |1\rangle\,, \quad X|1\rangle = |0\rangle$$



while Z flips signs

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

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



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

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

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.

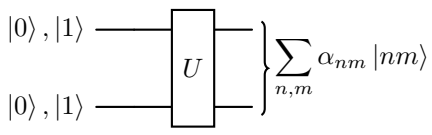
Example 1.11. The following state is separable

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

while the following states are entangled

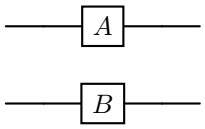
$$\frac{|00\rangle + |11\rangle}{\sqrt{2}}\,, \quad \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

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

U:



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

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

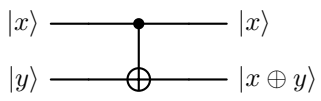
As for vectors, few 4×4 unitary matrices U are of the form $A\otimes B$. The other are actually more important for quantum computing: matrices that act on multiple qubits and not on individual ones independently. In classical computing there are several classical gates (e.g. AND, OR, XOR, etc) that cannot 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

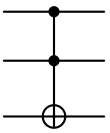


where $x, y = 0, 1$ and $\oplus$ denotes the XOR or, equivalently, the sum mod 2. In fact, if $x = 0$ nothing happens, but if $x = 1$ then

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

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

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

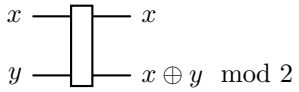
With many copies of the initial state $|\psi\rangle$, one may perform may measurements and extract the values of $|\alpha|^2$ and $|\beta|^2$.

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

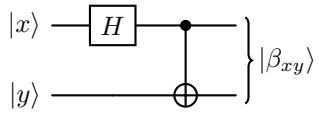
$$|\psi\rangle = \sum_{n,m=0}^{1} \alpha_{nm} |\beta_{nm}\rangle$$

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

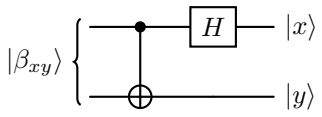
$$10: \quad (Z\otimes I)|\psi\rangle = \frac{|00\rangle-|11\rangle}{\sqrt{2}} = |\beta_{10}\rangle$$

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

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

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.

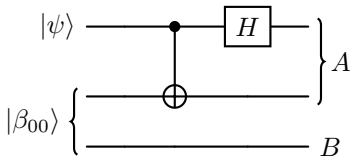
However, by sharing an entangled pair

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

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

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

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

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

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 $a = \pm 1$ 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 $C = \pm 2$. The mean value of C is

$$\begin{aligned} |\langle C \rangle| &= \left| \sum_{aa'bb'=\pm 1} P(a, a', b, b')[(a + a')b + (a - a')b] \right| \\ &\leq \sum_{aa'bb'=\pm 1} P(a, a', b, b')|(a + a')b + (a - a')b| = \langle |C| \rangle \\ &= 2 \sum_{aa'bb'=\pm 1} P(a, a', b, b') = 2 \end{aligned}$$

This is the CHSH inequality

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

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] 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

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

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

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

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

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 algorithms

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.

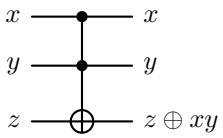
3.1 Quantum and classical circuits

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

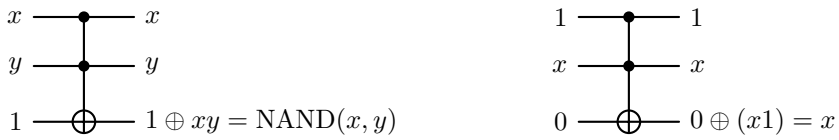
If U is unitary, so is its inverse U^{-1}. In general, classical circuits are not reversible. The NOT gate is reversible, but others are not, for example the AND and OR gates

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 result (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 Nielsen, 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.

3.3 Quantum parallelism

To encode a number x with a classical bit description

$$x = x_{n-1} x_{n-2} \cdots x_0 \,, \quad x_i = \pm 1 \,, \quad i = 0, \ldots, n-1$$

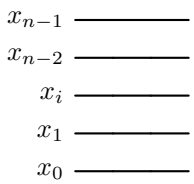
equivalent to

$$x = 2^{n-1} x_{n-1} + 2^{n-2} x_{n-2} + \cdots + x_0 = \sum_{i=0}^{n-1} 2^i x_i$$

for example $x = 1010_2 = 10$, one needs n qubits put in the state

$$|x\rangle = |x_{n-1}\rangle \otimes |x_{n-2}\rangle \otimes \cdots \otimes |x_0\rangle$$

or in circuit language with Mermin's ordering convention



In this way one describes all integers with $0 \leq x < N = 2^n$. The states $|x\rangle$ are a basis in the Hilbert space $\mathcal{H}$ of dimension $\dim \mathcal{H} = N = 2^n$ and are called the computational basis.

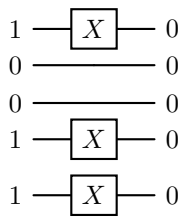
One may wonder how the state $|x\rangle$ can be achieved. If one starts with the state $|\psi\rangle$, one can make a measurement in the computational basis and obtain a random set of 0s and 1s. One may apply single gates X to flip the bits and put them in the desired form. Suppose one would like to start with

$$|0\rangle = |0\rangle \otimes \cdots \otimes |0\rangle$$

for a system of 5 qubits. Suppose that, by measuring the state $|\psi\rangle$, one finds

$$|1\rangle \otimes |0\rangle \otimes |0\rangle \otimes |1\rangle \otimes |1\rangle$$

To get to $|0\rangle$ one may apply



In quantum mechanics, measurements also serve as a means of preparation of a state: like the Stern–Gerlach experiment. Unless stated otherwise, it is assumed that the initial state of the computation is

$$|0\rangle = \bigotimes |0\rangle = |0\rangle^{\otimes n}$$

One may perform operations, for example write a circuit that computes a function

$$f(x) : \{0,1\}^{\otimes n} \to \{0,1\} \,, \quad x \mapsto f(x) = 0, 1$$

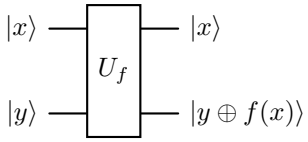
This should be done case by case, but all one may compute classically can also be computed, with similar techniques, quantum mechanically.

One important aspect is the use of a data register (or query) where x is stored and a target (or output) register where the result y is recorded

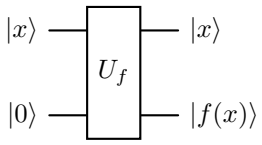
$$|x, y\rangle$$

One typically constructs a circuit that computes

$$|x, y\rangle \to |x, y \oplus f(x)\rangle$$



By starting with $y = 0$, one can read the value of $f(x)$ in the target register



The operator U_f is its own inverse

$$|x, y\rangle \to |x, y \oplus f(x)\rangle \to |x, y \oplus f(x) \oplus f(x)\rangle = |x, y\rangle$$

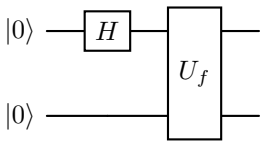
and also unitary because it just permutes basis states.

The interest of quantum mechanics is that one may start with the state $|x\rangle$ not in the computation basis, but in a superposition (which cannot be done in classical computing). Consider $n = 1$ and start with

$$|x, y\rangle = |x\rangle \otimes |y\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes |0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle)$$

This can be done by applying first an Hadamard gate

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



Then the result is

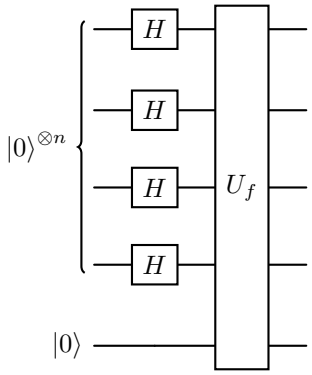
$$\frac{1}{\sqrt{2}}\Big[\,|0, f(0)\rangle + |1, f(1)\rangle\,\Big]$$

This state contains a superposition of the results $f(0)$ and $f(1)$, and contains information about both. One "evaluates" both $f(0)$ and $f(1)$ simultaneously: this is quantum parallelism.

The same can be done with multiple qubits by applying many Hadamard gates $H^{\otimes n}$ to the data register

$$\begin{aligned}
H^{\otimes n}\,|0\rangle^{\otimes n} &= (H \otimes \cdots \otimes H)\,|0\rangle \otimes \cdots \otimes |0\rangle = H\,|0\rangle \otimes \cdots \otimes H\,|0\rangle \\
&= \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \cdots \otimes \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\
&= \frac{1}{2^{\frac{n}{2}}}\Big[\,|00\cdots0\rangle + |10\cdots0\rangle + \text{all combinations of 0s and 1s}\Big] \\
&= \frac{1}{2^{\frac{n}{2}}}\sum_{x=0}^{2^n-1}|x\rangle
\end{aligned}$$

The number x can be thought of as a collection of digits or as an integer less than 2^n. The above sum is a superposition of all the numbers $0 \le x < N = 2^n$ with equal probability. Applying the operator U_f to it



gives

$$|\psi\rangle = \frac{1}{2^{\frac{n}{2}}} \sum_x |x, f(x)\rangle$$

All the values of $f(x)$ are computed in a single step. Beware: in quantum mechanics, having the state $|\psi\rangle$ is not sufficient to compute all the values of $f(x)$. A measurement gives a random x_0 and $f(x_0)$: just one random value. Moreover, the state $|\psi\rangle$ cannot be cloned, so one cannot repeat the measurements. However, one can extract some partial information about the function f that is classically not accessible.

3.4 Deutsch's algorithm

Consider a function

$$f(x) : \{0, 1\} \to \{0, 1\}$$

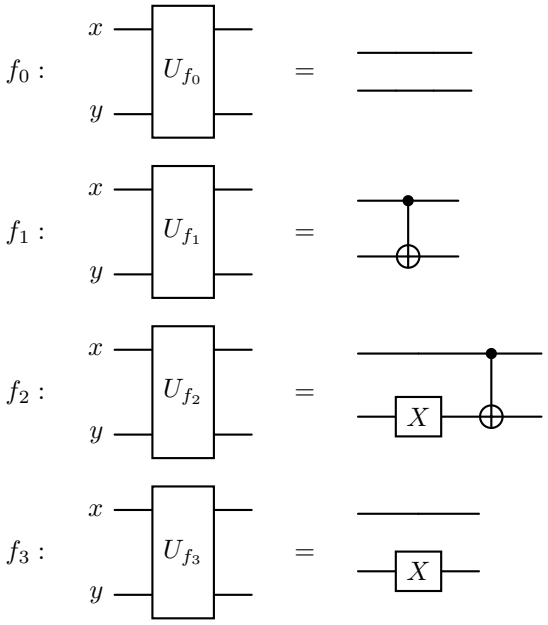
There are four such functions

$f(x)$	$x = 0$	$x = 1$
f_0	0	0
f_1	0	1
f_2	1	0
f_3	1	1

All of them can be implemented on a classical or quantum computer. On the latter, one would like to implement

$$U_f \, |x, y\rangle = |x, y \oplus f(x)\rangle$$

This can be done as follows



An example of the above circuits can be seen for f_2. Let $x = 0$, then y is flipped by X, but not the CNOT

$$y \to y \oplus f_2(0) = y \oplus 1$$

Let $x = 1$, then y is flipped twice, first by X then by CNOT, so it remains unchanged

$$y \to y \oplus f_2(1) = y \oplus 0 = y$$

Similarly, all the other functions f_j.

Lecture 6

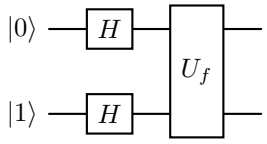
mer 23 ott 2024 16:30

Consider a circuit with an unknown f_j, an oracle[4]. In classical computing, in order to know which function is inside, one needs to compute the values $f(0)$ and $f(1)$ using the circuit: two queries are needed.

Deutsch's algorithm computes if $f(0) = f(1)$ or not. A classical algorithm needs two evaluations, but Deutsch's algorithm is quantum and gives extra information in a single computation. The algorithm makes use of a trick: it combines quantum parallelism with superposition in the register qubit. Starting from

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} = (H \otimes H)\,|01\rangle$$

where the first fraction corresponds to the parallelism, while the second is the superposition (which is also called interference), one applies U_f. Equivalently, one starts with $|01\rangle$ and applies



[4] A black box which is able to solve certain problems in a single operation.

The interference is useful because

$$|x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \overset{U_f}{\to} |x\rangle \otimes \frac{|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle}{\sqrt{2}} = \begin{cases} |x\rangle \otimes \dfrac{|0\rangle - |1\rangle}{\sqrt{2}}, & f(x) = 0 \\ |x\rangle \otimes \dfrac{|1\rangle - |0\rangle}{\sqrt{2}}, & f(x) = 1 \end{cases}$$

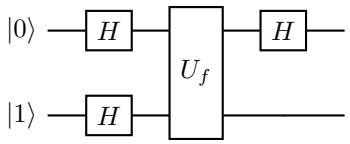
$$= (-1)^{f(x)} |x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

so that

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \to \frac{(-1)^{f(0)} |0\rangle + (-1)^{f(1)} |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

$$= \begin{cases} \dfrac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \dfrac{|0\rangle - |1\rangle}{\sqrt{2}} (-1)^{f(0)}, & f(0) = f(1) \\ \dfrac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes \dfrac{|0\rangle - |1\rangle}{\sqrt{2}} (-1)^{f(0)}, & f(0) \neq f(1) \end{cases}$$

Applying another Hadamard gate on the first qubit



gives the result

$$|01\rangle \to \begin{cases} (-1)^{f(0)} |0\rangle \otimes \dfrac{|0\rangle - |1\rangle}{\sqrt{2}}, & f(0) = f(1) \\ (-1)^{f(0)} |1\rangle \otimes \dfrac{|0\rangle - |1\rangle}{\sqrt{2}}, & f(0) \neq f(1) \end{cases}$$

By measuring the first qubit, one learns whether $f(0) = f(1)$ or $f(0) \neq f(1)$ with a single computation, which contrasts the classical case that requires at least two computations to know if the two outputs are the same. Since there are four functions f_j, one needs at least two queries to be able to extract the same information as with classical means. By giving up some information about the problem, one can perform less computations.

3.5 Other simple algorithms

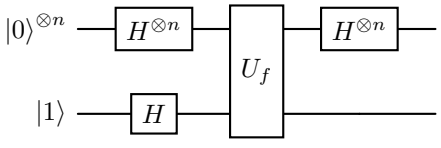
There are many other quantum algorithms that are more efficient than classical ones. In each, one learns some information about some function.

Deutsch–Jozsa algorithm. The Deutsch–Jozsa algorithm is the generalization of Deutsch's algorithm made by allowing multi-qubit functions with result 0 or 1

$$f(x) : \{0,1\}^{\otimes n} \to \{0,1\}$$

Consider a function $f(x)$ either constant or balanced (i.e. half of inputs are mapped to 0 and the other to 1). The algorithm is made to discern the two cases. The input x is one of 2^n possible combinations of n bits. In the classical case, one needs $2^{n-1}+1$ evaluations of the function $f(x)$ to determine whether the function is constant or balanced. In fact, after 2^{n-1} (i.e. half of the bits) one may get all 0s and still not know. The next computation discerns the two cases: if one obtains 0 then the function is constant, otherwise it is balanced. In quantum computing, one just needs one evaluation of the function f.

The algorithm is the same as Deutsch's



In particular

- Using $H^{\otimes n}$, one transforms
$$|0\rangle^{\otimes n} \to \sum_x |x\rangle$$
while using H one transforms
$$|1\rangle \to \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$
Therefore, before applying the oracle U_f, one has
$$(H^{\otimes n} \otimes H)\,|0\rangle^{\otimes n} \otimes |1\rangle = \sum_x \frac{|x\rangle}{2^{\frac{n}{2}}} \frac{|0\rangle - |1\rangle}{\sqrt{2}} = |\psi\rangle$$

- As before, by applying U_f on every $|x\rangle$ one has
$$|x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}} \to (-1)^{f(x)} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$
so one finds
$$U_f : |\psi\rangle \mapsto \sum_x (-1)^{f(x)} \frac{|x\rangle}{2^{\frac{n}{2}}} \frac{|0\rangle - |1\rangle}{\sqrt{2}} = |\phi\rangle$$

- Finally, one applies the last Hadamard $H^{\otimes n}$. For $n = 1$, one has
$$H\,|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}\,, \quad H\,|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}} \iff H\,|x\rangle = \sum_{z=0}^{1} \frac{(-1)^{xz}}{\sqrt{2}} |z\rangle$$
Thus
$$H^{\otimes n}\,|x\rangle = H^{\otimes n}\,|x_{n-1}, \ldots, x_0\rangle = \sum_{z_0 \cdots z_{n-1}} \frac{(-1)^{x_0 z_0 + \cdots + x_{n-1} z_{n-1}}}{2^{\frac{n}{2}}} |z_{n-1}, \ldots, z_0\rangle$$
$$= \sum_{z=0}^{2^n-1} \frac{(-1)^{x \cdot z}}{2^{\frac{n}{2}}} |z\rangle$$
where $x \cdot z$ is the bitwise product of x and z mod 2
$$x \cdot z \equiv x_0 z_0 + \cdots + x_{n-1} z_{n-1} \mod 2$$
Applying this to the previous point gives
$$H^{\otimes n}\,|\phi\rangle = \sum_{x,z=0}^{2^n-1} \frac{(-1)^{f(x)+x \cdot z}}{2^n} |z\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

- The data register contains all the possible states $|z\rangle$. However, the state $|z\rangle = |0\rangle^{\otimes n}$ appears with the coefficient
$$\sum_{x=0}^{2^n-1} \frac{(-1)^{f(x)}}{2^n} = \begin{cases} 1\,, & f(x) \text{ constant} \\ 0\,, & f(x) \text{ balanced} \end{cases}$$
This can be seen as follows. If $f(x)$ is constant, then there are 2^n copies of (-1) which cancel the 2^n in the denominator and the sum is 1. If $f(x)$ is balanced, then there are the same number of $+1$ and -1 which give zero.

Therefore, if $f(x)$ is constant, then the state is
$$|0\rangle^{\otimes n} \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$
because with amplitude 1 there is no room for other states. If $f(x)$ is balanced, then there is no state $|0\rangle^{\otimes n}$ in the sum. With just one measurement, one can learn if f is constant or balanced given that the result is $|0\rangle^{\otimes n}$ or otherwise.

Notice that one does not evaluate f, but learns some of its properties, not all of them and it is unclear which ones. There is yet no algorithm for each property.

This quantum algorithm has an exponential speed up: from $2^{n-1} + 1$ queries to 1. Though, there are statistical classical algorithms that are polynomial, so the speed up is minimal.

Bernstein–Vazirani algorithm. Consider the function

$$f(x) = a \cdot x\,, \quad 0 \le a, x < 2^n$$

which computes the bitwise inner product of a and x

$$a \cdot x = a_0 x_0 + \cdots + a_{n-1} x_{n-1} \mod 2$$

One would like to find the value of a. Classically, one needs n bits and computations to find every component a_j. With computing computing, one needs just one evaluation of the oracle U_f.

Consider the same circuit as before. It produces

$$\sum_{x,z=0}^{2^n-1} \frac{(-1)^{f(x)+x\cdot z}}{2^n} |z\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} = \sum_{x,z=0}^{2^n-1} \frac{(-1)^{x\cdot(a+z)}}{2^n} |z\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

The relevant sum is

$$\begin{aligned}\sum_x (-1)^{x\cdot(a+z)} &= \sum_{x_0\cdots x_{n-1}} (-1)^{x_0(a_0+z_0)} \cdots (-1)^{x_{n-1}(a_{n-1}+z_{n-1})} \\ &= \prod_{j=0}^{n-1} \Big[\sum_{x_j=0}^{1} (-1)^{(a_j+z_j)x_j} \Big]\end{aligned}$$

There are two cases:

- if $a_j + z_j = 0 \mod 2$, equivalently $z \equiv a \mod 2$, then

$$\prod_{j=0}^{n-1} [(-1)^{0\cdot 0} + (-1)^{0\cdot 1}] = \prod_{j=0}^{n-1} 2 = 2^n \neq 0$$

- If $a_j + z_j = 1 \mod 2$, then

$$\prod_{j=0}^{n-1} [(-1)^{1\cdot 0} + (-1)^{1\cdot 1}] = \prod_{j=0}^{n-1} [1-1] = 0$$

In the first case, the result is

$$|a\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

and the parameter a is read by measuring the data register.

Remark 3.2. Notice that all these problems are academic and of little interest. The Deutsch–Jozsa algorithm is interesting because it is an exponential speed-up: from classical $O(2^{n-1})$ to quantum $O(1)$ query. However, there is a classical probabilistic algorithm that is only polynomial in n, so the quantum version gains little. The interest and speed-up resurface in Shor's and Grover's algorithms.

3.6 Quantum Fourier transform

The quantum Fourier transform is the unitary transformation on the space of n qubits acting as

$$U_{\mathrm{FT}} |x\rangle = \frac{1}{2^{\frac{n}{2}}} \sum_{y=0}^{2^n-1} \mathrm{e}^{2\pi \mathrm{i} \frac{xy}{2^n}} |y\rangle$$

where xy is ordinary multiplication. The quantum Fourier transformation is at the core of Shor's algorithm.

When acting on a generic state

$$|\psi\rangle = \sum_x \gamma(x) |x\rangle$$

the transformation is

$$U_{\rm FT}(|\psi\rangle) = \sum_{xy} \frac{\gamma(x)}{2^{\frac{n}{2}}} \mathrm{e}^{2\pi \mathrm{i} \frac{xy}{2^n}} |y\rangle \equiv \sum_y \hat{\gamma}(y) |y\rangle$$

The coefficients $\gamma(x)$ are transformed into their classical discrete Fourier counterparts

$$\hat{\gamma}(y) = \frac{1}{2^{\frac{n}{2}}} \sum_x \mathrm{e}^{2\pi \mathrm{i} \frac{xy}{2^n}} \gamma(x)$$

To compute classically the transformed coefficients $\hat{\gamma}$, one needs 2^{2n} operations. There is famous classical algorithm, the fast Fourier transform, that does it in $O(n2^n)$ operations. The quantum version requires a circuit with $O(n^2)$ elementary gates. However, from a single copy of the transformed vector $U_{\rm FT}(|\psi\rangle)$ one cannot extract all the classical transformed coefficients $\hat{\gamma}(y)$, but just a random pair $(y_0, \hat{\gamma}(y_0))$. The quantum Fourier transform can be used in many algorithms (e.g. Shor's, factoring, discrete logarithm).

In the following, the operator $U_{\rm FT}$ is constructed explicitly. Since

$$H^{\otimes n} |0\rangle^{\otimes n} = \frac{1}{2^{\frac{n}{2}}} \sum_y |y\rangle$$

the operator $U_{\rm FT}$ can be written as

$$U_{\rm FT} |x\rangle = \mathcal{Z}^x H^{\otimes n} |0\rangle^{\otimes n}$$

where $\mathcal{Z}$ acts as

$$\mathcal{Z} |y\rangle = \mathrm{e}^{2\pi \mathrm{i} \frac{y}{2^n}} |y\rangle$$

For $n = 1$, one has

$$\mathcal{Z} = Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} \mathrm{e}^{\mathrm{i}0} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\pi} \end{bmatrix} = \mathrm{e}^{\mathrm{i}\pi n}, \quad n = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

so that, if

$$|y\rangle = |y_{n-1}\rangle \otimes \cdots \otimes |y_0\rangle, \quad y = \sum_{i=0}^{n-1} 2^i y_i$$

and

$$\left[\sum_{i=0}^{n-1} 2^i n_i \right] |y\rangle = \left[\sum_{i=0}^{n-1} 2^i y_i \right] |y\rangle = y |y\rangle$$

where n_i is the matrix n acting on the i-th qubit, then

$$\mathcal{Z} |y\rangle = \exp\left[\frac{2\pi \mathrm{i}}{2^n} \sum_{j=0}^{n-1} 2^j y_j \right] |y\rangle = \mathrm{e}^{2\pi \mathrm{i} \frac{y}{2^n}} |y\rangle, \quad \mathcal{Z}^x |y\rangle = \exp\left[\frac{2\pi \mathrm{i}}{2^n} \sum_{i=0}^{n-1} 2^i x_i \sum_{j=0}^{n-1} 2^j n_j \right] |y\rangle$$

For simplicity, consider $n = 4$

$$\mathcal{Z}^x = \exp\left[\frac{2\pi \mathrm{i}}{16} (8x_3 + 4x_2 + 2x_1 + x_0)(8n_3 + 4n_2 + 2n_1 + n_0) \right]$$

The matrices n_j have eigenvalues 0 and 1, so on $|0\rangle$ and $|1\rangle$ one finds

$$\mathrm{e}^{2\pi \mathrm{i} n} = \begin{cases} 1 \\ \mathrm{e}^{2\pi \mathrm{i}} = 1 \end{cases} = I$$

In fact

$$\mathrm{e}^{2\pi \mathrm{i} n} = (\mathrm{e}^{\mathrm{i}\pi n})^2 = Z^2 = I$$

Acting on $|y\rangle$, the terms in the exponent of $\mathcal{Z}^x$ that are integer multiples of $2\pi \mathrm{i} n$ give 1. Therefore, one is left with

$$\mathcal{Z}^x = \exp\left\{ \mathrm{i}\pi \left[n_3 x_0 + n_2 \left(\frac{x_0}{2} + x_1 \right) + n_1 \left(\frac{x_0}{4} + \frac{x_1}{2} + x_2 \right) + n_0 \left(\frac{x_0}{8} + \frac{x_1}{4} + \frac{x_2}{2} + x_3 \right) \right] \right\}$$

At this point, one would like to commute each matrix n_j with the respective Hadamard gate H_j. One notices that

$$\mathrm{e}^{\mathrm{i}\pi x n} H\left|0\right\rangle = H\left|x\right\rangle$$

in fact

$$\begin{cases} H\left|0\right\rangle = H\left|0\right\rangle\ , & x=0 \\ \mathrm{e}^{\mathrm{i}\pi n} H\left|0\right\rangle = ZH\left|0\right\rangle = Z\left|+\right\rangle = \left|-\right\rangle = H\left|1\right\rangle\ , & x=1 \end{cases}$$

The terms within $\mathcal{Z}^x$ without a fraction can be moved out and the others can be reorganized

$$\begin{aligned} \mathcal{Z}^x H^{\otimes n}\left|0\right\rangle^{\otimes n} &= H_3\left|x_0\right\rangle \otimes \mathrm{e}^{\mathrm{i}\pi n_2 \frac{x_0}{2}} H_2\left|x_1\right\rangle \otimes \mathrm{e}^{\mathrm{i}\pi n_1(\frac{x_0}{4}+\frac{x_1}{2})} H_1\left|x_2\right\rangle \otimes \mathrm{e}^{\mathrm{i}\pi n_0(\frac{x_0}{8}+\frac{x_1}{4}+\frac{x_2}{2})} H_0\left|x_3\right\rangle \\ &= H_3 \mathrm{e}^{\mathrm{i}\pi n_2 \frac{x_0}{2}} H_2 \mathrm{e}^{\mathrm{i}\pi n_1(\frac{x_0}{4}+\frac{x_1}{2})} H_1 \mathrm{e}^{\mathrm{i}\pi n_0(\frac{x_0}{8}+\frac{x_1}{4}+\frac{x_2}{2})} H_0\left|x_0\right\rangle\left|x_1\right\rangle\left|x_2\right\rangle\left|x_3\right\rangle \end{aligned}$$

where one notices that n_i and H_j commute if $i \neq j$. Notice also that the order of the states on the right has been reversed. The state $\left|x_0\right\rangle\left|x_1\right\rangle\left|x_2\right\rangle\left|x_3\right\rangle$ is an eigenstate of the number operators n_3, n_2, n_1 and n_0 with respective eigenvalues[5] x_0, x_1, x_2 and x_3:

$$n_{n-i}\left|x_i\right\rangle = x_i\left|x_i\right\rangle$$

Therefore one may safely substitute x_i for n_{n-i} to have

$$U_{\mathrm{FT}} = H_3(\mathrm{e}^{\mathrm{i}\frac{\pi}{2}n_2 n_3} H_2)(\mathrm{e}^{\mathrm{i}\frac{\pi}{2}n_1 n_2}\mathrm{e}^{\mathrm{i}\frac{\pi}{4}n_1 n_3} H_1)(\mathrm{e}^{\mathrm{i}\frac{\pi}{2}n_0 n_1}\mathrm{e}^{\mathrm{i}\frac{\pi}{4}n_0 n_2}\mathrm{e}^{\mathrm{i}\frac{\pi}{8}n_0 n_3} H_0)P$$

where P is a permutation

$$P\left|x_3\right\rangle\left|x_2\right\rangle\left|x_1\right\rangle\left|x_0\right\rangle = \left|x_0\right\rangle\left|x_1\right\rangle\left|x_2\right\rangle\left|x_3\right\rangle$$

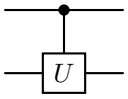
The operator U_{FT} is unitary. In fact, all operators with a matrix n are of the form

$$\exp\left[\mathrm{i}\pi \frac{n_i n_j}{2^{|i-j|}}\right]$$

where $|i-j|$ is the distance between the i-th qubit and the j-th. If the i-th qubit is $i = \left|0\right\rangle$ then $n_i = 0$ and the exponential is the identity. If $i = \left|1\right\rangle$ then $n_i = 1$ and the exponential is

$$\exp\left[\mathrm{i}\pi \frac{n_j}{2^{|i-j|}}\right]$$

So the exponential can be written as a controlled operator. One may introduce a controlled gate

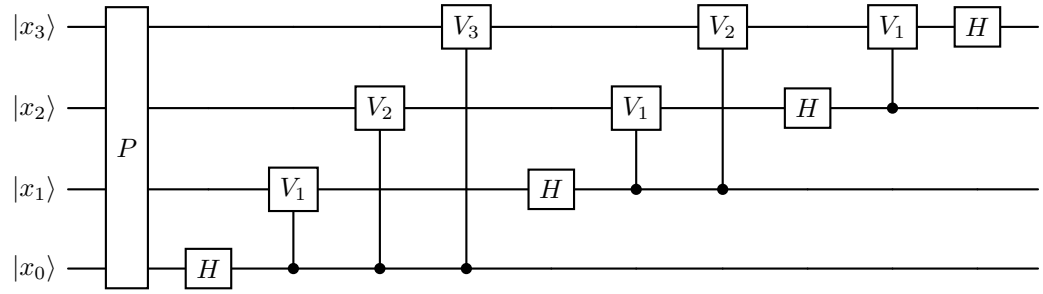


that does nothing if the upper qubit is $\left|0\right\rangle$, while applies U to the second qubit if the first is $\left|1\right\rangle$. The CNOT gate is $U = X$.

Let

$$V_k = \exp\left[\frac{\mathrm{i}\pi n}{2^k}\right]$$

then the quantum Fourier transform operator is



with a number of gates

$$\sum_{i=0}^{n-1}(n-i) = \sum_{i=1}^{n} i = \frac{1}{2}n(n+1) \sim O(n^2)$$

The operator P is linear in n, so the algorithm is still of the order $O(n^2)$ operations.

[5] The subscript of n_j corresponds to the position of the qubit in the sequence $\left|x_{n-1}\right\rangle \cdots \left|x_0\right\rangle$ and not specifically to $\left|x_j\right\rangle$ in any position. For example for $\left|x_1, x_3, x_2, x_0\right\rangle$, one has $n_0 \leftrightarrow x_0$, $n_1 \leftrightarrow x_2$, $n_2 \leftrightarrow x_3$, $n_3 \leftrightarrow x_1$.

Lecture 7

lun 28 ott 2024 16:30

3.7 Shor's algorithm — period finding

Consider a periodic bijective function acting on the integers

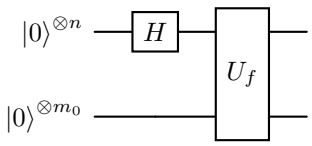
$$f : \mathbb{Z} \to \{0,1\}^{\otimes m_0}$$

One would like to find its period

$$r < N = 2^{n_0} \implies f(x) = f(x+r)$$

The best classical algorithm, the general number field sieve, requires exponential resources $O(\exp[n_0^{1/3}(\ln n_0)^{2/3}])$, while Shor's algorithm uses the quantum Fourier transform and has polynomial time complexity $O(n_0^2(\ln n_0)^2)$.

Consider a data register with n qubits and an output register containing m_0 qubits. The required qubits are of the order $n \sim 2n_0$ so that $2^n \sim N^2$. The circuit is



The two gates act as

$$U_f[(H^{\otimes n}|0\rangle^{\otimes n}) \otimes |0\rangle^{\otimes m_0}] = U_f\left(\sum_{x=0}^{2^n-1} \frac{1}{2^{\frac{n}{2}}}|x\rangle \otimes |0\rangle^{\otimes m_0}\right) = \frac{1}{2^{\frac{n}{2}}}\sum_{x=0}^{2^n-1}|x, f(x)\rangle$$

One measures the output register since the interest is on the values of $f(x)$. One obtains some value $f(x) = f_0$ and the data register is a sum of all the values $|x_i\rangle$ that satisfy $f(x_i) = f_0$. Since $2^n \sim N^2$ and $r < N$, one has many such values, at least of order N. Let m be the number of those values, then the data register state has collapsed into

$$|\psi\rangle_{\text{data}} = \frac{1}{\sqrt{m}}\sum_{k=0}^{m-1}|x_0 + kr\rangle$$

where x_0 is the smallest integer such that $f(x_0) = f_0$ and due to periodicity

$$f(x_0 + kr) = f(x_0) = f_0$$

Notice that, by the Born rule, after a measurement one has to normalize the data state $|\psi\rangle_{\text{data}}$. The number m is the smallest integer such that

$$x_0 + mr \geq 2^n \iff m = \left[\frac{2^n}{r}\right] \vee m = \left[\frac{2^n}{r}\right] + 1$$

where $[x]$ is the integer part of x. Since $r < N$ and $2^n \geq N^2$, then m is typically large. The state $|\psi\rangle_{\text{data}}$ is not useful. A measurement gives $x_0 + \bar{k}r$ with unknown x_0 and random $\bar{k}$. Two values $x_0 + k'r$ and $x_0 + \bar{k}r$ would help because their difference is proportional to r, but one cannot get two states in $|\psi\rangle_{\text{data}}$ with a single measurement and still no cloning is allowed. Repeating the setup gives a different x_0 and so the results are not useful.

However, a unitary transformation, the Fourier transform, can eliminate x_0

$$\begin{aligned} U_{\text{FT}}\left[\frac{1}{\sqrt{m}}\sum_{k=0}^{m-1}|x_0+kr\rangle\right] &= \frac{1}{2^{\frac{n}{2}}\sqrt{m}}\sum_{y=0}^{2^n-1}\sum_{k=0}^{m-1}\exp\left[2\pi\mathrm{i}\frac{(x_0+kr)}{2^n}y\right]|y\rangle \\ &= \sum_{y=0}^{2^n-1}\mathrm{e}^{\frac{2\pi\mathrm{i}}{2^n}x_0 y}\sum_{k=0}^{m-1}\frac{\mathrm{e}^{\frac{2\pi\mathrm{i}}{2^n}kry}}{2^{\frac{n}{2}}\sqrt{m}}|y\rangle \end{aligned}$$

If one measures the output register $|y\rangle$, one finds a particular value with probability

$$P(y) = \left| \mathrm{e}^{\frac{2\pi \mathrm{i}}{2^n} x_0 y} \sum_{k=0}^{m-1} \frac{\mathrm{e}^{\frac{2\pi \mathrm{i}}{2^n} kry}}{2^{\frac{n}{2}}\sqrt{m}} \right|^2 = \frac{1}{2^n m} \left| \sum_{k=0}^{m-1} \mathrm{e}^{\frac{2\pi \mathrm{i}}{2^n} kry} \right|^2$$

which is independent of x_0. If one plots the probability, one finds peaks at the values of y that are multiples of $2^n/r$: $y_j = j2^n/r$ for $j = 0, \ldots, r-1$. There are r peaks. The reason is that all the phases add up

$$\sum_{k=0}^{m-1} \mathrm{e}^{\frac{2\pi \mathrm{i}}{2^n} kry_j} = \sum_{k=0}^{m-1} \mathrm{e}^{2\pi \mathrm{i} kj} = \sum_{k=0}^{m-1} 1 = m$$

and the height of the peaks is

$$P(y_j) = \frac{m}{2^n}$$

which is about r^{-1} for large N since $m \sim 2^n/r$. With r peaks, each with height $P \sim r^{-1}$, one almost saturates the total probability. Indeed, for all other y, the phases tend to interfere and cancel in the sum over k. The more N is large, the sharper are the peaks on the values y_j. For $N \to \infty$, one can easily show that the distribution becomes a sum of Dirac delta functions.

Notice that y is a discrete variable and y_j as defined above may not be an integer. It turns out that $4/\pi^2 \approx 40\%$ of the distribution $P(y)$ is near one of the peaks y_j with a maximum error of $1/2$. One may study the probability $P(y)$ near $y = j2^n/r + \delta$. Using the geometric series one finds

$$P(y = j2^n/r + \delta) = P(\delta) = \frac{1}{2^n m} \left| \frac{\mathrm{e}^{2\pi \mathrm{i} mr\delta/2^n} - 1}{\mathrm{e}^{2\pi \mathrm{i} r\delta/2^n} - 1} \right|^2 = \frac{1}{2^n m} \frac{\sin^2(\pi mr\delta/2^n)}{\sin^2(\pi r\delta/2^n)}$$

Noting that

$$\delta \leq \frac{1}{2} \implies \pi\delta \leq \frac{\pi}{2}$$

also

$$x \geq \sin x \geq \frac{2}{\pi} x$$

and keeping in mind

$$\frac{r}{2^n} \sim \frac{1}{m} \ll 1\,, \quad \frac{mr}{2^n} \sim 1$$

then

$$P(\delta) \geq \frac{4}{\pi^2} \frac{1}{r}$$

Since there are r values, one finds

$$P(\delta) \geq \frac{4}{\pi^2} \approx 0.4053$$

Therefore, with 40% probability, one has a result y that satisfies

$$\left| y - j\frac{2^n}{r} \right| < \frac{1}{2} \iff \left| \frac{y}{2^n} - \frac{j}{r} \right| < \frac{1}{2^{n+1}}$$

and one finds r by comparing $y/2^n$ and j/r. In this case, one needs $n > n_0$ to be sure that j/r is unique. In fact, $y/2^n$ divides the interval $[0, 1]$ in 2^n subintervals. The rational number j/r is somewhere at a distance less than $1/2$ from some $y/2^n$. Two different rational numbers j_1/r_1 and j_2/r_2 differ by

$$\left| \frac{j_1}{r_1} - \frac{j_2}{r_2} \right| = \frac{|r_2 j_1 - r_1 j_2|}{r_1 r_2} \geq \frac{1}{N^2} = \frac{1}{2^n}\,, \quad r_1, r_2 < N$$

so they cannot be both near the same $y/2^n$. From number theory, a theorem states that if

$$\left| x - \frac{j}{r} \right| < \frac{1}{2r^2}$$

which is true since

$$\left| x - \frac{1}{r} \right| \leq \frac{1}{2 \cdot 2^n} = \frac{1}{2N^2} < \frac{1}{2r^2}\,, \quad r < N$$

then j/r appears as a partial fraction in the continued-fraction expansion of $x \in [0, 1)$

$$x = \cfrac{1}{x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cdots}}}$$

One needs to run the algorithm for the continued-fraction expansion (which a classical computer does efficiently), obtain the partial fractions, check the denominators and stop at the fraction with largest denominator below 2^{n_0} to find r. See Mermin, appendix K for two examples.

There are a few caveats to Shor's algorithm. In reality one finds

$$\frac{j_0}{r_0} = \frac{j}{r}$$

if j and r are not coprime. In such case, r is a multiple of r_0. If the coefficient between the two is small, it can be checked with a classical computer too, otherwise one may rerun the algorithm.

The algorithm can also be improved. If $r < N/2$, then the probability of finding a value near the peaks goes from 40% to 90%. One may also improve it by taking $n = 2n_0 + q$ with extra qubits.

3.8 Breaking RSA encryption

Shor's algorithm is useful to break the Rivest–Shamir–Adleman (RSA) encryption. The following is a superficial discussion.

Observer B has a large integer $N = pq$ which is the product of two large primes p and q, and has a number c with no factor in common with $(p-1)(q-1)$. The observer computes the inverse of c

$$cd = 1 \mod (p-1)(q-1)$$

This can be efficiently done with a classical algorithm, i.e. Euclid's. Observer B sends to Observer A only N and c. Observer A encodes a message a into the string

$$b = a^c \mod N$$

and sends it to B who decodes it with

$$a = b^d \mod N$$

This can be proven in number theory.

The public knows N, c and the encoded message, but not p nor q. Factoring $N = pq$ is a problem that takes exponential time, so the method is secure.

However, with a quantum computer, one is able to find the period of the function

$$f(x) = b^x \mod N$$

or the smallest r such that

$$f(x + r) = f(x)$$

or

$$b^r \mod N = 1$$

The period r is called the order of $b \mod N$ and one can prove that it exists and $r \leq N$. This can be done quantum mechanically with $O(\ln^2 N)$ operations. A quantum computer finds the period r such that $b^r = 1 \mod N$, while a classical computer finds d' such that $cd' = 1 \mod N$. Then one can prove that the message a is

$$a = b^{d'} \mod N$$

Order finding and factoring. There is no need to factor $N = pq$ if one is just interested in breaking RSA encryption. However, order finding is equivalent to factoring. Given $N = pq$, consider a number a that is coprime with N. The order is found using quantum computing

$$a^r = 1 \mod N$$

Its existence is guaranteed by Euler's theorem: if a, p and q are all coprime, then

$$a^{(p-1)(q-1)} = 1 \mod pq$$

Since $r = (p-1)(q-1) < N$ then r is the period of $f(x) = a^x \mod N$. The algorithm is inherently probabilistic and one is either lucky in the choice of a or has to repeat the process for another value of a. Suppose one finds the period r to be even. Let

$$x = a^{\frac{r}{2}} \mod N$$

One has

$$0 = a^r - 1 = x^2 - 1 = (x-1)(x+1) \mod N$$

At this point, $x - 1 \neq 0 \mod N$ because otherwise $a^{\frac{r}{2}} = 1 \mod N$ and the period is then $r/2$ instead of r. Suppose also that

$$x + 1 \neq 0 \mod N$$

Since $N = pq$ and $(x-1)(x+1)$ is divisible[6] by N, but neither $x-1$ nor $x+1$ are, then $x-1$ must be divisible by p and $x+1$ by q (or viceversa). However, the only divisors of N are p and q themselves, so p is the greatest common divisor of N and $x-1$, while q is the greatest common divisor of N and $x+1$. Finding the greatest common divisor is efficiently done using Euclid's algorithm.

There are other encryption procedures, for example Diffie–Hellman which uses the discrete logarithm, but Shor's algorithm can solve the discrete logarithm problem in polynomial time too.

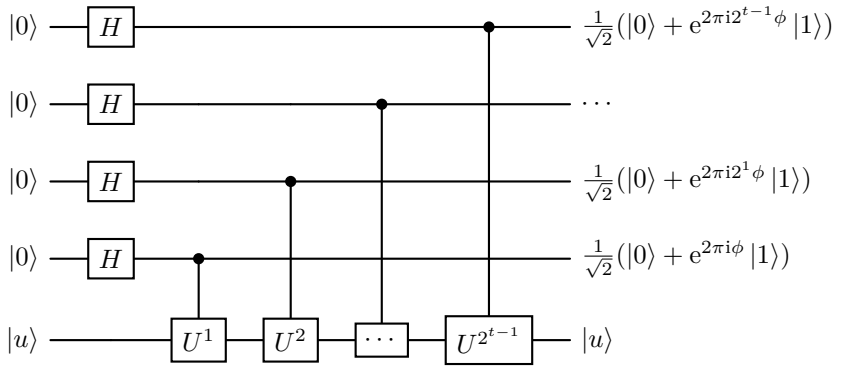
Lecture 8

mer 30 ott 2024 16:30

Variant of Shor's algorithm — phase estimation. See Nielsen, §§5.2, 5.3. Shor's algorithm can also be formulated as a phase estimation algorithm. The following is a way to estimate the eigenvalue (i.e. the phase) of a unitary operator U. Consider the operator

$$U|u\rangle = \mathrm{e}^{2\pi \mathrm{i}\phi}|u\rangle$$

and the following circuit with t qubits in the data register



One should understand the results on the right-hand side as coming from the operator U applied to its eigenvector $|u\rangle$ giving the corresponding eigenvalue. Since this eigenvalue is just a number, it is put in front of the control qubit $|1\rangle$. The data register is in the state

$$\frac{1}{2^{\frac{t}{2}}}(|0\rangle + \mathrm{e}^{2\pi \mathrm{i} 2^0 \phi}|1\rangle) \otimes (|0\rangle + \mathrm{e}^{2\pi \mathrm{i} 2^1 \phi}|1\rangle) \otimes \cdots \otimes (|0\rangle + \mathrm{e}^{2\pi \mathrm{i} 2^{t-1} \phi}|1\rangle) = \frac{1}{2^{\frac{t}{2}}} \sum_{k=0}^{2^t-1} \mathrm{e}^{2\pi \mathrm{i}\phi k}|k\rangle$$

[6] This is a consequence of $0 = (x-1)(x+1) \mod N$.

whose inverse quantum Fourier transform is exactly $|\phi\rangle$. If the phase is an integer, $0 \leq \phi < 2^t$, a quantum Fourier transform would give just the state $|\phi\rangle$. A measurement in the computational basis gives the value ϕ. Otherwise, it is an approximation and one can estimate the number t such that it works well. If the phase is not an integer, the quantum Fourier transform gives

$$\frac{1}{2^t} \sum_{k,j=0}^{2^t-1} \exp\left[2\pi \mathrm{i} k \left(\phi - \frac{j}{2^k}\right)\right] |j\rangle$$

If one parametrizes

$$\phi = \frac{j}{2^t} + \frac{\delta}{2^t}\,, \quad |\delta| \leq \frac{1}{2}$$

then the sum above is

$$\sum_j \left[\frac{1}{2^t} \sum_{k=0}^{2^t-1} \mathrm{e}^{2\pi \mathrm{i} \frac{\delta k}{2^t}}\right] |j\rangle$$

and the probability of measuring the state $|j\rangle$, which is the best integer approximation, is

$$P(j) = \frac{1}{2^{2t}} \left| \sum_{k=0}^{2^t-1} \mathrm{e}^{2\pi \mathrm{i} \frac{\delta k}{2^t}} \right|^2 \geq \frac{4}{\pi^2}$$

One should note that it is difficult to start with an eigenvector $|u\rangle$. If one begins with

$$|\psi\rangle = \sum_u c_u |u\rangle$$

the algorithm produces

$$\sum c_u |\phi_u\rangle$$

and a measurement gives just one eigenvalue ϕ_u.

Order finding can be expressed with the above formalism. One would like to find the period r such that $a^r = 1 \mod N$, for a and N coprime. The operator is

$$U|x\rangle = |ax \mod N\rangle$$

It implements a permutation so it is unitary. Its eigenvalues are

$$\mathrm{e}^{2\pi \mathrm{i} \frac{s}{r}}\,, \quad s = 0, \ldots, r-1$$

with eigenvectors

$$|u_j\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} \mathrm{e}^{-2\pi \mathrm{i} \frac{jk}{r}} \left|a^k \mod N\right\rangle$$

The target register can be initialized with

$$|1\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} |u_k\rangle$$

which is easier to construct than a particular eigenvector $|u\rangle$.

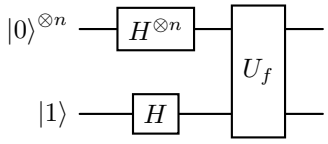
3.9 Grover's search algorithm

The other class of problems in quantum computing where one gains a speed-up is searching in a database. Consider N objects. One would like to search for a specific object a. Abstractly, one has a function $f(x)$ from $x = 0, \ldots, N-1$ to 0 or 1 such that

$$f(x) = \begin{cases} 1\,, & x = a \\ 0\,, & x \neq a \end{cases}$$

The database can be a text, an Amazon warehouse, or a mathematical problem. For example, suppose one knows that p is the sum of squares of two integers $p = x^2 + y^2$, one would like to find the two integers. For all x, one computes $\sqrt{p - x^2}$ until one finds an integer y. Classically, it takes $N/2$ searches (evaluating $f(x)$ for $x = 0, 1, 2, \dots$) before getting the right result with probability $1/2$. Quantum mechanically, it only takes $O(\sqrt{N})$ searches. This is not an exponential speed-up, but it is a speed-up when N is large.

Grover's algorithm works as follows. Consider the usual starting state with n initial qubits



$$\frac{1}{2^{\frac{n}{2}}} \sum_x (-1)^{f(x)} |x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

Notice that one starts with $H|1\rangle$ (equivalently $HX|0\rangle$) in the target register in order to compute phases

$$(-1)^{f(x)} = \begin{cases} -1\,, & x = a \\ +1\,, & x \neq a \end{cases}$$

The target register becomes irrelevant and is no longer used. The effect of applying the function on an element of the computational basis is the oracle

$$O|x\rangle = (-1)^{f(x)} |x\rangle = \begin{cases} |x\rangle\,, & x \neq a \\ -|x\rangle\,, & x = a \end{cases}$$

This is a reflection. If we split the Hilbert space $\mathcal{H}$ into the space spanned by $|a\rangle$ and orthogonal space

$$\{|a\rangle\}^{\perp} = \{x \in \mathcal{H} \mid \langle x|a\rangle = 0\}$$

the operator O does not act at all on the orthogonal space. The operator can also be written as

$$O = I - 2|a\rangle\langle a|$$

Indeed

$$O|a\rangle = |a\rangle - 2|a\rangle \langle a|a\rangle = |a\rangle - 2|a\rangle = -|a\rangle$$

and if $\langle x|a\rangle = 0$, then

$$O|x\rangle = |x\rangle - 2|a\rangle \langle a|x\rangle = |x\rangle$$

The initial computational state is the uniform superposition

$$|\phi\rangle = H^{\otimes n} |0\rangle^{\otimes n} = \frac{1}{2^{\frac{n}{2}}} \sum_x |x\rangle$$

which is the equal superposition state of all $|x\rangle$ with amplitude $1/2^{\frac{n}{2}}$. One may define another reflection operation G around $|\phi\rangle$:

$$G|\phi\rangle = |\phi\rangle$$

For orthogonal states $\langle x|\phi\rangle = 0$ gives

$$G|x\rangle = -|x\rangle$$

The operator's explicit form is then

$$G = 2|\phi\rangle\langle\phi| - I$$

notice the sign difference with respect to O. This operator can be constructed starting from a simpler one that reflects all states that are not $|0\rangle^{\otimes n}$

$$|0 \cdots 0\rangle \to |0 \cdots 0\rangle$$

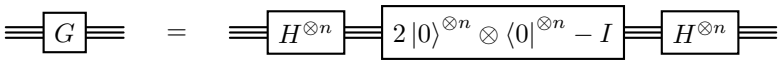
while, if at least one $x_i \neq 0$, it acts as

$$|x_1 \dots x_n\rangle \to -|x_1 \dots x_n\rangle$$

These two operations can be combined into the operator

$$2\left|0\right\rangle^{\otimes n}\otimes\left\langle 0\right|^{\otimes n}-I$$

which can be implemented on qubits (exercise: write a circuit). The operator G is then

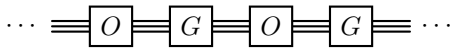


in fact

$$(H^{\otimes n})(2\left|0\right\rangle^{\otimes n}\left\langle 0\right|^{\otimes n}-I)(H^{\otimes n})=2\left|\phi\right\rangle\!\left\langle\phi\right|-I=G$$

since $\left|\phi\right\rangle=H^{\otimes n}\left|0\right\rangle^{\otimes n}$ and $H^2=I$.

Then, Grover's algorithm is the repeated application of the operator GO:



After $O(\sqrt{N})$ repetitions, the state is the object $\left|a\right\rangle$ with probability almost one.

Geometric interpretation. One may draw the operators O and G in the plane spanned by $\left|a\right\rangle$ and $\left|\phi\right\rangle$. Since the state $\left|a\right\rangle$ is a particular state within the superposition state $\left|\phi\right\rangle$, the two are almost perpendicular (supposing that n is large). Therefore, from the inner product, one notices that the angle θ between $\left|a\right\rangle^{\perp}$ and $\left|\phi\right\rangle$ is small:

$$\left\langle a|\phi\right\rangle=\frac{1}{2^{\frac{n}{2}}}$$

since $\left|\phi\right\rangle$ is equi-amplitude on all states and $N=2^n$ is large, so

$$\sin\theta=\cos\left(\frac{\pi}{2}-\theta\right)=\frac{1}{2^{\frac{n}{2}}}\implies\theta\sim\frac{1}{2^{\frac{n}{2}}}=\frac{1}{\sqrt{N}}.$$

The combined action of the operators O and G is a double reflection (i.e. a rotation) and rotates the state $\left|\phi\right\rangle$ by an angle 2θ to the state $\left|\phi'\right\rangle$. This state is at an angle of 3θ with respect to $\left|a\right\rangle^{\perp}$. Another iteration of GO rotates by a further 2θ so that the angle is now 5θ and so on. After r iterations, the angle is

$$2r\theta+\theta$$

and becomes $\pi/2$ (i.e. $\left|a\right\rangle$) with a good probability. If

$$2r\theta+\theta=\frac{\pi}{2}$$

then the number of iterations is

$$r=\frac{\pi}{4\theta}-\frac{1}{2}\approx\frac{\pi}{4}\theta^{-1}\approx\frac{\pi}{4}\sqrt{N}$$

After $O(\sqrt{N})$ iterations of the operators GO, one finds $\left|a\right\rangle$ with probability almost one. One may compute a by measuring the final state $(GO)^r\left|\phi\right\rangle$. With probability almost 1, one gets a. One may check it by computing $f(a)$ and see if it gives 1.

Even if $(GO)^r\left|\phi\right\rangle$ is not exactly $\left|a\right\rangle$, it is near

$$(GO)^r\left|\phi\right\rangle=\sqrt{1+\varepsilon^2}\left|a\right\rangle+\varepsilon\left|a\right\rangle^{\perp},\quad\varepsilon\ll 1$$

and a measurement almost always gives $\left|a\right\rangle$ and not $\left|a\right\rangle^{\perp}$. But, if one is unlucky and does not obtain a, one may run the algorithm again, maybe with few more iterations, and try again. After a limited number of tries, one finds a.

3.10 Another look at Grover's algorithm

There is another simple geometric interpretation of Grover's algorithm. Starting with a generic

$$|\psi\rangle = \sum_x \alpha_x |x\rangle$$

the operator O just flips the amplitude of the state with $x = a$:

$$\begin{cases} \alpha_a \to -\alpha_a \,, & x = a \\ \alpha_x \to \alpha_x \,, & x \neq a \end{cases}$$

The operator G is a reflection along the state $|\phi\rangle$ with equal amplitudes $\alpha_x = 1/\sqrt{N}$

$$|\phi\rangle = \frac{1}{\sqrt{N}} \sum_x |x\rangle = \frac{1}{2^{\frac{n}{2}}} \sum_x |x\rangle \,, \quad N = 2^n \,, \quad G = 2\,|\phi\rangle\langle\phi| - I$$

and its action on the state $|\psi\rangle$ is slightly more complicated

$$\begin{aligned} G\,|\psi\rangle &= 2\frac{1}{N} \sum_x |x\rangle \sum_y \langle y| \left[\sum_z \alpha_z |z\rangle \right] - \sum_x \alpha_x |x\rangle = \sum_x \left[2 \sum_y \frac{\alpha_y}{N} - \alpha_x \right] |x\rangle \\ &= \sum_x (2\langle\alpha\rangle - \alpha_x)\,|x\rangle \end{aligned}$$

At the first line one has used

$$\langle y|z\rangle = \delta_{yz}$$

At the second line, the average is

$$\langle\alpha\rangle = \frac{1}{N} \sum_y \alpha_y$$

Therefore, the operator G acts on each component as

$$\alpha_x \to 2\langle\alpha\rangle - \alpha_x$$

It is a reflection with respect to the average amplitude $\langle\alpha\rangle$.

Pictorially, in Grover's algorithm, one starts with equal amplitudes and flips the state $|a\rangle$ using the operator O

$$\alpha_a \to -\alpha_a .$$

so that the average $\langle\alpha\rangle = 1/\sqrt{N}$ is now smaller. Then one reflects all values with respect to this new average using G, so that α_a is amplified. Repeating the steps, all other states have smaller and smaller amplitudes, while the amplitude of the state $|a\rangle$ is amplified. One proceeds until one gets almost only $|a\rangle$. It works similarly for searches of multiple objects. Suppose one would like to find M objects in an N-dimensional database: one repeats the steps above where O acts on all such M objects simultaneously. This time one needs $\sqrt{N/M}$ runs to find one item.

Lecture 9

lun 04 nov 2024 16:30

4 Open systems

See Nielsen §§2.4, 8.2, 8.3, Preskill §3.4.

4.1 Density matrix

The systems discussed so far are closed (isolated). Real systems always interact with the environment: this leads to loss of coherence, loss of energy, etc. The correct formalism in this case is the density matrix, which describes a quantum system when, due to ignorance, one has just a probabilistic (in a classical, not quantum, sense) understanding of what is the state of the system. Real systems are open.

Consider a system that can be in one state of a collection $\{\psi_i\}$, but one only knows (due to ignorance) that this happens with (classical) probability p_i. One defines the density matrix[7]

$$\rho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \ , \quad \sum_i p_i = 1, \quad \langle\psi_i|\psi_i\rangle = 1$$

Notice that the states $|\psi_i\rangle$ are not necessarily orthogonal to each other. This is useful because, for a quantum state $|\psi_i\rangle$ and an observable A, the expectation value of A is

$$\langle A\rangle = \langle\psi_i|A|\psi_i\rangle$$

If one performs experiments, and one has partial knowledge of the system given by the probabilities p_i, then the measured expectation value is

$$\langle A\rangle = \sum_i p_i \, \langle\psi_i|A|\psi_i\rangle$$

This can be rewritten as

$$\boxed{\langle A\rangle = \mathrm{tr}(\rho A)}$$

in fact, using the cyclic property of the trace, one finds

$$\mathrm{tr}(\rho A) = \sum_i p_i \ \mathrm{tr}(|\psi_i\rangle\langle\psi_i| \, A) = \sum_i p_i \, \langle\psi_i|A|\psi_i\rangle = \langle A\rangle$$

Here, the trace means

$$A = \sum_n A_{nn} = \sum_n \langle n|A|n\rangle \ , \quad I = \sum_n |n\rangle\langle n|$$

with $\{|n\rangle\}$ being a complete orthonormal system; so the cyclic property is

$$\mathrm{tr}(|\psi\rangle\langle\psi| \, A) = \sum_n \langle n|\psi\rangle \ \langle\psi|A|n\rangle = \sum_n \langle\psi|A|n\rangle \, \langle n|\psi\rangle = \langle\psi|A|\psi\rangle$$

A quantum state in the old sense is called pure: the density matrix is a quantum state without the notion of classical probability p_i. The state $|\psi\rangle$ can be also written as a density matrix

$$\rho = |\psi\rangle\langle\psi|$$

An ensemble of quantum states in the new sense

$$\rho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i|$$

where at least two $p_i \neq 0$, is called a mixed state or mixture.

Example 4.1 (Statistical mechanics)**.** Perhaps the simplest example where one needs to use density matrices is statistical mechanics: a system with temperature T. If the possible energy states are $\{E_n, |n\rangle\}$, each can occur with classical probability

$$P(E_n) = \frac{1}{Z} \, \mathrm{e}^{-\frac{E_n}{k_\mathrm{B} T}}$$

where

$$Z = \sum_n \mathrm{e}^{-\frac{E_n}{k_\mathrm{B} T}} \implies \sum_n P(E_n) = 1$$

so that in the computations one uses the density matrix

$$\rho = \sum_n \frac{1}{Z} \, \mathrm{e}^{-\frac{E_n}{k_\mathrm{B} T}} \, |n\rangle\langle n|$$

From this, one defines the thermal vacuum expectation value at a temperature T as

$$\langle A\rangle_T = \mathrm{tr}(\rho A)$$

[7] Strictly speaking it is an operator.

Example 4.2 (Single qubit)**.** A pure state in $|0\rangle$ has density matrix

$$\rho = |0\rangle\langle 0| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

which is diagonal with just one non-zero entry. A mixture, $|0\rangle$ with probability p and $|1\rangle$ with probability $1-p$, has

$$\rho = p\,|0\rangle\langle 0| + (1-p)\,|1\rangle\langle 1| = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix}$$

which is still diagonal, but with more than one entry. The probability p can be, for example

$$p = \frac{1}{Z}\,\mathrm{e}^{-\frac{E_0}{k_\mathrm{B}T}}$$

or the situation in which one performed a measurement and sometimes forgot the result. The probability $p = 1$ (pure state) is the case with more information, while $p = \frac{1}{2}$ is the mixture with zero information possible: all states are equiprobable

$$\rho = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} = \frac{1}{2} I$$

Notice, however, that the density matrix ρ does not need to be diagonal, neither for pure states nor for mixed ones. Indeed, a pure state which is a superposition $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$ has matrix

$$\rho = |\psi\rangle\langle\psi| = \begin{bmatrix} a \\ b \end{bmatrix} \begin{bmatrix} a^* & b^* \end{bmatrix} = \begin{bmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{bmatrix}$$

where the off-diagonal terms characterize the interference (i.e. superposition), while the diagonal terms are the probabilities of obtaining a state of the basis when measuring the state $|\psi\rangle$. For a mixture, the states $|\psi_i\rangle$ need not be orthogonal, so a mixture of $|0\rangle$ and $|+\rangle$ in equal parts is

$$\rho = \frac{1}{2}\,|0\rangle\langle 0| + \frac{1}{2}\,|+\rangle\langle +| = \frac{1}{2}\begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix} + \frac{1}{4}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$$

Properties. Some general properties of density matrices are:

- The density matrix ρ is hermitian and positive.
  - ◇ Hermicity is a consequence of
    $$(|\psi\rangle\langle\psi|)^\dagger = |\psi\rangle\langle\psi|$$
    and p_i being real.
  - ◇ Positivity means
    $$\langle\phi|\,\rho\,|\phi\rangle \geq 0\,, \quad \forall\,|\phi\rangle$$
    in fact
    $$\langle\phi|\,\rho\,|\phi\rangle = \sum_i p_i\,\langle\phi|\psi_i\rangle\,\langle\psi_i|\phi\rangle = \sum_i p_i |\langle\phi|\psi_i\rangle|^2 \geq 0$$
- The trace is one, $\operatorname{tr}\rho = 1$.
  - ◇ This can be seen as follows
    $$\operatorname{tr}\left[\sum_i p_i\,|\psi_i\rangle\langle\psi_i|\right] = \sum_i p_i \operatorname{tr}(|\psi_i\rangle\langle\psi_i|) = \sum_i p_i = 1$$
- It holds $\operatorname{tr}\rho^2 \leq 1$ and, for pure states only, $\operatorname{tr}\rho^2 = 1$.

◇ The density matrix ρ is hermitian, so it can be diagonalized:

$$\rho = \sum_n p_n \, |n\rangle\langle n| \iff \rho = \mathrm{diag}(p_1, \ldots, p_n)$$

It holds

$$\mathrm{tr}\, \rho = 1 \,, \quad \sum_n p_n = 1 \implies \mathrm{tr}\, \rho^2 = \sum_n p_n^2$$

Since ρ is positive

$$p_n \geq 0 \,, \quad \sum_n p_n = 1 \implies 0 \leq p_n \leq 1$$

Therefore

$$\mathrm{tr}\, \rho^2 = \sum_n p_n^2 \leq \sum_n p_i = 1$$

The inequality is saturated if $p_n^2 = p_n$. However, since $p_n^2 \leq p_n$ and $0 \leq p_n \leq 1$ then it can only be $p_n = 1$. If, for one n, it holds $p_n = 1$, then all the others are zero $p_n = 0$ and

$$\rho = |n\rangle\langle n|$$

is pure.

The proof of the third property implies that the density matrix ρ has no unique interpretation. From a mixture of generic non-orthogonal states

$$\rho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i|$$

one can diagonalize the matrix to obtain a mixture of another set of orthogonal states

$$\rho = \sum_n p_n \, |n\rangle\langle n|$$

The complete information about the system is comprised of the probabilities p_i and the states $|\psi_i\rangle$.

Example 4.3 (Single qubit)**.** The most general density matrix ρ, being hermitian, is

$$\rho = a_0 I + \mathbf{a} \cdot \boldsymbol{\sigma}$$

The property $\mathrm{tr}\, \rho = 1$ implies $a_0 = 1/2$:

$$\rho = \frac{1}{2}(I + \mathbf{r} \cdot \boldsymbol{\sigma}) \,, \quad \mathbf{r} = \frac{1}{2}\mathbf{a}$$

The eigenvalues are

$$\lambda = \frac{1}{2} \pm \frac{|\mathbf{r}|}{2}$$

since one may rewrite $\mathbf{r} \cdot \boldsymbol{\sigma} = |\mathbf{r}|(\mathbf{n} \cdot \boldsymbol{\sigma})$, where $\mathbf{n}$ is a unit vector, and one notices that $(\mathbf{n} \cdot \boldsymbol{\sigma})^2 = I$ so its eigenvalues are ± 1. From positivity, one has $|\mathbf{r}| \leq 1$. From this it follows

$$\mathrm{tr}\, \rho^2 = \frac{1}{4} \mathrm{tr}\left[I + 2\mathbf{r} \cdot \boldsymbol{\sigma} + (\mathbf{r} \cdot \boldsymbol{\sigma})^2\right] = \frac{2 + 2|\mathbf{r}|^2}{4} = \frac{1 + |\mathbf{r}|^2}{2} \leq 1$$

and for a pure state it holds

$$\mathrm{tr}\, \rho^2 = 1 \iff |\mathbf{r}| = 1$$

For pure states, the vector $\mathbf{r}$ lives on the unit sphere. One may extend the Bloch sphere to the interior by placing the generic matrix ρ, using $\mathbf{r}$, into the ball of radius 1. The points on the surface are the Bloch sphere and are pure states. The points in the interior are mixed states.

The points on the z-axis interpolate between $|0\rangle$ and $|1\rangle$:

$$\rho = \frac{I + r\sigma_3}{2} = \frac{1}{2}\begin{bmatrix} 1 + r & 0 \\ 0 & 1 - r \end{bmatrix} \,, \quad \mathbf{r} = (0, 0, r) \,, \quad -1 \leq r \leq 1$$

These points are the ensemble with $|0\rangle$ with probability p and $|1\rangle$ with probability $1-p$ (the position on the z-axis is related to the probability). The points on the x-axis between $|+\rangle$ and $|-\rangle$ are

$$\rho = \frac{I + r\sigma_1}{2} = \frac{1}{2}\begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}, \quad \mathbf{r} = (r, 0, 0), \quad -1 \leq r \leq 1$$

which is a mixture of $|+\rangle$ and $|-\rangle$. The center of the ball is common to all segments and it is $\rho = I/2$, the most uncertain state. Since $I = \sum_n |n\rangle\langle n|$, for any orthonormal basis it holds

$$\rho = \frac{1}{2}I = \begin{cases} \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \\ \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -| \end{cases}$$

where one may notice again that the interpretation of the density matrix ρ is not unique.

4.2 Tracing over subsystems

The most useful application of the density matrix formalism concerns the case when the system is separable

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

and one chooses to perform an experiment only on the subspace A. The ignorance of the subspace B leads to a description of A in terms of mixed states, even if one has started with full knowledge of $\mathcal{H}$ (i.e. the pure states).

Consider a system in the Hilbert space $\mathcal{H}$ described by the density matrix ρ and an observable acting only in A: that is, an operator of the form

$$O = O_A \otimes I_B$$

Given a basis $\{|nm\rangle = |n\rangle_A \otimes |m\rangle_B\}$ in $\mathcal{H}$, the expectation value of the operator O is

$$\begin{aligned} \langle O\rangle = \mathrm{tr}(O\rho) &= \sum_{nm} \langle nm|O\rho|nm\rangle = \sum_{nmn'm'} \langle nm|O|n'm'\rangle \, \langle n'm'|\rho|nm\rangle \\ &= \sum_{nmn'm'} \langle n|O_A|n'\rangle \, \langle m|I_B|m'\rangle \, \langle n'm'|\rho|nm\rangle = \sum_{nmn'} \langle n|O_A|n'\rangle \, \langle n'm|\rho|nm\rangle \\ &= \sum_{nn'} \langle n|O_A|n'\rangle \, \langle n'|\rho_A|n\rangle = \sum_{n} \langle n|O_A\rho_A|n\rangle = \mathrm{tr}_A(O_A\rho_A) \end{aligned}$$

At the second line, one notices that

$$\langle m|I_B|m'\rangle = \delta_{mm'}$$

At the third line, one defines an operator ρ_A on $\mathcal{H}_A$ that has matrix elements

$$\langle n'|\rho_A|n\rangle = \sum_m \langle n'm|\rho|nm\rangle$$

This operator can be formally written as a partial trace over the degrees of freedom of B

$$\rho_A = \sum_m {}_B\langle m|\rho|m\rangle_B \equiv \mathrm{tr}_B\, \rho$$

by formally dropping $\langle n'|$ and $|n\rangle$. The trace $\mathrm{tr}_B\, \rho$ is also an operator. The operator ρ_A acts as an effective density matrix for the system A:

$$\langle O\rangle = \mathrm{tr}_A(O_A\rho_A)$$

All operations are done in the subspace $\mathcal{H}_A$ and one may forget about $\mathcal{H}_B$.

If one starts with a factorized system

$$\rho = \rho_A \otimes \rho_B$$

meaning things happen independently in A and B, then obviously

$$\operatorname{tr}_B \rho = \operatorname{tr}_B \rho_A \otimes \rho_B = \sum_m {}_B\langle m|\rho_A \otimes \rho_B|m\rangle_B = \rho_A \sum_m \langle m|\rho_B|m\rangle = \rho_A \operatorname{tr} \rho_B = \rho_A$$

However if one starts with a generic density matrix ρ, then by tracing over B one loses information, in general. This happens even if one starts with a pure state in $\mathcal{H}_A \otimes \mathcal{H}_B$ when such state is entangled.

Example 4.4 (Separable state of two qubits)**.** Consider a separable state of two qubits $|00\rangle$. The first qubit belongs to Observer A, while the second belongs to Observer B. Observer A has no information about what B is doing or even possessing. The physics in A's lab is well-described by

$$\rho_A = \operatorname{tr}_B \rho = \operatorname{tr}_B |00\rangle\langle 00| = \sum_{m_B} \langle m|00\rangle \langle 00|m\rangle = |0\rangle\langle 0|$$

which is again the pure state $|0\rangle$. The observers are not correlated and can work independently.

Example 4.5 (Entangled state of two qubits)**.** Consider the two observers to be sharing a Bell pair

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

Their measurements are correlated. The total density matrix ρ is

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|)$$

Tracing over $\mathcal{H}_B$

$$\rho_A = \sum_m \langle m|\rho|m\rangle = {}_B\langle 0|\rho|0\rangle_B + {}_B\langle 1|\rho|1\rangle_B$$

gives only the terms where B's state is 0 or 1 in both bra-kets because any other combination is zero

$$\langle x0|00\rangle \langle 11|y0\rangle = 0$$

so that

$$\rho_A = \frac{1}{2}|0\rangle_A {}_A\langle 0| + \frac{1}{2}|1\rangle_A {}_A\langle 1| = \frac{1}{2}I$$

This is a mixed state even if the starting state is pure. This is the most undetermined state.

Entanglement creates correlation between subsystems. One may give a description of the full system in terms of quantum mechanics with pure states, but if one wants to describe only a subsystem, then mixed states are necessary and they parameterize the ignorance about the entire system.

This can be reversed. If ρ_a is a density matrix in $\mathcal{H}_A$, then there is always a bigger Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ whose density matrix ρ is pure in a state $|\psi\rangle$ and

$$\rho_A = \operatorname{tr}_B \rho\,, \quad \rho = |\psi\rangle\langle\psi|$$

This procedure is called **purification**. This is a straightforward theorem (see Nielsen, §2).

Lecture 10

mer 06 nov 2024 16:30

4.3 Open systems

The qubits always interact with the environment, they are open systems. The Hilbert space is

$$\mathcal{H} = \mathcal{H}_q \otimes \mathcal{H}_E$$

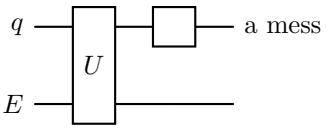
where $\mathcal{H}_q$ is for the qubit and $\mathcal{H}_E$ is for the environment. In the Hilbert space $\mathcal{H}$, we can apply standard quantum mechanics. There is a (pure) state

$$|\psi\rangle \in \mathcal{H}_q \otimes \mathcal{H}_E$$

that evolves unitarily

$$|\psi\rangle \to U\,|\psi\rangle$$

However, one cannot always keep the space $\mathcal{H}_E$, which has infinitely many degrees of freedom, so one needs to trace over E. As one has seen, even if the state $|\psi\rangle$ is pure, one obtains a density matrix ρ for $\mathcal{H}_q$. Moreover, restricted to $\mathcal{H}_q$, the operator U does not need to be unitary



All this generically happens, even if the dynamics of the universe is unitary. The qubits, even if they start in a well-defined state $a\,|0\rangle + b\,|1\rangle$ (initially pure), become entangled with the environment under a general U: the trace over E then gives a density matrix, as one has in the Bell pair example above. Moreover, a qubit in an excited level, e.g. $|1\rangle$, in an atom can decay by spontaneous emission, emitting a photon that is lost in the environment. In this example, $\mathcal{H}_E$ is the Hilbert space of radiations and $\mathcal{H}_q$ is Hilbert space of atomic two-level system. From the point of view of $\mathcal{H}_q$, the state $|1\rangle$ suddenly transforms into $|0\rangle$, while if it were in $|0\rangle$ it would stay there. Thus, spontaneous emission as seen by $\mathcal{H}_q$ is not unitary

$$U = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad U : \begin{cases} |1\rangle \to |0\rangle \\ |0\rangle \to |0\rangle \end{cases}$$

Of course, there is a well-defined unitary evolution in the coupled system of atom and radiation, $\mathcal{H}_q \otimes \mathcal{H}_E$, as discussed in many quantum mechanics textbooks.

This can be seen in more details. In a closed system, time evolution is achieved through a unitary operator U

$$|\psi\rangle \to U\,|\psi\rangle$$

so that on the density matrix one has

$$\rho = |\psi\rangle\langle\psi| \to U\,|\psi\rangle\langle\psi|\,U^\dagger = U\rho U^\dagger$$

The matrix evolves by conjugation. This is extended by linearity to all density matrices

$$\rho = \sum_i p_i\,|\psi_i\rangle\langle\psi_i| \to \sum_i p_i U\,|\psi_i\rangle\langle\psi_i|\,U^\dagger = U\rho U^\dagger$$

Consider the combined system $\tilde{\mathcal{H}} = \mathcal{H} \otimes \mathcal{H}_E$ with E the environment. In $\tilde{\mathcal{H}}$, the density matrix is $\tilde{\rho}$. This can be a pure state

$$\tilde{\rho} = |\psi\rangle\langle\psi|$$

or a mixture (after all, the system is at finite temperature and the environment is not generally the entire universe). Assume that the density matrix is

$$\tilde{\rho} = \rho \otimes \rho_E$$

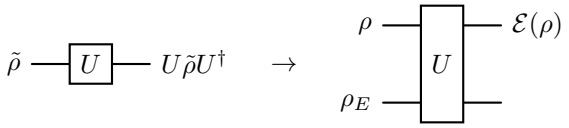
This is what happens with qubits: one prepares them in ρ, possibly pure and coherent, and one would like to know what happens after. Then, after some time evolution, one has

$$\tilde{\rho} \to U(\rho \otimes \rho_E)U^\dagger$$

and when one traces over E, one obtains

$$\mathcal{E}(\rho) = \mathrm{tr}_E[U(\rho \otimes \rho_E)U^\dagger]$$

which acts only on $\mathcal{H}$. One goes from the complete system to the reduced system



The matrix $\mathcal{E}(\rho)$ must still be a good density matrix, but the map

$$\rho \to \mathcal{E}(\rho)$$

is not a unitary operation in general. Let $\{|e_k\rangle\}$ be a basis of the environment E. Then, by definition

$$\mathcal{E}(\rho) = \sum_k \langle e_k| U(\rho \otimes \rho_E) U^\dagger |e_k\rangle$$

For simplicity and with no loss of generality, suppose that the environment E is initially in some pure state $|e_0\rangle$. If the matrix ρ_E is mixed, then one may take a bigger Hilbert space and perform a purification. So let

$$\rho_E = |e_0\rangle\langle e_0|$$

Then

$$\begin{aligned}\mathcal{E}(\rho) = \operatorname{tr}_E U(\rho \otimes \rho_E) U^\dagger &= \sum_k \langle e_k| U(\rho \otimes \rho_E) U^\dagger |e_k\rangle \\ &= \sum_k \langle e_k|U|e_0\rangle \, \rho \, \langle e_0|U^\dagger|e_k\rangle = \sum_K E_k \rho E_k^\dagger\end{aligned}$$

where $E_k = \langle e_k|U|e_0\rangle$ is still an operator acting on the Hilbert space $\mathcal{H}$. The operators E_k are called **operation elements** for $\mathcal{E}$, which is itself called **quantum operation**. The formula

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger$$

is called **operator-sum representation**. There is a general theory of it (see Nielsen, §8, Preskill, §3), but it is not needed in this course. One would like $\mathcal{E}(\rho)$ to still be a density matrix. In particular, it must hold $\operatorname{tr} \mathcal{E}(\rho) = 1$. Imposing this condition gives

$$1 = \operatorname{tr} \mathcal{E}(\rho) = \sum_k \operatorname{tr}\Big(E_k \rho E_k^\dagger\Big) = \sum_k \operatorname{tr}\Big(\rho E_k^\dagger E_k\Big) = \left[\sum_k E_k^\dagger E_k\right] \operatorname{tr} \rho\,, \quad \forall \rho$$

Since $\operatorname{tr} \rho = 1$ it must hold

$$\sum E_k^\dagger E_k = I$$

This condition can be relaxed by allowing measurements.

The interpretation of the operator-sum representation is the following. The representation is the combination of different possible processes where, in each

- the system executes a (not necessarily unitary) transformation E_k

  $$\rho \to E_k \rho E_k^\dagger$$

  which is a non-unitary evolution.
- The environment jumps from $|e_0\rangle \to |e_k\rangle$.

Only the entire evolution on $\mathcal{H} \otimes \mathcal{H}_E$ is unitary, so singularly

$$E_k^\dagger E_k \neq I$$

4.4 Bit-flip and phase-flip

One may first study single-qubit quantum operations. By interacting with the environment, one qubit can lose coherence, be flipped, acquire a phase: the coefficients of the linear combination

$$a\,|0\rangle + b\,|1\rangle$$

may change. In the following one studies some basic operations and how they act on the Bloch sphere.

Bit-flip channel. This is one of the typical errors occurring in a quantum computer: the state $|0\rangle$ is flipped into $|1\rangle$ and viceversa. The operation E is just $X = \sigma_1$ or, better, one would like to describe a situation where this can happen or not, due to some unknown interaction with the environment. One uses two operation elements $E = I, \sigma_1$. The only general requirement is

$$I = \sum_k E_k^\dagger E_k$$

so one puts

$$E_0 = \sqrt{p} I = \sqrt{p} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad E_1 = \sqrt{1-p}\, X = \sqrt{1-p} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

In this way

$$E_0^\dagger E_0 + E_1^\dagger E_1 = pI + (1-p)X^2 = I$$

The parameter $p \in (0,1)$ can be interpreted as the probability that nothing happens and $1-p$ is the probability of a qubit flip. The initial density matrix ρ of the system is sent into

$$\rho \to \mathcal{E}(\rho) = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger = p\rho + (1-p) X \rho X$$

Since for a qubit, the density matrix can be expressed as $2\rho = I + \mathbf{r} \cdot \boldsymbol{\sigma}$, one finds

$$\begin{aligned} \mathcal{E}(\rho) &= p \frac{I + r_1\sigma_1 + r_2\sigma_2 + r_3\sigma_3}{2} + (1-p) \frac{I + r_1\sigma_1 - r_2\sigma_2 - r_3\sigma_3}{2} \\ &= \frac{1}{2}[I + r_1\sigma_1 + (2p-1) r_2\sigma_2 + (2p-1) r_3\sigma_3] \end{aligned}$$

or equivalently

$$(r_1, r_2, r_3) \to (r_1, (2p-1)r_2, (2p-1)r_3)$$

For $p = 1$ there is no flip, while for $p < 1$ the Bloch sphere shrinks towards the x-axis. The extreme case is $p = 1/2$ where one obtains a segment connecting $|+\rangle$ and $|-\rangle$: a point on such segment is a probabilistic superposition (a mixture) of the two states. Pure states, in general, become mixtures.

Phase-flip channel. One may repeat the same argument with

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

This is the error corresponding to a flip in the sign of the superposition

$$a\,|0\rangle + b\,|1\rangle \to a\,|0\rangle - b\,|1\rangle\,, \quad -1 = \mathrm{e}^{\mathrm{i}\pi}$$

The relative phase has changed. In this case, the operation elements are

$$E_0 = \sqrt{p} I = \sqrt{p} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad E_1 = \sqrt{1-p}\, Z = \sqrt{1-p} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Similarly, one obtains

$$(r_1, r_2, r_3) \to ((2p-1)r_1, (2p-1)r_2, r_3)$$

The Bloch sphere shrinks towards the z-axis. Again, for $p = 1/2$, the sphere is degenerate and reduces to a segment along the z-axis. In such case, the density matrix is

$$\rho = \frac{1}{2}(I + r_3\sigma_3) = \frac{1}{2} \begin{bmatrix} 1 + r_3 & 0 \\ 0 & 1 - r_3 \end{bmatrix}$$

This is a mixed state

$$\frac{1 + r_3}{2}\,|0\rangle\langle 0| + \frac{1 - r_3}{2}\,|1\rangle\langle 1|$$

This also happens if one starts with a pure state ($|\mathbf{r}| = 1$) on the Bloch sphere: it gets mapped into a mixed state. This is an effect of the interaction with the environment (i.e. errors in quantum computing). For example, a pure state $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$ becomes

$$(a\,|0\rangle + b\,|1\rangle)(a^*\,\langle 0| + b^*\,\langle 1|) \to \frac{1+r_3}{2}\,|0\rangle\langle 0| + \frac{1-r_3}{2}\,|1\rangle\langle 1|$$
$$\begin{bmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{bmatrix} \to \frac{1}{2}\begin{bmatrix} 1+r_3 & 0 \\ 0 & 1-r_3 \end{bmatrix}$$

The off-diagonal elements (i.e. the interference terms) are lost. This is de-coherence.

Bit-phase flip channel. This is the above argument with Y. The operation elements are

$$E_0 = \sqrt{p}I = \sqrt{p}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad E_1 = \sqrt{1-p}\,Y = \sqrt{1-p}\begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}$$

The Bloch sphere shrinks towards the y-axis.

Depolarizing channel. The density matrix of the depolarizing channel is

$$\mathcal{E}(\rho) = (1-p)\rho + \frac{1}{3}\rho(X\rho X + Y\rho Y + Z\rho Z)$$

The density matrix is untouched with probability $1-p$ and it is flipped by X, Y, or Z each with probability $p/3$. Using[8]

$$\frac{1}{2}I = \frac{1}{4}(\rho + X\rho X + Y\rho Y + Z\rho Z)$$

one may write

$$\mathcal{E}(\rho) = \frac{2}{3}\rho I + \left(1 - \frac{4}{3}p\right)\rho = \frac{1}{2}\tilde{p}I + (1-\tilde{p})\rho\,, \quad \tilde{p} \equiv \frac{4}{3}p$$

This has the interpretation that the qubit is untouched with a probability $1-\tilde{p}$ or depolarized (i.e. transformed into a completely mixed state $I/2$) with probability $\tilde{p}$. In this case the Bloch sphere shrinks equally in all directions towards the origin.

4.5 Amplitude and phase damping

A convenient characterization of typical quantum processes affecting qubits is given in terms of amplitude damping (i.e. loss of energy to the environment, like spontaneous emission) and phase damping (i.e. loss of phase, de-coherence due to scattering with the environment). Consider a two-level system, $|0\rangle$ and $|1\rangle$, with different energy — by convention $|1\rangle$ is the excited state.

Amplitude damping. Amplitude damping is a loss of energy to the environment and consequent decrease in amplitude, typically due to spontaneous emission. The excited state transitions to the ground state, $|1\rangle \to |0\rangle$, emitting a photon which is lost to the environment, so that in the state $a\,|0\rangle + b\,|1\rangle$ after some time one finds $|b|^2 \to 0$. One models the channel using a spontaneous emission, so the environment is the quantized electromagnetic field plus other things. Any system is coupled to the electromagnetic field, even if there is no electric or magnetic field, because in QED there are vacuum fluctuations, photons can be created (i.e. spontaneous emission).

The environment has a state $|0\rangle_E$ with no photon and a state $|1\rangle_E$ with a photon and other states. The unitary evolution of the system coupled to the environment is

$$U: \begin{cases} |0\rangle \otimes |0\rangle_E \to |0\rangle \otimes |0\rangle_E \\ |1\rangle \otimes |0\rangle_E \to \sqrt{1-p}\,|1\rangle \otimes |0\rangle_E + \sqrt{p}\,|0\rangle \otimes |1\rangle_E \end{cases}$$

[8] This can be proven by observing that X changes sign to $\sigma_{2,3}$, while Y to $\sigma_{1,3}$ and Z to $\sigma_{1,2}$. All the Pauli matrices cancel out since each appears twice with positive sign and twice with negative sign. Also recall $2\rho = I + \mathbf{r}\cdot\boldsymbol{\sigma}$.

where, if $|0\rangle_E$ and $|1\rangle_E$ are normalized, then the operator U preserves scalar products and it is unitary. One has assumed that the environment is initially in the vacuum state $|0\rangle_E$. If the system is in the ground state $|0\rangle$, nothing happens. If the system is in the excited state $|1\rangle$, there is a probability $1-p$ that it remains there and a probability p that it decays to $|0\rangle$ by emitting a photon and therefore moving the environment from $|0\rangle_E$ to $|1\rangle_E$.

By tracing over E, one may read the quantum processes $E_k = {}_E\langle k|U|0\rangle_E$

$$E_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{bmatrix}, \quad E_1 = \begin{bmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{bmatrix}$$

To compute the entries, one has to calculate

$$\langle n|E_k|m\rangle = \langle nk|U|m0\rangle \ , \quad n,m = 0,1$$

applying U as above. Notice that

$$E_0^\dagger E_0 + E_1^\dagger E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1-p \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & p \end{bmatrix} = I$$

as it should. The operation element E_1 is the quantum jump $|1\rangle \to |0\rangle$. The density matrix ρ evolves as

$$\rho = \begin{bmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{bmatrix} \to \mathcal{E}(\rho) = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger = \begin{bmatrix} \rho_{00} + p\rho_{11} & \sqrt{1-p}\,\rho_{01} \\ \sqrt{1-p}\,\rho_{10} & (1-p)\rho_{11} \end{bmatrix}$$

If Γ is the rate of decay (probability per unit time), then for small Δt, the probability is $p = \Gamma\,\Delta t$. At finite t, the exponential behavior $\mathrm{e}^{-\Gamma t}$ is expected. After n iterations, $t = n\,\Delta t$, one finds

$$(1-p) \rightsquigarrow (1-p)^n = (1-\Gamma\,\Delta t)^n = \left(1 - \frac{\Gamma t}{n}\right)^n \to \mathrm{e}^{-\Gamma t}, \quad n \to \infty$$

so that the density matrix is

$$\rho = \begin{bmatrix} \rho_{00} + (1-\mathrm{e}^{-\Gamma t})\rho_{11} & \mathrm{e}^{-\Gamma t/2}\rho_{01} \\ \mathrm{e}^{-\Gamma t/2}\rho_{10} & \mathrm{e}^{-\Gamma t}\rho_{11} \end{bmatrix}$$

where $T = 1/\Gamma$ is the decay time. Notice the presence of $\Gamma/2$ instead of Γ in the off-diagonal terms: this is typical of all simple quantum processes. After some time, the off-diagonal terms go to zero (one always loses interference, coherence), and also the ρ_{11} component goes to zero (this is the amplitude damping). Eventually, the system decays and the final state is $|0\rangle$ and the density matrix is $p(t \to \infty) = |0\rangle\langle 0|$, a pure state

$$\rho(t \to \infty) = \begin{bmatrix} \rho_{00} + \rho_{11} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

One obtains the ground state. The Bloch sphere shrinks towards $|0\rangle$.

Lecture 11

lun 11 nov 2024 16:30

Phase damping. A system always scatters particles (e.g. photons) in the environment without necessarily changing its state. This invariably results in a loss of coherence (loss of relative phase). One models this phenomenon with the unitary operator

$$U : \begin{cases} |0\rangle\,|0\rangle_E \to \sqrt{1-p}\,|0\rangle\,|0\rangle_E + \sqrt{p}\,|0\rangle\,|1\rangle_E \\ |1\rangle\,|0\rangle_E \to \sqrt{1-p}\,|1\rangle\,|0\rangle_E + \sqrt{p}\,|1\rangle\,|1\rangle_E \end{cases}$$

which is again isometric in $\mathcal{H}_S \otimes \mathcal{H}_E$. With probability $1-p$ nothing happens (the environment is assumed to initially be in $|0\rangle_E$), while with probability p the environment is kicked into two possible states $|0\rangle_E$ or $|1\rangle_E$ according to the state $|0\rangle$ or $|1\rangle$ of the system. The state of the system does not change. Notice that the channel chooses a preferred basis for the Hilbert space $\mathcal{H}$: $|0\rangle$ and $|1\rangle$ is the basis in which nothing happens in $\mathcal{H}$, there is no bit flip. These two states

are called pointer states of de-coherence theory. By tracing over E, with $E_k = {}_E\langle k|U|0\rangle$, one finds the operation elements to be

$$E_0 = \sqrt{1-p}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \quad E_1 = \sqrt{p}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \quad E_2 = \sqrt{p}\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$$

and the density matrix to be

$$\begin{aligned}\rho \to \mathcal{E}(\rho) &= E_0\rho E_0^\dagger + E_1\rho E_1^\dagger + E_2\rho E_2^\dagger = (1-p)\rho + p\frac{1+\sigma_3}{2}\rho\frac{1+\sigma_3}{2} + p\frac{1-\sigma_3}{2}\rho\frac{1-\sigma_3}{2}\\ &= (1-p)\rho + \frac{1}{2}p\rho + \frac{1}{2}p\sigma_3\rho\sigma_3 = \left(1-\frac{1}{2}p\right)\rho + \frac{1}{2}pZ\rho Z\\ &= \tilde{p}\rho + (1-\tilde{p})Z\rho Z\,, \quad \tilde{p} = 1-\frac{1}{2}p\end{aligned}$$

so that this is equivalent to the phase flip channel. With probability $1-\tilde{p} = p/2$, one flips the relative phase:

$$a\,|0\rangle + b\,|1\rangle \to a\,|0\rangle - b\,|1\rangle$$

The effect on the density matrix ρ is

$$\rho \to \mathcal{E}(\rho) = \begin{bmatrix}\rho_{00} & (1-p)\rho_{01}\\ (1-p)\rho_{10} & \rho_{11}\end{bmatrix} = \begin{bmatrix}\rho_{00} & \mathrm{e}^{-\Gamma t}\rho_{01}\\ \mathrm{e}^{-\Gamma t}\rho_{10} & \rho_{11}\end{bmatrix}$$

and as previously discussed the Bloch sphere shrinks towards the z-axis

$$(r_1, r_2, r_3) \to ((2\tilde{p}-1)r_1, (2\tilde{p}-1)r_2, r_3) = ((1-p)r_1, (1-p)r_2, r_3)$$

In this channel, the off-diagonal terms in the density matrix ρ go to zero for $t \gg 1$ and the matrix remains diagonal, but there is no amplitude damping on ρ_{00} and ρ_{11}.

Remark 4.6. Notice that all processes of this type (channels that may depend on continuous parameters, random phase kicks $R_z(\theta)$) are actually equivalent to other simpler discrete Z channels. This issue of continuous versus discrete is conceptually important for error-correction that acts by correcting only discrete errors and automatically works for all. To appreciate better this point, consider a process where the state $|\psi\rangle$ is given a random phase using

$$R_z(\theta) = \exp\left[-\frac{\mathrm{i}}{2}\theta\sigma_3\right]$$

The state is

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle \to a\mathrm{e}^{-\frac{\mathrm{i}}{2}\theta}\,|0\rangle + b\mathrm{e}^{\frac{\mathrm{i}}{2}\theta}\,|1\rangle$$

where the phase is shifted by θ. The density matrix is

$$\rho = |\psi\rangle\langle\psi| \to R_z(\theta)\,|\psi\rangle\langle\psi|\,R_z(\theta)$$

If one averages over the angle θ with a Gaussian distribution of variance 2λ

$$G(0, 2\lambda) = \frac{1}{\sqrt{2\pi\lambda}}\mathrm{e}^{-\frac{\theta^2}{4\lambda}}$$

one finds exactly

$$\rho = \frac{1}{\sqrt{4\pi\lambda}}\int_{-\infty}^{+\infty}\mathrm{d}\theta\,R_z(\theta)\,|\psi\rangle\langle\psi|\,R_z(\theta) = \begin{bmatrix}|a|^2 & ab^*\mathrm{e}^{-\lambda}\\ a^*b\mathrm{e}^{-\lambda} & |b|^2\end{bmatrix}$$

which is the effect discussed above.

Remark 4.7. Notice also that this effect explains de-coherence in macroscopic systems: a dust particle interacting with CMB photons through scattering loses coherence and relative phase in a time $T = 1/\Gamma$, with no relative phase, no double-slit experiment or Schrödinger cat. Typically, for macroscopic objects the de-coherence time T is much smaller than the damping amplitude time (friction, etc) so de-coherence happens quickly. De-coherence theory (Zurek) explains the Born rule and the wave function collapse with a unitary evolution of system/environment.

Combined damping. Putting the two channels together gives the general parametrization

$$\rho = \begin{bmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{bmatrix} \to \begin{bmatrix} \rho_{00} + (1 - \mathrm{e}^{-\Gamma_1 t})\rho_{11} & \mathrm{e}^{-\Gamma_2 t}\rho_{01} \\ \mathrm{e}^{-\Gamma_2 t}\rho_{10} & \mathrm{e}^{-\Gamma_1 t}\rho_{11} \end{bmatrix}$$

where Γ_1 is the longitudinal relaxation rate (along the z-axis) and Γ_2 is the transverse relaxation rate. They are expressed in terms of the amplitude damping rate Γ_a and the phase damping rate Γ_φ according to

$$\Gamma_1 = \frac{1}{T_1} \equiv \Gamma_a\,, \quad \Gamma_2 = \frac{1}{T_2} = \frac{1}{2}\Gamma_a + \Gamma_\varphi$$

since Γ_a always produces also phase damping. The times T_1 and T_2 are standard notation in quantum theory.

5 Quantum error correction

See Nielsen §§10.1, 10.2, Mermin §§5.1, 5.2, 5.3, 5.4, 5.6.

5.1 Correcting errors

Classical and quantum bits are subject to errors (wires, gates and in quantum computing also interaction with the environment and de-coherence). The problem looks serious in quantum computing because it is almost impossible to isolate completely a qubit and coherence (i.e. the relative phase) is important for computations. One needs to develop methods for detecting and correcting errors.

Classical computing. In classical computing, to avoid errors, one encodes a bit with repetitions, e.g.

$$0 \to 000\,, \quad 1 \to 111$$

The trade-off is using more bits. The basic assumptions are that the typical error is the flip of one bit $0 \leftrightarrow 1$, that errors are independent and that is is unlikely that two errors affect the same triplet of encoded bits. Un these assumptions, one needs to send the bits 000 instead of 0 and 111 instead of 1: the bits 000 and 111 are the logical bits. The receiver can decode the error by looking at the bits. If they get the logical bits, they know that the message has not been corrupted. If they get any other combination, for example 100, they know that a mistake has occurred, but may also deduce that the original bit is 000 and not 111 (which would require two flips). The receiver uses a majority rule.

This works well if errors are statistically independent. The receiver can still misinterpret the result if two or more errors occurred simultaneously. If p is the probability of one error, then the probability of two errors is $3p^2(1-p)$, of three is p^3. Therefore, the probability of mistaking the bit is the sum of the two previous probabilities $3p^2 - 2p^3$. Before the encoding, having just one bit, the probability of mistaking it is p. With the majority rule, the probability is lower and better

$$p < \frac{1}{2} \implies 3p^2 - 2p^3 < p$$

So it is good for a small probability of mistakes.

Quantum computing. One may use the same rules in quantum computing

$$|0\rangle \to |000\rangle\,, \quad |1\rangle \to |111\rangle$$

However, in quantum mechanics, one expects some formidable problems:

- the no cloning theorem does not allow to replicate the state to be encoded;
- errors are continuous, not discrete, how would one deal with all of them;
- to check for corruption, one must perform a measurement, but in quantum computing a measurement destroys the original state.

Luckily, all these problems can be overcome. One may start with a simple model where one would like to correct bit-flip errors.

First problem. Consider three qubits to encode

$$|0\rangle \to |000\rangle = |\bar{0}\rangle = |0\rangle_{\mathrm{L}} \ , \quad |1\rangle \to |111\rangle = |\bar{1}\rangle = |1\rangle_{\mathrm{L}}$$

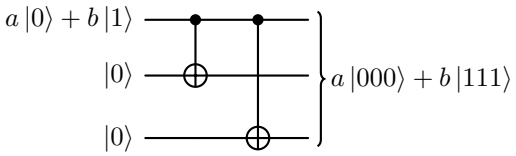
where L stands for logical. The generic state becomes

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle \to a\,|000\rangle + b\,|111\rangle$$

This can be done while avoiding cloning because the state above is not

$$(a\,|0\rangle + b\,|1\rangle) \otimes (a\,|0\rangle + b\,|1\rangle) \otimes (a\,|0\rangle + b\,|1\rangle)$$

It is straightforward to make a circuit that creates such state with just two CNOTs

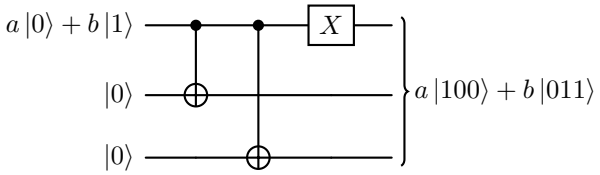


It would seem that one is violating the no-cloning theorem when one of the two coefficients is zero, for example $a = 0$, since

$$|100\rangle \to |111\rangle \iff |\psi, \tilde{\psi}, \tilde{\psi}\rangle \to |\psi, \psi, \psi\rangle$$

However, recall that the theorem fails when the states $|\psi\rangle$ and $|\widetilde{\psi}\rangle$ are orthogonal. Therefore, the first problem is solved.

Third problem. Suppose that a bit-flip occurs in one of the three qubits, e.g. the first



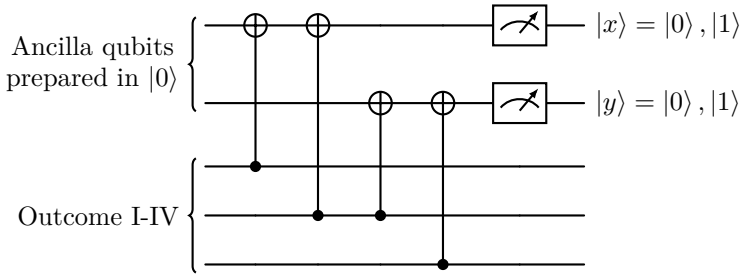
One would like to find a way to detect the error without changing the state. The three possible one-flip errors that can occur are

$$\mathrm{II} : a\,|100\rangle + b\,|011\rangle \ , \quad \mathrm{III} : a\,|010\rangle + b\,|101\rangle \ , \quad \mathrm{IV} : a\,|001\rangle + b\,|001\rangle$$

One assume that two or three simultaneous errors do not happen, so the last possibility is that nothing happens

$$\mathrm{I} : a\,|000\rangle + b\,|111\rangle$$

The four possible outcomes can be distinguished with a circuit. It makes use of ancilla bits that do not affect the system



It is straightforward to check for arbitrary coefficients a and b that

$$\begin{aligned} \text{I} &\to (x, y) = (0, 0) \\ \text{II} &\to (x, y) = (1, 0) \\ \text{III} &\to (x, y) = (1, 1) \\ \text{IV} &\to (x, y) = (0, 1) \end{aligned}$$

Therefore, by reading the tuple (x, y), one knows what error has occurred without changing the state of the system. This solves the third problem.

An example is the following. Consider the outcome III. In the state $a\,|010\rangle$, the second qubit flips both x and y to $x = y = 1$. In the state $b\,|101\rangle$, the first qubit flips x, while the third qubit flips y, both to $x = y = 1$. The final state, before the measurement, is

$$|11\rangle \otimes (a\,|010\rangle) + |11\rangle \otimes (b\,|101\rangle) = |11\rangle \otimes (a\,|010\rangle + b\,|101\rangle)$$

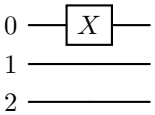
5.2 Bit-flip code

Suppose one would like to correct bit-flip errors, not caring about others. The 3-qubit encoding suffices

$$|0\rangle \to |000\rangle\,, \quad |1\rangle \to |111\rangle$$

The procedure is the following. Send the state $a\,|000\rangle + b\,|111\rangle$ along the circuit, measure it in a non-destructive way (called syndrome[9] measurement) and read the result (x, y). If the result is $(x, y) = (0, 0)$, then do nothing. If it is $(x, y) = (1, 0)$, then a flip occurred on the first qubit, therefore restore the qubit to its original form by acting on the first qubit with X. If the result is $(x, y) = (1, 1)$, a flip occurred on the second qubit and one applies the gate X to it. If $(x, y) = (0, 1)$, a flip occurred on the third qubit and one applies X. This can be automated by replacing the measurement on ancilla qubits with CNOT and Toffoli and performing the correction.

This is the error-correcting code. It can be better interpreted by looking at the error syndrome: it can be encoded into two numbers, $x = 0, 1$ and $y = 0, 1$. One would like to one what one is actually measuring. To fix the conventions, let the encoded qubits be labelled with $i = 0, 1, 2$ (Mermin convention).



The above circuit is the unitary operation $X \otimes I \otimes I$ that can be written as X_0 for simplicity. Similarly X_1 is $I \otimes X \otimes I$ and X_2 is $I \otimes I \otimes X$. A similar notation is used for Y and Z.

The Hilbert space of 3 qubits has dimension 8. The codewords are $|000\rangle = |\bar{0}\rangle$ and $|111\rangle = |\bar{1}\rangle$, and the message spans a two-dimensional subspace:

$$C = \{a\,|000\rangle + b\,|111\rangle\} \subset \mathcal{H}$$

In the following, there are frequently operators of the form

$$M = M_1 \otimes M_2 \otimes M_3\,, \quad M_i \in \{I, X, Y, Z\}$$

i.e. products of identity and Pauli matrices. For example,

$$X \otimes I \otimes I\,, \quad X \otimes Z \otimes I\,, \quad Z \otimes Y \otimes X\,, \quad \ldots$$

only products, no sums. Obviously, since matrices acting on different qubits commute and $I^2 = \sigma_i^2 = I$, for all of these it holds

$$M^2 = I\,, \quad I \equiv I \otimes I \otimes I$$

[9] The error syndrome is the pattern of errors.

For example

$$M = Z \otimes Y \otimes I \implies M^2 = (Z \otimes Y \otimes I)(Z \otimes Y \otimes I) = Z^2 \otimes Y^2 \otimes I^2 = I \otimes I \otimes I$$

The operator M is Hermitian and can be diagonalized. Since $M^2 = I$, the eigenvalues of M are $\lambda = \pm 1$. If M is not the identity, the two eigenvalues are ± 1 and one may decompose the 3-qubit Hilbert space into the two eigenspaces corresponding to the eigenvalues

$$\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2\,, \quad M_1 = +1\,, \quad M_2 = -1$$

The two subspaces $\mathcal{H}_1$ and $\mathcal{H}_2$ have the same dimension. For example, for $M = X \otimes I \otimes I$, one may use the basis $\{|\pm\rangle\,, |\pm\rangle\,, |\pm\rangle\}$, then

$$\begin{aligned} \mathcal{H}_1 &= \{|+\pm\pm\rangle\}\,, \quad M = 1\,, \quad \dim \mathcal{H}_1 = 4 \\ \mathcal{H}_2 &= \{|-\pm\pm\rangle\}\,, \quad M = -1\,, \quad \dim \mathcal{H}_2 = 4 \end{aligned}$$

The reason why the Hilbert space can be decomposed into two spaces of the same dimension is that, for every non-trivial operator, there exists an invertible element, for example $Z \otimes I \otimes I$ that maps $\mathcal{H}_1$ into $\mathcal{H}_2$:

$$|\psi\rangle \in \mathcal{H}_1 \to (Z \otimes I \otimes I)\,|\psi\rangle \in \mathcal{H}_2$$

and viceversa, so they both have the same dimension. The point is that $Z \otimes I \otimes I$ flips $|+\rangle \to |-\rangle$ and $|-\rangle \to |+\rangle$. The reason is that $Z \otimes I \otimes I$ anticommutes with M

$$(Z \otimes I \otimes I)M = -M(Z \otimes I \otimes I)$$

So if $M\,|\psi_1\rangle = +\,|\psi_1\rangle$, then

$$M(Z \otimes I \otimes I)\,|\psi_1\rangle = -(Z \otimes I \otimes I)M\,|\psi_1\rangle = -(Z \otimes I \otimes I)\,|\psi_1\rangle \implies (Z \otimes I \otimes I)\,|\psi_1\rangle \in \mathcal{H}_2.$$

since the vector $(Z \otimes I \otimes I)\,|\psi_1\rangle$ has eigenvalue -1. The idea with $M = \pm 1$ cuts the Hilbert space into half. The codeword space

$$C = \{a\,|000\rangle + b\,|111\rangle\}$$

can be identified with the eigenspace of the two operators

$$Z_0Z_1 \equiv Z \otimes Z \otimes I\,, \quad Z_1Z_2 \equiv I \otimes Z \otimes Z$$

with eigenvalues $+1$ for both. Each Z_0Z_1 and Z_1Z_2 cuts the Hilbert space $\mathcal{H}$ in half:

$$8/(2 \cdot 2) = 2$$

which is the dimension of C. Since Z is 1 on $|0\rangle$ and -1 on $|1\rangle$, it is clear that $Z_0Z_1 = Z_1Z_2 = +1$ on C.

The measurement of x and y is actually a measurement of $Z_0Z_1(x)$ and $Z_1Z_2(y)$. The error subspaces have

Name	Set	Z_0Z_1	Z_1Z_2	(xy)
E II	$\{a'\,\|100\rangle + b'\,\|011\rangle\}$	-1	1	(10)
E III	$\{a''\,\|010\rangle + b''\,\|101\rangle\}$	-1	-1	(11)
E IV	$\{a'''\,\|001\rangle + b'''\,\|110\rangle\}$	1	-1	(01)

More formally, an error is X acting on one of the qubits. This operator anticommutes with one or both of Z_0Z_1 and Z_1Z_2, and flips the sign of the corresponding eigenvalue

	I	X_0	X_1	X_2
Z_0Z_1	$+$	$-$	$-$	$+$
Z_1Z_2	$+$	$+$	$-$	$-$

while it commutes with operators in different subspaces. This formalism of the syndrome error being detected by operators $M^2 = I$, is useful later as the base idea of stabilizer codes.

Notice that one needs four orthogonal subspaces: the code C, the three error subspaces $\mathrm{E\,II}, \mathrm{E\,III}, \mathrm{E\,IV}$, each of dimension two, and

$$2 + 2 + 2 + 2 = 8$$

All the Hilbert space is taken by C and the three errors on one qubit. The spaces are orthogonal because they are eigenspaces with different eigenvalues of Z_0Z_1 and Z_1Z_2: eigenvectors corresponding to different eigenvalues are orthogonal. To use this method to correct bit-flip errors by encoding one qubit into a sequence of n qubits, one needs a two-dimensional space for the code and n two-dimensional spaces for the bit-flip that can occur in n positions. Since with n qubits, the dimension is $\dim \mathcal{H} = 2^n$, then

$$2^n \geq 2 + 2n \implies 2^{n-1} \geq 1 + n$$

so $n = 3$ is the smallest number that works.

5.3 Shor's nine-qubit code

The first example of a code correcting all errors is due to Shor. It uses nine qubits. It is mostly historical, since the Steane code with seven qubits is used instead. Shor's code is intuitive.

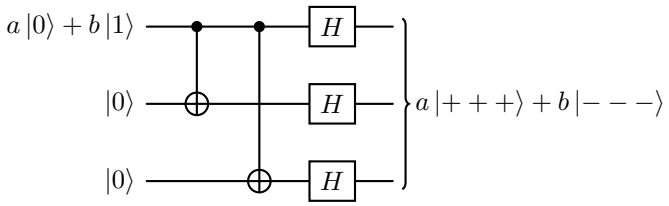
Phase-flip code. Phase-flip errors

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

can be corrected in the same ways as bit-flip errors. These latter are equivalent to the X gate, while phase-flip errors correspond to the Z gate. To go from the X basis $|\pm\rangle$ to the Z basis $|0\rangle$, $|1\rangle$, one uses the Hadamard gate. One just changes the basis and operates in the same way [r]

$$|0\rangle \to |+\,+\,+\rangle \ , \quad |1\rangle \to |-\,-\,-\rangle$$

It is simply done with



Errors in the first qubit

$$Z_0(a\,|+\,+\,+\rangle + b\,|-\,-\,-\rangle) = a\,|-\,+\,+\rangle + b\,|+\,-\,-\rangle$$

just flip $|+\rangle \to |-\rangle$. Indeed

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

So one can measure X_0X_1 and X_1X_2, and correct a mistake with the appropriate Z. This amounts as doing the same steps as the bit-flip but with the correspondences

$$|0, 1\rangle \leftrightarrow |+, -\rangle \ , \quad X \leftrightarrow Z \, , \quad Z \leftrightarrow X$$

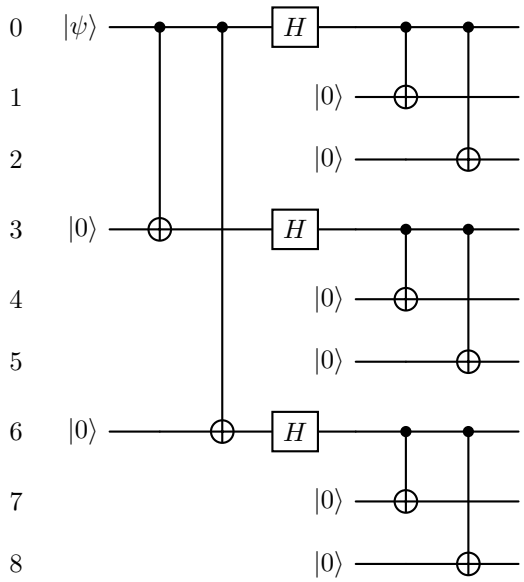
Shor's code. Shor's code applies first the phase-flip method then the bit-flip method to have

$$|0\rangle \to |+++\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}\frac{|0\rangle+|1\rangle}{\sqrt{2}}\frac{|0\rangle+|1\rangle}{\sqrt{2}} \to \frac{|000\rangle+|111\rangle}{\sqrt{2}}\frac{|000\rangle+|111\rangle}{\sqrt{2}}\frac{|000\rangle+|111\rangle}{\sqrt{2}}$$

so that the basis is

$$|0\rangle \to |\bar{0}\rangle = \frac{(|000\rangle+|111\rangle)(|000\rangle+|111\rangle)(|000\rangle+|111\rangle)}{2\sqrt{2}}$$
$$|1\rangle \to |\bar{1}\rangle = \frac{(|000\rangle-|111\rangle)(|000\rangle-|111\rangle)(|000\rangle-|111\rangle)}{2\sqrt{2}}$$

This can be done with concatenation



If a bit-flip occurs in the first three-qubit block, then one uses Z_0Z_1 and Z_1Z_2 to detect it and one may correct it with X as before. If it occurs in the second block, then one uses Z_3Z_4 and Z_4Z_5. If it occurs in the last block, then one uses Z_6Z_7 and Z_7Z_8.

If a phase-flip occurs in the first qubit, then

$$|000\rangle \pm |111\rangle \to |000\rangle \mp |111\rangle$$

so the value of $X_0X_1X_2$ changes. The same is true for the second and third qubits (notice that the code is degenerate). If the phase-flip occurs in the second three-qubit block, then $X_3X_4X_5$ changes. If the phase-flip occurs in the third three-qubit block, then $X_6X_7X_8$ changes. To detect a mistake, one measures $X_0X_1X_2X_3X_4X_5$ and $X_3X_4X_5X_6X_7X_8$ to obtain one of four results $(\pm1,\pm1)$ which identifies such mistake. Notice that one needs to use three matrices for each block to ? [r] $|000\rangle \leftrightarrow |111\rangle$ altogether. One corrects the mistake by applying the appropriate Z.

The syndrome measurements for the first qubit are

Operator	bit-flip X_0	phase-flip Z_0	bit-phase-flip Y_0
Z_0Z_1	$-$	$+$	$-$
Z_1Z_2	$+$	$+$	$+$
Z_3Z_4	$+$	$+$	$+$
Z_4Z_5	$+$	$+$	$+$
Z_6Z_7	$+$	$+$	$+$
Z_7Z_8	$+$	$+$	$+$
$X_0X_1X_2X_3X_4X_5$	$+$	$-$	$-$
$X_3X_4X_5X_6X_7X_8$	$+$	$+$	$+$

Since $Y = ZX$, bit- and phase-correcting codes can also deal with bit-phase errors. So Shor's code corrects bit-flips and phase-flips.

Second problem. Correcting discrete errors is sufficient to also correct continuous errors. The general logic is the following. Consider only a single error on a single qubit. The general errors is

$$\rho \to \sum_k E_k \rho E_k^\dagger$$

where

$$E_k |\psi\rangle = (\alpha I + \beta_1 X + \beta_2 Y + \beta_3 Z) |\psi\rangle$$

Each matrix acts on a single qubit

$$X = I \otimes \cdots \otimes X \otimes \cdots \otimes I$$

and the parameters α, β_i are continuous. If one is able to arrange an experiment such that

$$|\psi\rangle \,, \quad X |\psi\rangle \,, \quad Y |\psi\rangle \,, \quad Z |\psi\rangle$$

can be distinguished (e.g. the four states lie in orthogonal subspaces), then one performs a measurement to detect an error. The system collapses to one of the cases $S |\psi\rangle$, with $S \in \{I, X, Y, Z\}$, regardless of the continuous coefficients. The state can be restored by applying S a second time

$$S |\psi\rangle \to S^2 |\psi\rangle = |\psi\rangle$$

Therefore discrete error correction is enough to account also for continuous error correction. This solve the second problem in quantum computing.

5.4 Steane's seven-qubit code

Steane's code is more fault-tolerant. It is based on seven qubits. It uses the operators

$$M_0 = X_1 X_4 X_5 X_6 \,, \quad M_1 = X_1 X_3 X_5 X_6 \,, \quad M_2 = X_2 X_3 X_4 X_6$$

as well as

$$N_0 = Z_0 Z_4 Z_5 Z_6 \,, \quad N_1 = Z_1 Z_3 Z_5 Z_6 \,, \quad N_2 = Z_2 Z_3 Z_4 Z_6$$

to diagnose the error syndrome. Each operator satisfies

$$M_i^2 = N_i^2 = I$$

and they all commute because there is always an even number of X_i and Z_i with the same index in each operator

$$[M_i, N_j] = [M_i, M_j] = [N_i, N_j] = 0 \,, \quad \forall ij$$

Since the operators commute, they can be simultaneously diagonalized with an eigenspace of dimension

$$2^7/(2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2) = 2$$

This space spans the logical qubits $|\bar{0}\rangle$ and $|\bar{1}\rangle$.

Logical qubits. It is not necessary to write the logical qubits explicitly. They can be obtain by projection. Consider a generic state $|\psi\rangle$. One notices that

$$M_i(I + M_i) = M_i + M_i^2 = M_i + I$$

Therefore

$$M_i[(I + M_i) |\psi\rangle] = (I + M_i) |\psi\rangle$$

the state $(I + M_i) |\psi\rangle$ is an eigenstate of M_i with eigenvalue 1 regardless of the particular form of $|\psi\rangle$. Therefore, the logical qubits can be written as

$$|\bar{0}\rangle = \frac{I + M_0}{\sqrt{2}} \frac{I + M_1}{\sqrt{2}} \frac{I + M_2}{\sqrt{2}} |0000000\rangle$$
$$|\bar{1}\rangle = \frac{I + M_0}{\sqrt{2}} \frac{I + M_1}{\sqrt{2}} \frac{I + M_2}{\sqrt{2}} |1111111\rangle$$

and they are eigenvectors of M_i with eigenvalue 1. Since $[M_i, N_j] = 0$, one may also apply N_i to $|0000000\rangle$ or $|1111111\rangle$ and obtain the eigenvalue 1 since the operators N_i are products of four Zs. So the logical qubits are also eigenstates of N_i with eigenvalue 1. Moreover, since

$$(I + M)^2 = I + 2M + M^2 = 2(I + M)$$

the logical qubits are normalized

$$\langle\bar{0}|\bar{0}\rangle = \langle 0000000|\,(I + M_0)(I + M_1)(I + M_2)\,|0000000\rangle = \langle 0000000|\,I\,|0000000\rangle = 1$$

where one notices that the action of each M_i flips an even number of zeros to ones, $0 \to 1$, in the ket which is inevitably met with all zeros in the bra, e.g.

$$\langle 0000000|\,M_0 I M_2\,|0000000\rangle = \langle 0000000|0111010\rangle = 0$$

The logical qubits are orthogonal for a similar reason: the qubit $|\bar{0}\rangle$ has only terms with an even number of ones, while $|\bar{1}\rangle$ has only terms with an odd number of ones.

Steane's code. One would like to correct all single errors on the codeword

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

by applying $E_a\,|\psi\rangle$ with $E_a \in \{X_i, Y_i, Z_i\}$, $i = 0, \ldots, 6$. Each error E_a maps the code subspaces into an orthogonal subspace, but all subspaces are mutually orthogonal and they can be distinguished by the eigenvalues of M_i and N_i. Consider the following table detailing the presence of the operators X_i and Z_i inside M_i and N_i

Bit-flip	X_0	X_1	X_2	X_3	X_4	X_5	X_6
M_0	✓				✓	✓	✓
M_1		✓		✓		✓	✓
M_2			✓	✓	✓		✓
Phase-flip	Z_0	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
N_0	✓				✓	✓	✓
N_1		✓		✓		✓	✓
N_2			✓	✓	✓		✓

One may see a pattern. An error gives the same eigenvalue if it commutes with the operator M_i or N_i, and minus that same eigenvalue if it anti-commutes

$$[E, M] = 0\,, \quad M_i\,|\psi\rangle = |\psi\rangle \implies M_i E\,|\psi\rangle = E M_i\,|\psi\rangle = E\,|\psi\rangle \implies M = 1$$
$$\{E, M\} = 0\,, \quad M_i\,|\psi\rangle = |\psi\rangle \implies M_i E\,|\psi\rangle = -E M_i\,|\psi\rangle = -E\,|\psi\rangle \implies M = -1$$

The procedure to diagnose the errors is the following:

- Bit-flip error X_i. The operator X_i commutes with every M_j and every N_k that does not contain Z_i, in which case it anti-commutes. This anti-commutation is represented by the check mark symbol in the lower portion of the table. Therefore, an error corresponding to X_i gives a sequence of minus signs in the i-th column of the lower table where there are check marks.

- Phase-flip error Z_i. The operator Z_i commutes with every N_j and every M_k that does not contain X_i, in which case it anti-commutes. This anti-commutation is represented by the check mark symbol in the upper portion of the table. Therefore, an error corresponding to Z_i gives a sequence of minus signs in the i-th column of the upper table where there are check marks.

- Bit-phase-flip error Y_i. The operator Y_i commutes with every M_j and N_k that do not contain X_i nor Z_i, in which case it anti-commutes. This anti-commutation is represented by the check mark symbol in the both portions of the table. Therefore, an error corresponding to Y_i gives a sequence of minus signs in the i-th column of the whole table where there are check marks.

These results are all different, so one can diagnose them with a measurement and correct the mistake by applying the same operator X_i, Y_i or Z_i.

Notice that each error gives some minus signs only in one column. Other patterns correspond to the rest of the Hilbert space

$$\dim \mathcal{H} = 2^7 = 128\,, \quad \dim(C \oplus E_a C) = 2 + 2 \cdot 7 \cdot 3 = 44$$

The space on the right-hand side is the space of the codewords plus the space of one-qubit correcting codes (7 correcting qubits, 3 types of errors, 2 logical qubits). There are 84 missing states. They can be constructed as

$$X_i Z_j \left|\bar{0}\right\rangle\,, \quad X_i Z_j \left|\bar{1}\right\rangle\,, \quad i \neq j$$

which are indeed $7 \cdot 6 + 7 \cdot 6 = 84$. These are two-error states and are characterized by minus signs in two different columns of the table above.

Steane's code is historical and the favourite of textbooks and the literature due to its simplicity to be operated with gates. In general, after encoding, one needs to find a way to act with the Hadamard, CNOT, etc gates on the encoded blocks, but this is a mess with general M and N. Steane's code is easier because the operators acting on the logical qubits assume a very simple form (see Mermin, §5.7)

$$\bar{X} = X_0 X_1 X_2 X_3 X_4 X_5 X_6\,, \quad \bar{Z} = Z_0 Z_1 Z_2 Z_3 Z_4 Z_5 Z_6\,, \quad \bar{H} = H_0 H_1 H_2 H_3 H_4 H_5 H_6$$

The first two are a direct result of the (anti-)commutation relations of X and Z with the Ms of the projection operators inside the logical qubits. The third is a consequence of $HZ = XH$ and $ZH = HX$ which imply their barred counterparts: this essentially says that $\bar{H}$ exchanges the bases of $\bar{X}$ and $\bar{Z}$.

Steane's code is also fault-tolerant.

5.5 Other codes

An other interesting code is the perfect non-degenerate five-qubit code. Non-degenerate means that all single-qubit errors are distinguishable (Shor's code is degenerate) and perfect means that there is no redundancy in the Hilbert space.

When encoding one qubit with n qubits, one needs a Hilbert space of dimension $\dim \mathcal{H} = 2^n$. The codewords (i.e. the logical bits) and the single-qubit errors span a $2 + 2 \cdot 3 \cdot n$ dimensional subspace[10], if there is no degeneracy. Therefore

$$2^n \geq 6n + 2 \iff 2^{n-1} \geq 3n + 1$$

This is the quantum Hamming bound (for non-degenerate codes). The bound is saturated at $n = 5$. So there is a perfect code with $n = 5$ qubits. Notice that even if one includes degeneracy, then one still cannot go below five qubits $n < 5$ (see Preskill).

Five-qubit code.

6 Physical realization

6.1 About harmonic oscillators

There are many physical realizations of qubits for quantum computing or sensing as two-level states (ion-traps, semiconductor spins, superconducting circuits, NMR, etc). Spin-half and photon polarization are the two natural two-level systems in quantum mechanics, and they are indeed used in many experiments and technologies. Spins can be controlled with external magnetic fields:

$$H = -\mu \mathbf{S} \cdot \mathbf{B}$$

Photons can be manipulated with polarizers, beam-splitters and phase shifters (i.e. mediums with refractive index $n \neq 1$). Sometimes, the systems are more complicated. Two examples:

[10] There 2 logical qubits and each can have 3 errors.

- **Ion trap**. Ions are cooled down in an electromagnetic trap until there is just a vibrational mode of the center of mass that can be excited. Each ion has a nuclear spin (or some metastable state) that can be used as a single qubit: the two levels are the hyperfine splitting. A two-qubit state is obtained by tensoring the two-level spin system with the lowest excitation modes of the vibration (phonons) giving four states:

$$|\text{ion, vibration}\rangle = |0,0\rangle, |0,1\rangle, |1,0\rangle, |1,1\rangle$$

  Laser pulses are used to operate transitions between these states.

- **Superconducting qubits**. Superconducting qubits are LC circuits, macroscopic electromagnetic oscillators, that, at low temperature, act coherently and quantum due to superconductivity, acted upon by microwave drives and resonators. They are actually anharmonic oscillators (the lowest states are qubits), and also the electromagnetic cavity is a collection of quantum oscillators.

In both examples, one can give an effective description with just the physics of the harmonic oscillator. It describes both the electromagnetic field and the external couplings, and sometimes the qubit itself.

The following is a review of the basics. Consider a system where something oscillates at a frequency ω:

$$\ddot{x} + \omega^2 x = 0 \implies x(t) = A\mathrm{e}^{\mathrm{i}\omega t} + B\mathrm{e}^{-\mathrm{i}\omega t}$$

The Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$$

The procedure of canonical quantization is equivalent to the replacement of the canonical variables x and p with operators that obey the commutator $[x,p] = \mathrm{i}\hbar$. Introducing the ladder operators:

$$a = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x + \mathrm{i}p) \quad \text{and} \quad a^\dagger = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x - \mathrm{i}p)$$

or equivalently:

$$x = \sqrt{\frac{\hbar}{2m\omega}}(a^\dagger + a) \quad \text{and} \quad p = \mathrm{i}\sqrt{\frac{m\hbar\omega}{2}}(a^\dagger - a)$$

The Hamiltonian becomes

$$H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) = \hbar\omega\left(n + \frac{1}{2}\right)$$

where $\hat{n} = a^\dagger a$ is the number operator. The Hamiltonian has eigenstates $|n\rangle$

$$H|n\rangle = \hbar\omega\left(n + \frac{1}{2}\right)|n\rangle, \quad n \in \mathbb{N}_0$$

The ladder operators a and $a^\dagger$ act as lowering and raising operators:

$$a^\dagger a|n\rangle = n|n\rangle, \quad a|n\rangle = \sqrt{n}|n-1\rangle, \quad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$$

Their commutation relation is

$$[a, a^\dagger] = 1$$

The state $|n\rangle$ can be expressed as

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle$$

The spectrum is given by

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right)$$

and the energy levels are equidistant and infinite.

For a qubit, only two levels are needed. One could select two of them, e.g. $|0\rangle$ and $|1\rangle$, but one also needs to operate on them, for example, sending $|0\rangle \to |1\rangle$. This may be done with a photon of the same frequency ω:

$$E_1 - E_0 = \hbar\omega$$

However, the same photon has a chance to send $|1\rangle \to |2\rangle$, $|2\rangle \to |3\rangle$, etc. Equidistant levels are not good for quantum computing, the harmonic oscillator is not good at all.

In principle, one could realize some gates. For example, one may realize a CNOT gate by encoding two qubits into four states of a harmonic oscillator, e.g.

$$|00\rangle_{\rm L} = |0\rangle\,, \quad |01\rangle_{\rm L} = |2\rangle\,, \quad |10\rangle_{\rm L} = \frac{|4\rangle + |1\rangle}{\sqrt{2}}, \quad |11\rangle_{\rm L} = \frac{|4\rangle - |1\rangle}{\sqrt{2}}$$

Then, just by time evolution

$$|n\rangle \to {\rm e}^{-\frac{\rm i}{\hbar}E_n t}\,|n\rangle = {\rm e}^{-\frac{1}{2}{\rm i}\omega t}{\rm e}^{-{\rm i}n\omega t}\,|n\rangle$$

By ignoring the overall phase ${\rm e}^{-\frac{1}{2}{\rm i}\omega t}$ that is common to all states, one obtains

$$\begin{aligned}
|00\rangle_{\rm L} &\to |00\rangle_{\rm L} \\
|01\rangle_{\rm L} &\to {\rm e}^{-2{\rm i}\omega t}\,|01\rangle_{\rm L} \\
|10\rangle_{\rm L} &\to \frac{1}{\sqrt{2}}[{\rm e}^{-4{\rm i}\omega t}\,|4\rangle + {\rm e}^{-{\rm i}\omega t}\,|1\rangle] \\
|11\rangle_{\rm L} &\to \frac{1}{\sqrt{2}}[{\rm e}^{-4{\rm i}\omega t}\,|4\rangle - {\rm e}^{-{\rm i}\omega t}\,|1\rangle]
\end{aligned}$$

For $t = \pi/\omega$, this simplifies to

$$\begin{aligned}
|00\rangle_{\rm L} &\to |00\rangle_{\rm L} \\
|01\rangle_{\rm L} &\to |01\rangle_{\rm L} \\
|10\rangle_{\rm L} &\to \frac{|4\rangle - |1\rangle}{\sqrt{2}} = |11\rangle_{\rm L} \\
|11\rangle_{\rm L} &\to \frac{|4\rangle + |1\rangle}{\sqrt{2}} = |10\rangle_{\rm L}
\end{aligned}$$

The last two states are mapped into one another. This is precisely the CNOT gate. The problem is that the states $|nm\rangle_{\rm L}$ are not really tensor products of single qubits (one has done analog encoding, not digital) and it would be difficult to do multiple operations and use multiple qubits.

6.2 Electromagnetic field

The electromagnetic field in the vacuum is a collection of decoupled harmonic oscillators. One may describe the electromagnetic field in the vacuum with a vector potential $\mathbf{A}$

$$\mathbf{E} = -\partial_t \mathbf{A}\,, \quad \mathbf{B} = \nabla \times \mathbf{A}$$

in the Coulomb gauge $\nabla \cdot \mathbf{A} = 0$. The Maxwell equations become the wave equation for the potential

$$\Box\,\mathbf{A} = \frac{1}{c^2}\,\partial_t^2 \mathbf{A} - \nabla^2 \mathbf{A} = 0$$

with solution

$$\mathbf{A} = \mathbf{A}_{\mathbf{k}}{\rm e}^{{\rm i}\mathbf{k}\cdot\mathbf{x} - {\rm i}\omega t} + \mathbf{A}^*_{\mathbf{k}}{\rm e}^{-{\rm i}\mathbf{k}\cdot\mathbf{x} + {\rm i}\omega t}\,, \quad \omega = c|\mathbf{k}|$$

The gauge condition implies that there are just two transverse polarizations

$$\nabla \cdot \mathbf{A} = 0 \implies \mathbf{k}\cdot\mathbf{A}_{\mathbf{k}} = 0 \implies \mathbf{A}_{\mathbf{k}} = A_{\mathbf{k}1}\boldsymbol{\varepsilon}_1 + A_{\mathbf{k}2}\boldsymbol{\varepsilon}_2$$

where the polarization vectors are

$$\boldsymbol{\varepsilon}_i \cdot \mathbf{k} = 0\,, \quad |\boldsymbol{\varepsilon}_i| = 1\,, \quad i = 1, 2$$

The full electromagnetic field is a superposition of the elementary solutions

$$\mathbf{A}(\mathbf{x},t)=\sum_{\mathbf{k}}\sum_{i=1}^{2}A_{\mathbf{k}i}\boldsymbol{\varepsilon}_i\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)}+\text{c.c.}$$

These solutions are independent modes that oscillate in time with frequency ω

$$\mathcal{A}_{\mathbf{k}i}=A_{\mathbf{k}i}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega_{\mathbf{k}}t}\equiv(A_{\mathbf{k}i}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}})\mathrm{e}^{-\mathrm{i}\omega_{\mathbf{k}}t}\,,\quad\omega_{\mathbf{k}}=c|\mathbf{k}|$$

where the parenthesis is the spatial mode. Each solution solves the differential equation of the harmonic oscillator

$$\ddot{\mathcal{A}}_{\mathbf{k}i}+\omega_{\mathbf{k}}^2\mathcal{A}_{\mathbf{k}i}=0$$

Therefore the solutions are decoupled harmonic oscillators.