

Theory of Quantum Information and Quantum Computing

Maso*

February 2, 2025

Contents

1	Quantum mechanics	2
1.1	Qubits	3
1.2	Observables and measurements	4
1.3	Time evolution	6
1.4	Multipartite systems	9
1.5	No-cloning	11
2	Entanglement	13
2.1	Bell states	13
2.2	Applications	14
2.3	Bell inequalities	16
2.4	Quantum encryption	19
3	Quantum algorithms	20
3.1	Quantum and classical circuits	20
3.2	Universality	22
3.3	Quantum parallelism	22
3.4	Deutsch's algorithm	24
3.5	Other simple algorithms	26
3.6	Quantum Fourier transform	29
3.7	Shor's algorithm — period finding	31
3.8	Breaking RSA encryption	33
3.9	Grover's search algorithm	36
3.10	Another look at Grover's algorithm	38
4	Open systems	39
4.1	Density matrix	39
4.2	Tracing over subsystems	43
4.3	Open systems	44
4.4	Bit-flip and phase-flip	46
4.5	Amplitude and phase damping	48
5	Quantum error correction	51
5.1	Correcting errors	51
5.2	Bit-flip code	53
5.3	Shor's nine-qubit code	55
5.4	Steane's seven-qubit code	57
5.5	Other codes	59
5.6	Toric and surface codes	60
5.7	Topological quantum computing	66

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

6	Physical realization	67
6.1	About harmonic oscillators	67
6.2	Electromagnetic field	69
6.3	Driven qubit and Rabi oscillations	70
6.4	Single-qubit control	73
6.5	Cavity QED	74
6.6	Dressed states	76
6.7	Schrieffer–Wolff transformation	78
6.8	Coherent and squeezed states	79
6.9	Ion traps	82
6.10	Mølmer–Sørensen gate	85
6.11	Superconducting qubits — circuit QED (cQED)	88
6.12	Josephson effect	91
6.13	Transmon	93
6.14	cQED gates	95
6.15	Cross-resonance gate	101

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, defines 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 transferring 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 states 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 a 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. Therefore there are two real numbers that characterize 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 anything, 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 (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

$$|\psi\rangle \sim e^{i\alpha} |\psi\rangle, \quad \alpha \in \mathbb{R}$$

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

Remark. Notice that not all states in the Hilbert space make for a good quantum state. One needs it to be normalized.

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

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

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 linear combination of the two previous ones and can be identified by pair of complex numbers

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

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

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

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

$$A^\dagger = A$$

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

Theorem (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. 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] For a qubit, the Hilbert space is $\mathcal{H} \simeq \mathbb{C}^2$ and one may parametrize a self-adjoint matrix as

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

where σ_i are the Pauli matrices

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

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

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

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

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

with \mathbf{n} and $|\psi\rangle$ having the same angles.

Lecture 2

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

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

After the measurement, the state instantly becomes

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

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

mer 09 ott
2024 16:30

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 then 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

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

Notice that the normalization may give a phase, but one is interested in rays as quantum states. The probability amplitude 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 (IV, Time evolution). States evolve according to the Schrödinger equation

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

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

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

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

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

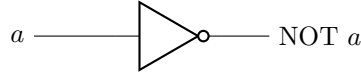
is conserved. The fact that $U(t)$ is unitary follows from the Schrödinger equation

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

this holds since the Schrödinger equation also implies

$$-i\hbar d_t \langle \psi(t) | = \langle \psi(t) | H^\dagger, \quad H^\dagger = H$$

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

$$\sigma_1 \sigma_3 = -i\sigma_2 \iff XZ = -iY$$

and $-iY$ is also unitary. Therefore one may use

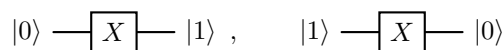
$$I, \quad X, \quad Z, \quad XZ$$

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

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

$$|0\rangle \xrightarrow{Z} |0\rangle, \quad |1\rangle \xrightarrow{Z} -|1\rangle$$

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

$$a|0\rangle + b|1\rangle \xrightarrow{Z} a|0\rangle - b|1\rangle$$

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

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

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

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

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

$$U = e^{-iHt} = e^{-iat} e^{-i\mathbf{b}\cdot\boldsymbol{\sigma}t} = e^{-iat} e^{-ib\mathbf{n}\cdot\boldsymbol{\sigma}t}, \quad \hbar = 1, \quad b = \|\mathbf{b}\|$$

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

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

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

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

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

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

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. Prove

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

Exercise. 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 (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 sum the squared values of all the coefficients containing 0 in the first qubit. After the measurement, the state collapses

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

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)

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

extended by linearity to all other vectors

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

Some 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.

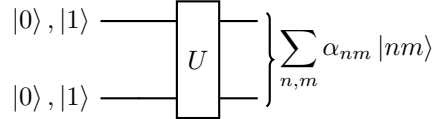
Example. The following state is separable

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

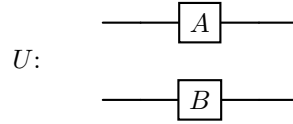
while the following states are entangled

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

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

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

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

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

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 separable 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

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

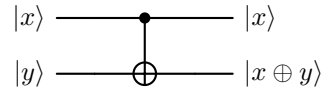
As for vectors, few 4×4 unitary matrices U are of the form $A \otimes B$. The other are actually more important for quantum computing: matrices that act on multiple qubits and not on individual ones independently. In classical computing there are several classical gates (e.g. AND, OR, XOR, etc) that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

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

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

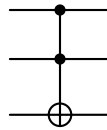


where $x, y = 0, 1$ and \oplus denotes the XOR or, equivalently, the sum mod 2. In fact, if $x = 0$ nothing happens, but if $x = 1$ then

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

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

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 many measurements and extract the values of $|\alpha|^2$ and $|\beta|^2$.

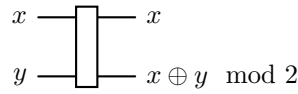
lun 14 ott
2024 16:30

In quantum computing, one typically has only one qubit in some particular 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

$$|\psi\rangle \rightarrow |0\rangle$$

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 coefficients 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 $|y\rangle = |0\rangle$, 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 $|y\rangle = |0\rangle$

$$|00\rangle \rightarrow |00\rangle$$

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

$$|10\rangle \rightarrow |11\rangle$$

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

$$|\psi\rangle \otimes |0\rangle \rightarrow |\psi\rangle \otimes |\psi\rangle$$

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

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

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 (No-cloning). There is no unitary operator that acts as

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

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

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 preserves the scalar product

$$\begin{aligned} (\langle\psi_2| \otimes \langle\psi_2|)(|\psi_1\rangle \otimes |\psi_1\rangle) &= (\langle\psi_2| \otimes \langle\phi|)(|\psi_1\rangle \otimes |\phi\rangle) \\ (\langle\psi_2|\psi_1\rangle)^2 &= \langle\psi_2|\psi_1\rangle \langle\phi|\phi\rangle \end{aligned}$$

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

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

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. \square

In quantum computing, one is only able to clone the states belonging to an orthonormal basis and 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

$$|\psi\rangle \rightarrow |01\rangle$$

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 with 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

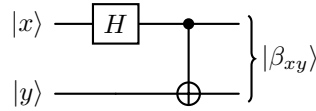
$$|\psi\rangle = \sum_{n,m=0}^1 \alpha_{nm} |\beta_{nm}\rangle$$

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

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

recalling that the Hadamard gate 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

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

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

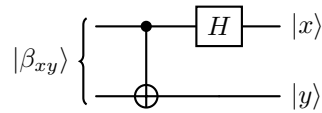
$$10 : (Z \otimes I) |\psi\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}} = |\beta_{10}\rangle$$

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

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

Observer A may encode two classical bits $xy \rightarrow |\beta_{xy}\rangle$. 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. Reading out the values of xy 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.

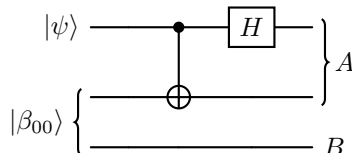
However, teleportation can be achieved by sharing an entangled pair

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

where 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

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

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. 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. 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 restrictions 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

2.3 Bell inequalities

mar 16 ott
2024 16:30

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 ignorance 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}

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

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

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

and the state

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

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

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

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 exists 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

$$|\langle C \rangle| \leq \langle |C| \rangle = 2$$

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 '82¹²³) still had loopholes:

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

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

Exercise. 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

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

¹A. 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.

²A. 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.

³A. 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.

where one uses

$$\sigma_i \sigma_j = \delta_{ij} I + i \varepsilon_{ijk} \sigma_k, \quad \text{Tr } \sigma_i = 0$$

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 messages. 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 then 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 chose 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 possible 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.

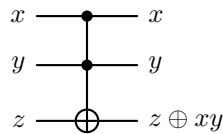
Reversibility. One difference is reversibility. Quantum circuits are reversible because gates are unitary

$$|\psi\rangle \xrightarrow{U} |\psi'\rangle \qquad |\psi'\rangle \xrightarrow{U^{-1}} |\psi\rangle$$

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

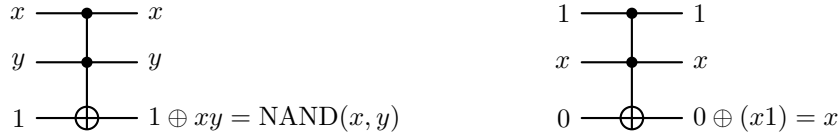


However, all classical computations can be written in a reversible way using multiple-bit gates. For example, the Toffoli gate is



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

$$U = e^{-iH}$$

in fact

$$U^\dagger U = e^{iH^\dagger} e^{-iH} = e^{iH} e^{-iH} = I$$

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

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

where one recalls

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

The 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. One may instead check the explicit form of $R_{\mathbf{n}}(\lambda)$ in terms of sine and cosine. Notice that

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

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

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

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 applies.

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

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

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

Exercise. 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 lots 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 $U(2^n)$ n -qubit gate. It is also difficult to work with more than two 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

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

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 $U \in U(2^n)$ 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 $U \in U(2)$, there exists a string of H s and T s such that the error is arbitrarily small

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

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

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

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

$$\begin{array}{c} x_{n-1} \text{ —————} \\ x_{n-2} \text{ —————} \\ x_i \text{ —————} \\ x_1 \text{ —————} \\ x_0 \text{ —————} \end{array}$$

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

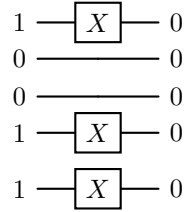
One may wonder how the state $|x\rangle$ can be prepared. If one starts with the state $|\psi\rangle$, one can make a measurement in the computational basis and obtain a random set of 0s and 1s. One may apply single gates X to flip the bits and put them in the desired form. Suppose one would like to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

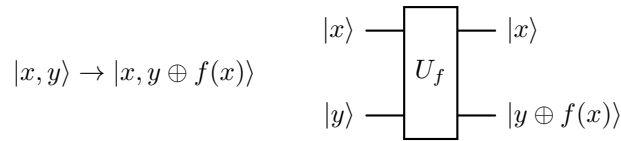
$$f(x) : \{0, 1\}^{\otimes n} \rightarrow \{0, 1\}, \quad x \mapsto f(x) = 0, 1$$

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

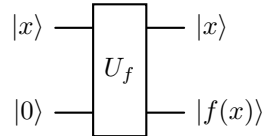
One important aspect is the use of a data register (or query) where x is stored and a target (or output) register where the result y is recorded

$$|x, y\rangle$$

One typically constructs a circuit that computes



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

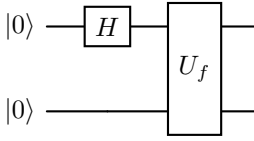
$$|x, y\rangle \rightarrow |x, y \oplus f(x)\rangle \rightarrow |x, y \oplus f(x) \oplus f(x)\rangle = |x, y\rangle$$

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 computational 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

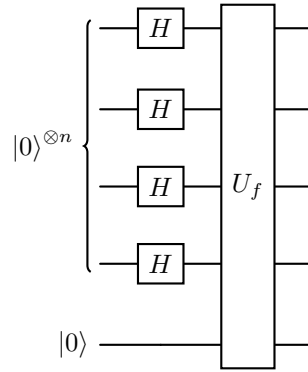
$$\frac{1}{\sqrt{2}} \left[|0, f(0)\rangle + |1, f(1)\rangle \right]$$

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

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

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 $0 \leq x < N = 2^n$ 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

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

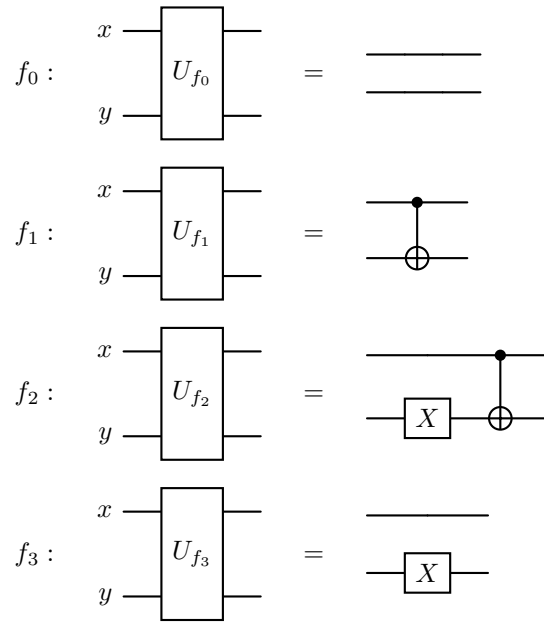
There are four such functions

$f(x)$	$x = 0$	$x = 1$
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

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

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

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

Similarly, all the other functions f_j .

Lecture 6

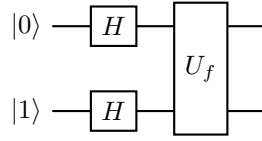
Consider a circuit with an unknown f_j , an oracle⁴. In classical computing, in order to know which function is inside, one needs to compute the values $f(0)$ and $f(1)$ using the circuit: two queries are needed.

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

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



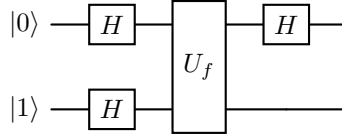
The interference is useful because

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

so that

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

Applying another Hadamard gate on the first qubit



gives the result

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

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

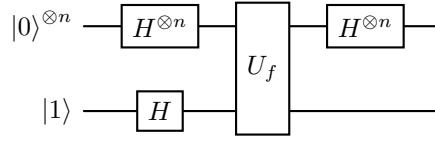
There are many other quantum algorithms that are more efficient than classical ones. In each, one learns some information about some function.

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

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

Consider a function $f(x)$ either constant or balanced (i.e. half of the space of inputs is 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

$$|0\rangle^{\otimes n} \rightarrow \sum_x |x\rangle$$

while using H one transforms

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

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

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

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

$$\begin{aligned} H^{\otimes n} |x\rangle &= H^{\otimes n} |x_{n-1}, \dots, x_0\rangle = \sum_{z_0 \dots z_{n-1}} \frac{(-1)^{x_0 z_0 + \dots + x_{n-1} z_{n-1}}}{2^{\frac{n}{2}}} |z_{n-1}, \dots, z_0\rangle \\ &= \sum_{z=0}^{2^n-1} \frac{(-1)^{x \cdot z}}{2^{\frac{n}{2}}} |z\rangle \end{aligned}$$

where $x \cdot z$ is the bitwise product of x and z mod 2

$$x \cdot z \equiv x_0 z_0 + \dots + x_{n-1} z_{n-1} \pmod{2}$$

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

$$\sum_{x=0}^{2^n-1} \frac{(-1)^{f(x)}}{2^n} = \begin{cases} \pm 1, & f(x) \text{ constant} \\ 0, & f(x) \text{ balanced} \end{cases}$$

This can be seen as follows. If $f(x)$ is constant, then there are 2^n copies of $(-1)^{f(x)}$ which cancel the 2^n in the denominator and the sum is ± 1 . If $f(x)$ is balanced, then there is 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

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

which computes the bitwise inner product of a and x

$$a \cdot x = a_0x_0 + \cdots + a_{n-1}x_{n-1} \pmod{2}$$

One would like to find the value of a . Classically, one needs n bits and n computations to find every component a_j . With quantum 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

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

There are two cases:

- if $a_j + z_j = 0 \pmod{2}$ for all j (equivalently $z \equiv a \pmod{2^n}$), 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 there exists a j for which $a_j + z_j = 1 \pmod{2}$, then the whole product is zero

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

In the first case, the result is

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

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

Remark. 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

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

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

When acting on a generic state

$$|\psi\rangle = \sum_{x=0}^{2^n-1} \gamma(x) |x\rangle$$

the transformation is

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

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

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

To compute classically the transformed coefficients $\hat{\gamma}$, one needs 2^{2n} operations. There is a 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 $U_{\text{FT}}(|\psi\rangle)$ 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_{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_{FT} can be written as

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

where \mathcal{Z} acts as

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

For $n = 1$, one has

$$\mathcal{Z} = Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} e^{i0} & 0 \\ 0 & e^{i\pi} \end{bmatrix} = e^{i\pi n}, \quad n = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

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

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

For simplicity, consider $n = 4$

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

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

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

In fact

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

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

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

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

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

in fact

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

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

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

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 $|x_0\rangle |x_1\rangle |x_2\rangle |x_3\rangle$ is an eigenstate of the number operators n_3, n_2, n_1 and n_0 with respective eigenvalues⁵ x_0, x_1, x_2 and x_3 :

$$n_{n-1-i} |x_i\rangle = x_i |x_i\rangle$$

Since every H_{n-1-i} is to the left of each x_i , one may safely substitute x_i for n_{n-1-i} to have

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

where P is a permutation

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

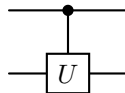
The operator U_{FT} is unitary. In fact, all operators with a matrix n are of the form

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

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

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

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



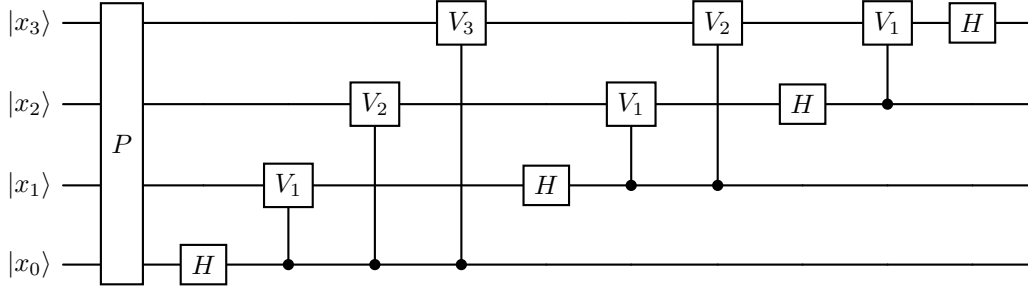
⁵The subscript of n_j corresponds to the position of the qubit in the sequence $|x_{n-1}\rangle \cdots |x_0\rangle$ and not specifically to $|x_j\rangle$ in any position. For example for $|x_1, x_3, x_2, x_0\rangle$, one has $n_0 \leftrightarrow x_0, n_1 \leftrightarrow x_2, n_2 \leftrightarrow x_3,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

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

Let

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

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.

Lecture 7

3.7 Shor's algorithm — period finding

lun 28 ott
2024 16:30

Consider a periodic bijective function acting on the integers

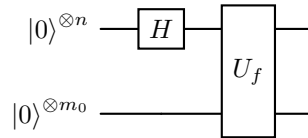
$$f : \mathbb{Z} \rightarrow \{0, 1\}^{\otimes m_0}$$

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

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

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 superposition of all the states $|x_i\rangle$ whose values x_i 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

$$x_0 + mr \geq 2^n \iff m = \left\lceil \frac{2^n}{r} \right\rceil \vee m = \left\lfloor \frac{2^n}{r} \right\rfloor + 1$$

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: 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

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

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

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

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 $j = 0, \dots, r-1$. There are r peaks. The reason is that all the phases add up

$$\sum_{k=0}^{m-1} e^{\frac{2\pi i}{2^n} k r y_j} = \sum_{k=0}^{m-1} e^{2\pi i k j} = \sum_{k=0}^{m-1} 1 = m$$

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 $N \rightarrow \infty$, 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 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

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

Noting that

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

also (which can be seen graphically)

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

and keeping in mind

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

then

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

Since there are r values, the probability of being close to any peak is

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

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 $1/2^{n+1}$ 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

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

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

$$x = \frac{1}{x_0 + \frac{1}{x_1 + \frac{1}{x_2 + \dots}}}$$

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 factors in common with $(p-1)(q-1)$. The observer computes the inverse of c

$$cd = 1 \pmod{(p-1)(q-1)}$$

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

$$b = a^c \pmod{N}$$

and sends it to B who decodes it with

$$a = b^d \pmod{N}$$

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

$$f(x) = b^x \pmod{N}$$

or the smallest r such that

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

or

$$b^r = 1 \pmod{N}$$

The period r is called the order of $b \pmod{N}$ 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 $b^r = 1 \pmod{N}$, then a classical computer finds d' such that $cd' = 1 \pmod{r}$. Then one can prove that the message a is

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

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

$$a^r = 1 \pmod{N}$$

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

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

Since $r = (p-1)(q-1) < N$ then r is the period of $f(x) = a^x \pmod{N}$.

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

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

One has

$$0 = a^r - 1 = x^2 - 1 = (x-1)(x+1) \pmod{N}$$

At this point, $x-1 \not\equiv 0 \pmod{N}$ because otherwise $a^{\frac{r}{2}} = 1 \pmod{N}$ and the period is then $r/2$ instead of r . Suppose also that

$$x+1 \not\equiv 0 \pmod{N}$$

Since $N = pq$ and $(x-1)(x+1)$ is divisible⁶ 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). Because the only divisors of N are p and q themselves, then 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.

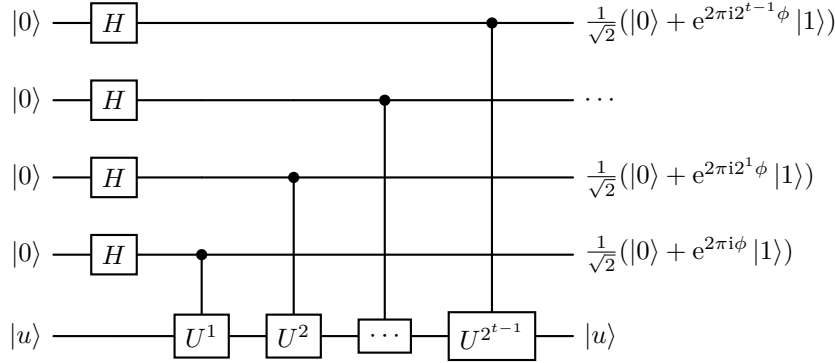
Lecture 8

 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

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

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

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

whose inverse quantum Fourier transform is exactly $|\phi\rangle$. The phase angle ϕ is a real number in $[0, 1)$:

$$\phi = 0.\phi_0\phi_1\dots\phi_{t-1}$$

since the integer part can be extracted and its exponential always gives 1. If the phase is exactly representable with t bits, then 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 exactly representable, then the quantum Fourier transform gives

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

If one parametrizes

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

then the amplitude corresponding to $j = l$ is

$$\frac{1}{2^t} \sum_{k=0}^{2^t-1} e^{2\pi i \frac{\delta k}{2^t}}$$

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

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

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

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

$$\sum_u c_u |\phi_u\rangle$$

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 $a^r = 1 \pmod{N}$, for a and N coprime. The operator is

$$U|x\rangle = |ax \pmod{N}\rangle$$

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

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

with eigenvectors

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

The target register can be initialized with

$$|1\rangle = \frac{1}{\sqrt{r}} \sum_{j=0}^{r-1} |u_j\rangle$$

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 $x = 0, \dots, N-1$ 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

$$\left. \begin{array}{l} |0\rangle^{\otimes n} \xrightarrow{H^{\otimes n}} \\ |1\rangle \xrightarrow{H} \end{array} \right\} U_f \left. \begin{array}{l} \\ \end{array} \right\} \frac{1}{2^{\frac{n}{2}}} \sum_{x=0}^{2^n-1} (-1)^{f(x)} |x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

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 oracle⁷

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

This is a reflection. If one splits the Hilbert space \mathcal{H} into the space spanned by $|a\rangle$ and its orthogonal space (thanks to the orthogonal decomposition theorem)

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

then 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 $|0\rangle^{\otimes n}$

$$|0\dots 0\rangle \rightarrow |0\dots 0\rangle$$

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

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

These two operations can be combined into the operator

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

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

$$\equiv \boxed{G} \equiv = \equiv \boxed{H^{\otimes n}} \equiv \boxed{2|0\rangle^{\otimes n}\langle 0|^{\otimes n} - I} \equiv \boxed{H^{\otimes n}} \equiv$$

in fact

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

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

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

$$\dots \equiv \boxed{O} \equiv \boxed{G} \equiv \boxed{O} \equiv \boxed{G} \equiv \dots$$

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

⁷Notice that this oracle is different from U_f . In particular, O is U_f when the target register (after the Hadamard gate) is $|-\rangle$: $U_f(|x\rangle \otimes |-\rangle) = (O|x\rangle) \otimes |-\rangle$.

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

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

since $|\phi\rangle$ 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 $|\phi\rangle$ by an angle 2θ to the state $|\phi'\rangle$. This state is at an angle of 3θ with respect to $|a\rangle^\perp$. 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. $|a\rangle$) 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 $|a\rangle$ with probability almost one. One may compute a by measuring the final state $(GO)^r |\phi\rangle$. With probability almost 1, one gets a . One may check it by computing $f(a)$ and see if it gives 1.

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

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

and a measurement almost always gives $|a\rangle$ and not $|a\rangle^\perp$. 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 state

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

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

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

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

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

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

$$\alpha_a \rightarrow -\alpha_a$$

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 a loss of coherence, a 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 $\{|\psi_i\rangle\}$, but one only knows (due to ignorance) that this happens with (classical) probability p_i . One defines the density matrix⁸

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

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

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

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

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

⁸Strictly speaking, it is an operator.

Here, the trace means

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

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

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

A quantum state as used until this moment 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

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

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

Example (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 microstates are $\{E_n, |n\rangle\}$, each can occur with classical probability

$$P(E_n) = \frac{1}{Z} e^{-\beta E_n}, \quad \beta = \frac{1}{k_B T}$$

where

$$Z = \sum_n e^{-\beta E_n} \implies \sum_n P(E_n) = 1$$

so that in the computations one uses the density matrix

$$\rho = \sum_n \frac{1}{Z} e^{-\beta E_n} |n\rangle\langle n|$$

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

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

Example (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

$$p = \frac{1}{Z} e^{-\beta E_0}$$

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.

◊ Hermiticity 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, $\text{tr } \rho = 1$.

◊ This can be seen as follows

$$\text{tr} \left[\sum_i p_i |\psi_i\rangle\langle\psi_i| \right] = \sum_i p_i \text{tr}(|\psi_i\rangle\langle\psi_i|) = \sum_i p_i = 1$$

- It holds $\text{tr } \rho^2 \leq 1$ and, for pure states only, $\text{tr } \rho^2 = 1$.

◊ The density matrix ρ is Hermitian, so it can be diagonalized:

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

It holds

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

Since ρ is positive

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

Therefore

$$\text{tr } \rho^2 = \sum_n p_n^2 \leq \sum_n p_n = 1$$

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 (Single qubit). The most general density matrix ρ , being Hermitian, is

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

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

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

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

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

and for a pure state it holds

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

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 $|1\rangle$:

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

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

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

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

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

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

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

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

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

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 (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

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

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

Example (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

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

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

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

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

Lecture 10

4.3 Open systems

mer 06 nov
2024 16:30

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

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

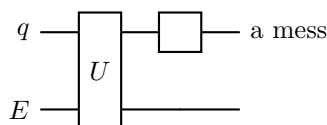
where \mathcal{H}_q is for the qubit and \mathcal{H}_E is for the environment. In the Hilbert space \mathcal{H} , one may apply standard quantum mechanics. There is a (pure) state

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

that evolves unitarily

$$|\psi\rangle \rightarrow U|\psi\rangle$$

However, one cannot always keep the space \mathcal{H}_E , which has infinitely many degrees of freedom, so one needs to trace over it. As one has seen, even if the state $|\psi\rangle$ is pure, one obtains a density matrix ρ of a mixed state 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 of a mixed state, as one has in the Bell pair example above. For example, a qubit in an excited level, e.g. $|1\rangle$, in an atom can decay by spontaneous emission, emitting a photon that is lost to the environment. In this case, \mathcal{H}_E is the Hilbert space of radiations and \mathcal{H}_q is Hilbert space of an 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

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

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

$$|\psi\rangle \rightarrow U|\psi\rangle$$

so that on the density matrix one has

$$\rho = |\psi\rangle\langle\psi| \rightarrow U|\psi\rangle\langle\psi|U^\dagger = U\rho U^\dagger$$

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

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

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

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

and, when one traces over E , one obtains

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

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

$$\tilde{\rho} \xrightarrow{U} U\tilde{\rho}U^\dagger \rightarrow \begin{array}{c} \rho \text{ --- } \boxed{U} \text{ --- } \mathcal{E}(\rho) \\ \rho_E \text{ --- } \boxed{U} \text{ --- } \end{array}$$

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

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

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

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

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 $\text{tr } \mathcal{E}(\rho) = 1$. Imposing this condition gives

$$1 = \text{tr } \mathcal{E}(\rho) = \sum_k \text{tr} (E_k \rho E_k^\dagger) = \sum_k \text{tr} (\rho E_k^\dagger E_k) = \text{tr} \left[\rho \sum_k E_k^\dagger E_k \right]$$

Since this relationship holds for all ρ and since $\text{tr } \rho = 1$, then

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

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

$$\rho \rightarrow E_k \rho E_k^\dagger$$

which is a non-unitary evolution.

- The environment jumps from $|e_0\rangle \rightarrow |e_k\rangle$.

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, a 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. A bit-flip 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

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

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

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

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. For a phase-flip, 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

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

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

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

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

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

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

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

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

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

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

The Bloch sphere shrinks towards the y -axis.

Depolarizing channel. The density matrix of the depolarizing channel is

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

The density matrix is untouched with probability $1-p$ and it is flipped by X , Y , or Z each with probability $p/3$. Using⁹

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

one may write

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

This has the interpretation that the qubit is untouched with a probability¹⁰ $1-\tilde{p}$ and 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, $|1\rangle \rightarrow |0\rangle$, 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 $|b|^2 \rightarrow 0$. 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

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

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 $E_k = {}_E\langle k|U|0\rangle_E$

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

⁹This 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 $2\rho = I + \mathbf{r} \cdot \boldsymbol{\sigma}$.

¹⁰Notice that $\tilde{p} \in [0, 1]$, while $p \in [0, 3/4]$. The argument here is presented backwards: the first form of $\mathcal{E}(\rho)$ is actually a different parametrization of the last form. See Nielsen, §8.3.4.

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, noting that n, m label the system, while $k, 0$ label the environment. 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 $|1\rangle \rightarrow |0\rangle$. The density matrix ρ evolves as

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

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 $e^{-\Gamma t}$ is expected. After n iterations, $t = n \Delta t$, one finds

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

so that the density matrix is

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

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 $\rho(t \rightarrow \infty) = |0\rangle\langle 0|$, a pure state

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

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

Lecture 11

Phase damping. A system always scatters particles (e.g. photons) in the environment without necessarily changing its own state. This invariably results in a loss of coherence (loss of relative phase). One models this phenomenon with the unitary operator¹¹

$$U : \begin{cases} |0\rangle \otimes |0\rangle_E \rightarrow \sqrt{1-p}|0\rangle \otimes |0\rangle_E + \sqrt{p}|0\rangle \otimes |1\rangle_E \\ |1\rangle \otimes |0\rangle_E \rightarrow \sqrt{1-p}|1\rangle \otimes |0\rangle_E + \sqrt{p}|1\rangle \otimes |2\rangle_E \end{cases}$$

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 $|1\rangle_E$ or $|2\rangle_E$ 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 $E_k = {}_E \langle k|U|0\rangle_E$, one finds the operation elements to be

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

for which

$$E_0^\dagger E_0 + E_1^\dagger E_1 + E_2^\dagger E_2 = I$$

¹¹This can be seen by noting that orthogonal bases are mapped into orthogonal bases.

and the density matrix to be

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

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

$$a|0\rangle + b|1\rangle \rightarrow a|0\rangle - b|1\rangle$$

The effect on the density matrix ρ is

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

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

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

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

Remark. 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

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

The state is

$$|\psi\rangle = a|0\rangle + b|1\rangle \rightarrow ae^{-\frac{i}{2}\theta}|0\rangle + be^{\frac{i}{2}\theta}|1\rangle$$

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

$$\rho = |\psi\rangle\langle\psi| \rightarrow R_z(\theta)|\psi\rangle\langle\psi|R_z^\dagger(\theta) = \begin{bmatrix} |a|^2 & ab^*e^{-i\theta} \\ a^*be^{i\theta} & |b|^2 \end{bmatrix}$$

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

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

one finds exactly

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

which is the effect discussed above.

Remark. In this case, the off-diagonal elements are suppressed, but this is just because one has chosen a specific basis of the Hilbert space, as previously said. There are other choices for which the diagonal elements are suppressed.

Remark. 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 and environment.

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

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

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

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

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.

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

The trade-off is using more bits. The basic assumptions are that

- the typical error is the flip of one bit $0 \leftrightarrow 1$,
- errors are independent,
- the probability of an error happening is small,
- and it is unlikely that two errors affect the same triplet of encoded bits.

Under 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

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

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

1. the no-cloning theorem does not allow to replicate the state to be encoded;
2. errors are continuous, not discrete, how would one deal with all of them;
3. 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 encoding qubits

$$|0\rangle \rightarrow |000\rangle = |\bar{0}\rangle = |0\rangle_L, \quad |1\rangle \rightarrow |111\rangle = |\bar{1}\rangle = |1\rangle_L$$

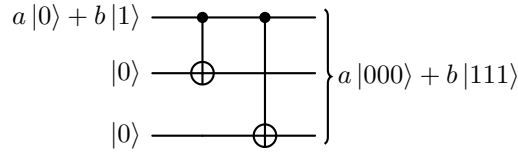
where L stands for “logical”. The generic state becomes

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

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

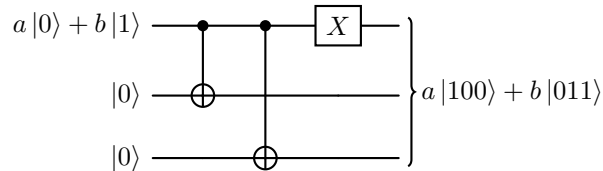


It would seem that one is violating the no-cloning theorem when one of the two coefficients is zero, for example $a = 0$, since

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

However, recall that the theorem fails when the states $|\psi\rangle$ and $|\tilde{\psi}\rangle$ 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



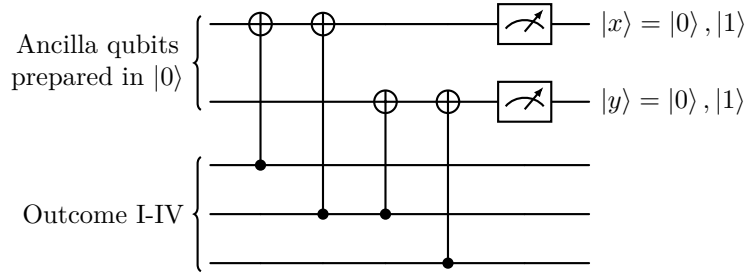
One would like to find a way to detect the error without changing the state. The three possible one-flip errors that can occur are

$$\text{II} : a|100\rangle + b|011\rangle, \quad \text{III} : a|010\rangle + b|101\rangle, \quad \text{IV} : a|001\rangle + b|110\rangle$$

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

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

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

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

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

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

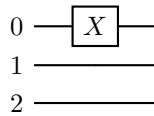
5.2 Bit-flip code

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

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

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 syndrome¹² 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 gates 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 know 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 $2^3 = 8$. The codewords¹³ 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

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

¹²The error syndrome is the pattern of errors.

¹³Also called logical elements, they are the linear combinations $a|000\rangle + b|111\rangle$.

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

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

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

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

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 three-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

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

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 :

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

and viceversa, so they both have the same dimension. The point is that $Z \otimes I \otimes I$ flips $|+\rangle \rightarrow |-\rangle$ and $|-\rangle \rightarrow |+\rangle$. The reason is that $Z \otimes I \otimes I$ anti-commutes 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 states with $M = \pm 1$ cut 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 (called stabilizers)

$$Z_0 Z_1 \equiv Z \otimes Z \otimes I, \quad Z_1 Z_2 \equiv I \otimes Z \otimes Z$$

with eigenvalues $+1$ for both. Each $Z_0 Z_1$ and $Z_1 Z_2$ 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 $Z_0 Z_1 = Z_1 Z_2 = +1$ on C .

The measurement of x and y is actually a measurement of $Z_0 Z_1(x)$ and $Z_1 Z_2(y)$. The error subspaces have

Name	Set	$Z_0 Z_1$	$Z_1 Z_2$	(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 error syndrome 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 E II, E III, 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.

Lecture 12

5.3 Shor's nine-qubit code

mer 13 nov
2024 16:30

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.

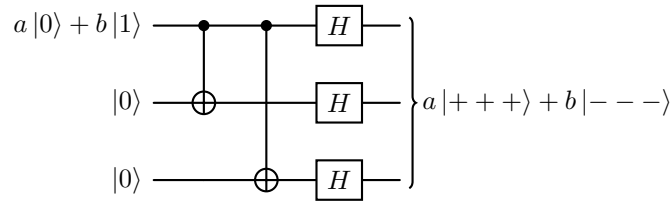
Phase-flip code. Phase-flip errors

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

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

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

This is simply done with



Errors in the first qubit

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

just flip $|+\rangle \rightarrow |-\rangle$. 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 corresponds to repeating the same steps as the bit-flip but with the correspondences

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

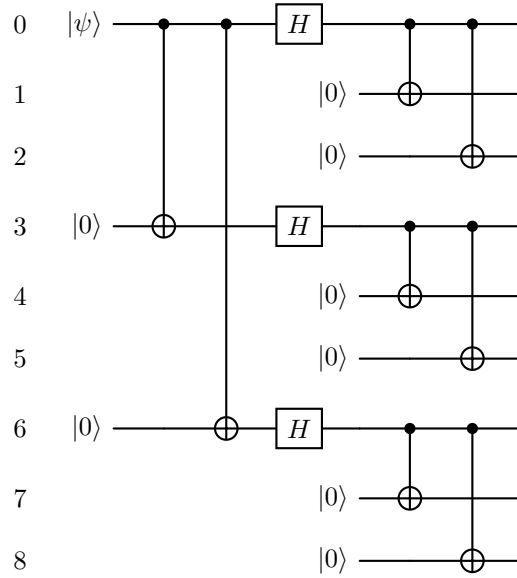
Shor's code. Shor's code applies first the phase-flip correction method then the bit-flip correction method to have

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

so that the basis is

$$\begin{aligned} |0\rangle &\rightarrow |\bar{0}\rangle = \frac{(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}} \\ |1\rangle &\rightarrow |\bar{1}\rangle = \frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}} \end{aligned}$$

This can be done with concatenation



Starting from

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

one obtains the state

$$a|\bar{0}\rangle + b|\bar{1}\rangle$$

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

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

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, one cannot distinguish which qubit has phase flipped). 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 in which block a mistake happened, one measures $X_0X_1X_2X_3X_4X_5$ and $X_3X_4X_5X_6X_7X_8$ to obtain one of four results $(\pm 1, \pm 1)$ which identifies such block. Notice that one needs to use three matrices for each block to read $|000\rangle \leftrightarrow |111\rangle$ 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
$Z_0 Z_1$	—	+	—
$Z_1 Z_2$	+	+	+
$Z_3 Z_4$	+	+	+
$Z_4 Z_5$	+	+	+
$Z_6 Z_7$	+	+	+
$Z_7 Z_8$	+	+	+
$X_0 X_1 X_2 X_3 X_4 X_5$	+	—	—
$X_3 X_4 X_5 X_6 X_7 X_8$	+	+	+

Since $iY = ZX$, 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 error is

$$\rho \rightarrow \sum_k E_k \rho E_k^\dagger$$

where

$$E_k |\psi\rangle = (\alpha I + \beta_1 X_i + \beta_2 Y_i + \beta_3 Z_i) |\psi\rangle$$

Each matrix acts on the single qubit i

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

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

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

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

5.4 Steane's seven-qubit code

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

$$M_0 = X_0 X_4 X_5 X_6, \quad M_1 = X_1 X_3 X_5 X_6, \quad M_2 = X_2 X_3 X_4 X_6$$

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

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

Since the operators commute, they can be simultaneously diagonalized with a codeword 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 obtained 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

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

and they are eigenvectors of M_i with eigenvalue 1. Since $[M_i, N_j] = 0$, one may apply N_i to $|\bar{0}\rangle$ or $|\bar{1}\rangle$ and obtain the eigenvalue 1 since the operators N_i are products of four Z s. 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, $0 \rightarrow 1$, in the ket which is inevitably met with all zeros in the bra, e.g.

$$\langle 0000000 | M_0 I M_2 | 0000000 \rangle = \langle 0000000 | 1011010 \rangle = 0$$

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 $E_a \in \{X_i, Y_i, Z_i\}$, $i = 0, \dots, 6$. Each error E_a maps the code subspace 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

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

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 qubit states). There are 84 missing states. They can be constructed as

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

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 M s of the projection operators inside the logical states. 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

Another 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 subspace¹⁴, 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).

¹⁴There are 2 logical qubit states and each qubit can have 3 errors.

Five-qubit code. The dimension of the Hilbert space is $\dim \mathcal{H} = 2^5$. One needs four projection operators to get a unique subspace of dimension $2^5 / (2 \cdot 2 \cdot 2 \cdot 2) = 2$. These are

$$M_0 = Z_1 X_2 X_3 Z_4, \quad M_1 = Z_2 X_3 X_4 Z_0, \quad M_2 = Z_3 X_4 X_0 Z_1, \quad M_3 = Z_4 X_0 X_1 Z_2$$

They all commute and square to the identity, $[M_i, M_j] = 0$ and $M_i^2 = I$. Notice that the indices are not quite symmetric, but the last term is not independent

$$M_4 = Z_0 X_1 X_2 Z_3 = M_0 M_1 M_2 M_3$$

The codewords are

$$\begin{aligned} |\bar{0}\rangle &= \frac{1+M_0}{\sqrt{2}} \frac{1+M_1}{\sqrt{2}} \frac{1+M_2}{\sqrt{2}} \frac{1+M_3}{\sqrt{2}} |00000\rangle \\ |\bar{1}\rangle &= \frac{1+M_0}{\sqrt{2}} \frac{1+M_1}{\sqrt{2}} \frac{1+M_2}{\sqrt{2}} \frac{1+M_3}{\sqrt{2}} |11111\rangle \end{aligned}$$

They are orthogonal and normalized for reasons identical to the seven-qubit case. The errors $E_a C$ with $E_a \in \{X_i, Y_i, Z_i\}$ (i.e. $3 \cdot 5 = 15$ of them) span an orthogonal subspace of the Hilbert space \mathcal{H} with different eigenvalues of the projection operators M_i given in the following table

Operator	$X_0 Y_0 Z_0$	$X_1 Y_1 Z_1$	$X_2 Y_2 Z_2$	$X_3 Y_3 Z_3$	$X_4 Y_4 Z_4$	I
M_0	+++	--+	+--	+--	--+	+
M_1	--+	+++	--+	+-	+-	+
M_2	+-	--+	+++	--+	+-	+
M_3	+-	+-	--+	+++	--+	+

All possible errors of signs are in the table and they are all different. Indeed

$$C \oplus E_a C = \mathcal{H}, \quad 2 + 2 \cdot 3 \cdot 5 = 2^5 = 32$$

Fault tolerance. There is a general theory of error-correction and a long list of codes in the literature. For example

- Calderbank–Shor–Steane (CSS) codes which generalize classical constructions to quantum computing;
- stabilizer codes;
- toric codes, introduced by Kitaev, or also called surface codes.

Some of these are covered in Nielsen, Preskill. There are also codes to correct multiple-qubit errors.

An important question is fault tolerance. One encodes some qubits, but then there can be other mistakes from gates and measurements, and errors can propagate, etc. Fault tolerance means that, under the assumption that each component fails with a probability p , one can construct a circuit with not too many other components that maintains the total error below a certain threshold. There is a classical theory of fault tolerance and also a quantum one which works (one needs to generate no more than one error in each block at each step). Steane's and five-qubit codes can be implemented in a fault-tolerant way. Fault tolerance is also the reason why the operator S is included in the set of universal gates $\{H, S, T, \text{CNOT}\}$ even if $S = T^2$.

Another possibility to minimize error is the use of topological quantum computation with anyons (thanks to Kitaev) that are topologically isolated from the environment.

Lecture 13

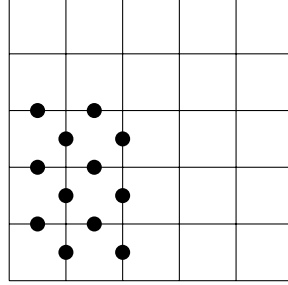
5.6 Toric and surface codes

The toric code, written by Kitaev and widely studied in quantum computing in its planar form (surface code), is important for large-scale computing due to its tolerance of errors. It is also

lun 18 nov
2024 16:30

associated with the idea of topological computing and is related to an interesting topological phase of matter with long-range entanglement (LRE).

The toric code is an $L \times L$ lattice of qubits with periodic boundary conditions (i.e. a torus), with qubits sitting on the links



There are two independent qubits per plaquette (to avoid double counting). The dimension of the Hilbert space is

$$\dim \mathcal{H} = 2^n = 2^{2L^2}, \quad n = 2L^2$$

There exists a planar version (surface code) with suitable boundary conditions, more useful for a practical implementation. It is a stabilizer code with the following projection operators (the stabilizers). For every vertex, there is

$$A_v = \prod_{j \in v} X_j$$

This operator acts on the four qubits sitting around the vertex. For every plaquette there is

$$B_p = \prod_{j \in p} Z_j$$

They commute between themselves and each other, since A_v and B_p have either zero or two pairs of $X_i Z_i$ in common and

$$(X_i Z_i)(X_j Z_j) = (X_j Z_j)(X_i Z_i)$$

Moreover, $A_v^2 = B_p^2 = I$. There are L^2 operators A and L^2 operators B , but they are not all independent. In fact

$$\prod_v A_v = \prod_p B_p = I$$

All X and Z appear twice in the product. So the independent combinations are $L^2 - 1$ for each type of operator. The common eigenspace (the codewords)

$$C = \{|\psi\rangle \in \mathcal{H} \mid A_v |\psi\rangle = B_p |\psi\rangle = +|\psi\rangle\}$$

has dimension

$$2^{2L^2} / (2^{L^2-1} \cdot 2^{L^2-1}) = 4$$

One can encode two qubits (logical qubits). For the surface code there is just one logical qubit. One may identify a state in the code as usual

$$|\bar{0}\bar{0}\rangle = \prod_v \frac{1 + A_v}{\sqrt{2}} |0 \cdots 0\rangle$$

It is eigenstate of both A_v and B_p . In fact, for the operators A_v holds

$$A_v^2 = 1 \implies A_v \frac{1 + A_v}{\sqrt{2}} = \frac{1 + A_v}{\sqrt{2}}$$

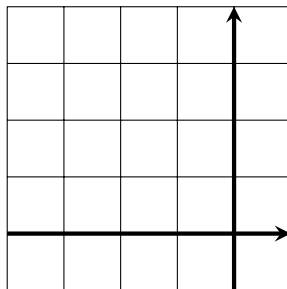
while for the operators B_p holds

$$[B_p, A_v] = 0 \implies B_p |\bar{0}\bar{0}\rangle = \prod_v \frac{1 + A_v}{\sqrt{2}} B_p |0 \cdots 0\rangle = |\bar{0}\bar{0}\rangle$$

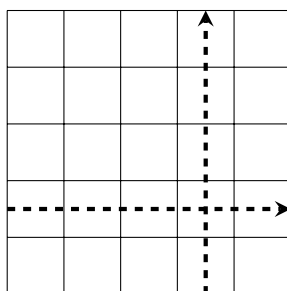
One may find the other logical states by identifying the logical operators

$$\bar{X}_1, \quad \bar{X}_2, \quad \bar{Z}_1, \quad \bar{Z}_2$$

acting on the pair of qubits: these are the operators that commute with all A_v and B_p . The logical \bar{Z}_1 and \bar{Z}_2 are loops of Z s acting on the links, where qubits live (they are the products of such Z s). The operator \bar{Z}_1 is a vertical loop, while \bar{Z}_2 is a horizontal loop (position does not matter)



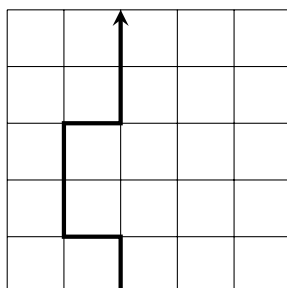
The two operators commute with A_v and B_p . With B_p is straightforward; while with A_v , one may notice that they share zero or two pairs (XZ) . The logical \bar{X}_1 and \bar{X}_2 are dual loops of X s passing through the links. The operator \bar{X}_1 is a horizontal dual loop, while \bar{X}_2 is a vertical dual loop



By applying \bar{X}_1 and \bar{X}_2 to $|\bar{0}\bar{0}\rangle$, one can reconstruct the other states in the code C

$$|\bar{1}\bar{0}\rangle = \bar{X}_1 |\bar{0}\bar{0}\rangle$$

and so on. It would seem that there are many options for logical operators. For example

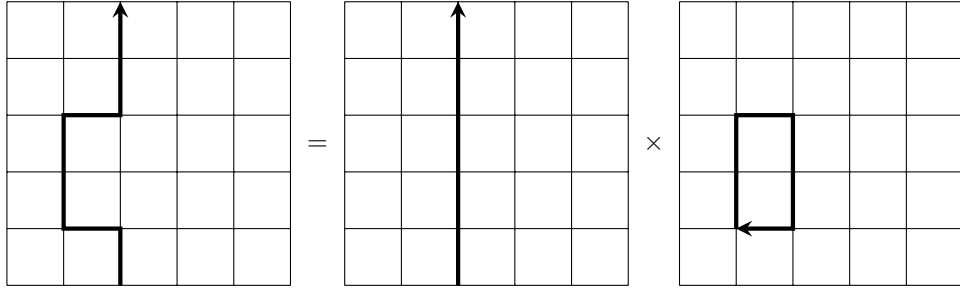


this also commutes with all A_v and B_p . However, homotopic¹⁵ loops are equivalent when acting on the code

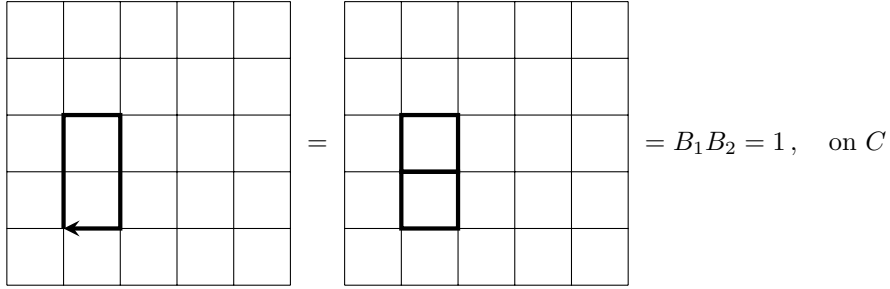
$$C = \{A_v = B_p = 1\}$$

¹⁵Two mathematical objects are said to be homotopic if one can be continuously deformed into the other.

For example



since $Z_i^2 = I$ and



The second loop is a product of Z_i along the links. Therefore, the example operator above is equivalent to a straight line. More generally, homotopic loops and dual loops differ by products of A_v or B_p that are 1 on the code C . Loops that can be deformed with continuity one into another are equivalent on C : the toric code is topological. A closed contour is homotopic to zero and it corresponds to the identity acting on C .

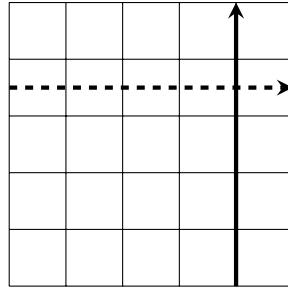
Exercise. Check that $\bar{X}_1, \bar{X}_2, \bar{Z}_1, \bar{Z}_2$ have the right commutation relations. The first

$$[\bar{Z}_1, \bar{Z}_2] = [\bar{X}_1, \bar{X}_2] = 0$$

is straightforward since \bar{Z} contains only Z s and so on. The second

$$\bar{Z}_1 \bar{X}_1 = -\bar{X}_1 \bar{Z}_1$$

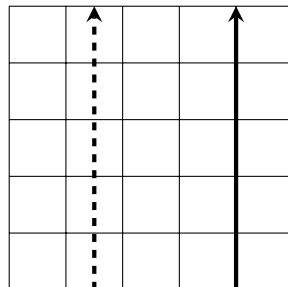
holds since the two operators are products of Z_i and X_i , and they only have one site in common



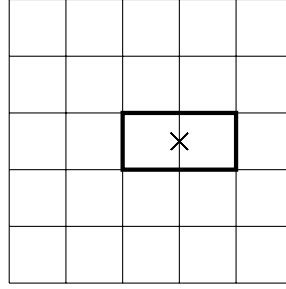
The third is

$$\bar{Z}_1 \bar{X}_2 = \bar{X}_2 \bar{Z}_1$$

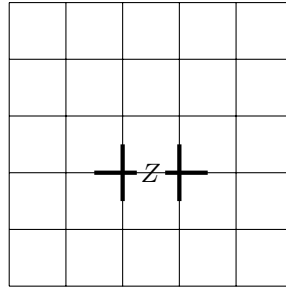
since they act on different qubits, which means that there is no link in common



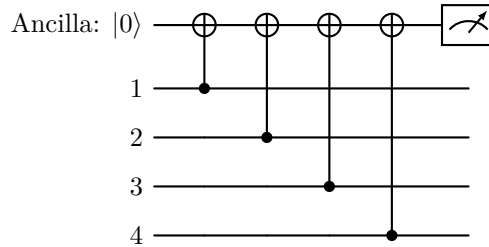
Error correction. Suppose that one error occurs: a bit flip (X acting on a link) is



The errors are the operators that anti-commute with the stabilizers. The operator X commutes with all A_v and anti-commutes with the plaquettes B_p containing the link: the operator A_v remains 1, and exactly two adjacent B_p change sign. The error and its position are so identified. A phase flip (Z acting on a link) is



All B_p remain 1 and exactly two adjacent A_v change sign. Bit-phase flip is the combination of the two. Once one recognizes the error, one can correct it. The analysis and error detection can be done efficiently using ancilla qubits sitting on the vertices (X -type) and at the center of the faces (Z -type). The ancilla qubits are used to measure B_p (Z -type) and A_v (X -type). See fig. 1 in this reference¹⁶. More precisely, in each plaquette, the four Z operators measure $B_p = Z_1 Z_2 Z_3 Z_4$ for the corresponding face. This can be done easily with the circuit



The result of the measurement is $|0\rangle$ or $|1\rangle$ which respectively correspond to $B_p = +1$ or $B_p = -1$. If the qubits are an even number of $|1\rangle$, then $B_p = +1$ and the CNOT acts an even number of times, doing nothing. Similarly, around each vertex, the four operators X measure $A_v = X_1 X_2 X_3 X_4$ for the corresponding vertex. The circuit is done by adding a few Hadamard gates.

The ancilla qubits allow to perform the syndrome measurement. If one finds the wrong values of B (A) for adjacent plaquettes (vertices), then one knows that a simple error has occurred and one can correct it by applying X (Z) to the qubit. Other results for B and A correspond to multiple errors. Consider the case where two B s are -1 , but not adjacent. This error can occur for a string of X -errors connecting two plaquettes with $B = -1$, but one cannot exactly say which string occurred because all open paths connecting the two plaquettes have the same effect: the only B s that change are the endpoints, while the ones along the error path are connected to two errors each, so they change twice back to $+1$. The error can be corrected by applying a string

¹⁶Fowler, A., et al. "Surface codes: Towards practical large-scale quantum computation", in Physical Review A, vol. 86, no. 3, 2012. arXiv: 1208.0928. DOI: <https://doi.org/10.48550/arXiv.1208.0928>.

of X s along any open path between the plaquettes. Indeed, the composition of the error and the correction is a closed contractible path (in the dual lattice), which can be written as the product of all A_v of the vertices enclosed by the closed path, since the internal X s in the product come up in pairs and give the identity. Since all $A_v = 1$ on the code C , the composition of error and correction is trivial on C : any correction fixes the mistake. Again, the toric code is topological. All open paths with the same endpoints which are homotopic are equivalent. Contractible loops in the dual lattice are trivial on C .

Single errors are local. Strings of errors are more difficult to generate and less likely to occur. Errors in the code C (e.g. mistaking $|\bar{1}0\rangle$ for $|\bar{0}0\rangle$) are even more difficult: one needs to create a string of errors that goes all along the torus, one needs L of them. This is suppressed for large L : the toric code is good for fault-tolerant computations.

Statistical mechanical interpretation. There is a nice statistical mechanical interpretation of the toric code. Consider the toric code as a real lattice model of spins with Hamiltonian

$$H = -J \sum_v A_v - J \sum_p B_p, \quad J > 0$$

where the sums are over all vertices and plaquettes. It is an Ising-like model. The ground states (i.e. the states of minimal energy) have all $A_v = B_p = 1$ and correspond to the code C , four-fold degenerate (the logical states). Excitations correspond to errors. In terms of the original spin variables, the ground states are highly entangled:

$$|\bar{0}\bar{0}\rangle = \prod_v \frac{1 + A_v}{\sqrt{2}} |000 \dots 0\rangle$$

corresponding to a superposition of loops made of spins that are flipped (with respect to the “all up” configuration).

Errors are operators acting on the links. A local operator Z flips the sign of the eigenvalue of two adjacent A_v , this costs $2J$ in energy and can be interpreted as creating two “electric” quasi-particles. Similarly, a continuous string of Z s creates two quasi-particles at a distance with the same cost in energy. The local operator X similarly creates pairs of “magnetic” quasi-particles with energy $2J$. Notice that one can always create pairs of quasi-particles and that one can move them around with no extra cost in energy.

These quasi-particles have peculiar properties. Consider a pair of adjacent electric quasi-particles and a pair of adjacent magnetic ones. Now loop one of the electric ones around one of the magnetic quasi-particles by adding a closed loop of Z s and bring it back to the original point. Ordinarily, without X , the closed loop of Z would act trivially. But now, it intersects a site with an action of X , and $XZ = -ZX$, and the state gets a phase $-1 = e^{i\pi}$. This is an Aharonov–Bohm effect: after a closed contour, one goes back to the initial state with an extra phase. It is said that the two types of quasi-particles have mutually non-trivial statistics.

The robustness of the toric code against errors translates into a robustness of the ground states against perturbations. The four-fold degeneracy cannot be lifted by any local perturbation, and one cannot even distinguish the ground states with local operators O (i.e. finite products of X, Y, Z in a small region of the lattice) since

$$\langle \bar{X}\bar{Y} | O | \bar{X}'\bar{Y}' \rangle = \delta_{X X'} \delta_{Y Y'}, \quad X, Y = 0, 1$$

since the operator O (unless it winds along the torus) is either the identity on C or an error which maps a state in C into something orthogonal.

Conclusions. The toric code is an example of a phase of matter with topological order. These phases are characterized by long-range entanglement (LRE) and a set of common properties:

- A robust ground state degeneracy on compact manifolds (e.g. a torus). This is associated with the topology of the manifold.
- Non-local excitations with non-local properties, like non-trivial statistics.

- The low-energy theory on a compact space depends only on the topology: a topological quantum field theory (TQFT). For example, the amplitudes of processes depend only on the topology of the particle history, not on shape or velocity. TQFT, introduced by Witten, captures the degeneracy of the ground states and the properties of quasi-particles.

5.7 Topological quantum computing

This discussion is just a sketch of the idea. Consider again the toric code. By encircling a magnetic quasi-particle with an electric one, one obtains a phase $-1 = e^{i\pi}$

$$|\psi\rangle \rightarrow -|\psi\rangle$$

The action of encircling corresponds to exchanging twice the quasi-particles. Now consider a condensed matter system with a set of quasi-particles such that the state of four¹⁷ quasi-particles is degenerate and has excitations of the form

$$\{\alpha |\psi_1\rangle + \beta |\psi_2\rangle\}$$

This degenerate state has two quantum states $|\psi_1\rangle$ and $|\psi_2\rangle$ corresponding to the same four particles in the same position. Then, when one exchanges two of them

$$|\psi\rangle = \alpha |\psi_1\rangle + \beta |\psi_2\rangle \rightarrow |\psi'\rangle = \alpha' |\psi_1\rangle + \beta' |\psi_2\rangle$$

with

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \rightarrow \begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} = U \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

one returns to the initial condition, but the state can be transformed into another one in the same degenerate Hilbert space. The only quantum mechanical constraint is that U is unitary. These states have a non-abelian statistic (the toric code has just one state: abelian statistic corresponds to a phase). When the particles are identical, they are not fermions nor bosons. They are called anyons. They can exist in two dimensions (not in three dimensions — in three dimensions a double exchange must be trivial because of topology, so a single exchange can give only ± 1). Non-abelian anyons can be used for quantum computations: just arrange an array of quasi-particles and braid them, each braid is a non-abelian unitary transformation.

If one can realize enough transformations U using the multiple exchanges, then one can perform computations. These are robust and fault-tolerant because noise and other local disturbances are unlikely to locally affect the world line of the particles: to make an error they should unwind a braiding and this globally affects the process. The most common type of (non-abelian) anyons are:

- The Ising anyons. They are not universal in the sense of quantum computing, but appear theoretically in many condensed matter systems:
 - ◊ Fractional quantum Hall effect ($\nu = 5/2$, $SU(2)_2$, TQFT);
 - ◊ Majorana zero-modes in two-dimensional topological superconductors (the simplest model is the one-dimensional Kitaev chain);

These are difficult to realize and handle experimentally.

- The Fibonacci anyons, which are universal. They arise in theory models that generalize the toric code, for example, and perhaps in the fractional quantum Hall effect at $\nu = 12/5$ ($SU(3)_3$).

The two most prominent figures who proposed and initiated the field of topological quantum computing are:

- Alexei Kitaev, who proposed in 1997 a topological quantum computer with anyons.

¹⁷Consider an even number of them, e.g. four, because in these models they are usually created in pairs.

- Michael Freedman, a Fields medalist for the Poincaré conjecture in four dimensions in 1986, who was interested in computing knots invariants (Jones polynomial) — an exponentially difficult problem in classical computing — with quantum physics when he learned that topological QFT computes the same. He leads the Microsoft Station Q in Santa Barbara, devoted to topological quantum computing. They announced to have seen and created experimentally Majorana zero-modes in 2018, but his paper was withdrawn in 2021.

Lecture 14

mer 20 nov
2024 16:30

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

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

where $\boldsymbol{\mu}$ is the magnetic moment, which is proportional to the spin. 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:

- **Ion traps.** 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 temperatures, act coherently and quantum due to superconductivity, and are acted upon by microwave drives and resonators (a cavity). The circuits 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 with an oscillating object at a frequency ω :

$$\ddot{x} + \omega^2 x = 0 \implies x(t) = Ae^{i\omega t} + Be^{-i\omega t}$$

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 $[x, p] = i\hbar$. One may introduce the ladder operators:

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

or equivalently:

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

The Hamiltonian becomes

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

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 $|0\rangle \rightarrow |1\rangle$. 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 $|1\rangle \rightarrow |2\rangle$, $|2\rangle \rightarrow |3\rangle$, 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.

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

Then, just by time evolution

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

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

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

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

$$\begin{aligned} |00\rangle_L &\rightarrow |00\rangle_L \\ |01\rangle_L &\rightarrow |01\rangle_L \\ |10\rangle_L &\rightarrow \frac{|4\rangle - |1\rangle}{\sqrt{2}} = |11\rangle_L \\ |11\rangle_L &\rightarrow \frac{|4\rangle + |1\rangle}{\sqrt{2}} = |10\rangle_L \end{aligned}$$

The last two states are mapped into one another. This is precisely the CNOT gate. The problem is that the states $|nm\rangle_L$ 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 can be described as 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

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

with solution

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

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

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

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

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

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. One knows how to quantize each of them: one substitutes the Fourier coefficients $A_{\mathbf{k}i}$ with the ladder operators $a_{\mathbf{k}i}$ and $a_{\mathbf{k}i}^\dagger$. The Hamiltonian of each is

$$H = \hbar \omega_{\mathbf{k}i} \left[a_{\mathbf{k}i}^\dagger a_{\mathbf{k}i} + \frac{1}{2} \right]$$

and the spectrum is $\{|n_{\mathbf{k}i}\rangle\}$.

The Hilbert space of the electromagnetic field is the tensor product of the Hilbert space of the single oscillators

$$\mathcal{H}_{\text{em}} = \bigotimes_{\mathbf{k}, i} \mathcal{H}_{\mathbf{k}, i}$$

called Fock space. The Hamiltonian of the electromagnetic field is

$$H_{\text{em}} = \sum_{\mathbf{k}, i} \hbar \omega_{\mathbf{k}i} \left[a_{\mathbf{k}i}^\dagger a_{\mathbf{k}i} + \frac{1}{2} \right]$$

The interpretation of a state $|n_{\mathbf{k}i}\rangle$ is that it is a state with $n_{\mathbf{k}i}$ photons with wave number \mathbf{k} and polarization i .

The ground state $|0\rangle$ is the vacuum: no photons. It has non-zero energy, but the energy of the vacuum is conventional (only the differences in the energy of the states matter): it can be set to zero and the Hamiltonian may be substituted with

$$H = \hbar \omega_{\mathbf{k}i} \left[a_{\mathbf{k}i}^\dagger a_{\mathbf{k}i} + \frac{1}{2} \right] \rightarrow \hbar \omega_{\mathbf{k}i} a_{\mathbf{k}i}^\dagger a_{\mathbf{k}i}$$

From this, the energy $E_1 - E_0 = \hbar \omega_{\mathbf{k}i}$ is the energy of a photon with frequency $\omega_{\mathbf{k}i}$. Similarly, the energy $E_2 - E_0 = 2\hbar \omega_{\mathbf{k}i}$ corresponds to two photons of the same frequency and polarization.

Since, regardless of the subscripts holds

$$|1\rangle = a^\dagger |0\rangle, \quad a|1\rangle = 0, \quad |2\rangle = \frac{1}{\sqrt{2}} a^\dagger |1\rangle$$

then $a_{\mathbf{k}i}^\dagger$ creates a photon of wave number \mathbf{k} and polarization i (i.e. frequency $\omega_{\mathbf{k}i}$), while $a_{\mathbf{k}i}$ destroys a photon with those same properties. The general state of the electromagnetic field is

$$|n_{\mathbf{k}_1, i_1}, n_{\mathbf{k}_2, i_2}, \dots, n_{\mathbf{k}_k, i_k}, \dots\rangle$$

is a collection of photons with various momenta and polarizations. In the present case, only a few modes of the electromagnetic field suffice. Photons are used to excite the energy levels. Their polarization is fixed by the atomic matrix elements and selection rules.

Necessarily then the energy of a photon must be $E_1 - E_0 = \hbar\omega$, where E_0, E_1 are the energies of the two-level qubit. Everything here collapses to just one oscillator, one pair of ladder operators, with Hamiltonian

$$H = \hbar\omega a^\dagger a$$

where the zero-point energy has been absorbed into a redefinition of the reference level of the energy. The field is an operator itself

$$\mathbf{A} = \sum_{\mathbf{k}, i} A_0 \boldsymbol{\varepsilon}_i(\mathbf{k}) a_{\mathbf{k}i} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} + A_0^* \boldsymbol{\varepsilon}_i(\mathbf{k}) a_{\mathbf{k}i}^\dagger e^{-i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$$

In quantum mechanics, one may use the Heisenberg picture ($\hat{O}_H = e^{iHt} \hat{O}_S e^{-iHt}$), where operators depend on time and satisfy the classical equations of motion. In this case, it is obvious that each component of the potential \mathbf{A} satisfies the wave equation

$$\square \mathbf{A} = 0$$

as an operator (notice that a and a^\dagger do not depend on t).

The electric field is

$$\mathbf{E} = -\partial_t \mathbf{A} = \sum_{\mathbf{k}, i} i\omega A_0 \boldsymbol{\varepsilon}_i a_{\mathbf{k}i} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} - i\omega A_0^* \boldsymbol{\varepsilon}_i a_{\mathbf{k}i}^\dagger e^{-i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$$

Most of the time, A_0 is chosen real or imaginary so that

$$\begin{aligned} \mathbf{E} &= \sum_{\mathbf{k}, i} iE_0 [\boldsymbol{\varepsilon}_i a_{\mathbf{k}i} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} - \boldsymbol{\varepsilon}_i a_{\mathbf{k}i}^\dagger e^{-i(\mathbf{k} \cdot \mathbf{x} - \omega t)}], \quad E_0 = \omega A_0, \quad A_0 \text{ real} \\ \mathbf{E} &= \sum_{\mathbf{k}, i} E_0 [\boldsymbol{\varepsilon}_i a_{\mathbf{k}i} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} + \boldsymbol{\varepsilon}_i a_{\mathbf{k}i}^\dagger e^{-i(\mathbf{k} \cdot \mathbf{x} - \omega t)}], \quad E_0 = i\omega A_0, \quad A_0 \text{ imaginary} \end{aligned}$$

6.3 Driven qubit and Rabi oscillations

Qubits can be driven and controlled by coupling them to external fields (spins with magnetic fields, atoms with electromagnetic fields, etc). The external field can be treated most of the time as a classical external field (lasers emit coherent radiation; one should use coherent states — see later subsection — but, in good approximation, they behave classically, so the classical description is good enough).

The qubit can be modeled as a two-level system $|0\rangle, |1\rangle$ interacting with the electromagnetic field with the typical atomic dipole interaction

$$H_{\text{int}} = -\mathbf{d} \cdot \mathbf{E}$$

The spatial dependence of the electric field \mathbf{E} is typically irrelevant: its wavelength is much greater than the size of the qubit. For example, the visible spectrum is of the order $\lambda \sim 10^{-7}$ m, while microwaves are of the order of centimeters to meters, both of which are much greater than the Bohr radius 10^{-10} m. So one may just eliminate the dependence on the position \mathbf{x} by evaluating the electric field $\mathbf{E}(\mathbf{x})$ at the position of the qubit. The electric field \mathbf{E} does not vary

on the size of the qubit. Then, if the frequency of the signal is ω_d (d for drive), one can take, with no loss of generality,

$$\mathbf{E} = \mathbf{E}_0 \cos(\omega_d t + \phi_0)$$

In an atomic qubit, the dipole is $\mathbf{d} = e\mathbf{x}$, where \mathbf{x} is position of the qubit. One is only interested in the Hamiltonian restricted to the levels $|0\rangle, |1\rangle$

$$\langle w | H_{\text{int}} | w' \rangle, \quad w, w' = 0, 1$$

For an atom, one computes the position operator on its states $\langle w | \mathbf{x} | w' \rangle$: the matrix element depends on the levels, selection rules, etc. Also, the coefficient \mathbf{E}_0 can be oriented arbitrarily. Depending on all these details, the matrix element $\langle w | \mathbf{E}_0 \cdot \mathbf{x} | w' \rangle$ can be an arbitrary Hermitian two-by-two matrix that may be written by expanding it in Pauli matrices

$$H_{\text{coupling}} = -(aI + b_i \sigma_i) \cos(\omega_d t + \phi_0)$$

The important point is that this is the most general form and it applies not only to atoms and dipole interaction, but to everything: for example, spins couple to magnetic fields, precisely as

$$H_{\text{coupling}} = -\boldsymbol{\mu} \cdot \mathbf{B}(t) \sim \boldsymbol{\sigma} \cdot \mathbf{B}_0 \cos(\omega_d t + \phi_0)$$

The free Hamiltonian also contains the qubit energy term: the state $|1\rangle$ has energy E_1 , the state $|0\rangle$ has energy E_0 , with the difference being $E_1 - E_0 = \hbar\omega_q$ (q for qubit). By convention, the state $|1\rangle$ is the excited state, and the reference value of energy is fixed in the middle:

$$E_1 = \frac{1}{2}\hbar\omega_q, \quad E_0 = -\frac{1}{2}\hbar\omega_q$$

So that

$$H_0 = -\frac{\hbar}{2}\omega_q \sigma_3 = -\frac{\omega_q}{2}\sigma_3, \quad \hbar = 1$$

The minus sign is there because $|1\rangle$ is the excited state. The entire Hamiltonian is

$$H = -\frac{\omega_q}{2}\sigma_3 - (aI + b_i \sigma_i) \cos(\omega_d t + \phi_0)$$

The matrices I and σ_3 just change the energy levels separating them further, and are straightforward to study (because they are diagonal on the states). One may focus on σ_1 and σ_2 , setting

$$a = b_3 = 0, \quad b_1 + ib_2 = Ae^{-i\phi_1}, \quad \phi_1, A \in \mathbb{R}$$

Therefore

$$(b_1 \sigma_1 + b_2 \sigma_2) = \begin{bmatrix} 0 & b_1 - ib_2 \\ b_1 + ib_2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & Ae^{i\phi_1} \\ Ae^{-i\phi_1} & 0 \end{bmatrix}$$

from which the Hamiltonian becomes

$$H = -\frac{\omega_q}{2}\sigma_3 - (Ae^{i\phi_1}\sigma_+ + Ae^{-i\phi_1}\sigma_-) \cos(\omega_d t + \phi_0)$$

where

$$\sigma_+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \frac{\sigma_1 + i\sigma_2}{2}, \quad \sigma_- = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \frac{\sigma_1 - i\sigma_2}{2}$$

The time evolution associated with the Hamiltonian H is a problem with no simple analytical solution. It can be attacked by solving differential equations or going to a “rotating frame”, as done in the following, and using the rotating wave approximation (RWA).

Going to a “rotating frame” simply means to change basis with a unitary time-dependent transformation

$$|\psi(t)\rangle \rightarrow |\tilde{\psi}(t)\rangle = U(t) |\psi(t)\rangle, \quad U^\dagger(t)U(t) = I$$

where $U(t)$ is not the time evolution operator. In some cases, it is related. The interaction picture in quantum mechanics is an example:

$$U(t) = e^{iH_0 t}$$

where H_0 is a time-independent unperturbed Hamiltonian (but notice the positive sign at the exponent). Probabilities are the same in all frames. The change of frame gives a new Hamiltonian. In fact, from the Schrödinger equation

$$i d_t |\psi(t)\rangle = H |\psi(t)\rangle$$

one finds

$$i d_t |\tilde{\psi}(t)\rangle = U i d_t |\psi(t)\rangle + i \dot{U} |\psi(t)\rangle = (UH + i\dot{U}) |\psi(t)\rangle = (UHU^\dagger + i\dot{U}U^\dagger)U |\psi(t)\rangle \equiv \tilde{H} |\tilde{\psi}(t)\rangle$$

Therefore

$$\tilde{H} = UHU^\dagger + i\dot{U}U^\dagger$$

The typical trick to evaluate \tilde{H} is using the identity

$$[T, A] = \alpha A \implies e^{iT} A e^{-iT} = e^{i\alpha} A$$

Proof. The proof is as follows. Let

$$A(t) = e^{itT} A e^{-itT}$$

Then

$$\dot{A}(t) = iT A(t) - i A(t) T = i e^{itT} [T, A] e^{-itT} = i \alpha e^{itT} A e^{-itT} = i \alpha A(t)$$

which can be solved to obtain

$$A(t) = A(0) e^{i\alpha t}$$

But

$$A(1) = e^{iT} A e^{-iT} = e^{i\alpha} A$$

since $A(0) = A$. □

One may change frame with $U(t) = e^{-\frac{i}{2}\omega_d \sigma_3 t}$ to rewrite the Hamiltonian H . One finds

$$\begin{aligned} U \sigma_3 U^\dagger &= \sigma_3 \\ U \sigma_+ U^\dagger &= e^{-i\omega_d t} \sigma_+ \\ U \sigma_- U^\dagger &= e^{i\omega_d t} \sigma_- \\ i \dot{U} U^\dagger &= \frac{\omega_d}{2} \sigma_3 \end{aligned}$$

since

$$[\sigma_3, \sigma_\pm] = \pm 2\sigma_\pm$$

Finally

$$\tilde{H} = -\frac{\omega_q - \omega_d}{2} \sigma_3 - (A e^{-i(\omega_d t - \phi_1)} \sigma_+ + A e^{i(\omega_d t - \phi_1)} \sigma_-) \frac{e^{i(\omega_d t + \phi_0)} + e^{-i(\omega_d t + \phi_0)}}{2}$$

which cannot yet be solved exactly. One may use the rotating wave approximation (after expanding out the product). It amounts to neglecting the fast oscillating terms in the Hamiltonian H since one expects that they average to zero¹⁸ (this can be checked by using time-dependent perturbation theory: resonant terms are more important because of denominators of the type $1/(E_f - E_i \pm \hbar\omega)$). Thus

$$\tilde{H} \sim -\frac{\Delta}{2} \sigma_3 - \frac{1}{2} (A e^{i\phi} \sigma_+ + A e^{-i\phi} \sigma_-), \quad \phi = \phi_0 + \phi_1$$

where $\Delta = \omega_q - \omega_d$ is the detuning frequency. Notice how the phase ϕ_0 is absorbed in the coefficient $A e^{i\phi}$. Since \tilde{H} is now time-independent (in the rotating frame), the states evolve with $e^{-i\tilde{H}t}$. Writing

$$\tilde{H} = -\frac{1}{2} (\Delta \sigma_3 + A \cos \phi \sigma_1 - A \sin \phi \sigma_2) \equiv -\frac{\Omega}{2} \mathbf{n} \cdot \boldsymbol{\sigma}$$

¹⁸Thanks to the Riemann–Lebesgue lemma.

with the Rabi frequency

$$\Omega = \sqrt{A^2 + \Delta^2}$$

and

$$\mathbf{n} = \left(\frac{A}{\Omega} \cos \phi \quad -\frac{A}{\Omega} \sin \phi \quad \frac{\Delta}{\Omega} \right), \quad \mathbf{n} \cdot \mathbf{n} = 1$$

the time evolution in the rotating frame is just a “rotation” of the qubit on the Bloch sphere

$$e^{-i\tilde{H}t} = e^{\frac{i}{2}\Omega(\mathbf{n} \cdot \boldsymbol{\sigma})t} = R_{\mathbf{n}}(-\Omega t)$$

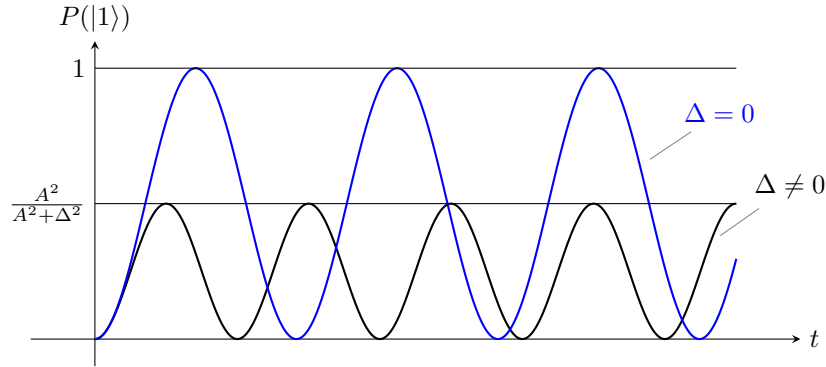
In particular, if one starts with the state $|0\rangle$:

$$|\tilde{\psi}(t=0)\rangle = |0\rangle$$

then the probability of being in the state $|1\rangle$ at time t is

$$\begin{aligned} P(|1\rangle) &= \left| \langle 1 | e^{-i\tilde{H}t} | 0 \rangle \right|^2 = \left| \langle 1 | \left(I \cos \frac{\Omega t}{2} + i(\mathbf{n} \cdot \boldsymbol{\sigma}) \sin \frac{\Omega t}{2} \right) | 0 \rangle \right|^2 \\ &= \left| i(n_1 + in_2) \sin \frac{\Omega t}{2} \right|^2 = \left| i \frac{A}{\Omega} e^{-i\phi} \sin \frac{\Omega t}{2} \right|^2 \\ &= \frac{A^2}{A^2 + \Delta^2} \sin^2 \left[\frac{\sqrt{A^2 + \Delta^2}}{2} t \right] \end{aligned}$$

The probability oscillates in time (these are the Rabi oscillations) with an angular frequency $\Omega = \sqrt{A^2 + \Delta^2}$. In the special case $\Delta = 0$ (i.e. at resonance: $\omega_d = \omega_q$), the probability $P(|1\rangle)$ becomes 1 at certain times.



Lecture 15

6.4 Single-qubit control

lun 25 nov
2024 16:30

As one has seen, by adding to a qubit Hamiltonian

$$H = -\frac{\omega_q}{2} \sigma_3$$

a perturbation along σ_1 and σ_2 , one can operate time-dependent rotations:

$$|\psi(t)\rangle = U^\dagger(t) |\tilde{\psi}(t)\rangle = e^{\frac{i}{2}\omega_d \sigma_3 t} R_{\mathbf{n}}(-\Omega t) |\tilde{\psi}(0)\rangle$$

where ω_d is the external frequency of the perturbation, $\Omega = \sqrt{A^2 + \Delta^2}$ and \mathbf{n} can be read from the amplitude and phase of the perturbation, and the detuning frequency is

$$\Delta = \omega_q - \omega_d$$

This can be visualized as a rotation on the Bloch sphere. Recall that $R_{\mathbf{n}}(\gamma)$ acts by rotating a vector on the Bloch sphere by an angle γ around the direction \mathbf{n} . For example, for

$$R_z(\gamma) = e^{-i\frac{\gamma}{2}\sigma_3}, \quad |\psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle$$

one finds

$$R_z(\gamma) |\psi\rangle = \cos \frac{\theta}{2} e^{-i\frac{\gamma}{2}} |0\rangle + e^{i\phi+i\frac{\gamma}{2}} \sin \frac{\theta}{2} |1\rangle = e^{-i\frac{\gamma}{2}} \left[\cos \frac{\theta}{2} |0\rangle + e^{i(\phi+\gamma)} \sin \frac{\theta}{2} |1\rangle \right]$$

where the overall phase can be ignored. On the Bloch sphere, the vector goes from (θ, ϕ) to $(\theta, \phi + \gamma)$, it is a rotation about the z axis by an angle γ .

There is always a rotation about z (a precession) due to H_0 , even if there is no perturbation ($A = 0$):

$$|\psi(t)\rangle = e^{\frac{i}{2}\omega_d\sigma_3 t} e^{\frac{i}{2}\Delta\sigma_3 t} |\tilde{\psi}(0)\rangle = e^{\frac{i}{2}\omega_q\sigma_3 t} |\tilde{\psi}(0)\rangle$$

This has frequency ω_q , the natural frequency of a qubit. Sometimes it is useful to go to the qubit's rotating frame with $U(t) = e^{iH_0 t}$ to eliminate this effect. There are many choices of useful reference frames.

The important point is that, with a perturbation, one can realize evolutions induced by an effective Hamiltonian $H = \alpha(\mathbf{n} \cdot \boldsymbol{\sigma})$ with any α and \mathbf{n} just by choosing the amplitude, phase and time t of the evolution. One may rotate and control a single qubit. One may create in this way all single-qubit gates.

One may interpret Rabi oscillations in another way. Consider the state $|0\rangle$ and the vector

$$\mathbf{n} = \left(\frac{A}{\Omega} \cos \phi \quad -\frac{A}{\Omega} \sin \phi \quad \frac{\Delta}{\Omega} \right)$$

If there is resonance $\Delta = 0$, meaning $\omega_d = \omega_q$, the axis \mathbf{n} lies in the (x, y) plane. The state $|0\rangle$ moves along a great circle and, at a certain time, it becomes $|1\rangle$, i.e. the probability $P(|1\rangle)$ can be 1. If there is no resonance, $\Delta \neq 0$, the state moves along a non-great circle since \mathbf{n} does not lie in the (x, y) plane and the state can never pass through $|1\rangle$, i.e. the probability $P(|1\rangle)$ is always strictly less than 1. One may also realize the evolution by repeatedly applying short pulses, or by modulating the signal

$$f(t) \cos(\omega_d t + \phi)$$

which can be studied using time-dependent perturbation theory.

6.5 Cavity QED

Previously, the electromagnetic field has been treated classically. One may upgrade to a quantum description¹⁹. In quantum optics, one creates Fabry–Pérot cavities with a high-intensity field and just a few modes with frequency ω_c (c stands for cavity) or multiples. An atom/qubit in the cavity can exchange and absorb photons if $\omega_q \approx \omega_c$. If photons cannot leave (if they have the right frequency they are reflected by the walls), they form a quantum system. Of course, spontaneously emitted photons can escape, and there is always some loss and decoherence.

The full Hamiltonian can be written by coupling the atom/qubit to the quantized electromagnetic field. One expects a dipole interaction

$$H_{\text{int}} = -\mathbf{d} \cdot \mathbf{E}$$

where the electric dipole \mathbf{d} depends on the actual interaction and provides the matrix elements for the transitions $|0\rangle \rightarrow |1\rangle$ and $|1\rangle \rightarrow |0\rangle$; and the electric field \mathbf{E} for a single electromagnetic mode is proportional to $(a + a^\dagger)$. In the usual dipole approximation, where the wavelength of the radiation λ_{rad} is far greater than the size of the qubit, there is no spatial dependence, and one may model the interaction with

$$H = \omega_c \left(a^\dagger a + \frac{1}{2} \right) - \frac{1}{2} \omega_q \sigma_3 + g(a + a^\dagger) \sigma_1$$

where g is a coupling constant and one chooses the direction x (σ_1) for simplicity. All other choices give similar results. The difference with the previous section is that the electric field $E \sim (a + a^\dagger)$ is now quantized. Its dependence on time is removed by going to the Heisenberg picture.

¹⁹Note that lasers are still best described by a classical Hamiltonian since the photons in a laser are in coherent states.

Consider $\sigma_1 = \sigma_+ + \sigma_-$ with

$$\sigma_+ = \frac{\sigma_1 + i\sigma_2}{2} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \sigma_- = \frac{\sigma_1 - i\sigma_2}{2} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

corresponding to the qubit processes

$$\sigma_+ |1\rangle = |0\rangle, \quad \sigma_- |0\rangle = |1\rangle$$

where the \pm are not intuitively placed because one has chosen $|0\rangle$ to be the ground state (i.e. spin up). The four terms

$$g(a^\dagger \sigma_+ + a \sigma_- + a \sigma_+ + a^\dagger \sigma_-)$$

correspond to the processes:

- $a^\dagger \sigma_+$ is the emission of a photon and decay $|1\rangle \rightarrow |0\rangle$;
- $a \sigma_-$ is the absorption of a photon and an excitation $|0\rangle \rightarrow |1\rangle$;
- $a^\dagger \sigma_-$ is the emission of a photon with an excitation;
- $a \sigma_+$ is the absorption of a photon with a decay.

The last two processes are unlikely to occur. Even in resonance, they violate energy conservation:

$$\Delta E \sim \pm(\omega_c + \omega_q) \sim \pm 2\omega_c$$

At first order in perturbation theory, they are forbidden by the Fermi golden rule. In an exact treatment, they have a very small probability, at least if ω_c and ω_q are not too far apart. One may drop them. This is again the rotating wave approximation (RWA), which can also be seen by going into a rotating frame. This time, it is convenient to use

$$U(t) = e^{iH_0 t}$$

with the free Hamiltonian

$$H_0 = \omega_c \left(a^\dagger a + \frac{1}{2} \right) - \frac{1}{2} \omega_q \sigma_3$$

This is called the interaction picture in perturbation theory. Splitting the Hamiltonian into

$$H = H_0 + H_1$$

then the Hamiltonian in the rotating frame is given only by the interaction

$$\tilde{H} = U H U^\dagger + i\dot{U} U^\dagger = H_0 + U H_1 U^\dagger - H_0 = U H_1 U^\dagger$$

Using the identity to evaluate left-right conjugation

$$[H_0, a^\dagger] = \omega_c a^\dagger, \quad U a^\dagger U^\dagger = e^{i\omega_c t} a^\dagger, \quad [H_0, \sigma_+] = -\omega_q \sigma_+, \quad U \sigma_+ U^\dagger = e^{-i\omega_q t} \sigma_+$$

one finds

$$\tilde{H} = g[a^\dagger \sigma_+ e^{i(\omega_c - \omega_q)t} + a \sigma_- e^{-i(\omega_c - \omega_q)t} + a \sigma_+ e^{-i(\omega_c + \omega_q)t} + a^\dagger \sigma_- e^{i(\omega_c + \omega_q)t}]$$

where for $\omega_c \approx \omega_q$ the terms $a^\dagger \sigma_+$ and $a \sigma_-$ oscillate slowly (close to resonance), and the terms $a \sigma_+$ and $a^\dagger \sigma_-$ oscillate rapidly as $e^{-2i\omega_c t}$ and $e^{2i\omega_c t}$ respectively. The rapid oscillations are dropped in the rotating wave approximation. The final result is the Jaynes–Cummings Hamiltonian:

$$H = \omega_c \left(a^\dagger a + \frac{1}{2} \right) - \frac{\omega_q}{2} \sigma_3 + g(a^\dagger \sigma_+ + a \sigma_-)$$

where, after dropping the aforementioned terms, one rotates back to the original frame. This Hamiltonian: appears not only in QED, but in all cases where a qubit interacts with a quantized mode of the electromagnetic field: all interactions $\mathbf{E} \cdot \mathbf{d}$, $\boldsymbol{\mu} \cdot \mathbf{B}$, etc, lead to the same model.

One may explicitly diagonalize the Hamiltonian H , because it is block-diagonal. To avoid confusion, let $|g\rangle = |0\rangle$ and $|e\rangle = |1\rangle$ be the ground state and excited state of a qubit, and $|n\rangle$ be the state of the electromagnetic field. The block structure is due to the fact that

$$a^\dagger \sigma_+ |e, n\rangle \propto |g, n+1\rangle, \quad a^\dagger \sigma_+ |g, n+1\rangle = 0, \quad a \sigma_- |g, n+1\rangle \propto |e, n\rangle, \quad a \sigma_- |e, n\rangle = 0$$

One may divide the Hilbert space into $|g, 0\rangle$ alone and $\{|e, n\rangle, |g, n+1\rangle\}$ with $n \in \mathbb{N}_0$. So the Hamiltonian becomes

$$H = \frac{1}{2}\omega_c + \begin{bmatrix} -\frac{1}{2}\omega_q & & & \\ & \begin{pmatrix} \omega_c - \frac{1}{2}\omega_q & g \\ g & \frac{1}{2}\omega_q \end{pmatrix} & & \\ & & \ddots & \\ & & & \begin{pmatrix} (n+1)\omega_c - \frac{1}{2}\omega_q & \sqrt{n+1}g \\ \sqrt{n+1}g & n\omega_c + \frac{1}{2}\omega_q \end{pmatrix} \end{bmatrix} \begin{bmatrix} |g, 0\rangle \\ |g, 1\rangle \\ |e, 0\rangle \\ |g, n+1\rangle \\ |e, n\rangle \end{bmatrix}$$

since

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

and

$$\begin{aligned} H_0 |e, n\rangle &= \left[\frac{1}{2}\omega_q + \omega_c \left(n + \frac{1}{2} \right) \right] |e, n\rangle \\ H_0 |g, n+1\rangle &= \left[-\frac{1}{2}\omega_q + \omega_c \left(n + 1 + \frac{1}{2} \right) \right] |g, n+1\rangle \end{aligned}$$

The addendum before the matrix is the zero-point energy. Letting $\Delta = \omega_q - \omega_c$, the generic block on the bottom right-hand side (along with the zero-point energy) is

$$\begin{bmatrix} (n+1)\omega_c - \frac{1}{2}\Delta & \sqrt{n+1}g \\ \sqrt{n+1}g & (n+1)\omega_c + \frac{1}{2}\Delta \end{bmatrix}$$

whose eigenvalues are

$$E_\pm = (n+1)\omega_c \pm \frac{1}{2}\sqrt{\Delta^2 + 4g^2(n+1)}$$

and the eigenvectors are called dressed states

$$|n, -\rangle = \cos \theta_n |g, n+1\rangle - \sin \theta_n |e, n\rangle, \quad |n, +\rangle = \sin \theta_n |g, n+1\rangle + \cos \theta_n |e, n\rangle$$

where the angle is defined by

$$\tan 2\theta_n = \frac{2g\sqrt{n+1}}{\Delta}$$

In addition, there is the state $|g, 0\rangle$ with energy $E_0 = -\Delta/2$.

The time evolution operator e^{-iHt} performs Rabi oscillations on the pairs of dressed states (labeled by n): each pair behaves as a two-level system oscillating coherently in cycles of absorption and emission of photons $|g, n+1\rangle \leftrightarrow |e, n\rangle$. The Rabi frequency of each pair

$$\Omega = \sqrt{\Delta^2 + 4g^2(n+1)}$$

depends on the number of photons. When many different photons n are present, there can be interference (i.e. cancellations) between the oscillations and there can be a restoration of the original pattern at a later time.

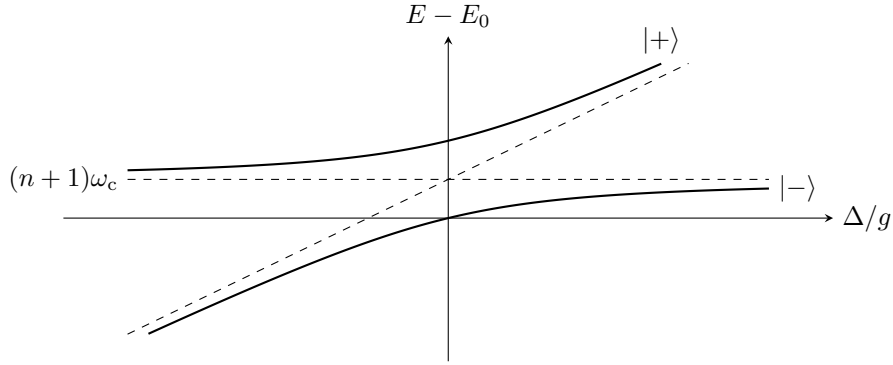
Lecture 16

6.6 Dressed states

Consider the energy of the dressed states as a function of the detuning frequency Δ . The difference with the energy $E_0 = -\Delta/2$ of the ground state $|g, 0\rangle$ is

$$E_\pm - E_0 = (n+1)\omega_c \pm \frac{1}{2}\sqrt{\Delta^2 + 4g^2(n+1)} + \frac{\Delta}{2}$$

mer 27 nov
2024 16:30



There are two interesting regimes.

Resonance. The resonance regime corresponds to $\Delta = 0$ ($\omega_c = \omega_q$). Here

$$\theta_n = \frac{\pi}{4}, \quad \cos \theta_n = \sin \theta_n = \frac{1}{\sqrt{2}}$$

There is maximum hybridization with

$$|n, \pm\rangle = \frac{|g, n+1\rangle \pm |e, n\rangle}{\sqrt{2}}$$

These states are called “polarons”. The splitting in energies is small (and depends on n):

$$E_{\pm} = (n+1)\omega_c \pm g\sqrt{n+1}$$

Dispersive regime. The dispersive regime corresponds to $\Delta \gg g$. Here $\theta_n \ll 1$, and the original states are almost unchanged

$$|n, -\rangle = |g, n+1\rangle + \dots, \quad |n, +\rangle = |e, n\rangle + \dots$$

This regime is near the free case ($g = 0$) with energies

$$E_- = (n+1)\omega_c - \frac{\Delta}{2} = \omega_c \left(n+1 + \frac{1}{2} \right) - \frac{\omega_q}{2}, \quad |g, n+1\rangle$$

$$E_+ = (n+1)\omega_c + \frac{\Delta}{2} = \omega_c \left(n + \frac{1}{2} \right) + \frac{\omega_q}{2}, \quad |e, n\rangle$$

One may take into account also the correction in g/Δ :

$$E_{\pm} = (n+1)\omega_c \pm \frac{1}{2}\Delta\sqrt{1 + \frac{4g^2(n+1)}{\Delta^2}} = (n+1)\omega_c \pm \left[\frac{\Delta}{2} + \frac{g^2(n+1)}{\Delta} + \dots \right]$$

One finds the same energies with the “effective” Hamiltonian:

$$H^{(2)} = \omega_c \left(a^\dagger a + \frac{1}{2} \right) - \frac{\omega_q}{2} \sigma_3 - \frac{g^2}{\Delta} \left(a^\dagger a + \frac{1}{2} \right) \sigma_3 + \frac{g^2}{2\Delta} I_2$$

where I_2 is the 2×2 identity matrix. Indeed

$$H^{(2)} |g, n+1\rangle = \omega_c \left(n+1 + \frac{1}{2} \right) - \frac{\omega_q}{2} - \frac{g^2}{\Delta} (n+1) = (n+1)\omega_c - \left(\frac{\Delta}{2} + \frac{g^2(n+1)}{\Delta} \right)$$

$$H^{(2)} |e, n\rangle = \omega_c \left(n + \frac{1}{2} \right) + \frac{\omega_q}{2} + \frac{g^2}{\Delta} \left(n + \frac{1}{2} \right) + \frac{g^2}{2\Delta} = (n+1)\omega_c + \left(\frac{\Delta}{2} + \frac{g^2(n+1)}{\Delta} \right)$$

This effective Hamiltonian $H^{(2)}$ arises by operating an infinitesimal change of reference frame and expanding up to $O(g^2/\Delta^2)$ the transformation operator

$$U = e^{-\frac{g}{\Delta}(a^\dagger \sigma_+ - a \sigma_-)}$$

This is a Schrieffer–Wolff transformation (see next section):

$$U = e^S, \quad S^\dagger = -S, \quad [S, H_0] = -H_{\text{int}} \implies U(H_0 + H_{\text{int}})U^\dagger = H_0 + \frac{1}{2}[S, H_{\text{int}}] + O(S^3)$$

Up to constant additive factors, the effective Hamiltonian is

$$H^{(2)} = (\omega_c - \chi\sigma_3) a^\dagger a - \frac{\tilde{\omega}_q}{2}\sigma_3$$

where $\chi = g^2/\Delta$ and $\tilde{\omega}_q = \omega_q + g^2/\Delta$ is the Lamb shift: the energy of the qubit is shifted, it has been renormalized, by the interaction with the vacuum fluctuations. From the parenthesis, one notices that the cavity has a frequency shifted by $\pm\chi$ depending on the qubit state. This can be used to perform quantum nondemolition (QND) measurements of the state of the qubit using the cavity.

One may write instead

$$H^{(2)} = \omega_c a^\dagger a - \frac{\sigma_3}{2} \left(\omega_q + \frac{g^2}{\Delta} + \frac{2g^2}{\Delta} a^\dagger a \right)$$

The first two addenda in the parenthesis are the Lamb shift, while the third addendum is the Autler–Townes effect (also called alternating current Stark effect). This addendum describes the number of photons that create noise in the frequency of the qubit.

6.7 Schrieffer–Wolff transformation

It is a trick in perturbation theory to write the Hamiltonian as

$$H = H_0 + H_{\text{int}}, \quad H_{\text{int}} \in O(\lambda), \quad \lambda \ll 1$$

The Schrieffer–Wolff transformation is a change of frame leading to the Hamiltonian

$$\tilde{H} = U H U^\dagger$$

where

$$U = e^S, \quad U^\dagger U = I \implies S^\dagger = -S$$

and

$$[S, H_0] = -H_{\text{int}} \implies S \in O(\lambda)$$

Therefore, expanding U in a Taylor series (or using the Campbell identity), one finds

$$\begin{aligned} \tilde{H} &= H + [S, H] + \frac{1}{2}[S, [S, H]] + \dots \\ &= H_0 + H_{\text{int}} + [S, H_0] + [S, H_{\text{int}}] + \frac{1}{2}[S, [S, H_0]] + \frac{1}{2}[S, [S, H_{\text{int}}]] + \dots \\ &= H_0 + \frac{1}{2}[S, H_{\text{int}}] + O(\lambda^3) \end{aligned}$$

where one uses the commutator above. In the particular case of the last section, one has

$$H_0 = -\frac{\omega_q}{2}\sigma_3 + \omega_c \left(a^\dagger a + \frac{1}{2} \right), \quad H_{\text{int}} = g(\sigma_+ a^\dagger + \sigma_- a), \quad S = -\frac{g}{\Delta}(\sigma_+ a^\dagger - \sigma_- a)$$

Then

$$[S, H_0] = -\frac{g}{\Delta}[(\omega_q - \omega_c)\sigma_+ a^\dagger - (-\omega_q + \omega_c)\sigma_- a] = -g(\sigma_+ a^\dagger + \sigma_- a) = -H_{\text{int}}$$

and

$$\begin{aligned} \frac{1}{2}[S, H_{\text{int}}] &= -\frac{g^2}{2\Delta}[\sigma_+ a^\dagger - \sigma_- a, \sigma_+ a^\dagger + \sigma_- a] = -\frac{g^2}{\Delta}[\sigma_+ a^\dagger, \sigma_- a] \\ &= -\frac{g^2}{\Delta}(\sigma_+ \sigma_- a^\dagger a - \sigma_- \sigma_+ a a^\dagger) = -\frac{g^2}{\Delta}\sigma_3 a^\dagger a + \frac{g^2}{\Delta} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= -\frac{g^2}{\Delta}\sigma_3 a^\dagger a + \frac{g^2}{2\Delta}(I_2 - \sigma_3) \end{aligned}$$

From this one sees that \tilde{H} is precisely the effective Hamiltonian $H^{(2)}$.

6.8 Coherent and squeezed states

A state with n photons is highly non-classical. The electric field

$$E \sim E_0(ae^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} + a^\dagger e^{-i(\mathbf{k} \cdot \mathbf{x} - \omega t)})$$

has zero expectation value, $\langle n|E|n\rangle = 0$, since $\langle n|a|n\rangle = \langle n|a^\dagger|n\rangle = 0$, which is not what one expects classically. The light in a laser, instead, is different, it is coherent and “more classical”. It is idealized with coherent states, superpositions of states with many photons that behave “almost” classically, and with an expectation value $\langle E \rangle$ that looks like a classical oscillating field.

By definition, a coherent state of an anharmonic oscillator $|\alpha\rangle$ is an eigenstate of the annihilation operator a (not of the number operator n):

$$a|\alpha\rangle = \alpha|\alpha\rangle, \quad \alpha \in \mathbb{C}$$

where α is complex because a is not Hermitian.

One would like to find such eigenstate $|\alpha\rangle$. Setting

$$|\alpha\rangle = \sum_{n=0}^{\infty} C_n |n\rangle$$

one has

$$a|\alpha\rangle = \sum_{n=0}^{\infty} C_n a|n\rangle = \sum_{n=1}^{\infty} C_n \sqrt{n} |n-1\rangle = \sum_{n=0}^{\infty} C_{n+1} \sqrt{n+1} |n\rangle \equiv \alpha \sum_{n=0}^{\infty} C_n |n\rangle = \alpha|\alpha\rangle$$

Equating the coefficients of $|n\rangle$ term by term, gives

$$C_{n+1} \sqrt{n+1} = \alpha C_n \implies C_{n+1} = \frac{\alpha}{\sqrt{n+1}} C_n = \dots = \frac{\alpha^{n+1}}{\sqrt{(n+1)!}} C_0$$

so that

$$|\alpha\rangle = \sum_{n=0}^{\infty} C_0 \frac{\alpha^n}{\sqrt{n!}} |n\rangle$$

the constant C_0 is obtained by normalizing the eigenstate $|\alpha\rangle$ to unity

$$1 = \langle \alpha | \alpha \rangle = |C_0|^2 \sum_{n=0}^{\infty} \frac{|\alpha|^{2n}}{n!} = |C_0|^2 e^{|\alpha|^2} \iff C_0 = e^{-\frac{1}{2}|\alpha|^2}$$

It is chosen real since global phases do not matter in quantum mechanics. Therefore

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle$$

Properties of coherent states. A few properties of coherent states are the following.

- Coherent states with $|\alpha| \ll 1$ described attenuated radiation, even single photons. The probability of having n photons is a Poisson distribution

$$P(|n\rangle) = \frac{|\alpha|^{2n}}{n!} e^{-|\alpha|^2}, \quad \langle n \rangle = |\alpha|^2$$

If $|\alpha| \ll 1$, then the average number of photons is small. Already for $\alpha = 1/2$, the average number of photons is $\langle n \rangle = 1/4$, so less than one photon on average. Around 78% of the time, there are no photons at all, while the probability of more than one photon is

$$P(n \geq 2 | \alpha = 1/2) \approx 2.6\%$$

Therefore, when there are photons, it is most likely just one.

- Coherent states are not an orthonormal basis (since they are eigenstates of a non-Hermitian operator), but are an overcomplete basis in the sense that each vector $|\psi\rangle$ of the Hilbert space can be written as

$$|\psi\rangle = \int d\alpha d\bar{\alpha} \psi(\alpha, \bar{\alpha}) |\alpha\rangle$$

and the relation of completeness is

$$\frac{1}{\pi} \int d\alpha d\bar{\alpha} |\alpha\rangle\langle\alpha| = I$$

The integrals are performed over the complex plane.

- The states can also be written using the Glauber displacement operator $D(\alpha)$:

$$|\alpha\rangle = D(\alpha) |0\rangle$$

with

$$D(\alpha) = e^{\alpha a^\dagger - \bar{\alpha} a}$$

Indeed, from the Campbell–Baker–Hausdorff formula

$$e^{X+Y} = e^X e^Y e^{-\frac{1}{2}[X,Y]} e^{-\frac{1}{12}([X,[X,Y]] + [Y,[Y,X]])} \dots$$

Nested commutators vanish because $[a, a^\dagger] = 1$. One finds

$$D(\alpha) = e^{\alpha a^\dagger} e^{-\bar{\alpha} a} e^{-\frac{1}{2}|\alpha|^2}$$

and using $a |0\rangle = 0$

$$D(\alpha) |0\rangle = e^{-\frac{1}{2}|\alpha|^2} e^{\alpha a^\dagger} |0\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} (a^\dagger)^n |0\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle = |\alpha\rangle$$

- The states evolve classically. First of all, the expectation values of position and momentum (or electric and magnetic fields for the electromagnetic case)

$$x \sim (a^\dagger + a), \quad p \sim i(a^\dagger - a), \quad E \sim (a^\dagger + a)$$

are not zero. In fact

$$\langle\alpha|(a + a^\dagger)|\alpha\rangle = \alpha + \bar{\alpha}$$

so that

$$\begin{aligned} \langle x \rangle &\sim \langle a + a^\dagger \rangle \sim \alpha + \bar{\alpha} \sim \text{Re } \alpha \\ \langle p \rangle &\sim \langle i(a^\dagger - a) \rangle \sim i(\bar{\alpha} - \alpha) \sim \text{Im } \alpha \end{aligned}$$

In the complex plane, α is the (averaged) position in phase space. Classically, the motion of a simple harmonic oscillator in phase space is a circle:

$$x = x_0 \cos(\omega t), \quad p = m\dot{x} = -m\omega x_0 \sin(\omega t)$$

up to a normalization. So the point $(x, p/m\omega)$ rotates with frequency $-\omega$. The same is true for α . If

$$|\psi(0)\rangle = |\alpha\rangle \implies |\psi(t)\rangle = e^{-iHt} |\psi(0)\rangle$$

and the annihilation operator acts on the state as

$$\begin{aligned} a |\psi(t)\rangle &= a e^{-iHt} |\psi(0)\rangle = e^{-iHt} (e^{iHt} a e^{-iHt}) |\alpha\rangle = e^{-iHt} (e^{-i\omega t} a) |\alpha\rangle \\ &= e^{-iHt} (e^{-i\omega t} \alpha) |\alpha\rangle = e^{-i\omega t} \alpha e^{-iHt} |\alpha\rangle = e^{-i\omega t} \alpha |\psi(t)\rangle \end{aligned}$$

where at the first line one has applied the formula for left-right conjugation; then, the coherent state evolves as

$$|\psi(t)\rangle = |e^{-i\omega t} \alpha\rangle$$

meaning $\alpha \rightarrow e^{-i\omega t} \alpha$. This demonstrates that the coherent state evolves as in classical physics.

- The states are minimal uncertainty states. Define the quadrature operators

$$X = \frac{a + a^\dagger}{2}, \quad Y = \frac{a - a^\dagger}{2i}$$

so that $\langle \alpha | X | \alpha \rangle = \text{Re } \alpha$ and $\langle \alpha | Y | \alpha \rangle = \text{Im } \alpha$. These correspond to the position in phase space. Since $[a, a^\dagger] = 1$, then

$$[X, Y] = \frac{i}{2}$$

This resembles the commutator $[x, p] = i\hbar$ up to a rescaling. The Heisenberg uncertainty principle gives

$$\Delta X \Delta Y \geq \frac{1}{4}$$

One may compare the states:

- ◊ For the eigenstates of the number operator $|n\rangle$:

$$\langle X \rangle = \langle n | \frac{a + a^\dagger}{2} | n \rangle = 0, \quad \langle Y \rangle = 0$$

and

$$\langle X^2 \rangle = \langle n | \frac{a^2 + a^{\dagger 2} + a^\dagger a + a a^\dagger}{4} | n \rangle = \langle n | \frac{2a^\dagger a + 1}{4} | n \rangle = \frac{2n + 1}{4} = \langle Y^2 \rangle$$

Then

$$\Delta X \Delta Y = \frac{2n + 1}{4}$$

The minimal uncertainty occurs for $n = 0$: the ground state is Gaussian (as known from quantum mechanics). For $n \gg 1$, one has $\Delta X \Delta Y \gg 1/4$: the minimal uncertainty is strongly violated for large n .

- ◊ For coherent states $|\alpha\rangle$:

$$\langle X \rangle = \text{Re } \alpha, \quad \langle Y \rangle = \text{Im } \alpha$$

and

$$\begin{aligned} \langle X^2 \rangle &= \langle \alpha | \frac{a^2 + a^{\dagger 2} + 2a^\dagger a + 1}{4} | \alpha \rangle = \frac{(\alpha + \bar{\alpha})^2}{4} + \frac{1}{4} \\ \langle Y^2 \rangle &= \langle \alpha | \frac{-a^2 - a^{\dagger 2} + 2a^\dagger a + 1}{4} | \alpha \rangle = -\frac{(\alpha - \bar{\alpha})^2}{4} + \frac{1}{4} \end{aligned}$$

So that

$$\Delta X^2 = \langle X^2 \rangle - \langle X \rangle^2 = \frac{1}{4}, \quad \Delta Y^2 = \langle Y^2 \rangle - \langle Y \rangle^2 = \frac{1}{4}$$

which mean

$$\Delta X \Delta Y = \frac{1}{4}, \quad \forall |\alpha\rangle$$

Coherent states have minimal uncertainty and also

$$\Delta X = \Delta Y = \frac{1}{2}, \quad \forall |\alpha\rangle$$

Squeezed coherent states. Minimal uncertainty is important in experiments. Besides noise and the environment, there is an intrinsic quantum uncertainty in measurements that can become important experimentally (it already is for gravitational waves). A coherent state has a minimum value for both ΔX and ΔY of a half. However, it is possible to make one of them smaller, without violating the Heisenberg uncertainty principle, if $\Delta X \Delta Y$ stays at a quarter: ΔX can be smaller than a half, if ΔY is larger. Squeezed coherent states achieve this

$$|\alpha, \zeta\rangle = S(\zeta) D(\alpha) |0\rangle, \quad S(\zeta) = e^{\frac{1}{2}\bar{\zeta} a^2 - \frac{1}{2}\zeta a^{\dagger 2}}, \quad \zeta = se^{i\theta} \in \mathbb{C}$$

Computations are more complicated, but in the end

$$\Delta X^2 = \frac{1}{4} \left[e^{-2s} \cos^2 \frac{\theta}{2} + e^{2s} \sin^2 \frac{\theta}{2} \right], \quad \Delta Y^2 = \frac{1}{4} \left[e^{-2s} \sin^2 \frac{\theta}{2} + e^{2s} \cos^2 \frac{\theta}{2} \right]$$

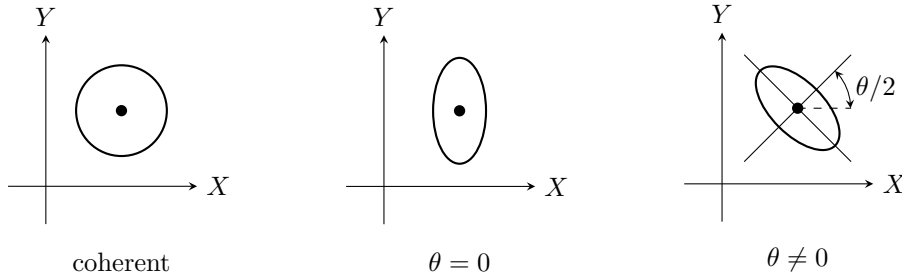
For $\theta = 0$, one has

$$\Delta X = \frac{1}{2} e^{-s}, \quad \Delta Y = \frac{1}{2} e^s$$

so that

$$\Delta X \Delta Y = \frac{1}{4}$$

but ΔX can be small. Plotted in the phase space, the uncertainty is



For the electric field, one has

$$E_C = \frac{\zeta_C}{2} (a e^{-i\omega t} + a^\dagger e^{i\omega t}) = \zeta_C (X \cos \omega t + Y \sin \omega t)$$

where the quadrature operators describe the position as it rotates in phase space.

Lecture 17

6.9 Ion traps

lun 02 dic
2024 16:30

Ion traps are a linear array of ions trapped by a time-dependent electromagnetic field given by the sum of two contributions

$$\Phi_{\text{dc}} = kU_0(z^2 - x^2 - y^2), \quad \Phi_{\text{rf}} = (V_0 \cos \Omega t + U_0) \left[1 - \frac{x^2 - y^2}{R^2} \right]$$

creating a harmonic oscillator potential

$$H_{\text{int}} = \sum_{i=1}^N \frac{M}{2} (\omega_x^2 x_i^2 + \omega_y^2 y_i^2 + \omega_z^2 z_i^2) + \sum_{i < j} \frac{e^2}{4\pi\epsilon_0 \|\mathbf{r}_i - \mathbf{r}_j\|}$$

where $\omega_z \ll \omega_x, \omega_y$ by design: little energy to wiggle along z and much energy to wiggle along x or y . The ions are cooled to their ground state using lasers (Doppler and sideband cooling) until $kT \ll \hbar\omega_z$ so that the corrections to the harmonic potential can be neglected, and the first excitation mode of the ions is the vibrational mode of the center of mass of the array along z with frequency ω_z . These modes are phonons that are also described with oscillators

$$z = z_0(a + a^\dagger), \quad z_0 = \sqrt{\frac{\hbar}{2NM\omega_z}}$$

Each one of the ions supports a two-level system. The qubits can be encoded in these ways (which are not exclusive):

- The ions have two special levels, $|0\rangle$ and $|1\rangle$, where one can encode qubits: they are either metastable excited states with long lifetime or a hyperfine splitting (which has an extremely long lifetime, $\Gamma_{\text{spont. emis.}} \sim \omega^3$, $\omega \sim \text{GHz}$, $\Gamma \sim 10^{-15} \text{ s}^{-1}$).

- Another qubit can be encoded in the vibrational mode, using $|0\rangle$ as the state of no phonons and $|1\rangle$ as the state of one phonon (this type of qubit is called “bus qubit”).

One operates with electromagnetic radiation. The ion two-level system interacts with the electromagnetic field through a magnetic dipole interaction which can be written as

$$H_{\text{int}} = \Omega \sigma_1 \cos(kz - \omega t + \phi)$$

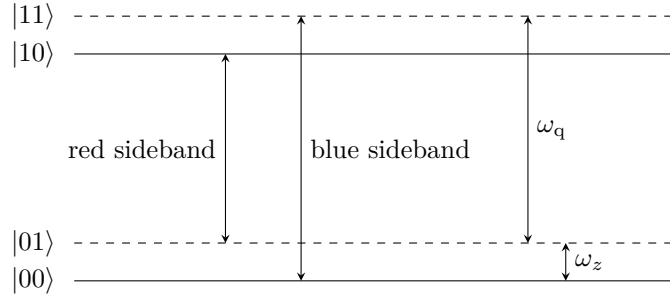
where one takes the electromagnetic radiation with frequency ω , direction z , and polarization along x . The parameter Ω is the amplitude of the radiation. One assumes that the wavelength λ is much bigger than the size of the qubit and also greater than the oscillation amplitude z_0

$$\eta = kz_0 = \frac{2\pi z_0}{\lambda} \ll 1$$

where η is the Lamb–Dicke parameter. Since kz is small, one may expand

$$\begin{aligned} H_{\text{int}} &= \Omega \sigma_1 \cos(-\omega t + \phi) - \Omega \sigma_1 \sin(-\omega t + \phi) kz + O(k^2 z^2) \\ &= \frac{\Omega}{2} (\sigma_+ + \sigma_-) (e^{-i(\omega t - \phi)} + e^{i(\omega t - \phi)}) + i \frac{\eta \Omega}{2} (\sigma_+ + \sigma_-) (a + a^\dagger) (e^{-i(\omega t - \phi)} - e^{i(\omega t - \phi)}) \\ &= H_I + H_{II} \end{aligned}$$

where one sets $z = z_0(a + a^\dagger)$. One may tune the radiation to use four frequencies $\pm\omega_q \pm \omega_z$



where, in the state $|nm\rangle$, n labels the qubit while m labels the phonon. One uses one laser ω_q to control the qubit, while two other lasers to pass from one qubit to the other (the sidebands). In the rotating wave approximation²⁰, using $\omega \sim \omega_q$, the first term H_I becomes

$$H_I = \frac{\Omega}{2} [\sigma_+ e^{i(\omega t - \phi)} + \sigma_- e^{-i(\omega t - \phi)}]$$

This term can be used to control the ion-qubit via Rabi oscillations.

Using the sideband frequency, in the RWA, the second term gives

$$\begin{aligned} H_{II} &= i \frac{\eta \Omega}{2} (-\sigma_+ a^\dagger e^{i(\omega t - \phi)} + \sigma_- a e^{-i(\omega t - \phi)}), \quad \omega \sim \omega_q - \omega_z, \quad \text{red sideband} \\ &+ i \frac{\eta \Omega}{2} (-\sigma_+ a e^{i(\omega t - \phi)} + \sigma_- a^\dagger e^{-i(\omega t - \phi)}), \quad \omega \sim \omega_q + \omega_z, \quad \text{blue sideband} \end{aligned}$$

as can be seen more formally, going to the rotating frame $U(t) = e^{iH_0 t}$ with

$$H_0 = -\frac{\omega_q}{2} \sigma_3 + \hbar \omega_z a^\dagger a$$

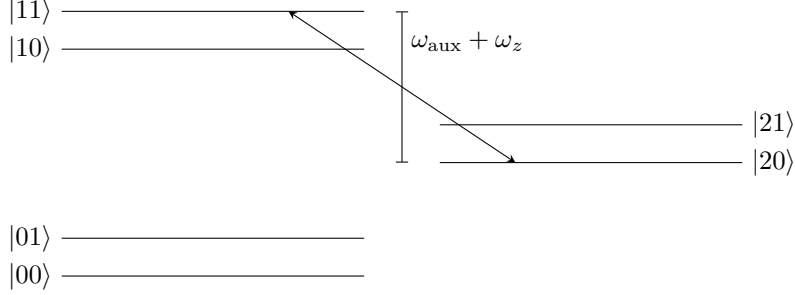
and using

$$U \sigma_+ U^\dagger = e^{-i\omega_q t} \sigma_+, \quad U a^\dagger U^\dagger = e^{i\omega_z t} a^\dagger$$

The term H_{II} can perform transitions between the ion-qubit and the vibrational one.

²⁰Notice that one has gone to the rotating frame, did the approximation, then rotated back.

Cirac–Zoller gate. The laser ω_q controls one qubit, so one may realize one-qubit gates. One may also wonder how to realize two-qubit gates. Cirac and Zoller proposed a simple CNOT gate using an ion with a third energy level [2]



where

$$E_1 - E_2 = \omega_{aux}$$

Then, it is simple to realize a controlled- Z gate on the states ion/phonon using a laser with frequency $\omega = \omega_{aux} + \omega_z$ to perform Rabi oscillations on $|20\rangle \leftrightarrow |11\rangle$ using H_I . One chooses the phase and time of the laser pulse to perform

$$R_x(2\pi) = e^{-\frac{i}{2}2\pi\sigma_1} = e^{-i\pi\sigma_1} = \cos \pi - i\sigma_1 \sin \pi = -I$$

that, ignoring the auxiliary $|20\rangle$, is just $|11\rangle \rightarrow -|11\rangle$, noting that the operator σ_1 acts on the space $\{|20\rangle, |11\rangle\}$. So

$$CZ = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix} \begin{matrix} |00\rangle \\ |01\rangle \\ |10\rangle \\ |11\rangle \end{matrix}$$

There are also other realizations without the third state (e.g. Mølmer–Sørensen gate, geometric phase gate, etc). To obtain a CNOT, one may add a Hadamard gate before and after the controlled- Z on the target qubit.

The Cirac–Zoller gate entangles the ion and the vibrational mode. It would be more useful to use the ions as qubits and the phonons as auxiliary qubit modes. This can be done by using the phonon as a bus to entangle ion-qubits. One can use a SWAP gate:

$$SWAP = \begin{bmatrix} 1 & & & \\ & 0 & 1 & \\ & -1 & 0 & \\ & & & 1 \end{bmatrix} \begin{matrix} |00\rangle \\ |01\rangle \\ |10\rangle \\ |11\rangle \end{matrix}$$

This acts non-trivially only on $|01\rangle$ and $|10\rangle$ as a matrix

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = i\sigma_2 = \cos \frac{\pi}{2} + i\sigma_2 \sin \frac{\pi}{2} = e^{i\frac{\pi}{2}\sigma_2} = R_y(-\pi)$$

and can be realized using the red sideband part of H_{II}

$$H_{II} \supset \frac{i\eta\Omega}{2}(-\sigma_+ a^\dagger e^{i(\omega t - \phi)} + \sigma_- a e^{-i(\omega t - \phi)})$$

using a laser with $\omega = \omega_q - \omega_z$ that performs rotations among $|01\rangle$ and $|10\rangle$. If one starts with a state with no phonons, then the SWAP gate acts as

$$(a|0\rangle + b|1\rangle) \otimes |0\rangle = a|00\rangle + b|10\rangle \rightarrow a|00\rangle + b|01\rangle = |0\rangle \otimes (a|0\rangle + b|1\rangle)$$

the information is swapped from the qubit to the bus.

If one labels the ions with $j = 1, 2, 3, \dots, N$, one has different CZ and SWAP gates for each index j . By swapping twice

$$(SWAP_k^{-1})(CZ_j)(SWAP_k)$$

one realizes a CZ gate on the ions j and k : it acts on k with j as control, while the vibrational mode is used to move information. The CNOT gate is as usual

$$\text{CNOT}_{jk} = H_k \text{SWAP}_k^{-1} \text{CZ}_j \text{SWAP}_k H_k$$

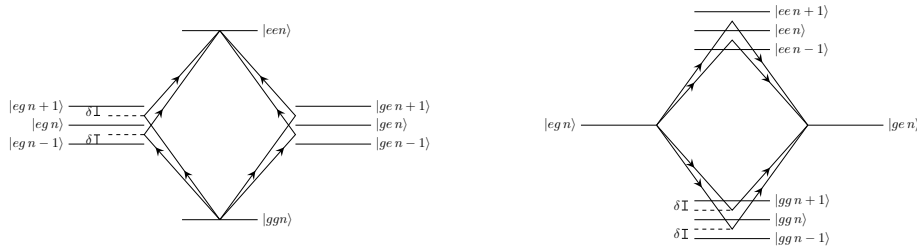
The Cirac–Zoller gate requires cooling the vibrational mode to the ground state. There are other gates where this is not required (e.g. Mølmer–Sørensen, geometric phase gate, etc.) and are more efficient. The Cirac–Zoller gate is mostly historical.

Lecture 18

6.10 Mølmer–Sørensen gate

mer 04 dic
2024 16:30

For example, the Mølmer–Sørensen gate uses a bi-chromatic laser to address two ion-qubits simultaneously and allows intermediate states with $n - 1$, n and $n + 1$ phonons with an n -independent reset. Denote $|0\rangle = |g\rangle$ and $|1\rangle = |e\rangle$ for the ion-qubit



Consider the figure on the left-hand side. To go from the lower state to the middle states, one has to use a red sideband laser (or blue sideband laser) and from there use the blue sideband laser (red sideband) to go to the upper state, independently of the number of phonons. For the figure on the right-hand side, the situation is similar but one applies the same type of laser. The two lasers have frequencies $\omega = \omega_q \pm (\omega_z - \delta)$ which have opposite detuning frequency δ , and one starts with a state with n phonons. The resonant transitions are two-step and read:

$$|ggn\rangle \leftrightarrow |een\rangle, \quad |egn\rangle \leftrightarrow |gen\rangle$$

at second order in perturbation theory. The effective Rabi frequency is n -independent. In fact, very roughly, one shows that at second order, the important terms are

$$\text{Rabi} \sim \sum_m \frac{\langle een | H_{\text{II}} | m \rangle \langle m | H_{\text{II}} | ggn \rangle}{E_m - E_{ggn} - \omega} \sim \frac{\sqrt{n+1}\sqrt{n+1}}{\delta} + \frac{\sqrt{n}\sqrt{n}}{-\delta} \sim \frac{1}{\delta}$$

where the intermediate states $|m\rangle$ are the middle states in the figure on the left-hand side. The Hamiltonian H_{II} has the same form as above, but with a contribution for each ion-qubit.

Exact solution. The problem is exactly solvable (in the RWA). One assumes that the two lasers couple with the same frequency Ω to all the ions, and they also have the same η . Then, from H_{II} , one gets two contributions: one for $\omega = \omega_q - (\omega_z - \delta)$ (red sideband) and one for $\omega = \omega_q + (\omega_z - \delta)$ (blue sideband). In the rotating frame ($U = e^{iH_0 t}$), the Hamiltonian is

$$H_{\text{II}} = \frac{i\eta\Omega}{2} [-S_+ a^\dagger e^{i(\delta t - \phi)} + S_- a e^{-i(\delta t - \phi)}] + \frac{i\eta\Omega}{2} [-S_+ a e^{-i(\delta t + \phi)} + S_- a^\dagger e^{i(\delta t + \phi)}]$$

where $\mathbf{S} = \boldsymbol{\sigma}^{(1)} \otimes I + I \otimes \boldsymbol{\sigma}^{(2)}$ is the total angular momentum. Setting $\phi = 0$ for simplicity, one finds

$$H_{\text{II}} = \frac{i\eta\Omega}{2} (S_- - S_+) [a e^{-i\delta t} + a^\dagger e^{i\delta t}]$$

The time evolution operator for a time-dependent Hamiltonian can be written in various forms

$$U_{\text{EV}}(t) = \mathcal{T} \exp \left[-i \int_0^t H(t_1) dt_1 \right] = \exp \left[\sum_{k=1}^{\infty} M_k(t) \right]$$

where M_k are the terms of the Magnus expansion

$$\begin{aligned} M_1(t) &= -i \int_0^t H(t_1) dt_1 \\ M_2(t) &= \frac{1}{2}(-i)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 [H(t_1), H(t_2)] \\ M_3(t) &= \frac{1}{6}(-i)^3 \int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 [H(t_1), [H(t_2), H(t_3)]] + [H(t_3), [H(t_2), H(t_1)]] \end{aligned}$$

Let $H = H_{\text{II}}$. Now, the first term is

$$\begin{aligned} M_1(t) &= \frac{\eta\Omega}{2}(S_- - S_+) \left[a \frac{e^{-i\delta t} - 1}{-i\delta} + a^\dagger \frac{e^{i\delta t} - 1}{i\delta} \right] \\ &= \frac{\eta\Omega}{\delta}(S_- - S_+) \left[ae^{-\frac{i}{2}\delta t} \sin \frac{\delta t}{2} + a^\dagger e^{\frac{i}{2}\delta t} \sin \frac{\delta t}{2} \right] \\ &= (S_- - S_+)[\alpha(t)a + \alpha^*(t)a^\dagger] \end{aligned}$$

where

$$\alpha(t) = \frac{\eta\Omega}{\delta} e^{-\frac{i}{2}\delta t} \sin \frac{\delta t}{2}$$

The commutator is

$$\begin{aligned} [H_{\text{II}}(t_1), H_{\text{II}}(t_2)] &= -\frac{\eta^2\Omega^2}{4}(S_- - S_+)^2([a, a^\dagger]e^{i\delta(t_2-t_1)} + [a^\dagger, a]e^{i\delta(t_1-t_2)}) \\ &= -\frac{\eta^2\Omega^2}{4}(S_- - S_+)^2(e^{i\delta(t_2-t_1)} - e^{i\delta(t_1-t_2)}) \\ &= -\frac{i\eta^2\Omega^2}{2}(S_- - S_+)^2 \sin[\delta(t_2 - t_1)] \end{aligned}$$

so that the second term is

$$M_2(t) = \frac{\eta^2\Omega^2}{4\delta} i(S_- - S_+)^2 \int_0^t dt_1 (-1 + \cos \delta t_1) = i(S_- - S_+)^2 \phi(t)$$

where

$$\phi(t) = \left(\frac{\eta\Omega}{2\delta} \right)^2 (\sin \delta t - \delta t)$$

This second expansion term $M_2(t)$ is independent of the operators a and a^\dagger , and also

$$[M_2(t), H_{\text{II}}(t')] = 0$$

which mean that

$$M_n(t) = 0, \quad n \geq 3$$

The expansion ends at second order

$$U_{\text{EV}}(t) = \exp[(S_- - S_+)[\alpha(t)a + \alpha^*(t)a^\dagger] + i(S_- - S_+)^2 \phi(t)]$$

To make the system insensitive to the number of phonons (i.e. independent of the ladder operators), one may choose a particular time

$$\delta \bar{t} = 2\pi, \quad \alpha(\bar{t}) = 0, \quad \phi(\bar{t}) = -\frac{\pi}{2} \left(\frac{\eta\Omega}{\delta} \right)^2 \equiv -\beta$$

and one finds

$$U_{\text{MS}} \equiv U_{\text{EV}}(\bar{t}) = e^{-i(S_- - S_+)^2 \beta} = e^{i\beta S_y^2}$$

Explicitly

$$S_y = \sigma_2^{(1)} \otimes I + I \otimes \sigma_2^{(2)} = \begin{bmatrix} 0 & -iI \\ iI & 0 \end{bmatrix} + \begin{bmatrix} \sigma_2 & 0 \\ 0 & \sigma_2 \end{bmatrix} = \begin{bmatrix} \sigma_2 & -iI \\ iI & \sigma_2 \end{bmatrix}$$

so that

$$S_y^2 = 2 \begin{bmatrix} I & -i\sigma_2 \\ i\sigma_2 & I \end{bmatrix} = 2 \begin{bmatrix} 1 & & & -1 \\ & 1 & 1 & \\ & 1 & 1 & \\ -1 & & & 1 \end{bmatrix} \begin{bmatrix} |gg\rangle \\ |ge\rangle \\ |eg\rangle \\ |ee\rangle \end{bmatrix}$$

It is block diagonal on the columns (23) and (14). The first block acts as $2(I + \sigma_1)$, while the second as $2(I - \sigma_1)$. The eigenvectors of S_y^2 are

$$\begin{aligned} |e_1\rangle &= \frac{|gg\rangle + |ee\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, & |e_2\rangle &= \frac{|gg\rangle - |ee\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \\ |e_3\rangle &= \frac{|ge\rangle + |eg\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, & |e_4\rangle &= \frac{|ge\rangle - |eg\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \end{aligned}$$

with respective eigenvalues

$$0, \quad 4, \quad 4, \quad 0$$

Including (an irrelevant) phase $e^{-2i\beta}$, the operator $e^{i\beta S_y^2}$ acts as

$$\begin{aligned} |gg\rangle &= \frac{|e_1\rangle + |e_2\rangle}{\sqrt{2}} \rightarrow \frac{e^{-2i\beta} |e_1\rangle + e^{2i\beta} |e_2\rangle}{\sqrt{2}} = \cos(2\beta) |gg\rangle - i \sin(2\beta) |ee\rangle \\ |ee\rangle &= \frac{|e_1\rangle - |e_2\rangle}{\sqrt{2}} \rightarrow \frac{e^{-2i\beta} |e_1\rangle - e^{2i\beta} |e_2\rangle}{\sqrt{2}} = -i \sin(2\beta) |gg\rangle + \cos(2\beta) |ee\rangle \\ |ge\rangle &= \frac{|e_3\rangle + |e_4\rangle}{\sqrt{2}} \rightarrow \frac{e^{2i\beta} |e_3\rangle + e^{-2i\beta} |e_4\rangle}{\sqrt{2}} = \cos(2\beta) |ge\rangle + i \sin(2\beta) |eg\rangle \\ |eg\rangle &= \frac{|e_3\rangle - |e_4\rangle}{\sqrt{2}} \rightarrow \frac{e^{2i\beta} |e_3\rangle - e^{-2i\beta} |e_4\rangle}{\sqrt{2}} = i \sin(2\beta) |ge\rangle + \cos(2\beta) |eg\rangle \end{aligned}$$

For $\beta = \pi/8$, one has

$$\cos(2\beta) = \sin(2\beta) = \frac{1}{\sqrt{2}}$$

and one obtains a basis of Bell states.

General case. The Mølmer-Sørensen gate construction is good for creating multi-particle entangled states, e.g. Greenberger-Horne-Zeilinger (GHZ) states, and the first paper was aimed just at that. For N ion-qubits, such GHZ state is

$$|\psi\rangle = \frac{1}{\sqrt{2}} (e^{i\phi_1} |ggg \cdots g\rangle + e^{i\phi_2} |eee \cdots e\rangle)$$

The choice

$$e^{i\phi_1} = e^{-i\frac{\pi}{4}}, \quad e^{i\phi_2} = (-1)^N e^{i\frac{\pi}{4}}$$

is useful for metrology and the tests of Bell's inequalities. For $N \geq 3$, there are inequivalent classes of entanglement. For example, for $N = 3$ there are two non-biseparable classes of three-qubit states: GHZ and W^{21} states

$$|\text{GHZ}\rangle = \frac{|000\rangle + |111\rangle}{\sqrt{2}}, \quad |W\rangle = \frac{|100\rangle + |010\rangle + |001\rangle}{\sqrt{3}}$$

The state $|\psi\rangle$ can be obtained with the operator $e^{i\beta J_y^2 t}$. Consider the spin notation $|JM\rangle$ of total angular momentum J and projection M along the z -axis (or a specified axis). The above state is

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left[e^{-i\frac{\pi}{4}} \left| \frac{N}{2}, \frac{N}{2} \right\rangle + (-1)^N e^{i\frac{\pi}{4}} \left| \frac{N}{2}, -\frac{N}{2} \right\rangle \right]$$

²¹W stands for Wolfgang Dür. See W. Dür, G. Vidal, J. Cirac. "Three qubits can be entangled in two inequivalent ways", in Physical Review A, vol. 62, no. 6, 2000. arXiv: <https://arxiv.org/abs/quant-ph/0005115>.

One may write

$$\left| \frac{N}{2}, \frac{N}{2} \right\rangle = \sum_M c_M \left| \frac{N}{2}, M \right\rangle_y$$

Up to a phase α , one finds

$$\left| \frac{N}{2}, -\frac{N}{2} \right\rangle = e^{i\alpha} e^{i\pi J_y} \left| \frac{N}{2}, \frac{N}{2} \right\rangle = (-1)^N \sum_M c_M e^{i\pi M} \left| \frac{N}{2}, M \right\rangle_y$$

The $(-1)^N$ can be checked with

$$e^{i\frac{\pi}{2}\sigma_2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = - \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The Mølmer-Sørensen gate has J_x and $(-1)^{\frac{N}{2}}$. The state can be rewritten as

$$|\psi\rangle = \sum_M c_M \frac{e^{-i\frac{\pi}{4}} + e^{i\frac{\pi}{4}}(-1)^M}{\sqrt{2}} \left| \frac{N}{2}, M \right\rangle_y = \sum_M c_M \left[\frac{1-i}{2} + (-1)^M \frac{1+i}{2} \right] \left| \frac{N}{2}, M \right\rangle_y$$

The coefficient in brackets depends on the parity of M

$$\frac{1-i}{2} + (-1)^M \frac{1+i}{2} = \begin{cases} 1, & M \text{ even} \\ -i, & M \text{ odd} \end{cases}$$

However, this same state can be obtained using the operator $e^{i\beta J_y^2 t}$ at $t = -\pi/2\beta$. In fact

$$|\psi\rangle = e^{i\beta J_y^2 t} |ggg \dots g\rangle = e^{-i\frac{\pi}{2} J_y^2} \left| \frac{N}{2}, \frac{N}{2} \right\rangle = \sum_M c_M e^{-i\frac{\pi}{2} M^2} \left| \frac{N}{2}, M \right\rangle_y$$

which is exactly the state above when one notices

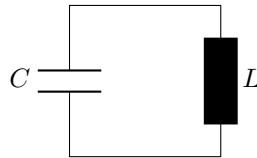
$$e^{-i\frac{\pi}{2} M^2} = \begin{cases} 1, & M \text{ even} \\ e^{-i\frac{\pi}{2}(1+2k)^2} = e^{-i\frac{\pi}{2}} = -i, & M \text{ odd} \end{cases}$$

Lecture 19

6.11 Superconducting qubits — circuit QED (cQED)

lun 09 dic
2024 16:30

At very low temperatures, a superconducting circuit is a quantum oscillator. Electrons form Cooper pairs that condense and give a supercurrent, which flows with no resistance and is separated from the electronic excitations (the typical current) by a gap Δ (i.e. the energy to break the Cooper pairs). It can be modelled with a circuit:



Recall that the charge of a capacitor with a voltage V is

$$Q = CV, \quad I = \frac{dQ}{dt}$$

while for an inductor holds

$$V = L \frac{dI}{dt}, \quad \Phi = LI$$

where the current is $I = d_t Q$ and the flux is

$$\Phi(t) = \int_{-\infty}^t V(t') dt' \iff V = \dot{\Phi}$$

The corresponding energy

$$E(t) = \int_{-\infty}^t V(t') I(t') dt', \quad \delta E = V \delta Q$$

for the capacitor is

$$E = \int V C \frac{dV}{dt} dt = \frac{1}{2} C V^2 = \frac{Q^2}{2C}$$

and for the inductor is

$$E = \int I L \frac{dI}{dt} dt = \frac{1}{2} L I^2 = \frac{\Phi^2}{2L}$$

where one assumes that all functions vanish at $t \rightarrow -\infty$.

The Lagrangian of the circuit above

$$\mathcal{L} = \frac{Q^2}{2C} - \frac{1}{2L} \Phi^2 = \frac{1}{2} C V^2 - \frac{1}{2L} \Phi^2 = \frac{C}{2} \dot{\Phi}^2 - \frac{1}{2L} \Phi^2$$

where the capacitor contributes to the kinetic energy and the inductor contributes to the potential energy, has the equation of motion of a simple harmonic oscillator

$$C \ddot{\Phi} + \frac{1}{L} \Phi = 0 \iff \ddot{\Phi} + \omega^2 \Phi = 0, \quad \omega = \frac{1}{\sqrt{LC}}$$

The conjugated momentum of the flux Φ is the charge

$$\Pi = \frac{\delta \mathcal{L}}{\delta \dot{\Phi}} = C \dot{\Phi} = C V = Q$$

so the Hamiltonian is

$$H = (\Pi \dot{\Phi} - \mathcal{L})|_{\dot{\Phi}=\dot{\Phi}(\dots)} = Q \frac{Q}{C} - \left(\frac{Q^2}{2C} - \frac{\Phi^2}{2L} \right) = \frac{Q^2}{2C} + \frac{\Phi^2}{2L}$$

which is the sum of the energies of the capacitor C and the inductor L .

Now, a superconductor at a temperature T less than the temperature equivalent to the Cooper pair energy Δ is really quantum. One may quantize it. From the Hamiltonian of the simple harmonic oscillator

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

one may identify

$$Q \leftrightarrow p, \quad \Phi \leftrightarrow x, \quad C \leftrightarrow m, \quad \frac{1}{\sqrt{LC}} \leftrightarrow \omega$$

to obtain

$$H = \frac{Q^2}{2C} + \frac{C}{2} \omega^2 \Phi^2$$

Translating the observables to operators, and introducing the ladder operators, one may rewrite the Hamiltonian as

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

where

$$\omega = \frac{1}{\sqrt{LC}}, \quad \Phi = \sqrt{\frac{\hbar}{2\omega C}} (a^\dagger + a), \quad Q = i\sqrt{\frac{\hbar \omega C}{2}} (a^\dagger - a), \quad [\Phi, Q] = i\hbar, \quad [a, a^\dagger] = 1$$

It is often written as

$$H = 4E_C n^2 + \frac{1}{2} E_L \phi^2, \quad \hbar \omega = \sqrt{8E_L E_C}$$

with the charging energy of the capacitor to add one electron e (i.e. half Cooper pair) being

$$E_C = \frac{e^2}{2C}$$

the number of Cooper pairs being

$$n = \frac{Q}{2e}$$

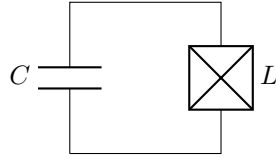
the reduced flux and the flux quantum being

$$\phi = 2\pi \frac{\Phi}{\Phi_0} = \frac{2e}{\hbar} \Phi, \quad \Phi_0 = \frac{h}{2e}$$

and finally the inductive energy being

$$E_L = \left(\frac{\Phi_0}{2\pi} \right)^2 \frac{1}{L}$$

A simple harmonic oscillator has equidistant energy levels: it is not a good qubit. One needs to obtain anharmonicity which is provided by the Josephson junction: when two superconductors are separated by a thin layer of insulator, there is a direct current crossing the insulator in the absence of any external electromagnetic field, owing to the tunneling of Cooper pairs. The Josephson junction is represented as a crossed square



The Josephson effect can be summarized with two equations

$$I = I_c \sin \phi, \quad V = \frac{d\Phi}{dt} = \frac{\hbar}{2e} \frac{d\phi}{dt} = \frac{\Phi_0}{2\pi} \frac{d\phi}{dt}$$

where I_c is the critical current. The second equation is tautological because it is a definition, while the first is non-trivial and follows from tunneling. The energy in a Josephson junction is

$$E = \int dt VI = \frac{\Phi_0}{2\pi} \int dt \frac{d\phi}{dt} I_c \sin \phi = -E_J \cos \phi$$

where the Josephson energy is

$$E_J = \frac{\Phi_0}{2\pi} I_c$$

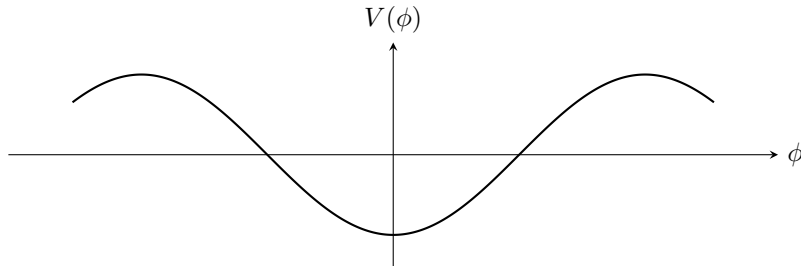
The Josephson junction has both a capacitance and an inductance. Therefore, the total capacitance of the circuit is

$$C_{\text{tot}} = C + C_J$$

and similarly for the inductance. The circuit above has Hamiltonian

$$H = 4E_C^{\text{tot}} n^2 - E_J \cos \phi$$

with a periodic anharmonic potential



6.12 Josephson effect

In a superconducting material, pairs of electrons (due to phonon interactions) condense and form bound states, called Cooper pairs, of charge $q = 2e > 0$ and bosonic statistics. At very low temperatures, bosons tend to stay all together in the ground state (differently from fermions that form the Fermi sea). The spectrum has a mass gap Δ which is the energy required to break a Cooper pair. In the ground state $|N\rangle$ there is a macroscopic number N of Cooper pairs.

At low temperatures, single electrons cannot be excited (this requires an energy Δ), and the current is due to Cooper pairs. While in a current of electrons, impurities and other effects disturb the electron motion, giving rise to finite resistance R , in a current of bosons, the Bose–Einstein statistics favours having all bosons in the same state, which also applies if they are in motion. Once the condensate is in motion, it is hard to disturb the single particles, and there is superconductivity: flow without resistance.

In a metal junction, electrons can tunnel from one electrode to the other. There is a current I due to a potential difference V related by Ohm's law $V = IR$. In a superconductor junction, only the Cooper pairs can tunnel, changing the number of pairs in each electrode

$$|n_1, n_2\rangle \rightarrow |n_1 \pm 1, n_2 \mp 1\rangle$$

Let $|m\rangle$ be the state of the first electrode with m pairs. A simple phenomenological Hamiltonian describing the tunneling is

$$H = -\frac{E_J}{2} \sum_m (|m\rangle\langle m+1| + |m+1\rangle\langle m|)$$

where E_J is the Josephson coupling energy. This Hamiltonian is symmetric in the two directions

$$\begin{aligned} (|m\rangle\langle m+1|) |m+1\rangle &= |m\rangle, & |m+1\rangle &\rightarrow |m\rangle \\ (|m+1\rangle\langle m|) |m\rangle &= |m+1\rangle, & |m\rangle &\rightarrow |m+1\rangle \end{aligned}$$

corresponding to a Cooper pair tunneling out to the second electrode and to a Cooper pair tunneling in from the second electrode. The eigenstates of the Hamiltonian H are

$$|\varphi\rangle = \sum_{m=-\infty}^{\infty} e^{im\varphi} |m\rangle, \quad \varphi \in [0, 2\pi)$$

with eigenvalue $-E_J \cos \varphi$. In fact

$$\begin{aligned} H |\varphi\rangle &= -\frac{E_J}{2} \sum_{\bar{m}, m} (|\bar{m}\rangle\langle \bar{m}+1| + |\bar{m}+1\rangle\langle \bar{m}|) e^{im\varphi} |m\rangle = -\frac{E_J}{2} \sum_m e^{i(m+1)\varphi} |m\rangle + e^{im\varphi} |m+1\rangle \\ &= -\frac{E_J}{2} \sum_m e^{i(m+1)\varphi} |m\rangle + e^{i(m-1)\varphi} |m\rangle = -E_J \cos \varphi \sum_m e^{im\varphi} |m\rangle = -E_J \cos \varphi |\varphi\rangle \end{aligned}$$

First Josephson relation. If one defines the number operator (of pairs)

$$\hat{n} = \sum_n n |n\rangle\langle n|$$

then the current across the junction is

$$\begin{aligned} \hat{I} &= \frac{d\hat{Q}}{dt} = q \frac{d\hat{n}}{dt} = q \frac{i}{\hbar} [H, \hat{n}] = q \frac{i}{\hbar} (H\hat{n} - \hat{n}H) \\ &= -\frac{E_J}{2} \frac{qi}{\hbar} \sum_{m,n} [|m\rangle\langle m+1| n |n\rangle\langle n| + |m+1\rangle\langle m| n |n\rangle\langle n| \\ &\quad - n |n\rangle\langle n| |m\rangle\langle m+1| - n |n\rangle\langle n| |m+1\rangle\langle m|] \\ &= -\frac{E_J}{2} \frac{qi}{\hbar} \sum_m [|m\rangle\langle m+1| (m+1) + |m+1\rangle\langle m| m - |m\rangle\langle m+1| m - |m+1\rangle\langle m| (m+1)] \\ &= -\frac{iq}{2\hbar} E_J \sum_m [|m\rangle\langle m+1| - |m+1\rangle\langle m|] \end{aligned}$$

At the first line, the derivative has been substituted using Ehrenfest's theorem of quantum mechanics (in the Heisenberg picture $A_H(t) = e^{\frac{i}{\hbar}Ht} A_S e^{-\frac{i}{\hbar}Ht}$, where A_S is time-independent).

Similarly as above, one finds

$$\hat{I}|\varphi\rangle = -\frac{iqE_J}{2\hbar} \sum_m [e^{i(m+1)\varphi} |m\rangle - e^{i(m-1)\varphi} |m\rangle] = \frac{qE_J}{\hbar} \sin \varphi |\varphi\rangle$$

So the state $|\varphi\rangle$ is also an eigenstate of the current with eigenvalue

$$\boxed{I(\varphi) = I_c \sin \varphi}, \quad I_c = \frac{qE_J}{\hbar}$$

where I_c is the critical current. This is the first Josephson relation.

Second Josephson relation. By adding a voltage difference V across a junction, the Hamiltonian becomes

$$H \rightarrow H - q\hat{n}V$$

where $\hat{Q} = q\hat{n} = 2e\hat{n}$ is the total charge. The Schrödinger equation gives

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = (H - q\hat{n}V) |\psi(t)\rangle$$

which can be solved

$$|\psi(t)\rangle = \exp \left[\frac{i}{\hbar} E_J \int_0^t \cos \left(\varphi_0 + \frac{qV\tau}{\hbar} \right) d\tau \right] |\varphi_0 + qV\hbar^{-1}t\rangle$$

where $|\psi(0)\rangle = |\varphi_0\rangle$ is the state at $t = 0$. In fact the expressions

$$\begin{aligned} i\hbar \frac{d}{dt} |\psi(t)\rangle &= -E_J \cos \left(\varphi_0 + \frac{qVt}{\hbar} \right) |\psi(t)\rangle \\ &+ \exp \left[\frac{i}{\hbar} E_J \int_0^t \cos \left(\varphi_0 + \frac{qV\tau}{\hbar} \right) d\tau \right] \sum_m e^{i(\varphi_0 + \frac{qV}{\hbar}t)m} (-qVm) |m\rangle \end{aligned}$$

and

$$\begin{aligned} (H - q\hat{n}V) |\psi(t)\rangle &= -E_J \cos \left(\varphi_0 + \frac{qVt}{\hbar} \right) |\psi(t)\rangle \\ &+ \exp \left[\frac{i}{\hbar} E_J \int_0^t \cos \left(\varphi_0 + \frac{qV\tau}{\hbar} \right) d\tau \right] (-qV) \sum_m e^{i(\varphi_0 + \frac{qV}{\hbar}t)m} m |m\rangle \end{aligned}$$

are equal when substituting the expression of $|\psi(t)\rangle$ above. Notice that

$$H|\varphi\rangle = -E_J \cos \varphi |\varphi\rangle$$

Apart from a phase factor (which is irrelevant in quantum mechanics), the state $|\varphi_0\rangle$ evolves linearly in time

$$\varphi_0 \rightarrow \varphi(t) = \varphi_0 + \frac{qVt}{\hbar} \iff \boxed{\frac{d\varphi}{dt} = \frac{qV}{\hbar}}$$

This is the second Josephson relation.

Final remarks. Notice that the system is really quantum. Only the states $|m\rangle$, eigenstates of a number operator

$$\hat{n} |m\rangle = m |m\rangle$$

like in the harmonic oscillator, contribute and reproduce the interaction from quantum mechanics.

The reason why the Hamiltonian is rewritten in terms of \hat{n} and φ is that the two are canonically conjugate

$$\hat{n} |\varphi\rangle = \sum_m e^{im\varphi} m |m\rangle = -i \frac{\partial}{\partial \varphi} |\varphi\rangle \implies \hat{n} = -i \frac{\partial}{\partial \varphi}, \quad [\varphi, \hat{n}] = i$$

where n and φ are the number of pairs and the phase of the condensate in the more complete microscopic theory, the Bardeen–Cooper–Schrieffer theory of superconductivity.

A real junction has also a capacitive energy, and the full Hamiltonian is

$$H = 4E_C \hat{n}^2 - E_J \cos \varphi$$

Lecture 20

mer 11 dic
2024 16:30

6.13 Transmon

The standard superconducting qubit is the transmon (as used by IBM, for example). Consider the Hamiltonian of the circuit above

$$H = 4E_C n^2 + E_J(1 - \cos \phi)$$

where one has added a constant term E_J . The transmon operates for $E_J/E_C \gg 1$ (differently from the first design of the charge qubit, the Cooper-pair box), so that, by expanding in a Taylor series

$$V(\phi) = E_J(1 - \cos \phi) = \frac{E_J}{2}\phi^2 - \frac{E_J}{4!}\phi^4 + \frac{E_J}{6!}\phi^6 + \dots = \frac{x^2}{2} - \frac{x^4}{4!E_J} + \frac{x^6}{6!E_J^2} + \dots$$

with $x = \sqrt{E_J}\phi$, for $E_J \gg 1$ one may truncate the expansion

$$H = 4E_C n^2 + E_J \frac{\phi^2}{2} - E_J \frac{\phi^4}{24} = H_0 - E_J \frac{\phi^4}{24}, \quad H_0 = 4E_C n^2 + E_J \frac{\phi^2}{2}$$

which is a harmonic oscillator with a quartic coupling: the eigenvalues are not equidistant. For example, the ground state and the first excited state are at an energy distance $\hbar\omega$, while the first and second excited states are at a distance $\hbar\omega + \alpha$ with $\alpha < 0$. If α is reasonably big (around 10% of $\hbar\omega$), then one may just focus on the lowest two states $|0\rangle$ and $|1\rangle$, supposing that one works with signals of frequency $\omega \sim \omega_{10}$ such that they cannot excite the system to the second level. For typical transmons, the values are

$$\omega_{10} \approx 3 - 4 \text{ GHz}, \quad \alpha \approx 300 \text{ MHz}, \quad E_J/E_C \approx 40$$

Recalling

$$E_C = \frac{e^2}{2C}, \quad E_J = \left(\frac{\hbar}{2e}\right)^2 \frac{1}{L}, \quad n = \frac{Q}{2e}, \quad \phi = \frac{2e}{\hbar}\Phi, \quad \hbar\omega = \frac{\hbar}{\sqrt{LC}} = \sqrt{8E_C E_J}$$

one obtains

$$\phi = \frac{2\pi}{\Phi_0}\Phi = \frac{2e}{\hbar}\Phi = \frac{2e}{\hbar}\sqrt{\frac{\hbar}{2\omega C}}(a + a^\dagger) = 2e\sqrt{\frac{1}{2\hbar\omega} \frac{2E_C}{e^2}}(a + a^\dagger) = 2\left[\frac{E_C}{8E_J}\right]^{\frac{1}{4}}(a + a^\dagger)$$

from which the Hamiltonian is

$$H = \hbar\omega \left(a^\dagger a + \frac{1}{2}\right) - \frac{E_C}{12}(a + a^\dagger)^4$$

Going in the rotating frame with the operator $U = e^{iH_0 t}$ (i.e. using the interaction picture), the ladder operators become

$$a \rightarrow e^{-i\omega t}a, \quad a^\dagger \rightarrow e^{i\omega t}a^\dagger$$

and most terms in $(a + a^\dagger)^4$ average to zero in the RWA. Only the terms with equal numbers of operators a and a^\dagger survive

$$(a + a^\dagger)^4 = a^4 + a^\dagger a a^\dagger a + \dots \rightarrow \cancel{a^4 e^{-i4\omega t}} + a^\dagger a a^\dagger a + \dots$$

So that

$$(a + a^\dagger)^4 \sim aaa^\dagger a^\dagger + aa^\dagger aa^\dagger + aa^\dagger a^\dagger a + a^\dagger a^\dagger aa + a^\dagger aaa^\dagger + a^\dagger aa^\dagger a$$

One may bring all the creation operators a^\dagger to the left (normal ordering):

$$\begin{aligned} (a + a^\dagger)^4 &\sim (aa^\dagger aa^\dagger + aa^\dagger) + (a^\dagger aaa^\dagger + aa^\dagger) + (a^\dagger aa^\dagger a + a^\dagger a) \\ &\quad + (a^\dagger a^\dagger aa) + (a^\dagger aa^\dagger a + a^\dagger a) + (a^\dagger a^\dagger aa + a^\dagger a) \\ &= (aa^\dagger a^\dagger a + 2aa^\dagger) + (a^\dagger aa^\dagger a + aa^\dagger + a^\dagger a) + (a^\dagger a^\dagger aa + 2a^\dagger a) \\ &\quad + (a^\dagger a^\dagger aa) + (a^\dagger a^\dagger aa + 2a^\dagger a) + (a^\dagger a^\dagger aa + a^\dagger a) \\ &= (a^\dagger aa^\dagger a + 2aa^\dagger + a^\dagger a) + (a^\dagger a^\dagger aa + aa^\dagger + 2a^\dagger a) + (a^\dagger a^\dagger aa + 2a^\dagger a) \\ &\quad + (a^\dagger a^\dagger aa) + (a^\dagger a^\dagger aa + 2a^\dagger a) + (a^\dagger a^\dagger aa + a^\dagger a) \\ &= [a^\dagger a^\dagger aa + 2(aa^\dagger + a^\dagger a)] + (a^\dagger a^\dagger aa + aa^\dagger + 2a^\dagger a) + (a^\dagger a^\dagger aa + 2a^\dagger a) \\ &\quad + (a^\dagger a^\dagger aa) + (a^\dagger a^\dagger aa + 2a^\dagger a) + (a^\dagger a^\dagger aa + a^\dagger a) \\ &= 6a^\dagger a^\dagger aa + 3aa^\dagger + 9a^\dagger a = 6(a^\dagger a^\dagger aa + 2a^\dagger a) + 3 \end{aligned}$$

The Hamiltonian of the transmon becomes

$$H = \hbar\omega a^\dagger a - \frac{E_C}{2} a^\dagger a^\dagger aa - E_C a^\dagger a + \text{const.}$$

It may be expressed as

$$H_{\text{tr}} = \hbar\tilde{\omega} a^\dagger a - \frac{E_C}{2} a^\dagger a^\dagger aa \equiv \hbar\tilde{\omega} a^\dagger a + \frac{\alpha}{2} a^\dagger a^\dagger aa$$

where

$$\hbar\tilde{\omega} = \hbar\omega - E_C = \sqrt{8E_J E_C} - E_C = \hbar\omega + \alpha, \quad \alpha = -E_C$$

and the constant term is dropped. The lowest three energy levels of the simple harmonic oscillator are also eigenstates of the transmon Hamiltonian but with different eigenvalues

$$H_{\text{tr}} |0\rangle = 0, \quad H_{\text{tr}} |1\rangle = \hbar\tilde{\omega} |1\rangle, \quad H_{\text{tr}} |2\rangle = (2\hbar\tilde{\omega} + \alpha) |2\rangle$$

with $\alpha = -E_C$ being negative and one has noticed that

$$a |1\rangle = |0\rangle, \quad a |2\rangle = \sqrt{2} |1\rangle, \quad a^\dagger |0\rangle = |1\rangle, \quad a^\dagger |1\rangle = \sqrt{2} |2\rangle$$

The differences between the energy levels are

$$E_1 - E_0 = \hbar\tilde{\omega}, \quad E_2 - E_1 = \hbar\tilde{\omega} + \alpha$$

If $E_J/E_C \approx 40$, then one finds

$$\hbar\omega \approx \sqrt{8E_J E_C} \sim 10E_C, \quad \alpha \approx -E_C$$

so the difference from the energy levels of the simple harmonic oscillator is about 10%.

The excited state $|2\rangle$ is not far: very often it needs to be included in computations to give a better result (but it is also needed for driving, coupling, and Rabi oscillations). For the moment, one may ignore it and truncate the Hilbert space to the sector spanned by $|0\rangle$ and $|1\rangle$. Then one may rewrite all operators as two-by-two matrices. As usual

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

thus

$$a^\dagger a = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \equiv -\frac{1}{2}\sigma_3 + \frac{1}{2}I$$

The quartic piece $a^\dagger a^\dagger aa$ is zero on $|0\rangle$ and $|1\rangle$ (but it needs to be taken into account when including corrections due to the excited state $|2\rangle$). So, up to a constant, one recovers the Hamiltonian of a qubit

$$H = -\frac{1}{2}\hbar\tilde{\omega}\sigma_3$$

The dictionary oscillator/qubit is

$$a|0\rangle = 0, \quad a|1\rangle = |0\rangle, \quad a^\dagger|0\rangle = |1\rangle, \quad a^\dagger|1\rangle = 0$$

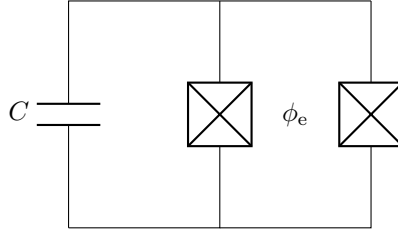
The last term is zero because one ignores $|2\rangle$. Thus

$$a = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \equiv \sigma_+, \quad a^\dagger = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \equiv \sigma_-$$

imply

$$\Phi \propto a + a^\dagger = \sigma_1, \quad Q \propto i(a^\dagger - a) = \sigma_2$$

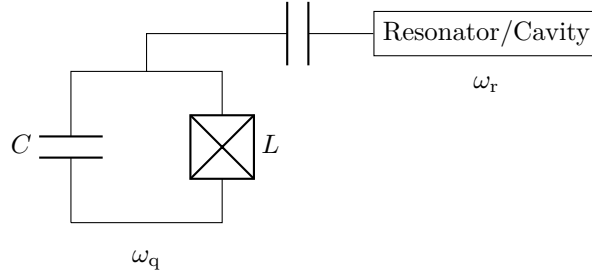
Tunable transmons. There are also tunable qubits, where $\tilde{\omega}$ can be varied by an external magnetic flux ϕ_e



6.14 cQED gates

Controlling and measuring qubits is done with external microwave fields (drives) and resonators (oscillating cavities). The logic is simple. However, to get the desired effects can be complicated.

Measuring and reading. To measure a qubit, one couples the qubit to an oscillator (resonator/cavity) with frequency ω_r



where $\omega_q = \tilde{\omega}$. This system can be described similarly to cavity QED. This is why the superconducting qubit is also called artificial atom. The coupling is given by the Jaynes–Cummings Hamiltonian ($\hbar = 1$):

$$H = -\frac{\omega_q}{2}\sigma_3 + \omega_r \left(a_r^\dagger a_r + \frac{1}{2} \right) + g(a_r^\dagger \sigma_+ + a_r \sigma_-)$$

where a_r, a_r^\dagger represent the resonator mode. In the dispersive limit $\Delta \gg g$, where the detuning frequency is $\Delta = \omega_q - \omega_r$, one has

$$H = -\frac{\tilde{\omega}_q}{2}\sigma_3 + (\omega_r - \chi\sigma_3)a_r^\dagger a_r, \quad \tilde{\omega}_q = \omega_q + \frac{g^2}{\Delta}$$

where the first term is the Lamb shift while

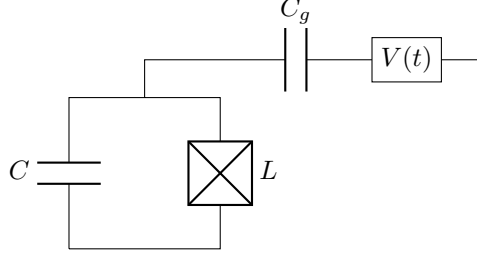
$$\chi = \frac{g^2}{\Delta}$$

is now corrected by higher levels since α is not too large

$$\chi = \frac{g^2}{\Delta} \rightarrow \frac{g^2}{\Delta(1 + \Delta/\alpha)}$$

Now the cavity frequency depends on the state of the qubit $\sigma_3 = \pm 1$: one can measure the cavity to learn the eigenvalue of σ_3 . This is a quantum nondemolition measurement (QND), it does not alter σ_3 .

Controlling qubits — single-qubit gates. To control qubits and construct single-qubit gates one has to couple the qubit to an external drive and use Rabi oscillations.



For a capacitive coupling, one expects the coupling energy between the qubit and the drive to be

$$C_g V_1 V_2 = \frac{C_g}{C} Q V_2$$

where V_1 and V_2 correspond to the qubit and drive, respectively. So that, recalling that Q is the conjugate momentum of the flux, $Q \propto i(a^\dagger - a) \equiv \sigma_2$, one obtains

$$H = -\frac{\omega_q}{2} \sigma_3 + ig(a^\dagger - a)V(t) = -\frac{\omega_q}{2} \sigma_3 + g\sigma_2 V(t)$$

which induces Rabi oscillations that have been already studied. If

$$V(t) = V_0 \cos(\omega_d t + \phi_0)$$

in the notation of the section about driven qubit and Rabi oscillations

$$Ae^{-i\phi_1} = b_1 + ib_2 = -igV_0 \implies A = gV_0, \quad \phi_1 = \frac{\pi}{2}$$

In the rotating frame, the Hamiltonian is

$$\tilde{H} = -\frac{1}{2}(\Delta\sigma_3 + A\sigma_1 \cos \phi - A\sigma_2 \sin \phi), \quad \phi = \phi_0 + \phi_1 = \phi_0 + \frac{\pi}{2}$$

and the evolution operator

$$U_{EV}(t) = e^{-i\tilde{H}t}$$

is the unitary operator on the qubit. At resonance, $\Delta = 0$, by choosing $\phi = 0$, one generates rotations around the x -axis on the Bloch sphere

$$U(t) = e^{\frac{i}{2}At\sigma_1} = R_x(-At)$$

and, choosing $\phi = \pi/2$, rotations about the y -axis

$$U(t) = e^{-\frac{i}{2}At\sigma_2} = R_y(At)$$

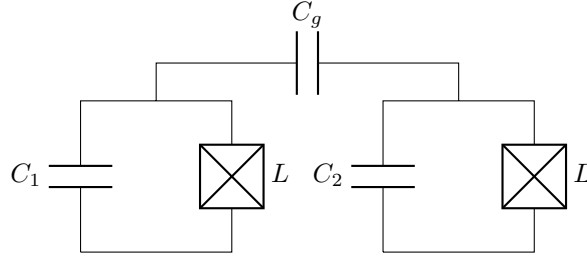
and, in the original frame, the qubit also has a precession around z . By tuning the amplitude V_0 , the phase ϕ_0 , and the time of evolution t , one may obtain all single-qubit gates.

The existence of the state $|2\rangle$ corrects the evolution and introduces mistakes. For example, starting from $|0\rangle$ with $\phi = 0$, $-At = \pi$, one obtains $|1\rangle$, but with the state $|2\rangle$, the target is missed. These can be corrected by changing the envelope $V_0 \rightarrow V_0(t)$ in order to undo the effect of excited levels (DRAG scheme: derivative reduction by adiabatic gate).

Lecture 21

Two-qubit gates. The simplest coupling is capacitive

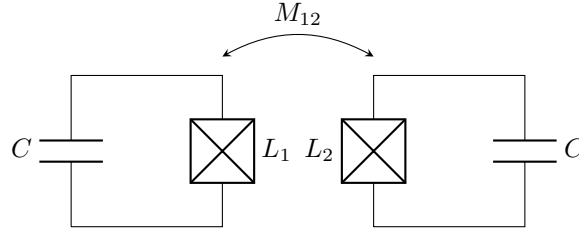
lun 16 dic
2024 16:30



whose Hamiltonian is

$$H_{\text{int}} = C_g V_1 V_2 = \frac{C_g}{C_1 C_2} Q_1 Q_2 \sim i(a_1^\dagger - a_1)(a_2^\dagger - a_2) \equiv \sigma_2^{(1)} \otimes \sigma_2^{(2)}$$

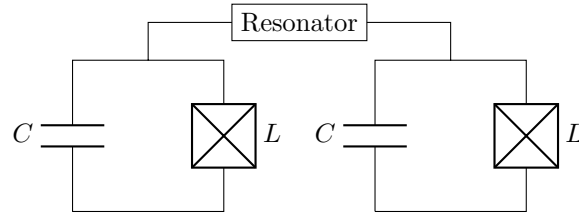
One may instead have inductive coupling



whose Hamiltonian is

$$H_{\text{int}} = M_{12} I_1 I_2 = \frac{M_{12}}{L_1 L_2} \Phi_1 \Phi_2 \sim (a_1 + a_1^\dagger)(a_2 + a_2^\dagger) \equiv \sigma_1^{(1)} \otimes \sigma_1^{(2)}$$

Or, instead, the coupling may be mediated by a resonator



whose Hamiltonian is

$$H_{\text{int}} = \omega_r \left(a_r^\dagger a_r + \frac{1}{2} \right) + g_1 (a_r^\dagger \sigma_+^{(1)} + a_r \sigma_-^{(1)}) + g_2 (a_r^\dagger \sigma_+^{(2)} + a_r \sigma_-^{(2)})$$

In any case the full Hamiltonian is

$$H = H_0 + H_{\text{int}} = -\frac{\omega_q^{(1)}}{2} \sigma_3^{(1)} - \frac{\omega_q^{(2)}}{2} \sigma_3^{(2)} + H_{\text{int}}$$

Transverse coupling. The transverse couplings, such as $\sigma_1^{(1)} \otimes \sigma_1^{(2)}$ or $\sigma_2^{(1)} \otimes \sigma_2^{(2)}$, are easy to obtain, for example, capacitively. With tunable qubits, one can construct a SWAP gate. Suppose the two qubits have tunable frequencies $\omega_q^{(1)}$ and $\omega_q^{(2)}$. Then, they oscillate at different rates. In a rotating frame $U = e^{iH_0 t}$

$$\sigma_\pm^{(i)} \rightarrow e^{\mp i\omega_q^{(i)} t} \sigma_\pm^{(i)}$$

and

$$\begin{aligned}\sigma_2^{(1)} \otimes \sigma_2^{(2)} &= -(\sigma_+^{(1)} - \sigma_-^{(1)}) \otimes (\sigma_+^{(2)} - \sigma_-^{(2)}) \\ &\rightarrow \sigma_+^{(1)} \sigma_-^{(2)} e^{-i(\omega_q^{(1)} - \omega_q^{(2)})t} + \sigma_-^{(1)} \sigma_+^{(2)} e^{i(\omega_q^{(1)} - \omega_q^{(2)})t} \\ &\quad - \sigma_+^{(1)} \sigma_+^{(2)} e^{-i(\omega_q^{(1)} + \omega_q^{(2)})t} - \sigma_-^{(1)} \sigma_-^{(2)} e^{i(\omega_q^{(1)} + \omega_q^{(2)})t}\end{aligned}$$

and if one tunes at resonance $\omega_q^{(1)} = \omega_q^{(2)}$, then the Hamiltonian is time-independent

$$\tilde{H}_{\text{int}} = g(\sigma_+^{(1)} \sigma_-^{(2)} + \sigma_-^{(1)} \sigma_+^{(2)})$$

for some coupling g , where the tensor product \otimes is understood. As a matrix, one has

$$\tilde{H}_{\text{int}} = g \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + g \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & g & 0 \\ 0 & g & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

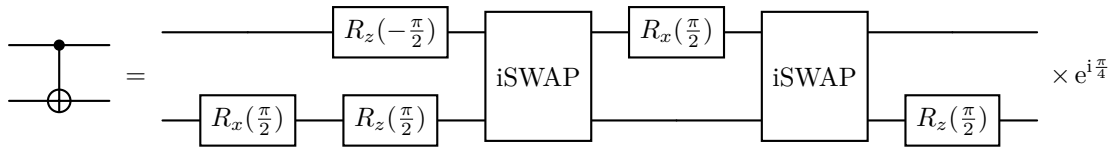
This gives the time-evolution operator

$$U_{\text{EV}}(t) = e^{-i\tilde{H}_{\text{int}}t} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(gt) & -i\sin(gt) & 0 \\ 0 & -i\sin(gt) & \cos(gt) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} |00\rangle \\ |01\rangle \\ |10\rangle \\ |11\rangle \end{matrix}$$

this form can be computed by noting that the central 2×2 block of \tilde{H}_{int} is just σ_1 and then one just expands the exponential using Euler's formula. For $t = \frac{3\pi}{2g}$, one obtains an iSWAP gate

$$U_{\text{EV}}\left(\frac{3\pi}{2g}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

It swaps $|01\rangle \leftrightarrow |10\rangle$ and adds a phase i . The iSWAP gate with the single-qubit gates H , T and S constitutes a universal set of gates. Though most algorithms use a CNOT, but realizing a CNOT gate is involved and expensive (it requires two iSWAP gates):



and requires tunable qubits.

Parametric coupling. The iSWAP gate can be obtained also without tunable qubits by modulating the coupling constant

$$g(t)\sigma_1 \otimes \sigma_1, \quad g(t) \sim \cos \omega_d t, \quad \omega_d = \omega_q^{(1)} - \omega_q^{(2)}$$

For example, with inductive couplings, one may consider three qubits where the middle one has a varying magnetic field. [r] image

Controlled-Z gate. There are other gates made with tunable qubits that use higher energy levels, for example, the controlled-Z gate (also called CPHASE). A CZ gate is an effective ZZ coupling

$$\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix} = \exp \left[-\frac{i\pi}{4} (\sigma_3 \otimes \sigma_3 - \sigma_3 \otimes I - I \otimes \sigma_3 + I \otimes I) \right]$$

Notice how the coupling is given by the first term in the exponential; the other three are one-qubit operations.

This gate can be realized by using also the energy level $|2\rangle$ and tuning the frequency $\omega_q^{(2)}$. Let $\omega_q^{(i)} + \alpha_i$ be the energy difference between this level and $|1\rangle$. The Hamiltonian of the system is

$$H = -\frac{\omega_q^{(1)}}{2}a_1^\dagger a_1 + \frac{\alpha_1}{2}a_1^\dagger a_1^\dagger a_1 a_1 - \frac{\omega_q^{(2)}}{2}a_2^\dagger a_2 + \frac{\alpha_2}{2}a_2^\dagger a_2^\dagger a_2 a_2 - g(a_1^\dagger - a_1)(a_2^\dagger - a_2)$$

where the last addendum is the interaction Hamiltonian H_{int} of a capacitive coupling. If $\omega_q^{(1)} \sim \omega_q^{(2)}$, one may still approximate

$$H_{\text{int}} \sim g(a_1 a_2^\dagger + a_2 a_1^\dagger)$$

but now the differences in energies compared to $|00\rangle$ are

$$H_0 = \begin{bmatrix} \omega_q^{(1)} & & & & \\ & \omega_q^{(2)} & & & \\ & & \omega_q^{(1)} + \omega_q^{(2)} & & \\ & & & 2\omega_q^{(1)} + \alpha_1 & \\ & & & & 2\omega_q^{(2)} + \alpha_2 \end{bmatrix} \begin{matrix} |10\rangle \\ |01\rangle \\ |11\rangle \\ |20\rangle \\ |02\rangle \end{matrix}$$

where $\alpha_i < 0$. The interaction Hamiltonian connects the various states

$$\tilde{H}_{\text{int}} = \begin{bmatrix} 0 & & & & \\ & 0 & g & & \\ & g & 0 & & \\ & & & 0 & \sqrt{2}g & \sqrt{2}g \\ & & & \sqrt{2}g & 0 & 0 \\ & & & \sqrt{2}g & 0 & 0 \end{bmatrix} \begin{matrix} |00\rangle \\ |10\rangle \\ |01\rangle \\ |11\rangle \\ |20\rangle \\ |02\rangle \end{matrix} = \sum_{nm} c_{nm} H_{nm}$$

where H_{nm} is the matrix with 1 in position (n, m) and zero in all others. For $\omega_q^{(2)} \neq \omega_q^{(1)}$, each coefficient c_{nm} gets a time dependence. In fact, noting

$$[H_0, H_{nm}] = (E_n - E_m)H_{nm}$$

in the rotating frame, one obtains

$$U H_{nm} U^\dagger = e^{i(E_n - E_m)t} H_{nm}, \quad U = e^{iH_0 t}$$

so that

$$\tilde{H}_{\text{int}} \rightarrow \sum_{nm} e^{i(E_n - E_m)t} c_{nm} H_{nm}$$

Reading the values of E_n and E_m from the matrix representation of H_0 , then when the commutator is zero, one can use the remaining coupling to implement gates. If $\omega_q^{(1)} = \omega_q^{(2)}$, then the entries of the lower-right sub-matrix of the interaction are oscillating fast and the interaction Hamiltonian becomes again

$$\tilde{H}_{\text{int}} = \begin{bmatrix} 0 & & & & \\ & 0 & g & & \\ & g & 0 & & \\ & & & 0 & \\ & & & & 0 \\ & & & & & 0 \end{bmatrix}$$

When computing the time-evolution operator, this matrix leads to a rotation between $|01\rangle$ and $|10\rangle$, similar to the one above. If one tunes $E_{|02\rangle} = E_{|11\rangle}$, then the interaction Hamiltonian is

$$\tilde{H}_{\text{int}} = \begin{bmatrix} 0 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & 0 & \sqrt{2}g \\ & & & \sqrt{2}g & 0 \end{bmatrix}$$

The tuning can be done through $\omega_q^{(2)}$ so that

$$\omega_q^{(1)} + \omega_q^{(2)} = 2\omega_q^{(2)} + \alpha_2 \implies \omega_q^{(2)} = \omega_q^{(1)} - \alpha_2 > \omega_q^{(1)}$$

This matrix leads to a rotation between $|11\rangle$ and $|02\rangle$. However, this last state is not of interest. If one tunes g in such a way that the unitary operator (i.e. the time-evolution operator at a particular time) is

$$R_x(2\pi) = -I, \quad |11\rangle \rightarrow -|11\rangle, \quad |02\rangle \rightarrow -|02\rangle$$

then this gives exactly the CZ gate (since the only relevant energy levels for the qubits are $|0\rangle$ and $|1\rangle$).

There are several ways to implement this gate. As shown above, one may tune $E_{|02\rangle} = E_{|11\rangle}$. Another method is the adiabatic excursion: instead of tuning the frequencies to be equal, they are made almost equal. One modifies the system by placing it in a cavity and the spectrum of the system is as follows [r] image, figs. 16a, 16b of 2009.08021

Consider a system with a tunable qubit $\omega_q^{(2)}$, starting in a configuration with

$$\omega_q^{(2)} > \omega_q^{(1)}$$

One lowers the tunable frequency until it reaches $\omega_q^{(1)}$. Due to the cavity, the energy levels $|10\rangle$ and $|01\rangle$ are almost equal (but not exactly due to the presence of other couplings). At such point, one operates an iSWAP. This gate can be obtained as follows. One lowers very slowly (i.e. in the adiabatic approximation) the tunable frequency until the point where the frequencies are almost equal, then one goes back to the starting frequency. One moves adiabatically along the black arrow lines and each qubit gets a phase that can be fixed by choosing the time t . If one chooses the phases and the time appropriately, then one obtains an iSWAP.

The same can be done with the excited level. This gives a CZ gate. Notice how this gate is on the left of the iSWAP (meaning it happens for higher values of $\omega_q^{(2)}$) since one tunes

$$\omega_q^{(2)} = \omega_q^{(1)} - \alpha_2 > \omega_q^{(1)}$$

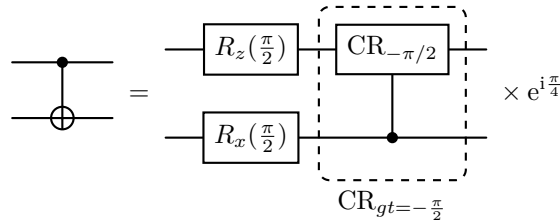
Cross-resonance gate. The coupling

$$\tilde{H}_{\text{int}} = \frac{g}{2} \sigma_3^{(1)} \otimes \sigma_1^{(2)} = \frac{g}{2} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

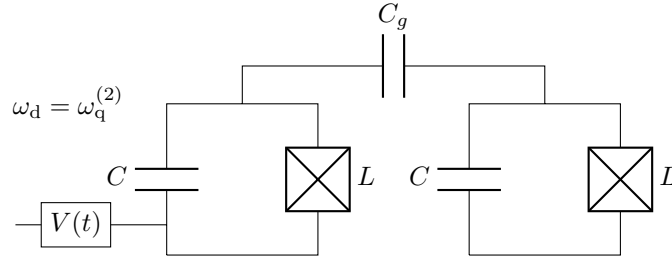
would lead to a simpler gate, called cross-resonance (CR) gate. The evolution operator is

$$U_{\text{EV}} = e^{-i\tilde{H}_{\text{int}}t} = \begin{bmatrix} \cos \frac{gt}{2} & -i \sin \frac{gt}{2} & 0 & 0 \\ -i \sin \frac{gt}{2} & \cos \frac{gt}{2} & 0 & 0 \\ 0 & 0 & \cos \frac{gt}{2} & i \sin \frac{gt}{2} \\ 0 & 0 & i \sin \frac{gt}{2} & \cos \frac{gt}{2} \end{bmatrix} \equiv \text{CR}_{gt}$$

It rotates separately between $|00\rangle \leftrightarrow |01\rangle$ and $|10\rangle \leftrightarrow |11\rangle$. This leads to the CNOT



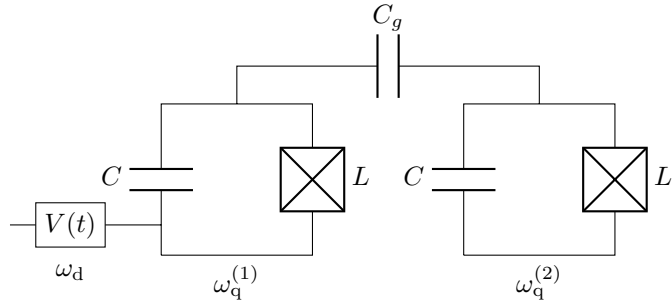
and can be obtained with transmons without tuning the frequency, but with cross-resonance (CR), driving qubit 1 with frequency equal to $\omega_q^{(2)}$ (tricky, see next section)



where $V(t)$ is a generator of microwaves, a driver with $\omega_d = \omega_q^{(2)}$.

6.15 Cross-resonance gate

No exam? Not done. Consider two transmons, driving the first with an external force with frequency ω_d . The transmons are capacitively coupled.



The Hamiltonian is

$$H = -\frac{\omega_q^{(1)}}{2}\sigma_3^{(1)} - \frac{\omega_q^{(2)}}{2}\sigma_3^{(2)} + J(\sigma_+^{(1)}\sigma_-^{(2)} + \sigma_-^{(1)}\sigma_+^{(2)}) + \mathcal{E}_d(t)(\sigma_+^{(1)}e^{-i\omega_d t} + \sigma_-^{(1)}e^{i\omega_d t})$$

$$= H_0 + H_{\text{int}} + H_{\text{ext}}$$

The first two addenda are the free Hamiltonian H_0 . The third addendum is H_{int} and it is $\sigma_2^{(1)} \otimes \sigma_2^{(2)}$ after applying the rotating wave approximation (RWA). The last addendum is an external coupling H_{ext} to qubit 1. One assumes for

$$J \ll \Delta_{\text{qq}} = \omega_q^{(1)} - \omega_q^{(2)}$$

a sort of dispersive limit. One may study such limit using the Schrieffer–Wolff transformation

$$U = e^{-\frac{J}{\Delta_{\text{qq}}}(\sigma_+^{(1)}\sigma_-^{(2)} - \sigma_-^{(1)}\sigma_+^{(2)})} = e^S, \quad S^\dagger = -S$$

where $[H_0, S] = H_{\text{int}}$. Indeed

$$[\sigma_3, \sigma_\pm] = \pm 2\sigma_\pm$$

and

$$[H_0, S] = \frac{J}{\Delta_{\text{qq}}}[(\omega_q^{(1)} - \omega_q^{(2)})\sigma_+^{(1)}\sigma_-^{(2)} - (-\omega_q^{(1)} + \omega_q^{(2)})\sigma_-^{(1)}\sigma_+^{(2)}] = J(\sigma_+^{(1)}\sigma_-^{(2)} - \sigma_-^{(1)}\sigma_+^{(2)}) = H_{\text{int}}$$

Then for $J \ll \Delta_{\text{qq}}$, one finds

$$\begin{aligned} \tilde{H} &= U H U^\dagger = S H e^{-S} = H + [S, H] + \frac{1}{2}[S, [S, H]] + \dots \\ &= H_0 + H_{\text{int}} + H_{\text{ext}} + [S, H_0] + [S, H_{\text{int}}] + [S, H_{\text{ext}}] \\ &\quad + \frac{1}{2}[S, [S, H_0]] + \frac{1}{2}[S, [S, H_{\text{int}}]] + \frac{1}{2}[S, [S, H_{\text{ext}}]] + \dots \\ &= H_0 + H_{\text{int}} + H_{\text{ext}} - H_{\text{int}} + [S, H_{\text{int}}] + [S, H_{\text{ext}}] - \frac{1}{2}[S, H_{\text{int}}] + O(J^2/\Delta_{\text{qq}}^2) \\ &\approx H_0 + H_{\text{ext}} + \frac{1}{2}[S, H_{\text{int}}] + [S, H_{\text{ext}}] \end{aligned}$$

The commutators are

$$\frac{1}{2}[S, H_{\text{int}}] = -\frac{J^2}{2\Delta_{\text{qq}}} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = -\frac{J^2}{\Delta_{\text{qq}}} \frac{1}{2} \sigma_3^{(1)} \otimes I + \frac{J^2}{\Delta_{\text{qq}}} \frac{1}{2} I \otimes \sigma_3^{(2)}$$

and

$$[S, H_{\text{ext}}] = -\mathcal{E}_d(t) \frac{J}{\Delta_{\text{qq}}} \sigma_3^{(1)} \otimes (\sigma_+^{(2)} e^{-i\omega_d t} + \sigma_-^{(2)} e^{i\omega_d t})$$

The last parenthesis drives the qubit 2. The full Hamiltonian is

$$\begin{aligned} \tilde{H} = & -\left(\omega_q^{(1)} + \frac{J^2}{\Delta_{\text{qq}}}\right) \frac{\sigma_3^{(1)}}{2} - \left(\omega_q^{(2)} + \frac{J^2}{\Delta_{\text{qq}}}\right) \frac{\sigma_3^{(2)}}{2} + \mathcal{E}_d(t) (\sigma_+^{(1)} e^{-i\omega_d t} + \sigma_-^{(1)} e^{i\omega_d t}) \\ & - \mathcal{E}_d(t) \frac{J}{\Delta_{\text{qq}}} \sigma_3^{(1)} \otimes (\sigma_+^{(2)} e^{-i\omega_d t} + \sigma_-^{(2)} e^{i\omega_d t}) \end{aligned}$$

Making a final rotation with

$$U = e^{i\hat{H}_0 t}, \quad \hat{H}_0 = \frac{1}{2} \omega_d (\sigma_3^{(1)} + \sigma_3^{(2)})$$

one obtains

$$\begin{aligned} H' = U \tilde{H} U^\dagger + i\dot{U} U^\dagger = & -\left(\omega_q^{(1)} + \frac{J^2}{\Delta_{\text{qq}}} + \omega_d\right) \frac{\sigma_3^{(1)}}{2} - \left(\omega_q^{(2)} - \frac{J^2}{\Delta_{\text{qq}}} + \omega_d\right) \frac{\sigma_3^{(2)}}{2} \\ & + \mathcal{E}_d(t) (\sigma_+^{(1)} + \sigma_-^{(1)}) - \mathcal{E}_d(t) \frac{J}{\Delta_{\text{qq}}} \sigma_3^{(1)} \otimes (\sigma_+^{(2)} + \sigma_-^{(2)}) \\ = & -\left(\omega_q^{(1)} + \frac{J^2}{\Delta_{\text{qq}}} + \omega_d\right) \frac{\sigma_3^{(1)}}{2} - \left(\omega_q^{(2)} - \frac{J^2}{\Delta_{\text{qq}}} + \omega_d\right) \frac{\sigma_3^{(2)}}{2} \\ & + \mathcal{E}_d(t) \sigma_1^{(1)} - \mathcal{E}_d(t) \frac{J}{\Delta_{\text{qq}}} \sigma_3^{(1)} \otimes \sigma_1^{(2)} \end{aligned}$$

Setting

$$\omega_d = -\omega_q^{(2)} + \frac{J^2}{\Delta_{\text{qq}}}$$

the only two-qubit gate operation is the last term, the CR term

$$H'' = \frac{1}{2} g_{\text{CR}} \sigma_3^{(1)} \otimes \sigma_1^{(2)}$$