

# Theory of Quantum Information and Quantum Computing

Maso\*

October 25, 2024

## Contents

<b>1</b>	<b>Quantum mechanics</b>	<b>2</b>
1.1	Qubits . . . . .	3
1.2	Observables and measurements . . . . .	4
1.3	Time evolution . . . . .	6
1.4	Multipartite systems . . . . .	8
1.5	No-cloning . . . . .	11
<b>2</b>	<b>Entanglement</b>	<b>12</b>
2.1	Bell states . . . . .	12
2.2	Applications . . . . .	14
2.3	Bell inequalities . . . . .	15
2.4	Quantum encryption . . . . .	18
<b>3</b>	<b>Quantum algorithms</b>	<b>19</b>
3.1	Quantum and classical circuits . . . . .	19
3.2	Universality . . . . .	21
3.3	Quantum parallelism . . . . .	22
3.4	Deutsch's algorithm . . . . .	24
3.5	Other simple algorithms . . . . .	26
3.6	Quantum Fourier transform . . . . .	28
3.7	Shor's algorithm . . . . .	31
3.8	Breaking RSA encryption . . . . .	33
3.9	Grover's search algorithm . . . . .	35
<b>4</b>	<b>Open systems</b>	<b>36</b>
4.1	Density matrix . . . . .	36

## Lecture 1

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

**Knowledge required.** Quantum mechanics from the bachelor.

---

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

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

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

Currently there are several quantum computers. The first quantum computer was based on quantum annealing which is good for optimization problems, but this is not treated in the course. The course studies the superconducting qubits at IBM and Google. Right now, the order of qubits is about a thousand and not much can be done. The majority of those qubits are allocated for error correction. The problem with qubits is transfer the classical information to the two levels in an isolated way. It is difficult to keep stable an excitation and it is even more difficult to maintain a superposition with a given relative phase. With time, the phases get washed out by the interactions with the environment. Keeping coherence of a two-level quantum system is difficult.

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

## 1 Quantum mechanics

See Nielsen §§1.2, 1.3, 2.1, 2.2.

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

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

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

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

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

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

If the two above are probabilities, then it must hold

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

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

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

## 1.1 Qubits

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

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

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

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

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

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

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

The length is defined as

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

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

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

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

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 pair of complex numbers linear combination of the two previous ones

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

The scalar product between two arbitrary vectors

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

is the typical Hermitian scalar product

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

One identifies

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

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

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

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

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

There is a convenient parametrization of the parameters

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

Applying phase invariance, one may eliminate a phase

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

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

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

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

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

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

## 1.2 Observables and measurements

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

$$A^\dagger = A$$

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

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

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

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

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

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

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

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

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

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

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

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

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

where  $\sigma_i$  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  $\sigma_3$ :

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

[r] If  $\sigma_3$  is an observable, there must be an eigenbasis. One may check that the eigenbasis of  $\sigma_1$  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  $\sigma_2$  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  $\sigma_i$  lie on the axis  $i$ . For a generic point on the sphere, one may define the spin in a generic direction

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

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

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

## Lecture 2

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

mer 09 ott  
2024 16:30

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

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.

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

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

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

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

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

The probability is just

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

The state collapses to the “normalized” state

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

Notice that the normalization may give a phase, but one is interested in rays as quantum states. The probability  $\alpha_n$  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  $\sigma_1$  and  $\sigma_3$  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  $\sigma_3$  gives 0, but now the spin along the  $x$  direction is undetermined.

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

### 1.3 Time evolution

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

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

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

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

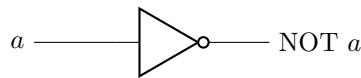
is conserved. The fact that  $U(t)$  is unitary follows from the Schrödinger equation

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

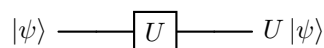
this holds since the Schrödinger equation also implies

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

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

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

$$|0\rangle \text{ --- } \boxed{X} \text{ --- } |1\rangle, \quad |1\rangle \text{ --- } \boxed{X} \text{ --- } |0\rangle$$

while  $Z$  flips signs

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

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

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

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

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

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

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  $e^{-i\alpha} R$  generate 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 1.8.** 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 1.9.** Prove that  $R_{\mathbf{n}}(x)$  acts as a rotation on a state on the Bloch sphere. See Nielsen, exercise 4.6.

## 1.4 Multipartite systems

One would like to study multiple qubits.

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

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

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

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

There must exist four states

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

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

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

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

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

The probability of getting 0 in the first qubit is

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

One sees that one has to 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

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

Vectors in the tensor product space can be written as

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

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

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

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

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

**Example 1.11.** The following state is separable

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

while the following states are entangled

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

A system of two qubits is depicted as

$$\begin{array}{c} |0\rangle, |1\rangle \\ |0\rangle, |1\rangle \end{array} \longrightarrow \boxed{U} \longrightarrow \left. \begin{array}{c} \\ \\ \end{array} \right\} \sum_{n,m} \alpha_{nm} |nm\rangle$$

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

$$U: \begin{array}{c} \text{---} \boxed{A} \text{---} \\ \text{---} \boxed{B} \text{---} \end{array}$$

In other words, there exist  $2 \times 2$  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  $2 \times 2$  description of the single qubits to a  $4 \times 4$  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 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  $\sigma_1$  and  $\sigma_3$  as the operators and the  $4 \times 4$  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  $4 \times 4$  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

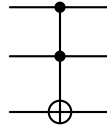
$$\begin{array}{c} |x\rangle \text{---} \bullet \text{---} |x\rangle \\ |y\rangle \text{---} \oplus \text{---} |x \oplus y\rangle \end{array}$$

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

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

Therefore, the gate CNOT is indeed a XOR.

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



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

## Lecture 3

lun 14 ott  
2024 16:30

### 1.5 No-cloning

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

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

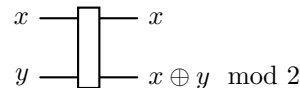
With many copies of the initial state  $|\psi\rangle$ , one may perform many measurements and extract the values of  $|\alpha|^2$  and  $|\beta|^2$ .

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

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



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

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

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

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 1.12** (No-cloning). Given a normalized state  $|\phi\rangle$ , 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  $|\psi\rangle$ .

*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

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

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 a basis, but this is enough.

## 2 Entanglement

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

### 2.1 Bell states

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

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

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

A simple example of a two-qubit entangled state is

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

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

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

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

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 in another sub-luminal method. Therefore, information cannot travel between them using the qubits only.

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

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

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

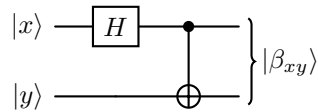
$$|\psi\rangle = \sum_{n,m=0}^1 \alpha_{nm} |\beta_{nm}\rangle$$

and the probability is given by

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

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

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



The various inputs and outputs are

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

recalling that the Hadamard gates acts as

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

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

## 2.2 Applications

There are many applications of entanglement in physics.

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

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

Suppose that the second qubit is in possession of observer  $B$  without modification. Observer  $A$  may perform an operation on their entangled qubit and then send that qubit to observer  $B$  who may perform a measurement to determine which Bell state the entangled pair is in. Observer  $A$  can do four operations, each one corresponding to a different Bell state. To send  $|\beta_{00}\rangle$ , one may do nothing. To send  $|\beta_{01}\rangle$ , one applies  $X$

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

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

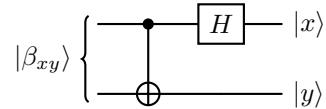
$$10 : (Z \otimes I) |\psi\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}} = |\beta_{10}\rangle$$

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



since

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

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

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

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

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

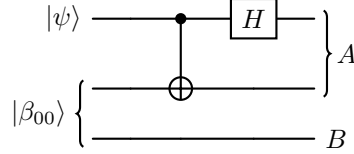
However, by sharing an entangled pair

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

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

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

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



to have

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

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

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

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

## Lecture 4

### 2.3 Bell inequalities

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

mar 16 ott  
2024 16:30

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  $(a, \lambda)$ : the result  $a$  and an experimentally inaccessible variable  $\lambda$ . To know the full theory, one needs to measure  $(a, \lambda)$ . 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 that 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  $\lambda$  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  $\lambda$  [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  $\pm 1$ . Observer  $B$  can measure  $b$  and  $b'$  with possible outcomes  $\pm 1$ . The two observers measure simultaneously and choose randomly whether to measure the primed or non-primed observable. The two cannot communicate. In a hidden variable theory, there would be a probability distribution  $P(a, a', b, b')$  and each variable existing in a given state  $\pm 1$  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  $\pm 1$ , the term  $a + a'$  is either 0 or  $\pm 2$  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  $\theta_{cd}$  is the angle between the two unit vectors  $\mathbf{c}$  and  $\mathbf{d}$ . If the two observers choose their axis to be separated by  $45^\circ$  going counter-clockwise in the sequence  $a'$ ,  $b$ ,  $a$ ,  $b'$ , then

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

This violates the CHSH inequality

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

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

Experiments have been done that support the violation. Physics is quantum. The most famous experiment (Aspect '82<sup>123</sup>) still had loopholes:

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

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

**Exercise 2.4.** Show that

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

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

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

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

**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

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

where one uses

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

## 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 message. Sending quantum states (e.g. polarized photons) along optical fibers in a protected way is not so difficult.

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

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

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

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

by adding them bitwise

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

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

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

**Protocol BB84.** Quantum mechanics may help solve the problem of creating a new random code. The quantum version, developed by Bennett and Brassard in 1984 (BB84), is the following. Observer  $A$  sends randomly two types of qubits: type  $C$  qubits  $|0\rangle$  or  $|1\rangle$ , or type  $H$  qubits  $|+\rangle$  or  $|-\rangle$ . The type and the state are both randomly chosen. Observer  $B$  receives the qubits, identifies them by the sequence in which they arrive and makes the choice of randomly measuring the result in the computational or Hadamard basis (measuring in the Hadamard basis is the same as applying the Hadamard gate and 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  $B$  chooses the same basis for the measurement, then they know that the result of the measurement is the same that  $A$  wanted to send. If the measurement is in a different basis,  $B$  gets a random result, uncorrelated with  $A$ 's. Observer  $B$  tells  $A$  over an insecure channel, which qubits were measured in the same basis and they both discard the others. What remains is a random series of 0s and 1s that can be used as a codepad.

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

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

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

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

### 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.

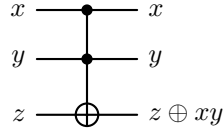
**Reversibility.** One difference is reversibility. Quantum circuits are reversible because gates are unitary

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

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

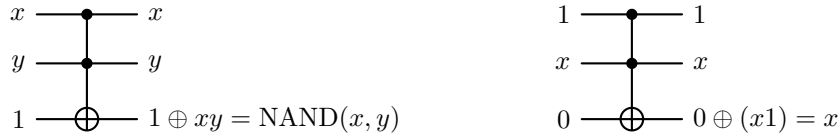


However, all classical computations can be written in a reversible way using multiple-bit gates. For example, the Toffoli gate is



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  $2 \times 2$  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  $2 \times 2$  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

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

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

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

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 holds.

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

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

with four real parameters  $\alpha, \beta, \gamma$  and  $\delta$ . Multiple qubits are messier: there are  $2^n$  states and the unitary matrices  $U \in U(2^n)$  contain  $2^{2n}$  real parameters.

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

### 3.2 Universality

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

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

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

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

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

In the proof, one notices that the generic unitary operator on  $n$  qubits  $U \in U(2^n)$  can be written as a product of  $2^{2n}$  2-qubit unitary operators. Then any  $4 \times 4$  unitary matrix  $U$  can be written as products of 1-qubit operators and CNOTs. For  $4 \times 4$  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  $\varepsilon$ . It is only a logarithmic loss and one can do all 2-qubits operations.

### 3.3 Quantum parallelism

To encode a number  $x$  with a classical bit description

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

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  $|x\rangle$  are a basis in the Hilbert space  $\mathcal{H}$  of dimension  $\dim \mathcal{H} = N = 2^n$  and are called the computational basis.

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

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

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

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

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

$$\begin{array}{c} 1 \text{ — } \boxed{X} \text{ — } 0 \\ 0 \text{ ————— } 0 \\ 0 \text{ ————— } 0 \\ 1 \text{ — } \boxed{X} \text{ — } 0 \\ 1 \text{ — } \boxed{X} \text{ — } 0 \end{array}$$

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

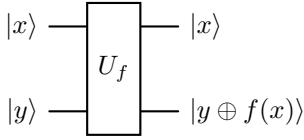
$$f(x) : \{0, 1\}^{\otimes n} \rightarrow \{0, 1\}, \quad x \mapsto f(x) = 0, 1$$

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

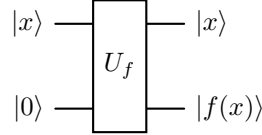
One important aspect is the use of a data register (or query) where  $x$  is stored and a target (or output) register where the result  $y$  is recorded

$$|x, y\rangle$$

One typically constructs a circuit that computes

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


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



The operator  $U_f$  is its own inverse

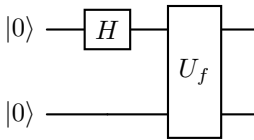
$$|x, y\rangle \rightarrow |x, y \oplus f(x)\rangle \rightarrow |x, y \oplus f(x) \oplus f(x)\rangle = |x, y\rangle$$

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 computation basis, but in a superposition (which cannot be done in classical computing). Consider  $n = 1$  and start with

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

This can be done by applying first an Hadamard gate

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


Then the result is

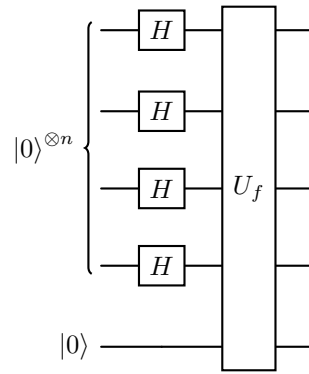
$$\frac{1}{\sqrt{2}} \left[ |0, f(0)\rangle + |1, f(1)\rangle \right]$$

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

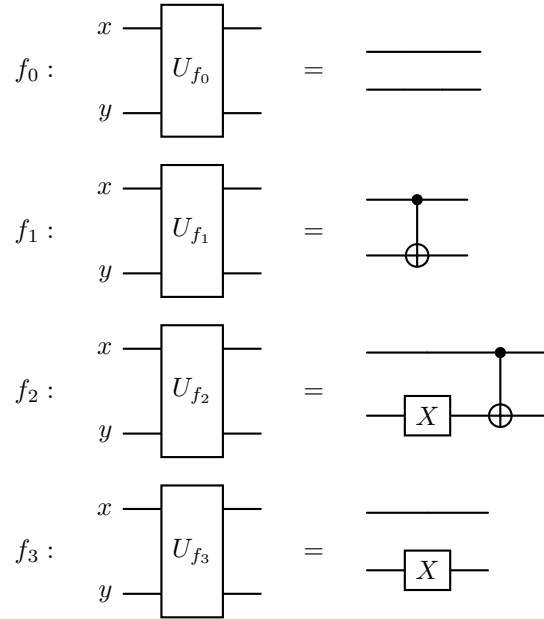
$f(x)$	$x = 0$	$x = 1$
$f_0$	0	0
$f_1$	0	1
$f_2$	1	0
$f_3$	1	1

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  $f_j$ .

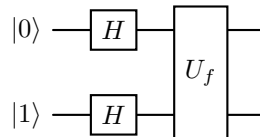
## Lecture 6

Consider a circuit with an unknown  $f_j$ , an oracle<sup>4</sup>. 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



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

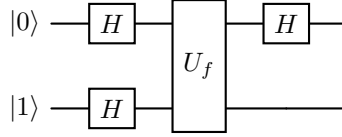
The interference is useful because

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

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

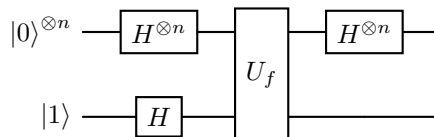
There are many other quantum algorithms that are more efficient than classical ones. In each, one learns some information about some function.

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

The algorithm is the same as Deutsch’s



In particular

- Using  $H^{\otimes n}$ , one transforms

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

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  $z \bmod 2$

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

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

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

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

Therefore, if  $f(x)$  is constant, then the state is

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

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

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

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

**Bernstein–Vazirani algorithm.** Consider the function

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

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 computations to find every component  $a_j$ . With computing computing, one needs just one evaluation of the oracle  $U_f$ .

Consider the same circuit as before. It produces

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

The relevant sum is

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

There are two cases:

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

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

- If  $a_j + z_j = 1 \pmod{2}$ , then

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

In the first case, the result is

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

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

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

### 3.6 Quantum Fourier transform

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

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

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

When acting on a generic state

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

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 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  $U_{\text{FT}}$  is constructed explicitly. Since

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

the operator  $U_{\text{FT}}$  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

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

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

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

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

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

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 eigenvalues<sup>5</sup>  $x_0, x_1, x_2$  and  $x_3$ :

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

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

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

where  $P$  is a permutation

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

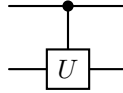
The operator  $U_{FT}$  is unitary. In fact, all operators with a matrix  $n$  are of the form

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

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

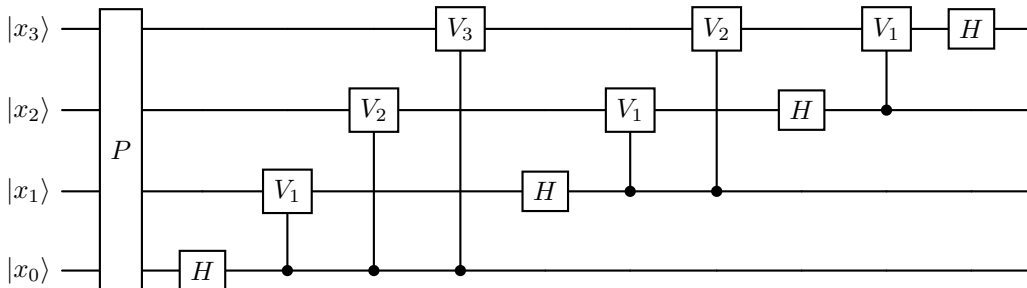


that does nothing if the upper qubit is  $|0\rangle$ , while applies  $U$  to the second qubit if the first is  $|1\rangle$ . The CNOT gate is  $U = X$ .

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.

<sup>5</sup>The subscript of  $n_j$  corresponds to the position of the qubit in the sequence  $|x_{n-1}\rangle \cdots |x_0\rangle$  and not to specifically  $|x_j\rangle$  in any position.

## Lecture 7

 lun 28 ott  
2024 16:30

### 3.7 Shor's algorithm

Consider a function acting on the integers

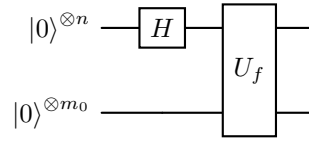
$$f : \mathbb{Z} \rightarrow \{0, 1\}^{\otimes m_0}$$

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 sum of all the values  $|x_i\rangle$  that satisfy  $f(x_i) = f_0$ . Since  $2^n \sim N^2$ , 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

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

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

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

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

If one measures  $|y\rangle$ , one finds a value with probability

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

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  $j = 1, \dots, r$ . 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  $m$ . 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.

**Exercise 3.3.** Study the probability  $P(y)$  near  $y = j2^n/r + \delta$ .

**Solution.** Using the geometric series one finds

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

Since

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

one has

$$P(\delta) \sim \frac{1}{r} \left[ \frac{\sin \pi \delta}{\pi \delta} \right]^2$$

and

$$\frac{1}{n} \left[ \frac{\sin n x}{\pi x} \right]^2 \rightarrow \delta(x), \quad n \rightarrow \infty$$

[r] ?

It turns out that  $4/\pi^2 \approx 40\%$  of the distribution  $P(y)$  is near one of the peaks  $y_j$  with a maximum error of  $1/2$ .

**Exercise 3.4.** Check that the maximum error is  $1/2$ .

**Solution.** Consider

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

also

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

Then

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

Since there are  $r$  values, one finds

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

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

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

and one finds  $r$  by comparing  $y/2^n$  and  $j/r$ . In this case, one needs  $n > n_0$  to be sure that  $j/r$  is unique. In fact,  $y/2^n$  divides the interval  $[0, 1]$  in  $2^n$  subintervals. The rational number  $j/r$



is somewhere at a distance less than  $1/2$  from some  $y/2^n$ . Two different rational numbers  $j_1/r_1$  and  $j_2/r_2$  differ by

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

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

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

which is true since

$$\left| x - \frac{1}{r} \right| \leq \frac{1}{22^n} = \frac{1}{2N^2} \leq \frac{1}{2r^2}, \quad r < N$$

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 factor 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

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

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  $b^r \equiv 1 \pmod{N}$ , while a classical computer finds  $d'$  such that  $cd' \equiv 1 \pmod{N}$ . 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

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

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

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

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  $a^{\frac{r}{2}} \equiv 1 \pmod{N}$  and the period is then  $r/2$  instead of  $r$ . Suppose also that

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

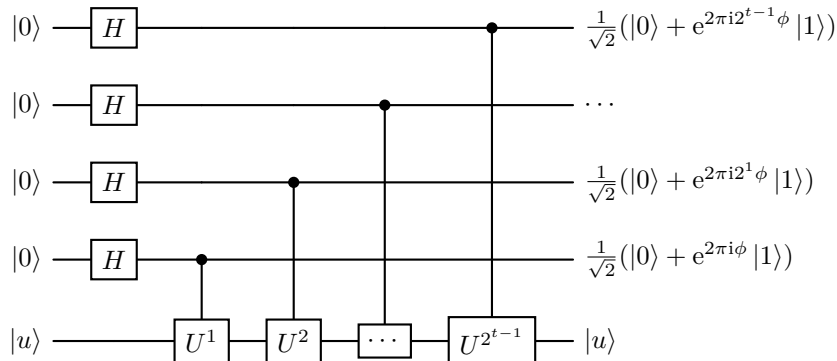
Since  $N = pq$  and  $(x-1)(x+1)$  is divisible<sup>6</sup> by  $N$ , but neither  $x-1$  nor  $x+1$  are, then  $x-1$  must be divisible by  $p$  and  $x+1$  by  $q$  (or viceversa). However, the only divisors of  $N$  are  $p$  and  $q$  themselves, so  $p$  is the greatest common divisor of  $N$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

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

**Phase estimation.** See Nielsen, §5.2. 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



<sup>6</sup>This is a consequence of  $0 = (x-1)(x+1) \pmod{N}$ .

The data register is in the state

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

whose quantum Fourier transform is exactly  $|\phi\rangle$ . If  $0 \leq \phi < 2^t$ , a quantum Fourier transform would give just  $|\phi\rangle$ . A measurement in the computational basis gives the value  $\phi$ . Otherwise, it is an approximation and one can estimate the number  $t$  such that it works well.

However, it is difficult to start with  $|u\rangle$ . If one begins with

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

one ends up with

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

and a measurement gives just one eigenvalue  $\phi_u$ .

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

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

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

### 3.9 Grover's search algorithm

The other class of problems we gain a speedup in is *searching in a database*. Suppose we have  $N$  objects and we search for a specific one. Abstractly, we have a function  $f(x)$  from  $x = 0, \dots, N-1$  to 0 or 1 such that

$$\begin{aligned} f(x) &= 1 & \text{if } x &= a \quad (\text{our object}), \\ f(x) &= 0 & \text{if } x &\neq a. \end{aligned}$$

The database can be a text, an Amazon warehouse, or a mathematical problem. Say, we know that  $p$  is the sum of squares of two integers  $p = x^2 + y^2$ ; find them! For all  $x$ , we compute  $p - x^2$  until we find an integer  $y$ . Classically, it takes  $N/2$  searches (evaluating  $f(x)$ ,  $x = 0, 1, 2, \dots$ ) 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. We start with the usual state

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

Notice that we start with  $H(x)$  in output in order to compute phase:

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

The output register will be irrelevant and we do not use it anymore. The "Oracle" is the operation

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

This is a reflection. If we split  $\mathcal{H}$  into  $|a\rangle$  and orthogonal space,  $O$  does not act on the first part:

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

Indeed,

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

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

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

Now, our initial computational state is the uniform superposition

$$|\psi\rangle = H^{\otimes n}|0\rangle^{\otimes n} = \frac{1}{\sqrt{N}} \sum_x |x\rangle.$$

The "equal" superposition state of all  $|x\rangle$  with amplitude  $1/\sqrt{N}$ . Let's define another reflection operation  $G$  around  $|\psi\rangle$ :

$$G|\psi\rangle = |\psi\rangle,$$

$$G|x\rangle = -|x\rangle \quad \text{if} \quad \langle x|\psi\rangle = 0,$$

or

$$G = 2|\psi\rangle\langle\psi| - I \quad (\text{note opposite sign}).$$

This can be easily constructed starting with a simpler one that reflects all states that are not  $|0\rangle^{\otimes n}$ .

$$|0\dots 0\rangle \rightarrow |000\dots 0\rangle,$$

$$|x_1\dots x_n\rangle \rightarrow -|x_1\dots x_n\rangle \quad \text{if at least one } x_i \neq 0.$$

This is easily implemented on qubits (exercise: write a circuit!) and then

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

Indeed,

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

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

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

$$\underbrace{GOGOGO\dots}_{O(\sqrt{N})}.$$

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

## 4 Open systems

### 4.1 Density matrix

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

Given a system that can be in a collection of states  $\{\psi_i\}$ , but we only know (due to ignorance) that this happens with (classical) probability  $P_i$ , we define the density matrix

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

$$\sum_i P_i = 1, \quad \langle\psi_i|\psi_i\rangle = 1$$

This is useful because, for a quantum state  $|\psi\rangle$  and an observable  $A$ , the expectation value of  $A$  is

$$\langle A \rangle = \langle\psi|A|\psi\rangle$$

If we perform experiments and we have the partial knowledge of the system given by the probabilities  $P_i$ , the measured expectation value will be

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

This can be rewritten as

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

Indeed, using a standard trick,

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

Here trace means for  $A = \sum_n A_{nm} |n\rangle\langle m|$ , and then

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

A quantum state in the old sense is called *pure*.  $|\psi\rangle$  can be also written as a density matrix

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

An ensemble of quantum states in the new sense

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

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

Perhaps the simplest example when we need to use density matrices is statistical mechanics: a system with temperature  $T$ . If we have possible energy states  $\{E_n, |n\rangle\}$ , each can occur with classical probability

$$P(E_n) = \frac{e^{-E_n/k_B T}}{Z}$$

where

$$Z = \sum_n e^{-E_n/k_B T} \quad \text{or} \quad \sum_n P(E_n) = 1$$

so that we will use

$$\rho = \sum_n \frac{e^{-E_n/k_B T}}{Z} |n\rangle\langle n|$$

in our computations.