Theory of Quantum Information and Quantum Computing

Maso[*]

October 11, 2024

Contents

Lecture 1

mer 02 ott 2024 16:30

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

Knowledge required. Quantum mechanics from the bachelor.

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

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

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

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

1 Quantum mechanics

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

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

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

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

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

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

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

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

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

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

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

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

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

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

The eigenbasis of σ_2 is given by

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

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

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

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

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

Lecture 2

mer 09 ott 2024 16:30

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

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

After the measurement, the state instantly becomes

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

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

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

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

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

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

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

The probability is just

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

The state collapses to the "normalized" state

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

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

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

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

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

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

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

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

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

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

1.3 Time evolution

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

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

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

The solution to the Schrödinger equation

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

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

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

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

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

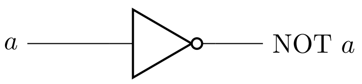
is conserved. The fact that $U(t)$ is unitary follows from the Schrödinger equation

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

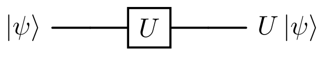
this holds since the Schrödinger equation also implies

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

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

$$X\left|0\right\rangle = \left|1\right\rangle \;, \quad X\left|1\right\rangle = \left|0\right\rangle$$

$$\left|0\right\rangle \;-\!\!\boxed{X}\!\!-\; \left|1\right\rangle \;, \qquad \left|1\right\rangle \;-\!\!\boxed{X}\!\!-\; \left|0\right\rangle$$

while Z flips signs

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

$$\left|0\right\rangle \;-\!\!\boxed{Z}\!\!-\; \left|0\right\rangle \;, \qquad \left|1\right\rangle \;-\!\!\boxed{Z}\!\!-\; -\left|1\right\rangle$$

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

$$a\left|0\right\rangle + b\left|1\right\rangle \;-\!\!\boxed{Z}\!\!-\; a\left|0\right\rangle - b\left|1\right\rangle$$

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

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

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

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

Other examples are

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

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

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

and the associated unitary matrix is

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

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

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

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

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

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

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

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

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

Exercise 1.8. Prove

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

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

1.4 Multipartite systems

One would like to study multiple qubits.

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

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

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

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

There must exist four states

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

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

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

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

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

The probability of getting 0 in the first qubit is

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

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

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

where the denominator is needed to normalize the state.

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

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

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

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

extended by linearity to all other vectors

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

Vectors in the tensor product space can be written as

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

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

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

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

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

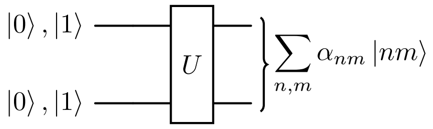
Example 1.11. The following state is separable

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

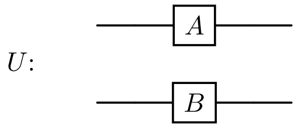
while the following states are entangled

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

A system of two qubits is depicted as



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



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

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

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

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

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

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

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

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

For vectors this is

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

which is just

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

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

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

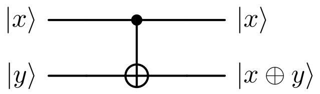
As for vectors, few 4×4 unitary matrices U are of the form $A \otimes B$. The other are actually more important for quantum computing: matrices that act on multiple qubits and not on individual ones independently. In classical computing there are several classical gates (e.g. AND, OR, XOR, etc) that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

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

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

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

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

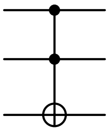


where $x, y = 0, 1$ and $\oplus$ denotes the XOR or, equivalently, the sum mod 2. In fact, if $x = 0$ nothing happens, but if $x = 1$ then

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

Therefore, the gate CNOT is indeed a XOR.

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



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