II Quantum Information and Computation

Haiyang (Rain) Zhu
hz430@srcf.net
Presented by Dr. Nilanjana Datta

Lent 2025

Contents

1 Principles of quantum mechanics **3**
1.1 Postulates of quantum mechanics . . . 4
1.2 Dirac notation for linear operators . . . 5
1.3 Quantum measurements . . . 10

2 Quantum entanglement and quantum gates **14**
2.1 Entanglement . . . 14
2.2 Quantum (logic) gates . . . 14
2.2.1 Single qubit gates . . . 15
2.2.2 2-qubit gates . . . 16
2.2.3 Quantum circuit for entanglement generation . . . 18

3 Quantum states as information carriers **19**
3.1 No-cloning theorem . . . 19
3.1.1 Proof of the no-cloning theorem . . . 19
3.1.2 Example: Misassumption of quantum cloning . . . 20
3.1.3 A special case of the no-signalling principle . . . 21
3.2 Distinguishing non-orthogonal states . . . 22
3.2.1 Helstrom-Holevo theorem . . . 22
3.3 Entanglement and its applications . . . 24
3.3.1 Superdense coding . . . 26
3.3.2 Quantum teleportation . . . 26

4 Quantum cryptography **28**
4.1 The one-time pad . . . 28
4.2 Quantum key distribution (QKD) . . . 29
4.2.1 BB84 . . . 29

5 Basic notions of classical computations and computational complexity **34**
5.1 Definitions . . . 34
5.1.1 Computational task and algorithm . . . 34

5.1.2 Model of classical computation . . . 34
5.2 Time-complexity of algorithms . . . 35
5.2.1 Time-complexity classes . . . 36
5.2.2 Black-box/Oracle promise problems . . . 37

6 Circuit model of quantum computation **38**
6.1 Poly-time quantum computations and BQP . . . 38
6.1.1 Black-box promise problem in quantum computation . . . 38
6.2 Computation via quantum parallelism . . . 39
6.2.1 Approximately universal set of quantum gates . . . 40

7 The Deutsch-Jozsa (DJ) algorithm **41**
7.1 Randomised classical algorithm . . . 44
7.2 Simon's problem/algorithm . . . 44

8 Quantum Fourier transform **46**
8.1 QFT mod N . . . 46
8.2 Periodicity determination . . . 47

9 Quantum algorithms for search problems **52**
9.1 The unstructured search problem . . . 52
9.2 Grover's algorithm . . . 53
9.3 Further details of Grover's algorithm . . . 57
9.3.1 Optimality . . . 57
9.3.2 Multiple good items . . . 58

10 Shor's factoring algorithm **59**
10.1 Factoring as a periodicity determination problem . . . 59
10.2 Computing period of modular exponential function . . . 61
10.3 Getting r from a good value of c . . . 63
10.4 Complexity of Shor's algorithm* . . . 66
10.5 Is NP equal to P? . . . 70

1 Principles of quantum mechanics

Consider a quantum system A. We may associate it with a *Hilbert space* $\mathcal{H}_A$, or simply $\mathcal{H}$.

Definition. A *Hilbert space* (HS) is a vector space equipped with an *inner product*.

Definition (Dirac bra-ket notation)**.** For $\mathcal{H} \cong \mathbb{C}^n$, we may write

$$|a\rangle = (a_1, a_2, \ldots, a_n)^T,\ |b\rangle = (b_1, b_2, \ldots, b_n)^T$$

where $a_i, b_i \in \mathbb{C}$. Then $|a\rangle, |b\rangle$ are *ket-vectors* or *kets*. For vectors

$$|u\rangle = (u_1, u_2, \ldots, u_n)^T,\ |v\rangle = (v_1, v_2, \ldots, v_n)^T$$

the *inner product* is defined as

$$(u, v) \equiv \langle u|v\rangle = \sum_{i=1}^{n} u_i^* v_i$$

We may split the inner product, and say $\langle u|$ is the *bra-vector* or *bra*, where $\langle u| = (u_1^*, u_2^*, \ldots, u_n^*)$. Then

$$\langle u|v\rangle = (u_1^* \ \ldots \ u_n^*) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_i u_i^* v_i$$

i.e. matrix multiplication. So bra-vectors are vectors in $\mathcal{H}^* : \mathcal{H} \to \mathbb{C}$ i.e. the *dual* to $\mathcal{H}$. Checking, $\langle u| : |v\rangle \mapsto \langle u|v\rangle \in \mathbb{C}$ is indeed such a map.

Proposition. *Inner products have the following properties:*

- *Positivity:* $\langle v|v\rangle \geqslant 0$ *with equality if and only if* $|v\rangle = 0$.
- *Linearity in second argument: If* $|v\rangle = c_1 |a\rangle + c_2 |b\rangle$ *where* $c_1, c_2 \in \mathbb{C}$ *then*

$$\langle u|v\rangle = c_1 \langle u|a\rangle + c_2 \langle u|b\rangle$$

- *Antilinearity in first argument: If* $\langle u| = r_1 |u_1\rangle + r_2 |u_2\rangle$ *then*

$$\langle u|v\rangle = r_1^* \langle u_1|v\rangle + r_2^* \langle u_2|v\rangle$$

- *Skew-symmetry:* $\langle u|v\rangle = \langle v|u\rangle^*$.

For $|v\rangle \in \mathcal{H}$ and $\langle v| \in \mathcal{H}^*$ then

$$\langle v| = (|v\rangle)^\dagger,\ |v\rangle = (\langle v|)^\dagger$$

and $(c_1 |v_1\rangle + c_2 |v_2\rangle)^\dagger = c_1^* \langle v_1| + c_2^* \langle v_2|$. Moreover, $((|v\rangle)^\dagger)^\dagger = |v\rangle$.

Definition. $|u\rangle, |v\rangle \in \mathcal{H}$ are *orthogonal* written $|u\rangle \perp |v\rangle$ if $\langle u|v\rangle = 0$.

Definition. The *length/norm* of a vector $|v\rangle$ is $\|v\| = \sqrt{\langle v|v\rangle}$.

Definition. A *basis* of $\mathcal{H}$ is a maximal set of pairwise orthogonal vectors of unit length $\{|e_i\rangle\}_{i=1}^n$ such that

$$\langle e_i|e_j\rangle = \delta_{ij}$$

where n is the *dimension* of $\mathcal{H}$. Any $|v\rangle \in \mathcal{H}$ can be written as $|v\rangle = \sum_i v_i |e_i\rangle$ where $v_j = \langle e_j|v\rangle \in \mathbb{C}$. We say $\{|e_i\rangle\}_{i=1}^n$ is an orthonormal basis (onb) of $\mathcal{H}$. If $\{\langle e_i|\}_{i=1}^n$ is an onb of $\mathcal{H}^*$ then for $\langle v| \in \mathcal{H}^*$ we have

$$\langle v| = \sum_i v_i^* \langle e_i|, \ v_i^* = \langle v|e_i\rangle$$

Definition. The *computational basis* of $\mathcal{H} \in \mathbb{C}^n$ is $\{|i\rangle\}_{i=0}^{n-1}$. Note that $|0\rangle \neq \mathbf{0}$.

Example. For $n = 2$, $\mathcal{H} \cong \mathbb{C}^2$ we have

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

1.1 Postulates of quantum mechanics

Postulate I: With any isolated (closed) quantum system which can be prepared in n perfectly distinguishable states, one can associate a Hilbert space $\mathcal{H}$ of dimension n such that the states of the system are given by $|\psi\rangle \in \mathcal{H}$ is of *unit length* ($\langle\psi|\psi\rangle = 1$). We may in fact represent them in a ray $\{e^{i\theta}|\psi\rangle\}$ for any $\theta \in \mathbb{R}$, the *global phase*, forming an equivalence class. This is allowed since the coefficients cancel when taking measurements i.e. physical irrelevant.

———————————[Lecture 1 finish]———————————

Example. Consider $\mathcal{H} \cong \mathbb{C}^2$ with computational basis $\{|0\rangle, |1\rangle\}$ where

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

Then the *superposition principle* states that, for any $|\psi\rangle \in \mathcal{H}$,

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

for some $\alpha, \beta \in \mathbb{C}$.

Definition (Qubits)**.** Any quantum system such that $\mathcal{H} \cong \mathbb{C}^2$ is a *qubit*, the fundamental unit of quantum information. $0, 1$ are called *bits*, so $|0\rangle, |1\rangle$ are *qubits* (basis).

Physically, a qubit can be

- electronic spin $|\uparrow\rangle$, $|\downarrow\rangle$;
- photon polarisation.

In quantum mechanics, only *mutually orthogonal* states are *perfectly distinguishable*. For $\mathcal{H} \cong \mathbb{C}^2$, there is an orthonormal basis $\{|0\rangle, |1\rangle\}$, but another can be $\{|+\rangle, |-\rangle\}$, where

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

Note. *Consider* $\mathcal{H} \cong \mathbb{C}^n$ *with orthonormal basis* $\{|e_i\rangle\}_{i=1}^n$. *For each state* $|\psi\rangle \in \mathcal{H}$ *such that* $\langle\psi|\psi\rangle = 1$, *if* $|\psi\rangle = \sum_i c_i |e_i\rangle$, *then*

$$1 = \langle\psi|\psi\rangle = \sum_i |c_i|^2, \ |c_i|^2 \geqslant 0$$

so $\{|c_i|^2\}_{i=1}^n$ *form a probability distribution.* $\langle e_i|\psi\rangle = c_i$ *are probability amplitudes.*

Postulate II (Composite system): For a *composite system* jointed with A_1, A_2 with Hilbert space $\mathcal{H}_1, \mathcal{H}_2$, the Hilbert space for this system is $\mathcal{H} \cong \mathcal{H}_1 \otimes \mathcal{H}_2$. States of A_1A_2 are given by unit vectors in $\mathcal{H}_1 \otimes \mathcal{H}_2$.

If $\mathcal{H}_1$ and $\mathcal{H}_2$ have, respectively, orthonormal bases $\{|a_i\rangle\}_{i=1}^{d_1}$, $\{|b_i\rangle\}_{j=1}^{d_2}$ where $d_1 = \dim \mathcal{H}_1$, $d_2 = \dim \mathcal{H}_2$, then

$$\{|a_i\rangle \otimes |b_i\rangle\}_{i=1\ldots d_1,\ j=1\ldots d_2}$$

is an orthonormal basis of $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Example. Consider a 2-qubit systems with $\mathcal{H}_1, \mathcal{H}_2 \cong \mathbb{C}^2$ and $\mathcal{H} \cong \mathcal{H}_1 \otimes \mathcal{H}_2$. $\{|0\rangle, |1\rangle\}$ is the orthonormal basis of $\mathbb{C}^2$, so the orthonormal basis of $\mathcal{H}$ is

$$\{|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle\}$$

We may compute this by

$$|a\rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \ |b\rangle = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \ \Rightarrow \ |a\rangle \otimes |b\rangle = \begin{pmatrix} a_1 \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\ a_2 \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} a_1b_1 \\ a_1b_2 \\ a_2b_1 \\ a_2b_2 \end{pmatrix}$$

we may write $|a\rangle \otimes |b\rangle = |a\rangle |b\rangle = |ab\rangle$. So the orthonormal basis of $\mathcal{H}$ is then

$$\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$$

1.2 Dirac notation for linear operators

Definition. A *linear map* $A : \mathcal{H} \to \tilde{\mathcal{H}}$ satisfies that if $|u\rangle = c_1 |u_1\rangle + c_2 |u_2\rangle$ then

$$A|u\rangle = c_1 A|u_1\rangle + c_2 A|u_2\rangle$$

If $A : \mathcal{H} \to \mathcal{H}$ is a *endomorphism*, then A is a *linear operator*.

Definition (Adjoint)**.** The *adjoint* of A is $A^\dagger$ where

$$\langle v|A^\dagger|u\rangle = \langle u|A|v\rangle^* \ \forall \ |u\rangle, |v\rangle \in \mathcal{H}$$

Definition. $\mathcal{L}(\mathcal{H})$ is the set of linear operators o n$\mathcal{H}$. If $A, B \in \mathcal{L}(\mathcal{H})$ then for all $|u\rangle \mathcal{H}$,

$$AB|u\rangle = A(B|u\rangle) \neq BA|u\rangle$$

where the last equation is satisfied only if $[A, B] = AB - BA = 0$.

Definition. Denote by I the *identity operator* on $\mathcal{H}$ where $I|u\rangle = |u\rangle$ for all $|u\rangle \in \mathcal{H}$. A^{-1} is the *inverse* of A where $A^{-1}A = I = AA^{-1}$. We also have $(AB)^\dagger = B^\dagger A^\dagger$ and $(\alpha A + \beta B)^\dagger = \alpha^* A^\dagger + \beta^* B^\dagger$.

Definition. A linear operator A is *normal* if $AA^\dagger = A^\dagger A$. It is *unitary* if $A^\dagger = A^{-1}$. It is *Hermitian* if $A^\dagger = A$.

Definition (Outer products of vectors)**.** If $\mathcal{H} \cong \mathbb{C}^n$ then $|a\rangle\langle b| \in \mathbb{M}_n(\mathcal{C})$.

Example. If $|a\rangle = (a_1, \ldots, a_n)^T$ and $|b\rangle = (b_1, \ldots, b_n)^T$ then

$$|a\rangle\langle b| = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix} = \begin{pmatrix} a_1 b_1^* & \cdots & a_1 b_n^* \\ \vdots & \ddots & \vdots \\ a_n b_1^* & \cdots & a_n b_n^* \end{pmatrix} \in \mathbb{M}_n(\mathbb{C})$$

Proposition. *Outer product is a linear operator on $\mathcal{H}$ or $\mathcal{H}^*$.*

Proof.

(i) $|u\rangle\langle v|$ on $|x\rangle \in \mathcal{H}$: $|u\rangle\langle v|x\rangle = |u\rangle\langle v|x\rangle = \langle v|x\rangle|u\rangle \in \mathcal{H}$.

(ii) $|v\rangle\langle v|$ on $\langle y| \in \mathcal{H}^*$: $\langle y|u\rangle\langle v| = \langle y|u\rangle\langle v| \in \mathcal{H}^*$.

□

Example (Product of outer products)**.** $A = |a\rangle\langle b|$, $B = |c\rangle\langle d|$ then

$$BA = |c\rangle\langle d| \ |a\rangle\langle b| = |c\rangle\langle d|a\rangle\langle b| = \langle d|a\rangle|c\rangle\langle b| \in \mathbb{M}_n(C)$$

Claim. Any $A \in \mathcal{L}(\mathcal{H})$ where $\mathcal{H} \cong \mathbb{C}^n$ can be represented as a matrix in $\mathbb{M}_n(\mathbb{C})$.

Proof. Let $\{|e_i\rangle\}_{i=1}^n$ be an orthonormal basis of $\mathcal{H}$, then for all $i = 1, \ldots, n$, $A|e_i\rangle = |f_i\rangle \in \mathcal{H}$. Then

$$A = \sum_{j=1}^n |f_j\rangle\langle e_j|$$

Why? We can check that

$$A|e_i\rangle = \sum_{j=1}^n |f_j\rangle \times \underbrace{\langle e_j|e_i\rangle}_{\delta_{ij}} = |f_i\rangle$$

□

Claim. If $\{|f_i\rangle\}_{i=1}^n$ is also an orthonormal basis of $\mathcal{H}$ then A is an *unitary* operator (preserves inner products).

Proof. If $A|e_i\rangle = |f_i\rangle$ and $A|e_j\rangle = |f_j\rangle$ then

$$\delta_{ij} = \langle f_i|f_j\rangle = \langle e_i|A^\dagger A|e_j\rangle \ \Rightarrow \ A^\dagger A = I$$

Note that if $|f_i\rangle = |e_i\rangle$ then $I = \sum_i |e_i\rangle\langle e_i|$. □

Remark. If we have an orthonormal basis $\{|e_i\rangle\}$, then

- Orthogonality: $\langle e_i|e_j\rangle = \delta_{ij}$.

- Completeness relation: $\sum_{i=1}^{n} |e_i\rangle \langle e_i| = I$.

If $\{|f_i\rangle\}_i$ and $\{|e_i\rangle\}_i$ are orthonormal bases of $\mathcal{H}$, then there exists a unitary operator U such that $|f_i\rangle = U|e_i\rangle$. To see this,

$$I = \sum_i |f_i\rangle \langle f_i| = \sum_i U|e_i\rangle \langle e_i| U^\dagger = U\left(\sum_i |e_i\rangle \langle e_i|\right) U^\dagger = UIU^\dagger = UU^\dagger$$

so U is unitary. If $\{|i\rangle\}_{i=1}^n$ is an orthonormal basis of $\mathcal{H}$ then any $A \in \mathcal{L}(\mathcal{H})$ can be written as

$$A = \sum_{i,j=1}^{n} A_{ij} |i\rangle \langle j|$$

where

$$A_{ij} = \langle i|A|j\rangle$$

To see this,

$$A|j\rangle = \sum_{i',j'} A_{i',j'} |i'\rangle \langle j'| |j\rangle = \sum_{i,j'} A_{i',j'} |i'\rangle$$

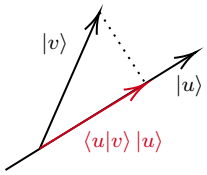
Then

$$\langle i|A|j\rangle = \sum A_{i'j} \langle i|i'\rangle = A_{ij}$$

---[Lecture 2 finish]---

Definition (Orthogonal projector)**.** An operator $A \in \mathcal{L}(\mathcal{H})$ is a *orthogonal projector* if $P^2 = P$ and $P^\dagger = P$. So $P^k|u\rangle = P|u\rangle$ for any $k \geqslant 1$ and $|u\rangle \in \mathcal{H}$. For any $|u\rangle \in \mathcal{H}$ such that $\langle u|u\rangle = 1$, let $P_u = |u\rangle \langle u|$, then $P_u^\dagger = P_u$. Moreover, $P_u^2 = |u\rangle \langle u| |u\rangle \langle u| = |u\rangle \langle u| = P_u$. Then:

- $P_u|v\rangle = |u\rangle \langle u|v\rangle = \langle u|v\rangle |u\rangle$, which is the component of $|v\rangle$ along $|u\rangle$.



- If $\{|e_i\rangle\}_{i=1}^n$ is an orthonormal basis of $\mathcal{H} \cong \mathbb{C}^n$, then

$$\sum_{i=1}^{k} |e_i\rangle \langle e_i|$$

  is the projector onto the subspace spanned by $|e_1\rangle, \ldots, |e_k\rangle$.

- $I = \sum_{i=1}^n |e_i\rangle \langle e_i|$ is the projection onto $\mathcal{H}$.

- For any $|v\rangle \in \mathcal{H}$,

$$|v\rangle = I|v\rangle = \sum_{i=1}^{n} |e_i\rangle \langle e_i| |v\rangle = \sum_{i=1}^{n} |e_i\rangle \langle e_i|v\rangle = \sum_i v_i |e_i\rangle$$

Definition (Trace)**.** The *trace* of an operator A is

$$\operatorname{Tr} A = \sum_k \langle e_k|A|e_k\rangle = A_{kk}$$

Let $\{|e_k\rangle\}_{i=1}^n$ be an orthonormal basis. We claim $\operatorname{Tr}(|a\rangle\langle b|) = \langle b|a\rangle$. To see this,

$$\begin{aligned}\operatorname{Tr}|a\rangle\langle b| &= \sum_k \langle e_k|a\rangle\langle b|e_k\rangle \\ &= \sum_k \langle b|e_k\rangle\langle e_k|a\rangle \\ &= \langle b|\sum_k |e_k\rangle\langle e_k|\,|a\rangle \\ &= \langle b|a\rangle\end{aligned}$$

$\operatorname{Tr} A$ is independent of the choice of orthonormal basis, since any orthonormal basis e.g. $\{|f_j\rangle\}_{j=1}^n$ also satisfies $I = \sum_j |f_j\rangle\langle f_j|$.

Definition (Pauli operators/matrices)**.** The Pauli matrices are the following 2x2 matrices in $\mathcal{L}(\mathcal{H}) = \mathbb{M}_2(\mathbb{C})$, $\mathcal{H} \cong \mathbb{C}^2$:

$$X = \sigma_x = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\ Y = \sigma_y = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix},\ Z = \sigma_z = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

- They are Hermitian and unitary.
- They anti-commute i.e. $\{X, Y\} = XY + YX = 0$.
- They are traceless.
- They also have the cyclic relation were $XY = iZ$.
- $X^2 = Y^2 = Z^2 = I$.
- Moreover, $X|0\rangle = |0\rangle$ and $X|1\rangle = |0\rangle$, so X is the *bit-flip* operator. $Z|0\rangle = |0\rangle$ and $Z|1\rangle = -|1\rangle$, so it is a *phase-flip* operator.
- Recall
$$|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}$$
which is an eigenbasis of X with $X|+\rangle = |+\rangle$ and $X|-\rangle = -|-\rangle$. $\{|0\rangle, |1\rangle\}$ is an eigenbasis of Z.
- $\{I = \sigma_0 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, X, Y, Z\}$ forms a basis for $\mathbb{M}_2(\mathbb{C})$ with respect to the *Hilbert-Schmidt inner product*:
$$\begin{gathered}A, B \in \mathbb{M}_2(\mathbb{C}), \langle A, B\rangle_{\text{HS}} = \operatorname{Tr}(A^\dagger B) \\ A, B \in \{I, X, Y, Z\},\ \langle A, B\rangle_{\text{HS}} = 0 \text{ for } A \neq B\end{gathered}$$
But note that $\langle A, A\rangle = 2 \neq 0$, since e.g. $\langle X, X\rangle = \operatorname{Tr}(X^\dagger X) = \operatorname{Tr}(I) = 2$. Hence any matrix $M \in \mathbb{M}_2(\mathbb{C})$ can be written as
$$M = \alpha X + \beta Y + \gamma Z + \delta I$$

Postulate III (Physical evolution of quantum systems): Any physically admissible evolution of a closed quantum system is represented by a *unitary* operator on $\mathcal{H}$.

If the system is in state $|\psi(t_1)\rangle$ at time t_1 and in $|\psi(t_2)\rangle$ at time t_2 then there exists a unitary operator $U(t_1, t_2)$ (assume $t_2 > t_1$) such that

$$|\psi(t_2)\rangle = U(t_1, t_2)\,|\psi(t_1)\rangle$$

Time evolution is given by the *Schrödinger equation*:

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

where $H = H^\dagger$, $H \in \mathcal{L}(\mathcal{H})$ is the *Hamiltonian* and $\hbar$ is *Planck's constant.*

- If H is time independent then

$$|\psi(t)\rangle = U(t)\,|\psi(t)\rangle\,,\ U(t) = e^{-\frac{i}{\hbar}Ht}$$

  where

$$e^{-\frac{i}{\hbar}Ht} = I - \frac{i}{\hbar}Ht + \frac{1}{2!}\left(-\frac{i}{\hbar}t\right)^2 H^2 + \ldots$$

- In fact, *any unitary operator* can be written as

$$U \equiv U(t) = e^{iAt} = \sum_{n=1}^{\infty}\left(\frac{it}{\hbar}\right)^n \frac{A^n}{n!}$$

  where $A = A^\dagger$ is real.

Remark. In this course, we will use basic unitaries as quantum logic gates to construct more complex unitaries.

Definition (Orthogonal subspaces)**.** Let $\{|e_i\rangle\}_{i=1}^n$ be an orthonormal basis. Then

(1) Orthogonality: $\langle e_i|e_j\rangle = \delta_{ij}$.

(2) Completeness: $\sum_{i=1}^n |e_i\rangle\langle e_i| = I$.

The notions (1) and (2) can be extended to subspaces of $\mathcal{H}$: $\mathcal{E}_1, \mathcal{E}_2 \subset \mathcal{H}$ are *mutually orthogonal* if for any $|v_1\rangle \in \mathcal{E}_1$ and $|v_2\rangle \in \mathcal{E}_2$, $\langle v_1|v_2\rangle = 0$. $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_k$ form an *orthogonal decomposition* of $\mathcal{H}$ written

$$\mathcal{H} = \mathcal{E}_1 \oplus \mathcal{E}_2 \oplus \cdots \oplus \mathcal{E}_k$$

and any $|v\rangle \in \mathcal{H}$ has a unique decomposition:

$$|v\rangle = \sum_{i=1}^k |v_i\rangle\,,\ |v_i\rangle \in \mathcal{E}_i$$

To see this, observe that

$$\begin{aligned} 1 = \langle v|v\rangle = \sum_j^{i=1} \langle v_j| \sum_i |v_i\rangle &= \sum_{ij} \langle v_j|v_i\rangle \\ &= \sum_{i=1}^{k} \langle v_i|v_i\rangle \text{ if } \langle v_j|v_i\rangle = \delta_{ij} \\ &= k \end{aligned}$$

Contradiction. So $\langle v_i|v_i\rangle \neq 1$ i.e. they are not normalised. In particular, $|v_i\rangle = \Pi_i |v\rangle$ where Π_i is the orthogonal projection onto $\mathcal{E}_i$, so

$$\langle v_i|v_i\rangle = \langle v|\Pi_i^\dagger \Pi_i|v\rangle = \langle v|\Pi_i|v\rangle \neq 1$$

We have $\sum_i \Pi_i = I$ (extension of completeness relation) and $\Pi_i\Pi_j = \Pi_i\delta_{ij}$ (no summation on RHS, orthogonality relation for projectors $\{\Pi_i\}$). If $\Pi_i = |e_i\rangle\langle e_i|$ are rank-1 projectors then the completeness relation reduces to $\sum_i |e_i\rangle\langle e_i| = I$.

1.3 Quantum measurements

Postulate IV (Measurement Postulate):

(a) A complete (projective) measurement on a quantum system with Hilbert space $\mathcal{H}$ is determined by the choice of an orthonormal basis $\{e_i\}_{i=1}^n$. (In fact, every such orthonormal basis in principle represents a possible measurement.)

- The outcome is the label of basis vector $\{|e_i\rangle\}$.
- If the system is in state $|\psi\rangle$, then by the *Born rule*,

$$p(e_k) = p(k) = \mathrm{Tr}(P_k |\psi\rangle\langle\psi|) = \langle\psi|P_k|\psi\rangle = \langle\psi|e_k\rangle\langle e_k|\psi\rangle = |\langle e_k|\psi\rangle|^2$$

  Recall that $|\psi\rangle = \sum_i \langle e_i|\psi\rangle |e_i\rangle$, where coefficients are probability amplitudes.
- If the outcome is e_k then

$$|\psi\rangle \mapsto |\psi'\rangle = |e_k\rangle = \frac{P_k|\psi\rangle}{\sqrt{p(k)}}$$

  where $P_k = |e_k\rangle\langle e_k|$.

[Lecture 3 finish]

(a) Complete projective (or von Neumann) measurement (continued):

Example. Consider $\mathcal{H} \cong \mathbb{C}^2$ and $|\psi\rangle = \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}|1\rangle$ then $\langle\psi|\psi\rangle = 1$. Measure in $\{|0\rangle, |1\rangle\}$ (standard basis) then outcomes are 0 or 1. Hence

$$p(0) = |\langle 0|\psi\rangle|^2 = \frac{1}{3};\ p(1) = |\langle 1|\psi\rangle|^2 = \frac{2}{3}$$

If the outcome is 0 then $|\psi\rangle \mapsto |\psi'\rangle = |0\rangle$.

Example. Consider $|\psi\rangle = \sqrt{\frac{2}{7}}\,|+\rangle + \sqrt{\frac{5}{7}}\,|-\rangle$ and let's measure in $\{|+\rangle, |-\rangle\}$. Then outcomes are $+$ or $-$. So $p(+) = \frac{2}{7}$ etc.

Measurement is equivalently specified by $\{P_i\}$ where $P_i = |e_i\rangle\langle e_i|$.

Example. Consider $\mathcal{H} \cong \mathbb{C}^2$ and $|\psi\rangle = a\,|+\rangle + b\,|-\rangle$ where $a, b \in \mathbb{C}$. Measure on $\{|+\rangle, |-\rangle\}$, and write $P_\pm = |\pm\rangle\langle\pm|$. Then

$$p(+) = \mathrm{Tr}(P_+\,|\psi\rangle\langle\psi|) = \langle\psi|P_+|\psi\rangle = |a|^2$$

If the outcome is $+$ then

$$|\psi\rangle \mapsto |\psi'\rangle = \frac{P_+\,|\psi\rangle}{\sqrt{p(+)}} = \frac{a\,|+\rangle}{|a|} = \frac{|a|e^{i\theta}|+\rangle}{|a|} = |+\rangle$$

where we could add in $e^{i\theta}$ term because it is just a global phase for some θ.

(b) Incomplete projective measurement: Let

$$\mathcal{H} = \bigoplus_{i=1}^{k} \mathcal{E}_i \qquad (*)$$

where $\{\mathcal{E}_i \subset \mathcal{H}\}$ are mutually orthogonal. Let Π_i be a orthogonal projection on $\mathcal{E}_i$, then $\Pi_i\Pi_j = \Pi_i\delta_{ij}$. An incomplete measurement of $|\psi\rangle \in \mathcal{H}$ w.r.t. the decomposition $(*)$ has outcomes $i \in \{1, 2, \ldots, k\}$ and

$$p(i) = \langle\psi|\Pi_i|\psi\rangle$$

Note. *A complete measurement is a special of an incomplete measurement where $\mathcal{E}_i$ are one-dimensional subspaces and Π_i are rank-one projections.*

The resulting state is then normalised:

$$|\psi\rangle \mapsto \frac{\Pi_i\,|\psi_i\rangle}{\sqrt{p(i)}}$$

An incomplete measurement can be complemented via a complete measurement. Consider the decomposition $(*)$ as before with $\mathcal{H} \cong \mathbb{C}^n$. We can choose an orthonormal basis

$$\mathcal{B} = \left\{|e_1^{(1)}\rangle, \ldots, |e_{m_1}^{(2)}\rangle, |e_1^{(2)}\rangle, \ldots, |e_{m_2}^{(2)}\rangle, \ldots, |e_1^{(k)}\rangle, \ldots, |e_{m_k}^{(k)}\rangle\right\}$$

where $m_i = \dim \mathcal{E}_i$ and this is consistent with $(*)$. So for each i, $\{e_j^{(i)}\}_{j=1}^{m_i}$ is an orthonormal basis of $\mathcal{E}_i$, and

$$\Pi^{(i)} = \sum_{j=1}^{m_i} |e_j^{(i)}\rangle\langle e_j^{(i)}|$$

is an orthogonal projection on $\mathcal{E}_i$. So we can perform complete measurement on $|\psi\rangle$ in the basis $\mathcal{B}$ and recover outcome probabilities of the incomplete measurement by summing relevant probabilities of complete measurement:

$$p(i) = \langle\psi|\Pi^{(i)}|\psi\rangle = \langle\psi|\sum_{j=1}^{m_i} |e_j^{(i)}\rangle\langle e_j^{(i)}|\,|\psi\rangle = \sum_{j=1}^{m_i} |\,\langle e_j^{(i)}|\psi\rangle\,|^2$$

Example (Parity measurement)**.** Consider the *parity* of a two-bit string $b_1b_2 \in \{0,1\}^2 = b_1 \oplus b_2$, where $0 \oplus 0 = 0 = 1 \oplus 1$. Then the parity measurement on a state of two qubits $\mathcal{H} = (\mathbb{C}^2)^{\otimes 2}$. The incomplete measurement has $\mathcal{H} = \mathcal{E}_0 \oplus \mathcal{E}_1$ where

$$\mathcal{E}_0 = \text{span}\{|00\rangle, |11\rangle\} \text{ (even parity)}; \ \mathcal{E}_1 = \text{span}\{|01\rangle, |10\rangle\} \text{ (odd parity)}$$

Let $|\psi\rangle \in \mathcal{H}$ be the initial state, where

$$|\psi\rangle = \sum_{i,j=0} a_{ij} |i\rangle |j\rangle$$

The possible outcomes of the parity measurement are 0 and 1. Then

$$\begin{aligned} & & p(0) &= \langle\psi|\Pi_0|\psi\rangle \\ \Pi_0 = |00\rangle\langle 00| + |11\rangle\langle 11| & \Rightarrow & p(0) &= |a_{00}|^2 + |a_{11}|^2 \end{aligned}$$

It is equivalent to doing a complete measurement in the basis

$$\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$$

then $p(0) = p_{00} + p_{11}$. But $p_{00} = \langle\psi|P_{00}|\psi\rangle = |a_{00}|^2$, $p_{11} = |a_{11}|^2$. So we get the same result. The post measurement state corresponding to even parity outcome is then

$$|\psi'\rangle = \frac{a_{00}|00\rangle + a_{11}|11\rangle}{\sqrt{p(0)}} = \frac{\Pi_0|\psi\rangle}{\sqrt{p(0)}}$$

(c) (From IB QM) measuring a quantum observable A (Hermitian operator): The outcomes are eigenvalues of A. The *spectral decomposition* (see supplement notes) of A is

$$A = \sum_{i=1}^{k} a_i \Pi_i$$

If $\{a_i\}_{i=1}^k$ is non-degenerate then $\Pi_i = |e_i\rangle\langle e_i|$. Otherwise, Π_i is the orthogonal projection on

$$\mathcal{E}_i = \text{span}\left\{|e_j^{(i)}\rangle : j = 1, \ldots, m_i\right\}$$

This is equivalent to the incomplete measurement corresponding to $\mathcal{H} = \otimes_{i=1}^k \mathcal{E}_i$. If a_i is non-degenerate then the eigenvalue equation is $A|e_i\rangle = a_i|e_i\rangle$, then $\Pi_i = |e_i\rangle\langle e_i| = P_i$.

(d) Extended Born Rule: Consider a measurement on only a part of a composite system e.g. S_1, S_2 with $\mathcal{H}_1 \otimes \mathcal{H}_2$ and orthonormal bases

$$\mathcal{B}_1 = \{|e_i\rangle\}_{i=1}^m, \ \mathcal{B}_2 = \{|f_j\rangle\}_{j=1}^n$$

of $\mathcal{H}_1 \cong \mathbb{C}^m$ and $\mathcal{H}_2 \cong \mathbb{C}^n$ respectively. We want to measure in the basis $\mathcal{B}_1$ only. This amounts to an incomplete measurement

$$\mathcal{H} \cong \mathcal{H}_1 \otimes \mathcal{H}_2 = \otimes_{i=1}^m \mathcal{E}_i, \ \mathcal{E}_i = \text{span}\{|e_i\rangle \otimes |\varphi\rangle \ \ \forall |\varphi\rangle \in \mathcal{H}_2\}$$

Check: $\Pi_i = P_i \otimes I$ is an orthogonal projection on $\mathcal{E}_i$.

Outcomes of the measurement are $k \in \{1, 2, \ldots, m\}$ where $P_k = |e_k\rangle \langle e_k|$. Then

$$p(k) = \langle \psi | P_k \otimes I | \psi \rangle$$

If $|\psi\rangle = \sum_{i,j=1}^{m,n} a_{ij} |e_i\rangle |f_j\rangle$, then

$$\begin{aligned} p(k) &= \sum_{i,j,i',j'} a_{ij}^* \langle e_i | \langle f_j | \Big(|e_k\rangle \langle e_k| \otimes I \Big) a_{i'j'} |e_{i'}\rangle |f_{j'}\rangle \\ &= a_{ij}^* a_{i'j'} \langle e_i | e_k \rangle \langle e_k | e_{i'} \rangle \langle f_j | f_{j'} \rangle \\ &= \sum_{j=1}^{n} |a_{kj}|^2 \end{aligned}$$

If the outcome is k, then

$$|\psi\rangle \mapsto |\psi'\rangle = \frac{(P_k \otimes I) |\psi\rangle}{\sqrt{p(k)}}$$

[Lecture 4 finish]

Summary (Complete and incomplete projective measurement)**.**

- Measurement outcomes are labels of mutually orthogonal subspaces of $\mathcal{H}$, or labels of orthogonal projection operators, or labels of basis vectors for a complete measurement.

  Example. A complete meaurement in $\{|+\rangle, |-\rangle\}$ characterised by $P_\pm = |\pm\rangle \langle \pm|$.

- States with *guaranteed* different outcomes lie in mutually orthogonal subspaces of $\mathcal{H}$.

Definition. Two states $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ are *perfectly distinguishable* if and only if

$$|\psi\rangle \perp |\phi\rangle \text{ i.e. } \langle \psi | \psi \rangle = 0$$

Perfect distinguishability of two states means that there exists a measurement which gives two distinct outcomes (e.g. $+, -$ or $0, 1$) with probability 1 when applied to the two states.

2 Quantum entanglement and quantum gates

2.1 Entanglement

In QIC, information is encoded in states of quantum systems. We ony consider closed i.e. isolated systems. We'll consider *product states* and *entangled states*.

Definition. Consider two systems A, B (e.g. 2 qubits) with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. Suppose AB is in state $|\psi_{AB}\rangle \equiv |\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$. Then $|\psi\rangle$ is a *product state* if it can be written as

$$|\psi\rangle = |\alpha\rangle \otimes |\beta\rangle\,,\ |\alpha\rangle \in \mathcal{H}_A,\ |\beta\rangle \in \mathcal{H}_B$$

If this is not possible, then $|\psi\rangle$ is an *entangled state*.

Example. Consider $\mathcal{H}_A, \mathcal{H}_B = \mathbb{C}^2$ with

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

We claim that this is entangled. To see this, assume there exists some $|\alpha\rangle = a\,|0\rangle + b\,|1\rangle$ and $|\beta\rangle = c\,|0\rangle + d\,|1\rangle$ such that $|\psi\rangle = |\alpha\rangle \otimes |\beta\rangle$ where $a, b, c, d \in \mathbb{C}$. Then

$$\begin{aligned}\frac{|00\rangle + |11\rangle}{\sqrt{2}} &= (a\,|0\rangle + b\,|1\rangle)(c\,|0\rangle + d\,|1\rangle) \\ &= ac\,|00\rangle + ad\,|01\rangle + bc\,|10\rangle + bd\,|11\rangle\end{aligned}$$

Then $ad = 0$ and $bc = 0$, contradiction. So $|\psi\rangle$ is entangled.

On Example Sheet 1, we will see that any $|\psi\rangle \in (\mathbb{C}^2)^{\otimes 2}$ such that $|\psi\rangle = a\,|00\rangle + b\,|01\rangle + c\,|10\rangle + d\,|11\rangle$ is entangled if and only if $ad - bc \neq 0$.

Note. *Such a charaterisation no longer suffices if A, B are not qubits, such as qudits $\mathbb{C}^d$, $\mathbb{C}^n$ where $d, n > 2$.*

Example. Is

$$|\psi\rangle = \frac{1}{\sqrt{2}}\,|00\rangle + \frac{1}{\sqrt{2}}\,|01\rangle$$

entangled? No, because

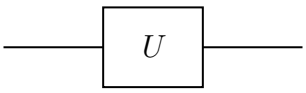
$$|\psi\rangle = |0\rangle\,\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |0\rangle \otimes |+\rangle$$

Entanglement is a valuable resource in QIC.

2.2 Quantum (logic) gates

Physically, it is a device which performs a fixed unitary operation on a selected set of qubits in a fixed period of time. In terms of mathematics, a quantum gate is a unitary operator. A *quantum circuit* is a device consisting of quantum gates.

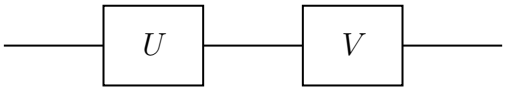
Example. Let U be a unitary operator on $\mathcal{H} \cong \mathbb{C}^2$.



We read from left to right. Each line (quantum wire) represents a qubit.

U can represent:

- a translation in spcace (e.g. atoms travelling through a cavity, or photons through an optical fibre);
- a translation in time (e.g. a sequence of operations performed on a qubit).



In matrix terms, the above diagram is represented by the matrix VU.

2.2.1 Single qubit gates

- Hadamard gate $H = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$. In Dirac notations,

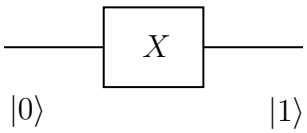
$$H = \frac{1}{\sqrt{2}}\left[|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| - |1\rangle\langle 1|\right]$$

  Then $H|0\rangle = |+\rangle$ and $H|1\rangle = |-\rangle$. Moreover, $H^\dagger = H$, $H^2 = I$ i.e. $H^{-1} = H$. Hence $H|+\rangle = |0\rangle$ and $H|-\rangle = |1\rangle$. H is a special case of a quantum Fourier transform (F_n). If $\mathbb{Z}_N = \{0, 1, \ldots, N-1\}$ then for $x, y \in \mathbb{Z}_n$ we have

$$F_n|x\rangle = \frac{1}{\sqrt{n}}\sum_{y\in\mathbb{Z}_n} e^{2i\pi xy/N}|y\rangle$$

  One can check that $H = F_2$.

- X-gate or NOT-gate: $X = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$. Then $X|k\rangle = |k \oplus 1\rangle$ for $k = 0, 1$, where $\oplus$ is addition mod 2. One can write, for example,

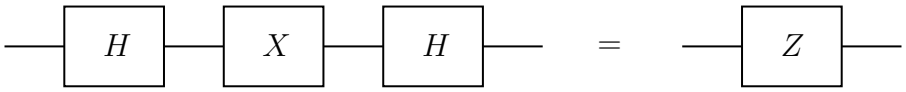


- Z-gate $Z = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$, then $Z|0\rangle = |0\rangle$ and $Z|1\rangle = -|1\rangle$. This is a phase flip gate.
- Phase gate $P_\theta = \begin{pmatrix}1 & 0\\ 0 & e^{i\theta}\end{pmatrix}$, then $Z = P_\pi$ and $P_\theta|0\rangle = |0\rangle$, $P_\theta|1\rangle = e^{i\theta}|1\rangle$.
- S-gate $S = \begin{pmatrix}1 & 0\\ 0 & i\end{pmatrix}$.
- T-gate $T = \begin{pmatrix}1 & 0\\ 0 & e^{i\pi/4}\end{pmatrix}$.

Remark. One can check that

$$HXH^\dagger = HXH = Z$$

or

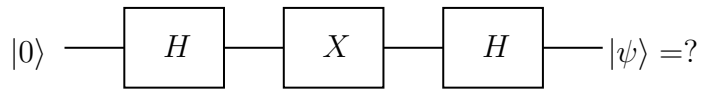


Also $HZH = X$ and $HZH = Y$.

Unitarities such as H which take Pauli operators, under conjugation, to Pauli operators are called *Clifford gates*.

Question. Which phase gate is in the Clifford group?

Example. Consider the following gates, what is $|\psi\rangle$?



We can do this step by step:

$$\begin{aligned}|0\rangle \overset{H}{\mapsto} |+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \overset{P_\theta}{\mapsto} \frac{|0\rangle + e^{i\theta}|1\rangle}{\sqrt{2}} \overset{H}{\mapsto} \frac{|+\rangle + e^{i\theta}|-\rangle}{\sqrt{2}} \\ = \frac{1}{\sqrt{2}}(1+e^{i\theta})|0\rangle + \frac{1}{\sqrt{2}}(1-e^{i\theta})|1\rangle = e^{i\theta}\left[\cos\frac{\theta}{2}|0\rangle - i\sin\frac{\theta}{2}|1\rangle\right]\end{aligned}$$

So $|0\rangle \mapsto |\psi\rangle = \cos\frac{\theta}{2}|0\rangle - i\sin\frac{\theta}{2}|1\rangle$, why? Quantum circuits are often called single qubit interference circuits.

2.2.2 2-qubit gates

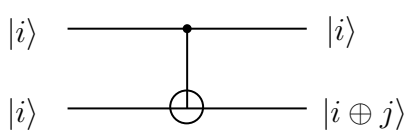
- CNOT or CX gate (C stands for controlled):

$$\text{CNOT} = \begin{pmatrix}[I] & [0] \\ [0] & [X]\end{pmatrix} = \begin{pmatrix}1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0\end{pmatrix}$$

  Then $\text{CNOT}|i\rangle|j\rangle = |i\rangle|i \oplus j\rangle$ for $i, j \in \{0, 1\}$. For any $|\psi\rangle \in \mathbb{C}^2$, $\text{CNOT}|0\rangle|\psi\rangle = |0\rangle|\psi\rangle$ and $\text{CNOT}|1\rangle|\psi\rangle = |0\rangle X|\psi\rangle$. Here the first qubit is the *control qubit* and second *target qubit*. This can be extended by linearity to arbitrary state of control qubit. For example,

$$\text{CNOT}|+\rangle|\psi\rangle = \frac{1}{\sqrt{2}}\text{CNOT}(|0\rangle + |1\rangle)|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle|\psi\rangle + |1\rangle X|\psi\rangle)$$

  In diagrams, this is represented by



If the roles of control and target qubits are changed, there would be a different gate. For example,

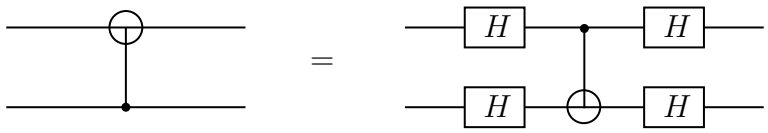
$$\mathrm{CNOT}_{12} |0\rangle_1 |1\rangle_2 = |0\rangle_1 |1\rangle_2$$
$$\mathrm{CNOT}_{21} |0\rangle_1 |1\rangle_2 = |1\rangle_1 |1\rangle_2$$

In the first line, 1 is control and 2 is target, they are then swapped.

Exercise. Check that

$$\mathrm{CNOT}_{21} = (H \otimes H)\mathrm{CNOT}_{12}(H \otimes H)$$

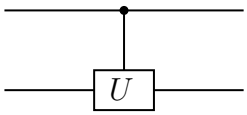
or



To see this, send it to basis vectors $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ and see.

[Lecture 5 finish]

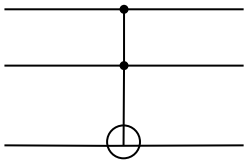
- Controlled-unitary (CU) gate:



Then $CU |0\rangle |\psi\rangle = |0\rangle |\psi\rangle$ and $CU |1\rangle |\psi\rangle = |1\rangle U |\psi\rangle$. If $U = Z$ then

$$CU = \begin{pmatrix} [I] & [0] \\ [0] & [U] \end{pmatrix}$$

- Toffoli gate

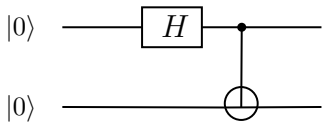


This maps

$$|\alpha\rangle |\beta\rangle |\psi\rangle \mapsto |\alpha\rangle |\beta\rangle |\psi\rangle \text{ if } \alpha, \beta \neq 1$$
$$|1\rangle |1\rangle |\psi\rangle \mapsto |1\rangle |1\rangle X |\psi\rangle$$

2.2.3 Quantum circuit for entanglement generation

Consider



Then

$$\begin{aligned}|0\rangle \otimes |0\rangle \equiv |0\rangle\,|0\rangle &\xmapsto{H\otimes I} |+\rangle\,|0\rangle \\ &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\,|0\rangle \\ &= \frac{1}{\sqrt{2}}(|0\rangle\,|0\rangle + |1\rangle\,|0\rangle) \\ &\xmapsto{\text{CNOT}} \frac{1}{\sqrt{2}}(|0\rangle\,|0\rangle + |1\rangle\,|1\rangle) = |\phi^+\rangle\end{aligned}$$

where $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is an entangled state.

In fact, the above circuit C transforms the orthonormal basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ into another orthonormal basis of $(\mathbb{C}^2)^{\otimes 2}$ of four entangled states, the *Bell basis*:

$$\begin{aligned}C\,|00\rangle &\mapsto |\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\ C\,|01\rangle &\mapsto |\psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \\ C\,|10\rangle &\mapsto |\phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \\ C\,|11\rangle &\mapsto |\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)\end{aligned}$$

Exercise. Check these and see that $\{|\phi^\pm\rangle, |\psi^\pm\rangle\}$ is an orthonormal basis.

Remark. If the circuit is reversed from the right to left, then it would tranform the Bell basis states to the computational basis states.

Definition. The *Bell measurement* measures on 2-qubits on the Bell basis. We can do this e.g. by 'rotating' the Bell basis to the computational basis and then do the measurement there.

3 Quantum states as information carriers

3.1 No-cloning theorem

Note. *The CNOT gate can copy a bit value in the first qubit:*

$$\text{CNOT}\,|x\rangle\,|0\rangle = |x\rangle\,|x\rangle\,,\ x \in \{0,1\}$$

The question is: can CNOT also copy superposition states? I.e. is it true that

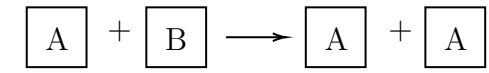
$$\text{CNOT}\,|\psi\rangle\,|0\rangle \overset{?}{\mapsto} |\psi\rangle\,|\psi\rangle$$

where $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$? The answer is NO. To see this, we compute

$$\begin{aligned}\text{CNOT}\,|\psi\rangle\,|0\rangle &= \text{CNOT}\,(a\,|0\rangle + b\,|1\rangle)\,|0\rangle \\ &= a\,|00\rangle + b\,|11\rangle \ \text{ an entangled state}\end{aligned}$$

So the CNOT gate would map a superposition in the control qubit to an entangled state of both qubits. This is known as the *No-cloning theorem.*

In the classical setting, copying is allowed. A photocopier works in the way:



original text blank sheet original text copy

In the quantum setting, there does not exist a universal quantum copier/cloner.

3.1.1 Proof of the no-cloning theorem

Quantum copying involves 3 qubit systems:

- A: the quantum information to be copied in A; $\mathcal{H}_A$.
- B: $\mathcal{H}_B = \mathcal{H}_A$ initially in a fixed state $|\varphi_0\rangle$ (or $|0\rangle$) of $\mathcal{H}_B$.
- M: the cloning machine, initially in, say $|M_0\rangle$ (ready state).

Theorem (No-cloning Theorem)**.** *Let S be any set of states of a quantum system A that contains at least one pair of non-orthogonal states. There there does not exist any unitary cloning process that achieves cloning for all states in S.*

Proof. Assume there exists a universal quantum cloner. Suppose $|\psi\rangle\,, |\phi\rangle$ such that $\langle\psi|\phi\rangle \neq 0$ (non-orthogonal) and $\langle\psi|\phi\rangle \neq 1$ (distinct). Then the quantum cloner behaves like

$$\begin{aligned}|\psi\rangle\,|0\rangle\,|M_0\rangle &\mapsto |\psi\rangle\,|\psi\rangle\,|M_\psi\rangle \\ |\phi\rangle\,|0\rangle\,|M_0\rangle &\mapsto |\phi\rangle\,|\phi\rangle\,|M_\phi\rangle\end{aligned}$$

Since the operator is unitary, it preserves inner products. Then

$$\begin{aligned}\langle\psi|\phi\rangle\,\underset{=1}{\langle 0|0\rangle}\underset{=1}{\langle M_0|M_0\rangle} &= \langle\psi|\phi\rangle\,\langle\psi|\phi\rangle\,\langle M_\psi|M_\phi\rangle \\ \langle\psi|\phi\rangle &= \langle\psi|\phi\rangle^2\,\langle M_\psi|M_\phi\rangle\end{aligned}$$

Let $x = \langle\psi|\phi\rangle$, then $x \neq 0, 1$. So x satisfies

$$\begin{aligned}x &= x^2|\,\langle M_\psi|M_\phi\rangle\,| \\ 1 &= x|\,\langle M_\psi|M_\phi\rangle\,| < |\,\langle M_\psi|M_\phi\rangle\,|\end{aligned}$$

Contradiction since the normalisation constraint on quantum states on $|M_\psi\rangle\,, |M_\phi\rangle$ states that their inner product must be $\leqslant 1$. So states of quantum systems cannot be cloned. □

3.1.2 Example: Misassumption of quantum cloning

Herbert proposed superluminal communication, which was physically impossible, because he assumed that quantum cloning was possible. In his proposal, consider the state

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Suppose Alice (A) and Bob (B) are distant from each other, and they share this state. Now suppose Alice wants to send a yes/no message. We can write

$$|\psi^+_{AB}\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B)$$

Claim. We can write $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|{+}{+}\rangle + |{-}{-}\rangle)$.

To see this, note that for any $M \in \mathbb{M}_2(\mathbb{C})$, it is true that

$$(I \otimes M)|\phi^+\rangle = (M^T \otimes I)|\phi^+\rangle \qquad (*)$$

where $M = \sum_{i,j=0}^{1} M_{ij} |i\rangle \langle j|$ and M^T is the transpose of M. Then

$$\begin{aligned}
\frac{1}{\sqrt{2}}(|{+}{+}\rangle + |{-}{-}\rangle) &= \frac{1}{\sqrt{2}}(H \otimes H)(|00\rangle + |11\rangle) \\
&= (H \otimes I)(I \otimes H)|\phi^+\rangle \\
&= (H \otimes I)(H^T \otimes I)|\phi^+\rangle \text{ by } (*) \\
&= (HH^T \otimes I)|\phi^+\rangle = |\phi^+\rangle \text{ since } HH^T = H^2 = I
\end{aligned}$$

Alice's message	Her action (measurement on A)	Outcomes and probabilities	Final state AB	Final state of B	
Yes	In basis $\mathcal{B}_0 = \{\|0\rangle, \|1\rangle\}$	0 w.p. $p_0 = \frac{1}{2}$ 1 w.p. $p_1 = \frac{1}{2}$	$\mapsto \|0\rangle\|0\rangle$ $\mapsto \|1\rangle\|1\rangle$	$\mapsto \|0\rangle$ $\mapsto \|1\rangle$	(i)
No	In basis $\mathcal{B}_1 = \{\|+\rangle, \|-\rangle\}$	$+$ w.p. $\frac{1}{2}$ $-$ w.p. $\frac{1}{2}$	$\mapsto \|+\rangle\|+\rangle$ $\mapsto \|-\rangle\|-\rangle$	$\mapsto \|+\rangle$ $\mapsto \|-\rangle$	(ii)

So Alice's measurement leads to a preparation of Bob's state B in scenarios (i) and (ii). We claim that these two preparations of the state of Bob are *completely indistinguishable* to Bob

(a) by any local measurement on B, or

(b) from the case in which Alice does no measurement.

[Lecture 6 finish]

Proof. Suppose Bob does measurement corresponding to Π_i (projection operator), and suppose B is in state $|\psi\rangle_B$ after Alice's measurement with

$$p(\text{outcome} = i) = \langle\psi|\Pi_i|\psi\rangle$$

Then

$$\begin{aligned} p_{\text{yes}}(i) &= p(\text{Alice} = \text{yes} \Rightarrow \text{Bob} = i) \\ &= \frac{1}{2}\langle 0|\Pi_i|0\rangle + \frac{1}{2}\langle 1|\Pi_i|1\rangle \text{ average since Bob doesn't know} \\ &= \frac{1}{2}\left[\operatorname{Tr}(\Pi_i |0\rangle\langle 0|) + \operatorname{Tr}(\Pi_i |1\rangle\langle 1|)\right] \\ &= \frac{1}{2}\operatorname{Tr}\left[\Pi_i(|0\rangle\langle 0| + |1\rangle\langle 1|)\right] \\ &= \frac{1}{2}\operatorname{Tr}\Pi_i \end{aligned}$$

and also

$$p_{\text{no}}(i) = \frac{1}{2}\langle +|\Pi_i|+\rangle + \frac{1}{2}\langle -|\Pi_i|-\rangle = \frac{1}{2}\operatorname{Tr}\Pi_i$$

If instead, Alice does no measurement then we can use the extended Born rule to find Bob's measurement:

$$p(i) = \langle \phi^+|I \otimes \Pi_i|\phi^+\rangle = \frac{1}{2}\operatorname{Tr}\Pi_i$$

So the results are the same regardless of Alice's message, or who does the measurement. □

Now we consider Herbert's proposal: to clone B. Suppose Alice does measurement at 12pm and Bob, just after 12pm, clones the qubit B, say, 10^6 times. Then he does a measurement on each copy of B in $\mathcal{B}_0 = \{|0\rangle, |1\rangle\}$. Then the measurement outcome would be

$$b = (b_1, b_2, \ldots, b_n) \in \{0,1\}^n,\ n = 10^6$$

If Alice's message was yes, then after cloning, Bob has

$$|0\rangle \ldots |0\rangle \text{ or } |1\rangle \ldots |1\rangle$$

each with probability $\frac{1}{2}$. Then

$$b = (0, \ldots, 0) \text{ or } (1, \ldots, 1)$$

each with probability $\frac{1}{2}$. If the message was no, then Bob has

$$b = |+\rangle \ldots |+\rangle \text{ or } |-\rangle \ldots |-\rangle$$

each with probability $\frac{1}{2}$. But now b is a uniformly random bit string with $b_i = 0$ or 1 each with probability $\frac{1}{2}$. Such a string can be distinguished from $(0, \ldots, 0)$ or $(1, \ldots, 1)$ except with probability $\frac{2}{2^{10^6}}$. To reduce this error probablity we can just clone more. If this were possible, then it would mean superluminal communication was possible.

3.1.3 A special case of the no-signalling principle

Consider the setup again, where Alice (A) and Bob (B) share a state $|\phi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, and $\mathcal{H}_A, \mathcal{H}_B \cong \mathbb{C}^n$. Then we claim that Alice cannot convey any information to Bob just by doing local measurement on A, that is, no local action on A alone can change the outcome probability distribution of any measurement by Bob on B.

Proof. (Basic case) Suppose Bob does a complete measurement on B in some orthonormal basis of $\mathcal{H}_B$: $\mathcal{B} = \{|b\rangle\}_{b=1}^{d_b}$ where $d_b = \dim \mathcal{H}_B$, and let $\{|a\rangle\}_{a=1}^{d_a}$, $d_a = \dim \mathcal{H}_A$ be an orthonormal basis of $\mathcal{H}_A$. Then we can write

$$|\phi_{AB}\rangle = \sum_{a,b} c_{ab} |a\rangle |b\rangle$$

Consider case 1: only Bob does measurement, then

$$\begin{aligned} p(b) &= \langle \phi_{AB} | I_A \otimes P_b | \phi_{AB} \rangle \text{ where } P_b = |b\rangle \langle b| \\ &= \sum_a |c_{ab}|^2 \end{aligned} \tag{$*$}$$

If the outcome is b then

$$|\phi_{AB}\rangle \mapsto \frac{(I_A \otimes P_b) |\phi_{AB}\rangle}{\sqrt{p(b)}}$$

Now consider case 2: Alice does measurement prior to Bob. Suppose she does measurement on A in $\{|a\rangle\}$, then

$$p(a) = \sum_b |c_{ab}|^2$$

If the outcome is a then

$$|\phi_{AB}\rangle \mapsto \frac{(P_a \otimes I_B) |\phi_{AB}\rangle}{\sqrt{p(a)}} = |\phi''_{AB}\rangle$$

Now consider the conditional probability on Alice measuring prior to Bob meauring in $\{|b\rangle\}$:

$$\begin{aligned} p(b|a) &= \langle \phi''_{AB} | I_A \otimes P_b | \phi''_{AB} \rangle \\ &= \frac{1}{p(a)} \langle \phi_{AB} | (P_a \otimes I_B)(I_A \otimes P_b)(P_a \otimes I_B) | \phi_{AB} \rangle \\ &= \frac{1}{p(a)} \langle \phi_{AB} | P_a \otimes P_b | \phi_{AB} \rangle \text{ since } P_a^2 = P_a \end{aligned}$$

So

$$p(a,b) = \langle \phi_{AB} | P_a \otimes P_b | \phi_{AB} \rangle = |c_{ab}|^2$$

But then

$$p(b) = \sum_a p(a,b) = \sum_a |c_{ab}|^2$$

which is the same as $(*)$. The same holds if Alice and Bob both do incomplete measurements. □

3.2 Distinguishing non-orthogonal states

3.2.1 Helstrom-Holevo theorem

Suppose we're given a quantum system in an unknown state $|\psi\rangle$. We're told that $|\psi\rangle = |\alpha_0\rangle, |\alpha_1\rangle$ each with probability $\frac{1}{2}$, where $|\alpha_0\rangle, |\alpha_1\rangle$ are distinct and non-orthogonal:

$$\langle \alpha_0 | \alpha_1 \rangle \neq 1, 0$$

Our aim is to determine whether $|\psi\rangle = |\alpha_0\rangle$ or $|\alpha_1\rangle$. What is the least we can do? For example, we can do nothing - just guess randomly. The probability of success is just $p_{\text{succ}} = \frac{1}{2}$. We now claim that we can do better by doing a measurement on the system: measure on $\{\Pi_0, \Pi_1\}$, each inferring $|\alpha_0\rangle$, $|\alpha_1\rangle$ respectively, and $\Pi_0 + \Pi_1 = I$ since they are projection operators. Now define the *average* probability of success:

$$\begin{aligned} p_{\text{succ}}(\Pi_0) &= \frac{1}{2}p(\text{outcome} = 0 |\, |\psi\rangle = |\alpha_0\rangle) + \frac{1}{2}p(\text{outcome} = 1 |\, |\psi\rangle = |\alpha_1\rangle) \\ &= \frac{1}{2}\left[\, \langle\alpha_0|\Pi_0|\alpha_0\rangle + \langle\alpha_1|\Pi_1|\alpha_1\rangle \,\right] \\ &= \frac{1}{2} + \frac{1}{2}\operatorname{Tr}\left[\Pi_0(|\alpha_0\rangle\langle\alpha_0| - |\alpha_1\rangle\langle\alpha_1|)\right] \text{ since } \Pi_1 = I - \Pi_0 \end{aligned}$$

[Lecture 7 finish]

We can then write

$$p_{\text{succ}}(\Pi_0) = \frac{1}{2} + \frac{1}{2}\operatorname{Tr}(\Pi_0\Delta)$$

where

$$\Delta = |\alpha_0\rangle\langle\alpha_0| - |\alpha_1\rangle\langle\alpha_1|$$

has the following properties:

(1) $\Delta^\dagger = \Delta$, so eigenvalues are real and eigenvectors form an orthonormal basis.

(2) For any $|\beta\rangle \in \mathcal{H}$ s.t. $\langle\beta|\alpha_0\rangle = \langle\beta|\alpha_1\rangle = 0$, we have $\Delta|\beta\rangle = 0$. So Δ acts non-trivially only on states in

$$\mathcal{V} = \operatorname{span}\{|\alpha_0\rangle, |\alpha_1\rangle\}$$

(3) $\operatorname{Tr}\Delta = 0$, so eigenvalues are $+\delta, -\delta$ for some δ, and let the corresponding eigenvectors be $|p\rangle, |m\rangle$. We can also define rank-1 projections $P_\delta = |p\rangle\langle p|$ and $P_{-\delta} = |m\rangle\langle m|$. So Δ has the spectral decomposition

$$\Delta = \delta P_\delta - \delta P_{-\delta} = \delta(|p\rangle\langle p| - |m\rangle\langle m|)$$

So in the basis $\{|p\rangle, |m\rangle\}$ of $\mathcal{V}$,

$$\Delta = \begin{pmatrix} \delta & 0 \\ 0 & -\delta \end{pmatrix}$$

We now determine δ in terms of $|\alpha_0\rangle, |\alpha_1\rangle$. Let $|\alpha_0^\perp\rangle \in \mathcal{V}$ be such that $\langle\alpha_0^\perp|\alpha_0\rangle = 0$, so $\{|\alpha_0\rangle, |\alpha_0^\perp\rangle\}$ is an orthonormal basis of $\mathcal{V}$. So we can write

$$|\alpha_1\rangle = c_0|\alpha_0\rangle + c_1|\alpha_0^\perp\rangle \tag{3.1}$$

$\mathcal{V} \simeq \mathbb{C}^2$ so wlog we can use the representation

$$|\alpha_0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle, \ |\alpha_0^\perp\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle \ \Rightarrow \ |\alpha_1\rangle = \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}$$

Then in this basis,

$$\Delta = \begin{pmatrix} |c_1|^2 & -c_0c_1^* \\ -c_0^*c_1 & -|c_1|^2 \end{pmatrix}$$

To find eigenvalues of Δ, we solve the equation $\det(\Delta - \lambda I) = 0$ to find $\lambda = \pm|c_1|$. Hence $\delta = |c_1|$. Let $|\alpha_0\rangle, |\alpha_1\rangle$ be such that $0 < |\langle\alpha_0|\alpha_1\rangle| < 1$, then $|\langle\alpha_0|\alpha_1\rangle| = \cos\theta$ for some $0 < \theta < \pi/2$. Then by (3.1),

$$|c_0| = \cos\theta,\ |c_1| = \sqrt{1-\cos^2\theta} = |\sin\theta|$$

Hence $\delta = |c_1| = \sin\theta$. Now we can find

$$\begin{aligned} p_{\text{succ}}(\Pi_0) &= \frac{1}{2} + \frac{1}{2}\operatorname{Tr}(\Pi_0\Delta) \\ &= \frac{1}{2} + \frac{1}{2}\operatorname{Tr}\left[\Pi_0(\delta\,|p\rangle\langle p| - \delta\,|m\rangle\langle m|)\right] \\ &= \frac{1}{2} + \frac{\sin\theta}{2}\left[\langle p|\Pi_0|p\rangle - \langle m|\Pi_0|m\rangle\right] \end{aligned}$$

Recall that Π_0, Π_1 are orthogonal projection operators. We claim that $\langle m|\Pi_0|m\rangle \geqslant 0$. To see this, consider

$$\begin{aligned} \langle m|\Pi_0|m\rangle &= \langle m|\Pi_0^2|m\rangle \ \text{ since } \Pi_0^2 = \Pi_0 \\ &= \langle m|\Pi_0^\dagger\Pi_0|m\rangle \\ &= \|\Pi_0\,|m\rangle\|^2 \geqslant 0 \end{aligned}$$

Hence

$$p_{\text{succ}}(\Pi_0) \leqslant \frac{1}{2} + \frac{\sin\theta}{2}\langle p|\Pi_0|p\rangle \leqslant \frac{1}{2} + \frac{\sin\theta}{2}$$

where the second inequality follows since $\langle p|\Pi_0|p\rangle \leqslant \langle p|p\rangle$. But can we find any $\{\Pi_0, \Pi_1\}$ such that equality holds? Indeed, if e.g. $\Pi_0 = |p\rangle\langle p|$ then

$$\begin{aligned} p^*_{\text{succ}}(\Pi_0) &= \max_{\Pi_0} p_{\text{succ}}(\Pi_0) \\ &\leqslant \frac{1}{2} + \frac{1}{2}\sin\theta,\ |\langle\alpha_0|\alpha_1\rangle| = \cos\theta \end{aligned}$$

and the equality is achieved i.e. there exists such measurement.

Theorem (Helstrom-Holevo theroem)**.** *Given any one of two equally likely (distinct and non-orthogonal) states $|\alpha_0\rangle, |\alpha_1\rangle$ with $|\langle\alpha_0|\alpha_1\rangle| = \cos\theta$ for some $0 < \theta < \pi/2$, the probability p_{succ} of correctly identifying the state by <u>any</u> quantum measurement satisfies*

$$p_{succ} \leqslant \frac{1}{2} + \frac{\sin\theta}{2}$$

and thus the bound is <u>tight</u>.

3.3 Entanglement and its applications

Recall the *Bell basis*, which is an orthonormal basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$. It consists of four maximally entangled states

$$|\phi^+_{AB}\rangle, |\phi^-_{AB}\rangle, |\psi^+_{AB}\rangle, |\psi^-_{AB}\rangle$$

where

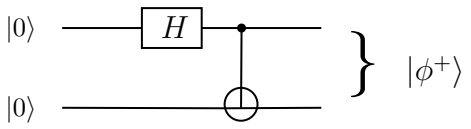
$$|\phi^{\pm}_{AB}\rangle = \frac{|00\rangle \pm |11\rangle}{\sqrt{2}},\ |\psi^{\pm}_{AB}\rangle = \frac{|01\rangle \pm |10\rangle}{\sqrt{2}}$$

These are known as Bell states or EPR states. If we do a local measurement on A (or on B), then

$$p(0) = \frac{1}{2} = p(1)$$

This implies that, even though we know the state of A, B completely, we know nothing about A (or B) individually. $|\Omega_{AB}\rangle$ is a pure state where $\Omega = \phi^{\pm}, \psi^{\pm}$.

In fact, the four bell states can be characterised by 2-bits. We know that



and the general cases are

$$C\,|00\rangle \mapsto |\phi^{+}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle),\ C\,|01\rangle \mapsto |\psi^{+}\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$$
$$C\,|10\rangle \mapsto |\phi^{-}\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle),\ C\,|11\rangle \mapsto |\psi^{-}\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$

We can define *phase bits* and *parity bits*, where

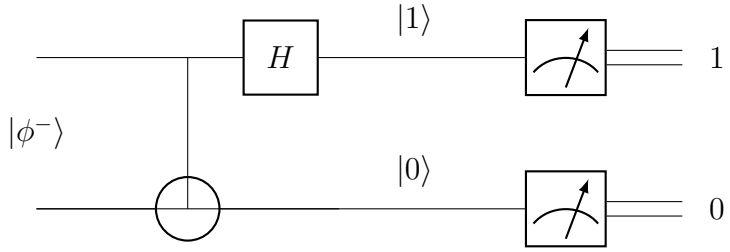
$$\text{Parity bit} = \begin{cases} 0 & \text{if constituent states in superposition have even parity} \\ 1 & \text{if they have odd parity} \end{cases}$$

and

$$\text{Phase bit} = \begin{cases} 0 & \text{if superposition has plus sign} \\ 1 & \text{if it has minus sign} \end{cases}$$

We can see that the first digit determines phase and the second determines parity, so for example,

$$\underset{\substack{\text{phase}\\\text{bit}}}{1}\ \underset{\substack{\text{parity}\\\text{bit}}}{0} \equiv |\phi^{-}\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}$$



If A and B are in the same location and we do a joint measurement on them with a complete projection measurement in the basis $\{|\phi^{+}\rangle, |\phi^{-}\rangle, |\phi^{+}\rangle, |\phi^{-}\rangle\}$, then the projection operators are

$$\left.\begin{aligned} P_{00} &= |\phi^{+}\rangle\langle\phi^{+}| \\ P_{01} &= |\psi^{+}\rangle\langle\psi^{+}| \\ P_{10} &= |\phi^{-}\rangle\langle\phi^{-}| \\ P_{11} &= |\psi^{-}\rangle\langle\psi^{-}| \end{aligned}\right\} \text{ with outcomes } \begin{cases} 00 \\ 01 \\ 10 \\ 11 \end{cases}$$

3.3.1 Superdense coding

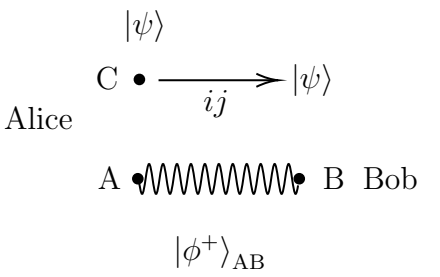
Consider the case where Alice and Bob are far apart from each other. The aim is for Alice to send 2 bits to Bob, but there is no way for classical communication. She can achieve the aim if they share a bell state $|\phi^+_{AB}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ to start with.

Alice's message	Act on A by a unitary operator	Final state of AB	Bob's action
00	I	$\|\phi^+\rangle$	Take Bell measurement
01	Z	$\|\phi^-\rangle$	
10	X	$\|\psi^+\rangle$	
11	XZ	$\|\psi^-\rangle$	

[Lecture 8 finish]

3.3.2 Quantum teleportation

Suppose Alice and Bob each hold one qubit in $|\phi^+\rangle_{AB}$, and Alice has another qubit C in state $|\psi\rangle = a|0\rangle + b|1\rangle$, which is not available to Bob. The aim is for Alice to send C to Bob, but she cannot do it physically to Bob and there isn't any quantum channel between them i.e. only classical communication is allowed.



The *protocol* is to use quantum teleportation. The initial state is CAB where Alice holds CA and Bob holds B. Then

$$\begin{aligned}
|\psi\rangle_C \otimes |\phi^+\rangle_{AB} &= (\alpha|0\rangle + \beta|1\rangle)_C \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{AB} \\
&\overset{!}{=} \frac{1}{2\sqrt{2}}(|00\rangle + |11\rangle)(\alpha|0\rangle + \beta|1\rangle) \mapsto \frac{1}{2}|\phi^+\rangle_{CA}|\psi\rangle_B \\
&+ \frac{1}{2\sqrt{2}}(|00\rangle - |11\rangle)(\alpha|0\rangle - \beta|1\rangle) \mapsto \frac{1}{2}|\phi^-\rangle_{CA}(Z|\psi\rangle)_B \\
&+ \frac{1}{2\sqrt{2}}(|01\rangle + |10\rangle)(\alpha|1\rangle + \beta|0\rangle) \mapsto \frac{1}{2}|\psi^+\rangle_{CA}(X|\psi\rangle)_B \\
&+ \frac{1}{2\sqrt{2}}(|01\rangle - |10\rangle)(\alpha|1\rangle - \beta|0\rangle) \mapsto \frac{1}{2}|\psi^-\rangle_{CA}(XZ|\psi\rangle)_B
\end{aligned}$$

Now Alice does a Bell measurement with outcome ij, $i, j \in \{0, 1\}$. Using classical communication, she sends ij to Bob.

- If $i = 0$ and $j = 1$, what is the post measurement state of CAB? The answer is $|\psi^+\rangle_{CA} \otimes (X|\psi\rangle)_B$. Then Bob's state is $(X|\psi\rangle)_B$. If he acts on it with X, then

$$X(X|\psi\rangle) = X^2|\psi\rangle = |\psi\rangle$$

- In general, if Alice's outcome was ij then Bob needs to act on his final state with $Z^i X^j$. To see this, one can check that his final state would be

$$X^j Z^i |\psi\rangle$$

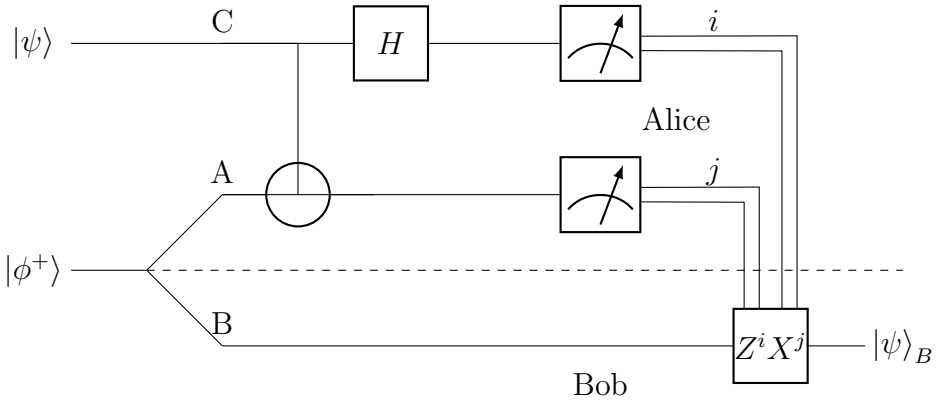
  then acting with $Z^i X^j$ leads to $|\psi\rangle$.

- Bob ends up with a copy of $|\psi\rangle$. Does this violate with the no-cloning theorem? The answer is no, because C is no longer in state $|\psi\rangle$.

- State transfer is unaffected by any physical process in the intervening space.

- Example of no-signalling: Consider the two preparations of Bob, before Alice's measurement he has a part of $|\phi^+\rangle_{AB}$, afterwards he has one of $|\psi\rangle, Z|\psi\rangle, X|\psi\rangle$ and $XZ|\psi\rangle$. Then Bob cannot distinguish between the two preparations, unless he knew the result of Alice's measurement. If he could, then it would violate the no-signalling principle.

The process can be illustrated diagramically:

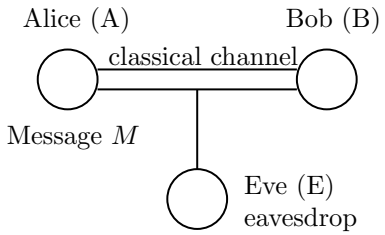


4 Quantum cryptography

The disadvantages of encoding information in quantum are:

- they cannot be raliably identified;
- to read a message, a measurement is to be taken, so the information is partial (probabilistic) and the state is damaged.

But these turn out to be the advantages of quantum cryptography. One task of quantum cryptography is to have secure/private communication.



We assume that the channel is *authenticated* i.e. Bob can verify that the message comes from Alice, and that Eve cannot modify the message. There exists a perfectly (provably) secrue cryptosystem to o this, known as the *One-time Pad*.

4.1 The one-time pad

The mechanism is for Alice and Bob to share a *private* key K, which is a sequence of random bits which Eve does not know. Further,

- K is created prior to this use.
- K is independent of the message M.
- the number of bits are equal: $|K| = |M| = n$.

The steps are the following:

(I) Alice computes $C = M \oplus K$. For example, if $M = \texttt{0110}$ and $K = \texttt{1010}$ then $C = \texttt{1100}$.

(II) Alice sends C to Bob through the channel.

(III) Bob does $C \oplus K = M \oplus K \oplus K = M$. So he receives the message.

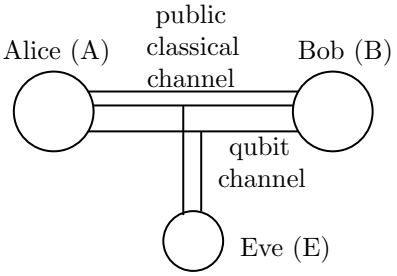
Eve knows C, but she cannot infer from it. The probability of Eve guessing right is 2^{-n}. However, the one-time pad is secure but inefficient due to choice of key. This is where QIC can help.

4.2 Quantum key distribution (QKD)

QKD allows Alice and Bob to generate a secret key (which they can use for the one-time pad) without having to meet or to trust a third party intermediary.

Note. *QKD is not used to encode the message itself, but only to generate the key.*

There are various protocols, such as BB84, B91 or E91. We will now focus on BB84, developed by Bennett and Brassard in 1984. We'll deviate from our previous setup a little bit:



[Lecture 9 finish]

Definition (Mutually unbiased bases (MUB))**.** Recall $\mathcal{B}_0 = \{|0\rangle, |1\rangle\}$ and $\mathcal{B}_1 = \{|+\rangle, |-\rangle\}$. If we measure any vector of one basis in the other basis, we get equally likely outcomes.

4.2.1 BB84

The steps are the following:

Step 1. Alice generates two m-bit strings uniformly at random:

$$\mathbf{x} = x_1 x_2 \dots x_m \in \{0,1\}^m$$
$$\mathbf{y} = y_1 y_2 \dots y_m \in \{0,1\}^m$$

and prepares m-qubits in the state

$$|\psi_{\mathbf{xy}}\rangle = |\psi_{x_1 y_1}\rangle \, |\psi_{x_2 y_2}\rangle \dots |\psi_{x_m y_m}\rangle$$
$$\text{where } |\psi_{00}\rangle = |0\rangle, \; |\psi_{10}\rangle = |1\rangle$$
$$|\psi_{01}\rangle = |+\rangle, \; |\psi_{11}\rangle = |-\rangle$$

If $y_i = 0$ she encodes x_i in the basis $\mathcal{B}_0 = \{|0\rangle, |1\rangle\}$. If $y_i = 1$ then in $\mathcal{B}_1 = \{|+\rangle, |-\rangle\}$. Then she sends $|\psi_{x_i y_i}\rangle$ $i = 1, \dots, m$ to Bob through m uses of qubit channel. Bob receives m-qubits *but* they need not be in the state $|\psi_{\mathbf{xy}}\rangle$ due to several reasons:

- Eve's tempering (measurement);
- noise in channel;
- bit flip $|0\rangle \mapsto |1\rangle$.

Let's first consider **Case I** where there's no disturbance. Then we can go on with

Step 2. Generate an m-bit string uniformly at random:

$$\mathbf{y}' = y_1' y_2' \dots y_m' \in \{0,1\}^m$$

If $y_i' = 0$ then he measures the i^{th} qubit in $\mathcal{B}_0 = \{|0\rangle, |1\rangle\}$, otherwise in $\mathcal{B}_1 = \{|+\rangle, |-\rangle\}$. Equivalently, since $H|\psi_{x_i y_i}\rangle \in \mathcal{B}_0$, then do measurement in $\mathcal{B}_0$. Let the outcomes of the measurement on m-qubit be

$$\mathbf{x}' = x_1' \dots x_m' \in \{0,1\}^m$$

Claim. If $y_i' = y_i$ then $x_i' = x_i$.

Proof. Suppose $y_i' = y_i = 0$ then $|\psi_{x_i y_i}\rangle = |\psi_{x_i 0}\rangle \in \mathcal{B}_0$. Since $y_i' = 0$, measurement by Bob in $\mathcal{B}_0$ yields the right come $x_i' = x_i$ with probability 1. On the other hand, if $y_i' = y_i = 1$ then $|\psi_{x_i 1}\rangle \in \mathcal{B}_1$, then for Bob, $H|\psi_{x_i 1}\rangle \in \mathcal{B}_0$ and thus measures in $\mathcal{B}_0$. Hence he determines x_i' unambigously i.e. $x_i' = x_i$. □

Step 3. Alice and Bob compare $\mathbf{y}$ and $\mathbf{y}'$ over the public classical channel. Then

- they discard all the x_i, x_i' for which $y_i \neq y_i'$.
- they do not reveal the other x_i, x_i''s.

Then they're left with shorter strings $\tilde{\mathbf{x}}, \tilde{\mathbf{x}}'$. If we are in **Case I**, then $\tilde{\mathbf{x}} = \tilde{\mathbf{x}}'$, so they now share a secret key.

Example ($m = 8$)**. Step 1.** Let's say Alice generates

$$\mathbf{x} = \texttt{01110100}$$
$$\mathbf{y} = \texttt{11010001}$$

Then Alice sends 8 qubits $|\psi_{x_1 y_1}\rangle \dots |\psi_{x_8 y_8}\rangle$ to Bob.

Step 2. Then Bob generates

$$\mathbf{y}' = \texttt{01110110}$$

Then

$$x_1 = 0,\ y_1 = 1 \ \Rightarrow\ |\psi_{01}\rangle = |+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

and since $y_1' = 0$, he measures in $\mathcal{B}_0$. Then outcomes are $0, 1$ each with probability $\frac{1}{2}$. Let's say $x_1' = 1$. For $x_2 = 1,\ y_0 = 1$ we have $|\psi_{11}\rangle = |-\rangle$ and $y_2' = 1$ so $x_2' = 1$ with probability 1. Suppose we get

$$\mathbf{x} = \texttt{01110100}$$
$$\mathbf{y} = \texttt{11010001}$$
$$\mathbf{y}' = \texttt{01110110}$$
$$\mathbf{x}' = \texttt{11010101}$$

Step 3. Comparing, we only keep the key `110`.

Claim. On average, the shared key $\sim \lfloor \frac{m}{2} \rfloor$ bits long. To see this, note that $p(y_i' = y_i) = p(y_i' \neq y_i) = \frac{1}{2}$, and so $p(\text{discard } x_i, x_i') = p(y_i' \neq y_i) = \frac{1}{2}$. So on average $\lfloor \frac{m}{2} \rfloor$ bits remain.

We now consider **Case II** where eavesdropping can happen, so $\tilde{\mathbf{x}}' \neq \tilde{\mathbf{x}}$ in general. Then they need to do

Step 4. Information Reconciliation (IR) Since $\tilde{\mathbf{x}}' \neq \tilde{\mathbf{x}}$ in general, Alice and Bob want to find the bit error rate (BER) of $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{x}}'$ i.e. the propotion of bits in $\tilde{\mathbf{x}}'$ which do not match with corresponding bits in $\tilde{\mathbf{x}}$. To do this, they *publicly compare* a sample of bits from their strings e.g.

$$\begin{aligned} &\text{(A)} \quad - * - * * - - - \in \tilde{\mathbf{x}} \\ &\text{(B)} \quad - * - * * - - - \in \tilde{\mathbf{x}}' \end{aligned}$$

They then determine BER in this sample. They discard *all* the bits x_i, x_i' in the samples, and they're left with $\tilde{\tilde{\mathbf{x}}}$ and $\tilde{\tilde{\mathbf{x}}}'$. They *assume* that $\tilde{\tilde{\mathbf{x}}}$ and $\tilde{\tilde{\mathbf{x}}}'$ have the *same* BER as in the considered samples. They then correct these errors (even though they do not know their locations) via IR. This then leaks info to Eve. So we have to sacrifice some more bits to get two strings that match.

Step 5. Privacy Amplification (PA) They now have a shared key about which Eve has no information. From the estimated BER they inferred how much information Eve has about the final strings.

Remark. BER depends on information that Eve gets - more information implies more disturbance due to measurement, which in turn implies higher BER. But the key point is MUBs for encoding. Noise (in qubit channel) also contributes to BER, but Alice and Bob can assume that all error arose from Eve's tempering i.e.

$$\text{BER} \gg \text{proportion of bits about which Eve has info}$$

Now we consider Eve's strategy. She is able to

- *intercept* and *resend attack*. She intercepts the qubits one-by-one, does measurement and resends (sends post-measurement state).
- use an auxiliary system (*ancilla*) to interact with some qubits as they pass through, changing the states, before measuring on the ancilla, giving information abuot multiple qubits at the same time.

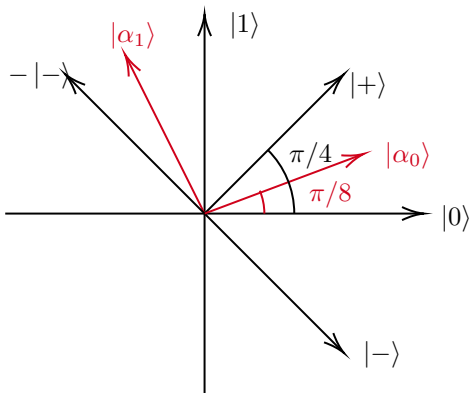
———————————[Lecture 10 finish]———————————

Let's first consider intercepting and resending attack.

Example (Measurement)**.** Eve may measure each qubit in the *Breidbart basis* $\{|\alpha_0\rangle, |\alpha_1\rangle\} \in \mathbb{C}^2$ given by

$$|\alpha_0\rangle = \cos\frac{\pi}{8}|0\rangle + \sin\frac{\pi}{8}|1\rangle, \ |\alpha_1\rangle = -\sin\frac{\pi}{8}|0\rangle + \cos\frac{\pi}{8}|1\rangle$$

or diagramically,



If a qubit (sent by A) is in $|0\rangle$ or $|+\rangle$ then Eve's measurement outcomes 0 (corresponding to $|\alpha_0\rangle$) and 1 (corresponding to $|\alpha_1\rangle$) has probabilities

$$p(0) = \cos^2 \frac{\pi}{8},\ p(1) = \sin^2 \frac{\pi}{8}$$

For each of the four encoding states

$$\begin{cases} |0\rangle, |+\rangle & y_i = 0 \\ |1\rangle, |-\rangle & y_i = 1 \end{cases}$$

What is the probability $p(x' \neq x)$ where $x \in \tilde{\mathbf{x}}, x' \in \tilde{\mathbf{x}}'$? If A sent $|0\rangle$ we know Bob measures in $\mathcal{B}_0 = \{|0\rangle, |1\rangle\}$ (since $\tilde{\mathbf{x}}, \tilde{\mathbf{x}}'$ consist of bits i for which $y_i' = y_i$). $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{x}}'$ are strings for which the *measurement basis* of B is the *encoding basis* for A.

Now suppose $x = 0, x' = 1$ and A sent $|0\rangle$ then

$$\begin{aligned} p(x = 0, x' = 1) &= p(x' = 1 | \text{A sent } |0\rangle) \\ &= p(x' = 1 | \text{E sent } |\alpha_0\rangle) P(\text{E gets outcome } 0 | \text{A sent } |0\rangle) \\ &+ p(x' = 1 | \text{E sent } |\alpha_1\rangle) P(\text{E gets outcome } 1 | \text{A sent } |0\rangle) \\ &= |\langle 1|\alpha_0\rangle|^2 |\langle 0|\alpha_0\rangle|^2 + |\langle 1|\alpha_1\rangle|^2 |\langle \alpha_1|0\rangle|^2 \\ &= 2 \sin^2 \frac{\pi}{8} \cos^2 \frac{\pi}{8} = \frac{1}{4} \end{aligned}$$

Similarly $p(x = 0, x' = 1) = \frac{1}{4}$ when A sends $|+\rangle$, and $p(x = 1, x' = 0) = \frac{1}{4}$ when A sends $|1\rangle$ or $|-\rangle$. Hence $p(x' \neq x) = \frac{1}{4}(4 \cdot \frac{1}{4}) = \frac{1}{4}$. Therefore Eve's action leads to a BER of $\frac{1}{4}$.

Suppose Alice and Bob have estimated the BER, what would they do now?

Example (A simple case of IR)**.** We are now in **Step 4**. Suppose A and B have $\tilde{\mathbf{x}}, \tilde{\mathbf{x}}'$ respectively, each having 7 bits, and suppose their estimated BER is $\frac{1}{7}$. We now write

$$\tilde{\mathbf{x}} = \mathbf{a} = a_1 a_2 \dots a_7;\ \tilde{\mathbf{x}}' = \mathbf{b} = b_1 b_2 \dots b_7$$

Using the classical channel they decide to act on $\mathbf{a}, \mathbf{b}$ by H, the check matrix of the Hamming code $[7, 4]$, where

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Then A computes the syndrome

$$\mathbf{s}^A = H\mathbf{a}^T = H\begin{pmatrix} a_1 \\ \vdots \\ a_7 \end{pmatrix} = \begin{pmatrix} s_1^A \\ s_2^A \\ s_3^A \end{pmatrix}$$

Then A sends the values s_i^A to B via the classical channel. On the other hand, B computes

$$\mathbf{s}^B = H\mathbf{b}^T = \begin{pmatrix} s_1^B \\ s_2^B \\ s_3^B \end{pmatrix}$$

So

$$\mathbf{s} = \mathbf{s}^B - \mathbf{s}^A = H(\mathbf{b} - \mathbf{a})^T = H\mathbf{e}^T \tag{4.1}$$

where $w(\mathbf{e}) = 1$, the *Hamming weight* or number of non-zero elements of $\mathbf{e}$. This is because the estimated BER implies at most one difference between $\mathbf{a}$ and $\mathbf{b}$. We know as a fact from the *classical error correcting code* (CECC) that there exists some bitstring $\mathbf{v} \in \{0,1\}^7$ with $w(\mathbf{v}) \leqslant 1$ such that

$$H\mathbf{v}^T = \mathbf{s} \tag{4.2}$$

Then combining (4.1) and (4.2), we have

$$\mathbf{v} = \mathbf{e}$$

Now B replaces $\mathbf{b}$ by $(\mathbf{b} - \mathbf{e})$. This is possible because he knows $\mathbf{s}$ and H. But in fact

$$\mathbf{b} - \mathbf{e} = \mathbf{a}$$

Hence they get the shared key $\mathbf{a}$.

Example (A 3-bit example of PA)**.** Now we are in **Step 5**. Suppose A and B end up with

$$\mathbf{a} = a_1a_2a_3 = \mathbf{b}$$

And suppose Eve knows at most one bit. Let

$$\mathbf{c} = (a_1 \oplus a_3, a_2 \oplus a_3) \in \{0,1\}^2$$

Claim. E knows nothing about $\mathbf{c}$.

Proof. To see this, we list the possible values of $\mathbf{a}$ and their corresponding $\mathbf{c}$:

$$\begin{aligned} \mathbf{a} &= \underline{0}00,\ \underline{0}01,\ \underline{0}10,\ \underline{0}11,\ 100, \ldots \\ \mathbf{c} &= 00,\ 11,\ 01,\ 10, \ldots \end{aligned}$$

Suppose Eve only know a_1, say $a_1 = 0$, then there are four possible values of $\mathbf{c}$ that she can get, but this gives no information. □

Hence $\mathbf{x}^* = \mathbf{x}'^* = \mathbf{c}$ is indeed a *private shared key* shared by A and B.

5 Basic notions of classical computations and computational complexity

5.1 Definitions

5.1.1 Computational task and algorithm

Definition. The *input* is a bit string, and the *input size* is the number of bits.

For example, given a n-bit string $\mathbf{b}$ for any $n \in \mathbb{N}$, is $\mathbf{b}$ prime?

Definition. The *output* is another bit string. If the output is either 0 or 1 then the task is called a *decision problem*.

Definition. $B = B_1 = \{0, 1\}$ is the *binary alphabet*. We'll call $B_n = \{0, 1\}^n$ and $B^* = \bigcup_{n=1}^{\infty} B_n$. A subset $L \subseteq B^*$ is called a *language* and a decision problem corresponds to the *recognition of a language*. Here recognition is recognising the binary strings which when used as in input yield yes/accept/0 as an answer.

More generally, the output is of length $n > 1$.

Example. The task FACTOR(x) has output y a non-trivial factor of x, or 1 if x is prime.

———————————[Lecture 11 finish]———————————

To solve a problem we use an *algorithm*: a precise set of instructions.

Definition. An *algorithm* is *efficient* if the number of elementary steps needed to execute it scales no faster than polynomially in n, the input size. The number of elementary steps is denoted $T(n)$, the *run-time*.

5.1.2 Model of classical computation

There are various known models of classical computation:

- Turing machine
- Cellular automata
- Circuit model (gate array)

We'll focus on the last one. In the circuit model, the input string

$$x = b_1 \dots b_n \in \{0, 1\}^n$$

is extended to

$$b_1 \dots b_n 0 \dots 0$$

The computational steps of this model are application of designated Boolean gates $f : B_n \to B_m$, resulting in the updated string. These fixed operatrions/gates should not become more complicated as n increases. We'll consider a *universal* set of gates

$$\{\text{AND}, \text{NOT}, \text{OR}\}$$

Any Boolean function/gate can be constructed from these. The output is some value of designated bits after the final step.

Definition (Circuit)**.** For each input size n, we have a *circuit* C_n, a prescribed sequence of computational steps. C_n only depends on n and *not* on the particular input. Then C_n is an algorithm or a computer programme.

Definition (Randomised classical computation)**.** In this model, the input string is extended to

$$b_1 \dots b_n \underbrace{r_1 \dots r_k}_{\text{random bits}} 0 \dots 0$$

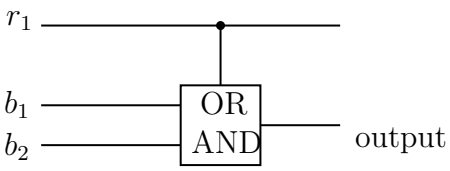
where the random bits $r_1 \dots r_n$ are chosen uniformly at random. If computation is repeated with the same $b_1 \dots b_n$, then $r_1 \dots r_k$ are generally different, so we'll get a different output. Hence

$$p(\text{any particular output}) \sim \frac{a}{2^k}$$

where a is the number of strings $r_1 \dots r_k$ that give the output.

We require $p(\text{output is correct}) > 1 - \delta$ for some chosen δ.

Example. Such a circuit can look like



5.2 Time-complexity of algorithms

In the circuit model,

$$T(n) = \text{total number of gates used in algorithm}$$

But is $T(n) < cn^k$ for all n large enough for some $c > 0$? If this is true, than $T(n)$ is a poly(n) algorithm or *poly-time* algorithm. Does $T(n)$ grow faster than a polynomial in n? For example, $2^n, 2^{\sqrt{n}}, n^{\log n}$.

Definition. For a positive function $T(n)$ we write

$$T(n) = O(f(n))$$

if there exists positive constants n_0, c such that

$$T(n) \leqslant cf(n) \ \forall n \geqslant n_0$$

Hence, 'T grows no faster than f' means

$$T(n) = O(\text{poly}(n))$$

if $T(n) = O(n^k)$ or some $k > 0$.

5.2.1 Time-complexity classes

Definition. P (poly-time) class contains languages whose membership can be decided (w.p. 1) by an poly(n) algorithm. In other words, a class of (decision) problems which can be solved on a classical computer with a deterministic algorithm in poly-time.

Definition. BPP (bounded error probabilistic poly-time) is a class of problems solvable in poly-time via randomised computation/algorithm w.p. $\geqslant \frac{2}{3}$.

Remark. The threshold $\frac{2}{3}$ can be increased to $1-\varepsilon$ for any $0 < \varepsilon < \frac{1}{2}$.
If there is a poly-time algorithm that succeeds w.p. $\frac{1}{2}+\delta$ for any chosen $\delta > 0$, then there exists a poly-time algorithm that succeeds w.p. $1-\varepsilon$ for any $0 < \varepsilon < \frac{1}{2}$.
To see this, suppose we repeat k times an algorithm that gives the correct answer w.p. $\frac{1}{2}+\delta$. The reasonable strategy is to take the majority voe. *Chernoff bound* tells us how well this strategy works.

Theorem (Chernoff bound)**.** *Let $X_1, X_2, \ldots, X_n$ be discrete independent identically distributed random variables where*

$$X_i = \begin{cases} 1 & w.p.\ \frac{1}{2}+\delta \\ 0 & w.p.\ \frac{1}{2}-\delta \end{cases}$$

or $X_i \sim \mathrm{Ber}(\frac{1}{2}+\delta)$. Then

$$P\left(\sum_{i=1}^{n} X_i \leqslant \frac{n}{2}\right) \leqslant e^{-2\delta^2 n}$$

We can interpret the values 1 as correct result and 0 wrong result. Then $\sum_{i=1}^{n} X_i$ is just the number of correct answers. So if we do k trials, then

$$\begin{aligned} P(\text{majority vote gives correct answer}) &= P\left(\sum_{i=1}^{k} X_i > \frac{k}{2}\right) \\ &= 1 - P\left(\sum_{i=1}^{k} X_i < \frac{k}{2}\right) \\ &\geqslant 1 - e^{2\delta^2 k} \end{aligned}$$

Hence the probability of making an error by taking the majority votes decreases exponentially in k. In the BPP definition, we chose $\frac{1}{2}+\delta = \frac{2}{3}$, then we just need to choose k to be a few hundred, to get the probability of error of 10^{-20}.

Example (FACTOR (N, M))**.** Given integers N, M with $M < N$, we need to decide if N has a non-trivial factor less than M. The best known algorithm has

$$T(n) = \exp\left[O\left(n^{1/3}(\log n)^{1/3}\right)\right]$$

In conclusion, the relation between classes is

$$\mathbf{P} \subseteq \mathbf{BPP} \subseteq \mathbf{EXP}$$

where **EXP** denotes class of all algorithms, which is non-empty as we saw in the example above.

[Lecture 12 finish]

5.2.2 Black-box/Oracle promise problems

Instead of input being a bit string of length n, we are given $f : B_n \to B_m$, known as a black box/oracle. We can query the oracle by giving it inputs and processing the outputs. Note that f is unknown, but we are given a *promise* about it. The task is to find some desired property of f by querying the oracle the least number of times. We can consider the oracle as another gate.

Definition. The *query complexity* if the number of times the oracle is used (as a function of its size).

Example.

(1) (Balanced v. constant problem) The input is a black box for $f : B_n \to B$ (one-bit output), and we are promised that f is either

- constant i.e. $f = 0$ for all $x \in B_n$ or $f = 1$ for all $x \in B_n$, or
- balanced i.e. $f = 0$ (or $f = 1$ respectively) for exactly half of the number of strings i.e. 2^{n-1} strings.

The problem is to find wheter f is constant or balanced with some desired probability (e.g. 0.99).

(2) (Search) The input is $f : B_n \to B$ with the promise that there is a unique $x \in B_n$ such that $f(x) = 1$ and $f = 0$ otherwise. The problem is then to find the special x.

(3) (Periodicity) The input is $f : \mathbb{Z}_n \to \mathbb{Z}_n$ where $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$. We are promised that f is periodic with period r i.e. $f(x + r \pmod n) = f(x)$ for all x. The problem is then to find r.

6 Circuit model of quantum computation

In the classical case, the input can be strings $b_1 \dots b_n 0 \dots 0$, while in the quantum space, we use qubits $|b_1\rangle \dots |b_n\rangle |0\rangle \dots |0\rangle$ where $b_i \in \{0,1\}$, $|b_i\rangle \in \mathbb{C}^2$.

Example. $H|0\rangle = |+\rangle$ then making measurement in $\{0,1\}$ gives outcomes $0,1$ each with probability $\frac{1}{2}$.

Definition. A *computational step* is the application of a quantum gate to a prescribed choice of qubits. These gates make up a *quantum circuit* (C_n). The *output* would then be the result of a measurement on some specified set of qubits. Measurement is done right at the end.

Definition. A *quantum computation* is defined by a family of circuits $(C_1, C_2, \dots)$.

Remark. Quantum gates are reversible but classical ones are not so in general.

Example. If x AND $y = 0$ then $xy \in \{01, 10, 00\}$. If $\text{NOT}(b) = 0$ then $b = 1$ and 0 otherwise.

6.1 Poly-time quantum computations and BQP

Definition (BQP). *Bounded-error quantum poly-time* is the class of (decision) problems that can be solved in poly-time with some fixed accuracy on a quantum computer. For each input size n we have a circuit C_n where

$$|C_n| = O(\text{poly}(n)) \text{ w.p. } \geqslant \frac{2}{3}$$

where $|C_n|$ is the number of gates in C_n.

Note. *We don't have to consider "QP" because the measurement at the end of a circuit gives an upper bound for error.*

Claim. $\mathbf{BPP} \subseteq \mathbf{BQP}$.

Proof. Any $\text{poly}(n)$ classical circuit can be replaced by an equivalent classical circuit of *reversible gates* and this is also a quantum circuit albeit consisting of gates which preserves the computational basis as a set. □

The question naturally arises: Is **BQP** strictly larger than **BQP**? Sadly, we don't know. We know that $\text{FACTOR}(M, N) \in \mathbf{BQP}$ but we don't know if it is in **BPP**.

6.1.1 Black-box promise problem in quantum computation

Consider $f : B_m \to B_n$. Its quantum analogue is a unitary operator U_f which is a reversible version of f $(\tilde{f})$.

Note. *Any $f : B_m \to B_n$ can be expressed equivalently in a reversible form: $\tilde{f} : B_{m+n} \to B_{m+n}$ where for $b \in B_m$ and $c \in B_n$,*

$$f(b, c) = (b, c \oplus f(b))$$

where $\oplus$ stands for addition $\pmod 2$. If we can compute f and do $\oplus$ then we can evaluate $\tilde{f}$. Conversely to find f from $\tilde{f}$ we can just set $c = 0 \dots 0$ and only read out the last n bits of $\tilde{f}$. Moreover, $\tilde{f}$ is reversible. To see this,

$$\tilde{f}(\tilde{f}(b, c)) = \tilde{f}(b, c \oplus f(b)) = (b, c \oplus f(b) \oplus f(b)) = (b, c)$$

Hence any classical algorithm using an oracle for f can equally well be done using an oracle for $\tilde{f}$.

Definition. The *quantum oracle* for $f : B_m \to B_n$ with $\tilde{f} : B_{m+n} \to B_{m+n}$ is defined as

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

where $|x\rangle$ is a m-qubit and $|y\rangle$ a n-qubit. Here $|x\rangle, |y\rangle$ are computational basis states i.e. $x \in \{0,1\}^m$ and $y \in \{0,1\}^n$.

Let $|\psi\rangle \in (\mathbb{C}^2)^{\otimes m+n}$ where $\{|x\rangle \otimes |y\rangle\}$ with $x \in B_m, y \in B_n$ is an orthonormal basis, can be written as

$$|\psi\rangle = \sum_{\substack{x \in B_m \\ y \in B_n}} c_{xy} |x\rangle |y\rangle$$

so

$$U_f |\psi\rangle \stackrel{!}{=} \sum c_{xy} U_f |x\rangle |y\rangle = \sum c_{xy} |x\rangle |y \oplus f(x)\rangle$$

Here $|x\rangle$ is the *input register* and $|y\rangle$ the *output register*.

[Lecture 13 finish]

Remark. From $\mathbf{P} \subseteq \mathbf{BPP} \subseteq \mathbf{BQP}$ it follows that any problem in $\mathbf{P}$ can be expressed using a circuit consisting of gates in $\{\text{AND}, \text{NOT}\}$. The quantum version of $\{\text{AND}, \text{NOT}\}$ are $\{\text{Toffoli}, X\}$ repectively. For U_f, it is important to have $|x\rangle$ at the output: classically no bits are lost. $\tilde{f}$ is a permuation of $(m+n)$-bit strings, so any U_f corresponding to a classical computation will be a permutation matrix, and randomness of $\mathbf{BPP}$ can be generated by superpositions and measurements.

6.2 Computation via quantum parallelism

U_f can act on a superposition of input registers (equal superposition states)

$$|\varphi_m\rangle = \frac{1}{\sqrt{2^m}} \sum_{x \in B_m} |x\rangle$$

by

$$\begin{aligned} U_f |\varphi_m\rangle |y\rangle &= U_f \left[\frac{1}{\sqrt{2^m}} \sum_{x \in B_m} |x\rangle |y\rangle \right] \\ &= \frac{1}{\sqrt{2^m}} \sum_{x \in B_m} U_f |x\rangle |y\rangle \\ &= \frac{1}{\sqrt{2^m}} \sum_{x \in B_m} |x\rangle |y \oplus f(x)\rangle \equiv |\psi_f\rangle \end{aligned}$$

$|\psi_f\rangle$ is generated by a single use of U_f, and it depends on $f(x)$ *for all* $x \in B_m$. This is known as *computation by quantum parallelism*. By further quantum processing of $|\psi_f\rangle$ (e.g. using other gates and measurement) we can get global information about f with just one use of U_f. In

contrast, with the classical case one use of f gives the output of only one input.

But how do we generate $|\varphi_m\rangle$? We can use the Hadamard gate:

$$\begin{aligned} H^{\otimes m}|0\rangle^{\otimes m} &= (H|0\rangle)^{\otimes m} = |+\rangle^{\otimes m} \\ &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \dots \ (m \text{ times}) \\ &= \frac{1}{\sqrt{2^m}} \sum_{x \in B_m} |x\rangle \end{aligned}$$

Remark. Hence we have created a superposition of 2^m terms using only a linear number of applicatons of H ($H^{\otimes m}$).

6.2.1 Approximately universal set of quantum gates

In the classical case we have $\{\text{AND}, \text{NOT}\}$ as a universal set. In the quantum case, we use unitary operators

$$U = e^{i\theta A}, \ A \text{ self adjoint}, \ \theta \in \mathbb{R}$$

U is then parameterised by a continuous parameter θ, so no finite set of quantum gates can generate all quantum gates exactly - even with very large circuits. But there exist *approximately universal sets* of quantum gates.

Definition (Closeness of unitary operators)**.** We say that the operator norm $\|U - V\| \leqslant \varepsilon$ if

$$\max_{\substack{|\psi\rangle \\ \langle\psi|\psi\rangle=1}} \|U|\psi\rangle - V|\psi\rangle\| \leqslant \varepsilon$$

Definition. A set $\mathcal{G}$ of quantum gates (acting on qubits) is *approximately universal* if for any unitary W acting on any number of qubits, and for any $\varepsilon \in [0, 1]$, there exists a circuit C, composed of gates from $\mathcal{G}$, whose overall unitary action satisfies

$$\|W - C\| \leqslant \varepsilon$$

Generally, $|C| = O(\exp(n))$ where n is the number of qubits on which W acts. But for some algorithms e.g. Quantum Fourier Transform, we have $|C| = O(\text{poly}(n))$.

Theorem (Solovay-Kitaev Theorem (Roughly))**.** *If $\mathcal{G}$ is an approximately universal set of quantum gates, and let*

$$\|W - \underbrace{g_{i_1} g_{i_2} \dots g_{i_k}}_{C}\| \leqslant \varepsilon, \ g_{i_j} \in \mathcal{G}$$

Then $k = |C| \sim (\log \varepsilon^{-1})^c$ for some constant $c \approx 2$.

Example. Let's take 2^{-12}, then $k \sim 12$ gates.

In Example Sheet 3 we will look that the quantum circuit $C = U_1 U_2 \dots U_m$ cannot be generated by a universal set, but can be generated by $C' = V_1 V_2 \dots V_m$ where $\|U_i - V_i\| \leqslant \varepsilon$. We will then prove that $\|C - C'\| \leqslant m\varepsilon$, so errors in the composition only add up, instead of multiplying.

7 The Deutsch-Jozsa (DJ) algorithm

This is an example of an algorithm showing expontential speed over classical one.

- **Problem:** Balanced v. constant.
- **Input:** Oracle for $f : B_n \to B$.
- **Promise:** f is either constant i.e. $f(x) = 0\ \forall x \in B_n$ or $f(x) = 1\ \forall x \in B_n$; or balanced i.e. $f(x) = 0$ (1 respectively) for exactly half of the number of inputs.

Classically, one needs $(2^{n-1} + 1)$ queries in the worse case scenario: We need to insepct output errors to the first 2^{n-1} outputs. If the outputs are the same for all those 2^{n-1} inputs, then we check the *next* (i.e. $(2^{n-1} + 1)^{\text{th}}$) output. If that is the same then we infer constant, else balanced. Hence $(2^{n-1} + 1)$ is sufficient, now we prove its necessary.

Necessity. Let's suppose we have an Adversary (A) who is in control of (the oracle of) f. Suppose we have a deterministic classical algorithm that solves the problem by making $k \leqslant 2^{n-1}$ queries. When the algorithm is applied, A hasn't yet chosen f, but simply gives output 0 for all queries. At the end of k queries the function f has been fixed on k inputs, but if $k \leqslant 2^{n-1}$ then A still has the freedom to choose the next output so as to contradict the outcome of the algorithm. So $(2^{n-1} + 1)$ queries are necessary. □

———————————[Lecture 14 finish]———————————

Our quantum oracle for $f : B_n \to B$ is U_f, where for $x \in B_n$ and $y \in B$,

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

Protocol of DJ algorithm

(i) Firstly, we *initialise* all qubits in state $|0\rangle$:

$$\left|x\right\rangle \left|y\right\rangle = \left|0\right\rangle^{\otimes n} \left|0\right\rangle$$

(ii) Second, we act on it with

$$\begin{aligned}(H^{\otimes n} \otimes HX)(\left|0\right\rangle^{\otimes n} \left|0\right\rangle) &= \left|+\right\rangle^{\otimes n} \otimes \left|-\right\rangle \\ &= \left(\frac{1}{\sqrt{2^n}} \sum_{x \in B_n} \left|x\right\rangle\right) \left|-\right\rangle \\ &= \left|A\right\rangle\end{aligned}$$

where in the second equality we used the notion of equal superposition states in §6.2.

(iii) Then, we use the oracle U_f:

$$\begin{aligned} U_f |A\rangle &= \frac{1}{\sqrt{2^n}} \sum_{x \in B_n} U_f |x\rangle |-\rangle \\ &= \frac{1}{\sqrt{2^n}} \sum_{x \in B_n} \left[U_f |x\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \right] \\ &= \frac{1}{\sqrt{2^{n+1}}} \sum_{x \in B_n} \left(|x\rangle |f(x)\rangle - |x\rangle |f(x)^C\rangle \right) \text{ where } f(x)^C = f(x) \oplus 1 \end{aligned}$$

Note that

$$\begin{aligned} U_f |x\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) &= \begin{cases} |x\rangle (|0\rangle - |1\rangle)/\sqrt{2} = |x\rangle |-\rangle & \text{if } f(x) = 0 \\ -|x\rangle (|0\rangle - |1\rangle)/\sqrt{2} = -|x\rangle |-\rangle & \text{if } f(x) = 1 \end{cases} \\ &= (-1)^{f(x)} |x\rangle |-\rangle \end{aligned}$$

Hence we can write

$$U_f |A\rangle = \frac{1}{\sqrt{2^n}} \left(\sum_{x \in B_n} (-1)^{f(x)} |x\rangle \right) |-\rangle$$

(iv) Note that the first n qubits are uncorrelated with the last qubit. So we can discard the last qubit and write

$$|f\rangle := \frac{1}{\sqrt{2^n}} \sum_{x \in B_n} (-1)^{f(x)} |x\rangle$$

Now we can consider what $|f\rangle$ looks like if

- f is constant; or
- f is balanced.

(a) **If f is constant**, then

$$|f\rangle = \pm \frac{1}{\sqrt{2^n}} \sum_{x \in B_n} |x\rangle = \pm H^{\otimes n} |0\rangle^{\otimes n}$$

since $f(x)$ has the same sign for all $x \in B_n$. If we apply $H^{\otimes n}$ on $|f\rangle$ then

$$H^{\otimes n} |f\rangle = \pm |0\rangle^{\otimes n}$$

(Recall that $H^2 = I$.)

(b) **If f is balanced**, then $|f\rangle$ has equal numbers of plus and minus signs (since exactly half of $x \in B_n$ has $f(x) = 1$ (or 0)) - at unknown locations. Now let

$$|\varphi_n\rangle = \frac{1}{\sqrt{2^n}} \sum_{x' \in B_n} |x'\rangle = H^{\otimes n} |0\rangle^{\otimes n}$$

and take the inner product of φ_n and $|f\rangle$:

$$\begin{aligned}\langle\varphi_n|f\rangle &= \frac{1}{2^n}\sum_{x,x'\in B_n}(-1)^{f(x)}\underbrace{\langle x'|x\rangle}_{\delta_{xx'}} \\ &= \frac{1}{2^n}\sum_{x\in B_n}(-1)^{f(x)} = 0\end{aligned}$$

where again that last equality follows since there are equally many plus and minus signs. Hence if f is balanced then

$$|f\rangle \perp |\varphi_n\rangle$$

Writing $H_n = H^{\otimes n}$, this means that

$$H_n|f\rangle \perp H_n|\varphi_n\rangle \;\Rightarrow\; H_n|f\rangle \perp |0\rangle^{\otimes n}$$

So if we write $H_n|f\rangle$ as a sum of n-qubits, the sum *must not* contain the $|00\ldots0\rangle$ term, so

$$H_n|f\rangle = \sum_{\substack{x\in B_n\\ x\neq 0\ldots0}} c_x|x\rangle$$

(Or equivalently $c_{0\ldots0} = 0$.)

Moving back to our protocol,

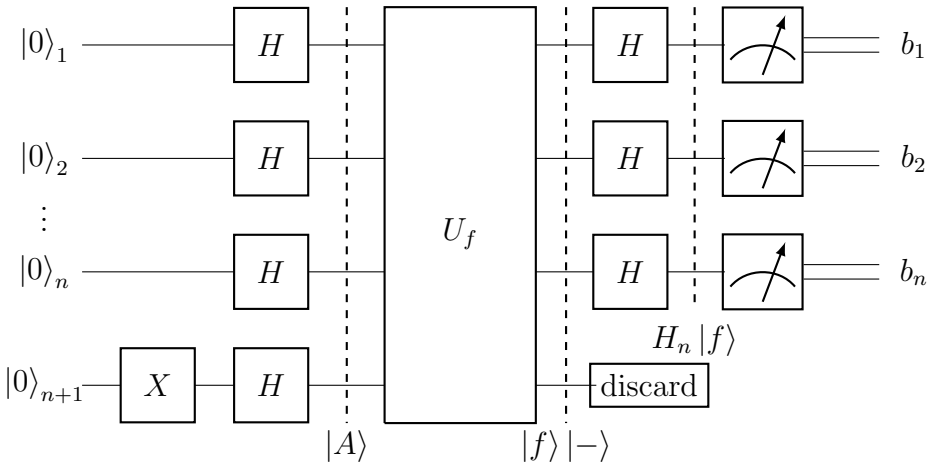
(v) Apply H_n on $|f\rangle$, as we discussed above.

(vi) Measure the n-qubits in the computational basis, then

- If f is constant, the outcome will be $0\ldots0$. Hence f is constant with probability 1.
- If f is balanced, then the outcome is *not* all zeros. Hence f is balanced with probability 1.

If we look at the number of queries used, we find that only **one** query is needed - the one use of U_f, along with $3n+2$ further operations:

- $(n+1)$ uses of H's and 1 use of X in step (ii);
- n uses of H's in step (iv); and
- n single qubit measurements.



What if classically one allows bounded error?

In the quantum case DJ works with one query *but* there is now a classical bounded error algorithm that solves the problem with only a *constant* number $O(\log 1/\varepsilon)$ of queries - depending only on ε and not on n. So the exponential speed-up of quantum case over classical case is lost.

7.1 Randomised classical algorithm

We choose k bits $x_1 \dots x_k$ uniformly at random. Then evaluate $f(x_1), \dots, f(x_k)$.

(a) If we get $f(x_i) = 0 \ \forall i$ or $f(x_i) = 1 \ \forall i$ then we infer that f is constant.

(b) If there exists $i \neq j$ such that $f(x_i) \neq f(x_j)$ then we infer that f is balanced.

In case (b) there is no error, but in case (a) there could be. Let p_e be the probability that f is balanced but we infer it to be constant. Then

$$p_e = \frac{2}{2^k} = \frac{1}{2^{k-1}}$$

which follows since each $f(x_i)$ has probability $1/2$ to take the value 0 or 1, and so there are two desired string (all 0 or all 1) among 2^k total possible strings. We would want $p_e < \varepsilon$, then we need

$$2^{k-1} > \frac{1}{\varepsilon} \ \Rightarrow \ k > \log \frac{1}{\varepsilon} + 1$$

Hence $k = O(\log 1/\varepsilon)$ suffices to guarantee $p_e < \varepsilon$ for all n.

Is there a black-box promise problem where we do get a quantum advantage even over bounded error classical problem? The answer is yes.

7.2 Simon's problem/algorithm

- **Input:** Oracle for $f : B_n \to B_n$.
- **Promise:** f is either *one-to-one* i.e. $f(x) = f(y)$ iff $x = y$, or *two-to-one* i.e. there exists $x \neq x'$ such that $f(x) = f(x') = f(y)$. Equivalently, we are given the promise that

  $$f(x) = f(y) \text{ if and only if } y = x \oplus \xi$$

  where $\xi = \texttt{0...0}$ corresponds to one-to-one, and $\xi \neq \texttt{0...0}$ two-to-one.
- **Problem:** Determine whether f is one-to-one or two-to-one, and in the latter case find ξ.

Note. *$f(x \oplus \xi) = f(x)$ and $f(x \oplus \xi \oplus \xi) = f(x)$ so f has a period.*

[Lecture 15 finish]

We'll see that in the quantum case $O(n)$ queries are needed, while in the classical case $O(\exp(n))$.

Example (Two-to-one function)**.** Consider $n = 3$:

x	000	001	010	011	100	101	110	111
$f(x)$	101	010	000	110	000	110	101	010

We see that $f(000) = f(110)$ and $f(010) = f(100)$. Since $000 \oplus 110 = 010 \oplus 100 = 110$, we conclude that $\xi = 110$.

Why is this hard classically? This is because we need to find two different inputs x, y such that $f(x) = f(y)$, but no structure of f is known. Suppose we make $2^{n/4}$ queries

$$x_1, x_2, \ldots, x_{2^{n/4}} \in B_n$$

Then the number of pairs of queries

$$\binom{2^{n/4}}{2} < (2^{n/4})^2$$

This is because

$$\binom{k}{2} = \frac{k!}{(k-2)!2!} = \frac{k(k-1)}{2} < k^2$$

The total number of pairs of $f(x)$ is 2^{n-1} (all values of $f(x)$ appear in pairs). Then the probability of picking a pair of inputs (x, y) such that $f(x) = f(y)$ is $2^{-(n-1)}$. Hence the total probability of successfully detecting $\xi \neq 0$ (i.e. f is two-to-one) is smaller than

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

Hence even as many as $2^{n/4}$ queries cannot help us detect that $\xi \neq 0 \ldots 0$ with better than *exponentially small probability.* It cannot form the basis of any bounded error algorithm.

8 Quantum Fourier transform

8.1 QFT mod N

We now generalise the Hadamard operator H from 2D ($H|0\rangle = |+\rangle$, etc.) to N-dimensions.

Definition. Let $\mathcal{H}_N$ be the N-dimensional Hilbert sapce with orthonormal basis $\mathcal{B}_N = \{|0\rangle, |1\rangle, \ldots, |N-1\rangle\}$ i.e. elements are labelled by elements of $\mathbb{Z}_N = \{0, 1, \ldots, N-1\}$. The quantum Fourier transform mod N (written QFT_N or simply QFT) is a unitary operator acting on $\mathcal{H}_N = \mathcal{B}_N$ by

$$\mathrm{QFT}\,|x\rangle = \frac{1}{\sqrt{N}} \sum_{y \in \mathbb{Z}_N} e^{i2\pi xy/N} |y\rangle = \frac{1}{\sqrt{N}} \sum_{y \in \mathbb{Z}_N} \omega^{xy} |y\rangle$$

where $\omega = e^{2\pi i/N}$ is the primitive N^{th} root of unity.

Note. *For $N = 2$ we return to qubits with $\mathcal{H}_2 \cong \mathbb{C}^2$ with orthonormal basis $\{|0\rangle, |1\rangle\}$ and $\mathrm{QFT}\,|0\rangle = |+\rangle$, $\mathrm{QFT}\,|1\rangle = |-\rangle$. Hence $\mathrm{QFT}_2 \equiv H$.*

We know an operator can be written in matrix form $\mathrm{QFT} \in \mathbb{M}_N(\mathbb{C})$, where

$$\begin{aligned}
(\mathrm{QFT})_{ij} &= \langle j|\,\mathrm{QFT}\,|k\rangle \\
&= \frac{1}{\sqrt{N}} \left\langle j \middle| \sum_{m \in \mathbb{Z}_N} e^{2\pi ikm/N} \middle| m \right\rangle \\
&= \frac{1}{\sqrt{N}} e^{2\pi ikj/N} = \frac{1}{\sqrt{N}} \omega^{jk}
\end{aligned}$$

So for example, the elements of the j^{th} row are

$$\frac{1}{\sqrt{N}}, \frac{\omega^j}{\sqrt{N}}, \frac{(\omega^j)^2}{\sqrt{N}}, \ldots, \frac{(\omega^j)^{N-1}}{\sqrt{N}}$$

We now denote

$$S_j = \sum_{k \in \mathbb{Z}_N} (\omega_j)^k$$

Recall

$$S := \sum_{k=0}^{N-1} \alpha^k = \begin{cases} N & \alpha = 1 \\ \dfrac{1-\alpha^N}{1-\alpha} & \alpha \neq 1 \end{cases}$$

Then setting $\alpha = \omega_j$ in the expression of S_j, we find

$$S_j = \begin{cases} N & \omega^j = 1 \\ \dfrac{1-(\omega^j)^N}{1-\omega^j} & \omega^j \neq 1 \end{cases}$$

But $\omega^j = 1$ if and only if $j = 0$. To see this, note that $\omega^j = e^{2\pi ij/N} = 1$ if and only if $j \equiv 0 \pmod{N}$, which only happens when $j = 0$. Therefore,

$$S_j = \begin{cases} N & j = 0 \\ 0 & j \neq 0 \end{cases}$$

where the case $j \neq 0$ follows from $(\omega^j)^N = 1$. With this, we can prove that QFT is indeed unitary i.e.

$$\mathrm{QFT}^\dagger \mathrm{QFT} = I = \mathrm{QFT}\, \mathrm{QFT}^\dagger$$

To see this, consider

$$(\mathrm{QFT}^\dagger \mathrm{QFT})_{km} = \sum_{j \in \mathbb{Z}_N} (\mathrm{QFT}^\dagger)_{kj} (\mathrm{QFT})_{jm}$$

and compute using our previous analysis to show this is δ_{km}.

8.2 Periodicity determination

We are given

- **Input:** Black-box for a function $f : \mathbb{Z}_N \to \mathbb{Z}_M$.
- **Promise:** f is periodic with period r, that is, r is the smallest number in $\mathbb{Z}_N$ such that $f(x+r \pmod{N}) = f(x)$. Also assume f is one-to-one in each period i.e. for any $x_1 \neq x_2$ and $x_1, x_2 \in \{0, \ldots, r-1\}$ we have $f(x_1) \neq f(x_2)$.
- **Task:** Find a method of determining r with some desired accuracy *independent of* N.

In the classical case, it can be shown that $O(N^{1/2})$ queries are necessary and sufficient. Note that

$$O(N^{1/2}) = O\left(2^{(\log N)/2}\right) = O\left((2^{1/2})^{\log N}\right) = O(\exp(\log N))$$

so it is *not* bounded by $\mathrm{poly}(\log N)$. In the quantum case, $O(\log \log N)$ queries suffice, along with $\mathrm{poly}(\log N)$ further processing steps. Hence the quantum algorithm is exponentially faster than the classical one.

Quantum algorithm for periodicity determination

- Construct

$$|\psi_N\rangle = \frac{1}{\sqrt{N}} \sum_{x \in \mathbb{Z}_N} |x\rangle \in \mathcal{H}_N$$

the uniform superposition states as before. Then consider the state

$$|\psi_N\rangle\, |0\rangle \in \mathcal{N} \otimes \mathcal{H}_N$$

and act on it with U_f:

$$\begin{aligned} U_f \left|\psi_N\right\rangle \left|0\right\rangle &= \frac{1}{\sqrt{N}} \sum_{x\in\mathbb{Z}_N} U_f \left|x\right\rangle \left|0\right\rangle \\ &= \frac{1}{\sqrt{N}} \sum_{x\in\mathbb{Z}_N} \left|x\right\rangle \left|0 + f(x) \pmod N\right\rangle \\ &= \frac{1}{\sqrt{N}} \sum_{x\in\mathbb{Z}_N} \left|x\right\rangle \left|f(x)\right\rangle \equiv \left|f\right\rangle \end{aligned}$$

Since r is the period of f, it follows that $r \mid N$. Then $A = N/r$ is the number of periods.

- Do a measurement on the second register in the basis $\mathcal{B}_M = \{|0\rangle, \ldots, |M-1\rangle\}$ and let the outcome be $y = f(x_0)$ where $x_0 \in \{0, 1, \ldots, r-1\}$ is the *least value* of x for which $f(x) = y$. We know $y = f(x_0) = f(x_0 + r) = \cdots = f(x_0 + (A-1)r)$, but we're only picking x_0. Note we stop at $x_0 + (A-1)r$ since $x_0 + Ar = x_0 + \frac{N}{r}r = x_0 + N \equiv x_0$.

[Lecture 16 finish]

- Let the probability of outcome $y = f(x_0)$ to be $p(y)$. Then the terms contributing to this outcome is

$$\frac{1}{\sqrt{N}} \left(\sum_{j=0}^{A-1} \left|x_0 + jr\right\rangle \right) \left|y\right\rangle$$

By the extended Born rule, we can find

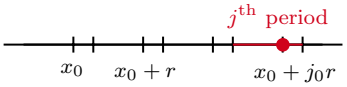
$$\begin{aligned} p(y) &= \left\| \frac{1}{\sqrt{N}} \sum_{j=0}^{A-1} \left|x_0 + jr\right\rangle \right\|^2 = \frac{1}{N} \sum_{j,j'} \left\langle x_0 + j'r | x_0 + j_r \right\rangle \\ &= \frac{1}{N} \sum_{j=0}^{A-1} 1 = \frac{A}{N} = \frac{1}{r} \end{aligned}$$

Hence $p(y = f(x_0)) = 1/r$ for any $x_0 \in \{0, \ldots, r-1\}$.

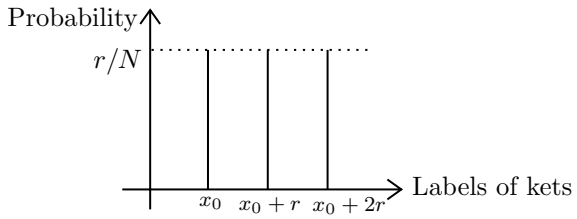
- The post-measurement state of the first register is

$$\left|\text{per}\right\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{A-1} \left|x_0 + jr\right\rangle / \sqrt{p(y)} = \frac{1}{\sqrt{A}} \sum_{j=0}^{A-1} \left|x_0 + jr\right\rangle$$

- Then, if we measure on the first register in $\mathcal{B}_N$, then the outcome is $(x_0 + j_0 r)$ for some $j_0 \in \{0, \ldots, A-1\}$ with probability $1/A$. Thus we have a random period (j_0^{th} period) and a random number in that period.



But this gives us *no* information about r. Since the outcomes x_0, $x_0 + r$, ... have a discrete distribution with probability r/N each.



Resolution with QFT

- Instead of measuring $|\text{per}\rangle$, we act on it by QFT($\equiv \text{QFT}_N$). For any $|x\rangle \in \mathcal{B}_N$ we have

$$\text{QFT}\,|x\rangle = \frac{1}{\sqrt{N}} \sum_{y \in \mathbb{Z}_N} \omega^{xy}\,|y\rangle\,,\ \omega = e^{2\pi i/N}$$

and so

$$\begin{aligned}
\text{QFT}\,|\text{per}\rangle &= \frac{1}{\sqrt{A}} \sum_{j=0}^{A-1} \text{QFT}\,|x_0 + jr\rangle \\
&= \frac{1}{\sqrt{NA}} \sum_{j=0}^{A-1} \sum_{j=0}^{N-1} \omega^{(x_0+jr)y}\,|y\rangle \\
&= \frac{1}{\sqrt{NA}} \sum_{y=0}^{N-1} \omega^{x_0 y} \left(\sum_{j=0}^{A-1} (\omega^{ry})^j \right) |y\rangle
\end{aligned}$$

Note that

$$S = \sum_{j=0}^{A-1} \omega^{jry} = \sum_{j=0}^{A-1} \alpha^j = \begin{cases} A & \alpha = 1 \\ \frac{1-\alpha^A}{1-\alpha} & \alpha \neq 1 \end{cases}$$

But $\alpha = \omega^{ry} = \exp(2\pi iry/N) = \exp(2\pi iy/A)$, so $\alpha = 1$ if $y = kA$ for $k = 0, \ldots, r-1$. If $\alpha \neq 1$ i.e. $y \neq kA$ for any $k \in \{0, \ldots, r-1\}$, we have $\alpha^A = 1$ and $S = 0$. In summary,

$$S = \sum_{j=0}^{A-1} \omega^{jry} = \begin{cases} A & y = kA \\ 0 & \text{otherwise} \end{cases}$$

Putting this back to our calculation of QFT,

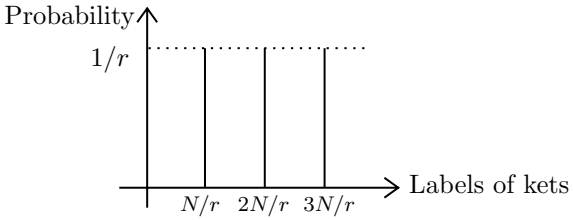
$$\begin{aligned}
\text{QFT}\,|\text{per}\rangle &= \frac{1}{\sqrt{NA}} \sum_{k=0}^{r-1} \omega^{x_0 kA} A\,|kA\rangle \\
&= \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} \omega^{x_0 kA}\,|kA\rangle \equiv |\Psi\rangle
\end{aligned}$$

Comparing this with our previous expression of $|\text{per}\rangle$, it follows that the phase shift x_0 has gone out of the state. But that's exactly what we want, since the phase does not matter after measurement.

- We now measure this state in $\mathcal{B}_N$, with outcome $c \equiv k_0 A = k_0 N/r$ for some $k_0 \in \{0, \ldots, r-1\}$ with probability

$$p(c) = \frac{1}{r} \left| \omega^{x_0 k N/r} \right|^2 = \frac{1}{r}$$

Hence the probability has no dependence on x_0.



Out aim now is to find r.

Periodicity determination

We have the relation

$$\frac{k_0}{r} = \frac{c}{N}$$

We know c and N, but how do we find r?

- **Case I:** k_0 is coprime to r. We cancel c/N down to the lowest term and read off the denominator $\tilde{r}$ which is equivalent to r. For example, if $k_0 = 3, r = 4$ and $c = 18, N = 24$, then

$$\frac{c}{N} = \frac{18}{24} = \frac{3}{4}$$

and we recover $r = 4$.

- **Case II:** k_0 is not coprime to r. Then cancellation leads to a denominator less than r. For example, if $k_0 = 3, r = 9$ and $c = 15, N = 45$ then

$$\frac{c}{N} = \frac{15}{45} = \frac{1}{3}$$

and $\tilde{r} = 3 < r$.

What do we do? We just take f and compare $f(0)$ with $f(\tilde{r})$. If $f(0) = f(\tilde{r})$ then $\tilde{r} = r$, otherwise $\tilde{r} < r$. Now comes the question: How do we know if k_0 if coprime to r? k_0 was uniformly chosen (by measurement) at random from $\{0, \ldots, r-1\}$. What is the probability that such uniformly random k_0 is coprime with the (unknown) period r? This is answered by the following theorem from number theory:

Theorem (Coprimality theorem)**.** *The number of integers less than $r \in \mathbb{Z}$ that are coprime to r grows as*

$$\Omega\left(\frac{r}{\log \log r}\right)$$

as r increases.

Note. *We write $f(n) = \Omega(g(n))$ if there is some constant α and $n_0 \in \mathbb{N}$ such that $f(n) \geqslant \alpha g(n)$ for all $n \geqslant n_0$.*

Hence if $k_0 < r$ is chosen uniformly at random, then

$$p(k_0 \text{ is coprime to } r) = \Omega\left(\frac{r/\log\log r}{r}\right) = \Omega\left(\frac{1}{\log\log r}\right)$$

which follows since r is the number of values that k_0 can take.

Claim. If we repeat the process $O(\log\log r) < O(\log\log N)$ times, we will obtain a coprime k_0 in at least one case with any fixed constant level of probability.

[Lecture 17 finish]

This follows from

Lemma 1. *Suppose a single trial has probability of success p and we repeat the trial m times independently, then for any constant $\varepsilon \in (0,1)$, the probability for at least one trial to be successful satisfies*

$$p_{\text{succ}} > 1 - \varepsilon$$

if e.g. $m = \lceil -(\log\varepsilon)/p \rceil = O(1/p)$.

In our case, $p = 1/\log\log r$ and hence $O(\log\log r)$ suffice to get one success.

Proof. We have

$$p_{\text{succ}} = 1 - p(\text{all fail}) = 1 - (1-p)^m \equiv 1 - \varepsilon$$
$$\Rightarrow\ (1-p)^m = \varepsilon\ \Rightarrow\ m\log(1-p) = \log\varepsilon\ \Rightarrow\ m = \frac{-\log\varepsilon}{-\log(1-p)}$$

But note that $p < -\log(1-p)$ for all $p \in (0,1)$, and so $m = \lceil -(\log\varepsilon)/p \rceil$ suffices. □

Query complexity

At each trial we need one use of U_f, one query of $f(0)$ and one query of $f(\tilde{r})$. Hence 3. So if we repeat $O(\log\log N)$ times, we use $O(\log\log N)$ queries.

What about implementing QFT? We only need $O(\log N)^2$ computational steps.

Note we haven't included other computations, including the cancelling of c/N to the lowest term - this can be done using Euclid's algorithm. In fact, all other steps can be implement in $O(\text{poly}(\log N))$ steps.

Periodicity determination for periodic functions on $\mathbb{Z}_N$ can be extended to an arbitrary group G, known as the *hidden subgroup problem.*

9 Quantum algorithms for search problems

Many problems can be cast as search problems. For example, the problem of factoring N is equivalent to searching among all integers less than N for one that divides N exactly.

9.1 The unstructured search problem

- **Problem:** Given a large database of N items. The aim is to locate a *particular* "good" item.
- **Assumption:** The database is unstructured, but given any item it is easy to check whether it is the good one.
- **Requirement:** The algorithm should locate the good item with probability of success $1-\varepsilon$ for some fixed $\varepsilon \in (0,1)$ independent of N.
- Each access to the database is a *query*.

In the classical case, $O(N)$ queries are necessary and sufficient. But in the quantum case, $O(\sqrt{N})$ suffices. We then have a *quadratic* speedup. The algorithm for this is known as *Grover's algorithm* (Lov Grover, 1996).

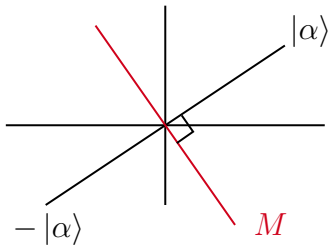
Definition. For any $|\alpha\rangle \in \mathcal{V} \equiv \mathcal{H}$ such that $\langle\alpha|\alpha\rangle = 1$, define

(1) $\Pi_{|\alpha\rangle} = |\alpha\rangle\langle\alpha|$ to be the rank-1 projection operator.

(2) $I_{|\alpha\rangle} = I - 2|\alpha\rangle\langle\alpha|$ to be the reflection operator. Note that

$$I_{|\alpha\rangle}|\alpha\rangle = |\alpha\rangle - 2|\alpha\rangle\langle\alpha|\alpha\rangle = -|\alpha\rangle$$

For example, if $\mathcal{V} \equiv \mathbb{R}^2$,



If $|\psi\rangle$ satisfies $\langle\alpha|\psi\rangle = 0$ then $I_{|\alpha\rangle}|\psi\rangle = |\psi\rangle$. For any $|\Psi\rangle \in \mathcal{V}$ we can write

$$|\Psi\rangle = a|\alpha\rangle + \sum_{i=1}^{d} b_i|\beta_i\rangle$$

where

$$d = \dim \mathcal{S}^{\perp}_{|\alpha\rangle},\ \mathcal{S}^{\perp}_{|\alpha\rangle} = \mathrm{span}\{|v\rangle \in \mathcal{V} :\ \langle v|\alpha\rangle = 0\}$$

and $\langle\beta_i|\alpha\rangle = 0$, $\langle\beta_i|\beta_j\rangle = \delta_{ij}$. We have

$$I_{|\alpha\rangle}|\Psi\rangle = -a|\alpha\rangle + \sum_{i=1}^{d} b_i|\beta_i\rangle$$

Hence $I_{|\alpha\rangle}$ only fips the sign of the amplitude of $|\alpha\rangle$.

Note. *For any unitary U,*

- $U\Pi_{|\alpha\rangle}U^\dagger = U\,|\alpha\rangle\,\langle\alpha|\,U^\dagger = \Pi_{U|\alpha\rangle}$.
- $UI_{|\alpha\rangle}U^\dagger = U(I - 2\,|\alpha\rangle\,\langle\alpha|)U^\dagger = I - 2U\,|\alpha\rangle\,\langle\alpha|\,U^\dagger = I_{U|\alpha\rangle}$.

If $\mathcal{V} \equiv \mathbb{C}^2$ or $\mathcal{V} \equiv \mathbb{R}^2$, for any $|v\rangle \in \mathcal{V}$ we can write $|v\rangle = a\,|\alpha\rangle + b\,|\alpha^\perp\rangle$ where $\langle\alpha|\alpha^\perp\rangle = 0$. Then

- $I_{|\alpha\rangle}\,|v\rangle = -a\,|\alpha\rangle + b\,|\alpha^\perp\rangle$.
- $I_{|\alpha^\perp\rangle}\,|v\rangle = a\,|\alpha\rangle - b\,|\alpha^\perp\rangle = -(-a\,|\alpha\rangle + b\,|\alpha^\perp\rangle) = -I_{|\alpha\rangle}\,|v\rangle$. Since this is true for any $|v\rangle \in \mathcal{V}$, we conclude that $I_{|\alpha^\perp\rangle} = -I_{|\alpha\rangle}$ if $\mathcal{V}$ is 2D.

9.2 Grover's algorithm

This algorithm needs $O(\sqrt{N})$ queries. By choosing $N = 2^n$, we can label the items by n-bit strings. The search problem then becomes a black-box problem where we replace the database by a black-box for $f : B_n \to B$ where $f(x_0) = 1$ for a particular x_0 (the good item) and $f(x) = 0$ otherwise.

———————————[Lecture 18 finish]———————————

Let U_f be the quantum oracle which acts by

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

where the input register $|x\rangle$ is a n-qubit state and the output register $|y\rangle$ is 1-qubit. In fact, instead of using U_f we use a related operator $I_{x_0} \equiv I_{|x_0\rangle}$ where

$$I_{x_0}\,|x\rangle = \begin{cases} -\,|x_0\rangle & x = x_0 \\ |x\rangle & x \neq x_0 \end{cases}$$

If $x_0 = 0\ldots0$ then we write $I_{x_0} = I_0$. But can we relate I_{x_0} to U_f? Note that if we set $|y\rangle = |-\rangle$, then

$$\begin{aligned} U_f\,|x\rangle\,|-\rangle &= \frac{1}{\sqrt{2}}U_f\big(\,|x\rangle\,|0\rangle - |x\rangle\,|1\rangle\,\big) \\ &= \frac{1}{\sqrt{2}}\big(\,|x\rangle\,|f(x)\rangle - |x\rangle\,|f(x)^C\rangle\,\big) \\ &= \begin{cases} \frac{1}{\sqrt{2}}\,|x_0\rangle\,\big(\,|1\rangle - |0\rangle\,\big) = -\,|x_0\rangle\,|-\rangle & \text{if } x = x_0 \\ |x\rangle\,|-\rangle & \text{if } x \neq x_0 \end{cases} \\ &= I_{x_0}\,|x\rangle\,|-\rangle \end{aligned}$$

Hence

$$U_f\,|x\rangle\,|-\rangle = I_{x_0}\,|x\rangle\,|-\rangle$$

We want to find $x_0 \in B_n$ with the least number queries (of U_f i.e. of I_{x_0}). We start with the equal superposition state

$$|\psi_0\rangle = H_n\,|0\rangle^{\otimes n} = |+\rangle^{\otimes n};\ H_n = H^{\otimes n}$$

and consider acting on it by the *Grover iteration operator*

$$Q = -H_n I_0 H_n I_{x_0}$$

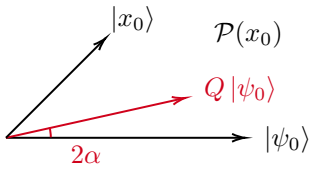
Note that

$$|\psi_0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in B_n} |x\rangle$$

and H, I_{x_0} all have real entries, so everything is real. To analyse the action of Q on $|\psi_0\rangle$, we use geometrical properties of I_{x_0} (as the reflection operator) described in terms of *real* Euclidean geometry.

Proposition 2. *In the plane $\mathcal{P}(x_0)$ spanned by $|x_0\rangle$ and $|\psi_0\rangle$, Q causes a rotation through an angle 2α where*

$$\sin\alpha = \frac{1}{\sqrt{2^n}} = \frac{1}{\sqrt{N}}$$



Proposition 3. *In the plane perpendicular to $\mathcal{P}(x_0)$, $Q \equiv -I$.*

But how do we use this? The answer is to repeatedly apply Q to $|\psi_0\rangle$ to get $|\psi'\rangle$, where $|\psi'\rangle \in (\mathbb{C}^2)^{\otimes n}$ is close to $|x_0\rangle$. Then we can measure in the computational basis of $(\mathbb{C}^2)^{\otimes n}$, which would give us $|x_0\rangle$ with a high probability.

- If N is large, $|\psi_0\rangle$ is almost orthogonal to $|x_0\rangle$. This can be seen from the expression for $|\psi_0\rangle$,

  $$\langle x_0|\psi_0\rangle = \frac{1}{\sqrt{2^n}} \xrightarrow{n\to\infty} 0$$

  Hence for the rotation through 2α,

  $$\sin\alpha = \frac{1}{\sqrt{N}} = \frac{1}{\sqrt{2^n}} \approx 0$$

  and $2\alpha \approx 2\sin\alpha \approx 2/\sqrt{N}$. So how many iterations of Q are needed to rotate $|\psi_0\rangle$ close to $|x_0\rangle$? Let β be the initial angle between $|\psi_0\rangle$ and $|x_0\rangle$, then for large N, $\beta \approx \pi/2$ and so the number of rotations m that is needed is satisfies

  $$m = \frac{\beta}{2\alpha} = \frac{\pi}{2}\frac{1}{2\alpha} = \frac{\pi}{2}\frac{\sqrt{N}}{2} = \frac{\pi}{4}\sqrt{N}$$

  Hence $O(\sqrt{N})$ suffices.

- For finite N, we have

$$|\psi_0\rangle = \frac{1}{\sqrt{N}} \sum_{x \in B_n} |x\rangle$$

and

$$\cos\beta = \langle x_0|\psi_0\rangle = \frac{1}{\sqrt{N}} \;\Rightarrow\; \beta = \cos^{-1}\left(\frac{1}{\sqrt{N}}\right)$$
$$\sin\alpha = \frac{1}{\sqrt{N}} \;\Rightarrow\; \alpha = \sin^{-1}\left(\frac{1}{\sqrt{N}}\right)$$

and hence

$$m = \frac{\beta}{2\alpha} = \frac{\cos^{-1}(1/\sqrt{N})}{2\sin^{-1}(1/\sqrt{N})}$$

which is independent of x_0. This is because we started with an equal superposition state.

Example (Search for one in four)**.** We have $N = 4$, so $n = 2$. Then $\sin\alpha = 1/2$ so $\alpha = \pi/6$. The state $|\psi_0\rangle$ is given by

$$|\psi_0\rangle = \frac{1}{2}\big[\,|00\rangle + |01\rangle + |10\rangle + |11\rangle\,\big]$$

and so for any $x_0 \in \{00, 01, 10, 11\}$ we have $\langle x_0|\psi_0\rangle = \cos\beta = 1/2$, so $\beta = \pi/3$. Then the number of iterations required is

$$m = \frac{\beta}{2\alpha} = \frac{\pi/3}{\pi/3} = 1$$

Hence for one-in-four we need only one query to locate the good item with probability 1. This follows since one use of Q needs one use of I_{x_0} i.e. one use of U_f.

Geomerical properties of Q

Combining Propositions 2 and 3, we claim that

$$Q = -H_n I_0 H_n I_{x_0} = -I_{|\psi_0\rangle} I_{x_0}$$

i.e. we have

$$H_n I_0 H_n = I_{|\psi_0\rangle}$$

The LHS is

$$\begin{aligned} H_n\big[I - 2\,|0\ldots0\rangle\,\langle 0\ldots0|\,\big]H_n &= I^{\otimes n} - 2(H\,|0\rangle)^{\otimes n}(\langle 0|\,H)^{\otimes n} \\ &= I^{\otimes n} - 2\big(\,|+\rangle\,\langle +|\,\big)^{\otimes n} \\ &= I^{\otimes n} - 2\,|\psi_0\rangle\,\langle\psi_0| = I_{|\psi_0\rangle} \end{aligned}$$

and the claim follows.

Let $|\psi_0^\perp\rangle$ be such that $\langle\psi_0|\psi_0^\perp\rangle = 0$, then we can also write

$$Q = I_{|\psi_0^\perp\rangle} I_{x_0}$$

To see this, note that if $|u\rangle = a\,|v\rangle + b\,|v^\perp\rangle$, then

$$\begin{aligned} I_{|v\rangle}\,|u\rangle &= -a\,|v\rangle + b\,|v^\perp\rangle \\ I_{|v^\perp\rangle}\,|u\rangle &= a\,|v\rangle - b\,|v^\perp\rangle \end{aligned}$$

and so $I_{|v^\perp\rangle} = -I_{|v\rangle}$.

Proof of Proposition 2. For any $|v\rangle \in \mathcal{P}(x_0)$, $Q\,|v\rangle$ is also in $\mathcal{P}(x_0)$ and so Q preserves $\mathcal{P}(x_0)$. Moreover, for all $|v\rangle \in \mathcal{P}(x_0)$,

$$\begin{aligned} I_{x_0}\,|v\rangle &= \left(I - 2\,|x_0\rangle\,\langle x_0|\,\right)|v\rangle = |v\rangle - 2\,\langle x_0|v\rangle\,|x_0\rangle \\ I_{|\psi_0\rangle}\,|v\rangle &= \left(I - 2\,|\psi_0\rangle\,\langle\psi_0|\,\right)|v\rangle = |v\rangle - 2\,\langle\psi_0|v\rangle\,|\psi_0\rangle \end{aligned}$$

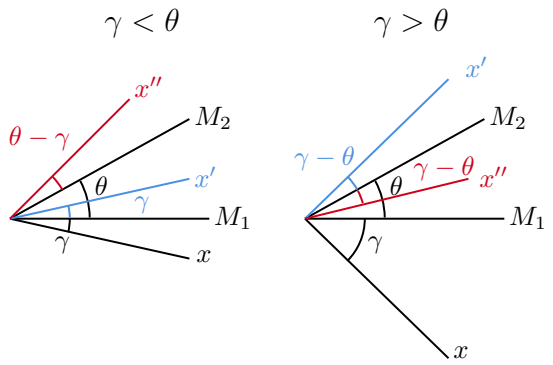
We'll show that Q causes a rotation with angle 2α. □

[Lecture 19 finish]

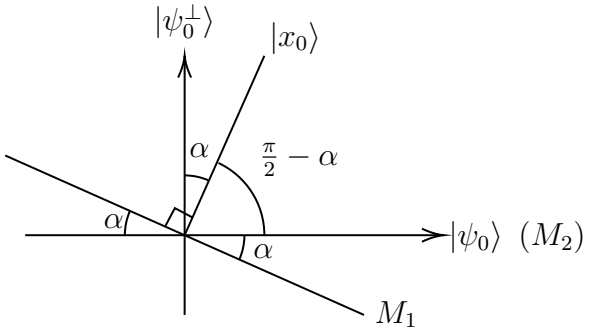
Proof. (Continued) α is the angle between two mirror lines M_1 (orthogonal to $|x_0\rangle$) and M_2 (orthogonal to $|\psi_0^\perp\rangle$).

Lemma. *Let M_1, M_2 be two mirror lines in $\mathbb{R}^2$, intersecting at the origin with angle θ. Then a reflection about M_1 followed by a reflection about M_2 is equivalent to an anticlockwise rotation with angle 2θ.*

To see this, two drawings suffice:



We have $Q = I_{|\psi_0^\perp\rangle} I_{x_0}$, then



Hence it follows that Q rotates through angle 2α, where

$$\langle x_0|\psi_0\rangle = \cos\beta = \cos(\pi/2-\alpha) = \sin\alpha$$

But also LHS is $1/\sqrt{N} = 1/\sqrt{2^n}$, and so the lemma tells us that $\sin\alpha = 1/\sqrt{N}$. □

Proof of Proposition 3. We use $Q = -I_{|\psi_0\rangle} I_{x_0}$. Suppose $|v\rangle \in \mathcal{P}^\perp(x_0)$ (plane orthogonal to x_0), then $I_{x_0}|v\rangle = |v\rangle$ by definition of I, and $I_{|\psi_0\rangle}|v\rangle = |v\rangle$ which follows as ψ_0 is on the plane. Hence $Q|v\rangle = -|v\rangle$ and so $Q = -I$. □

9.3 Further details of Grover's algorithm

9.3.1 Optimality

Let T be the number of iterations needed for large but finite N, and δ_T the angle between $|x_0\rangle$ and the rotated vector $|\psi_0^T\rangle$. Then from the previous figure we have

$$\begin{aligned}\delta_T &= \frac{\pi}{2} - \alpha - 2\alpha T \\ &= \frac{\pi}{2} - \alpha(1+2T) \\ &= \frac{\pi}{2} - (1+2T)\sin^{-1}\left(\frac{1}{\sqrt{N}}\right)\end{aligned}$$

If we do a measurement on $|\psi_0^T\rangle$ in the computational basis of $(\mathbb{C}^2)^{\otimes n}$, then

$$\begin{aligned}p(x_0,T) \equiv \text{prob(outcome is } x_0;T) &= |\langle x_0|\psi_0^T\rangle|^2 \\ &= \cos^2\delta_T \\ &= \cos^2\left(\frac{\pi}{2} - (1+2T)\sin^{-1}\left(\frac{1}{\sqrt{N}}\right)\right) \\ &= \sin^2\left((1+2T)\sin^{-1}\left(\frac{1}{\sqrt{N}}\right)\right)\end{aligned}$$

To maximise this, we choose T to be the closest integer for which

$$(1+2T)\sin^{-1}\left(\frac{1}{\sqrt{N}}\right) \approx \frac{\pi}{2}$$

Using the approximation $\sin^{-1}x = x + O(x^3)$ for small x,

$$T = \frac{\pi}{4\sin^{-1}(1/\sqrt{N})} - \frac{1}{2} = \frac{\pi}{4}\sqrt{N} - \frac{1}{2} - O\left(\frac{1}{N}\right)$$

and so $T = O(\sqrt{N})$. Note if we put $N = 4$ as in the one-in-four example, we get $T = \pi/(4\sin^{-1}(1/2)) - 1/2 = \pi/(4\pi/6) - 1/2 = 3/2 - 1/2 = 1$. So $T = 1$ suffices with $p(x_0, 1) = 1$.

9.3.2 Multiple good items

Suppose there are $r \geqslant 1$ good items, the task is to find one or all of the good item. Wlog we can label the good items as $x_1, \ldots, x_r$ followed by the "bad" items, and our black-box now becomes

$$f : B_n \to B,\ f(x_i) = \begin{cases} 1 & i \in \{1, \ldots, r\} \\ 0 & \text{otherwise} \end{cases}$$

Analogously to the one good item case where $I_{x_0} = I - 2\left|x_0\right\rangle\left\langle x_0\right|$, we define

$$I_G = I - 2\sum_{i=1}^{r} \left|x_i\right\rangle\left\langle x_i\right|$$

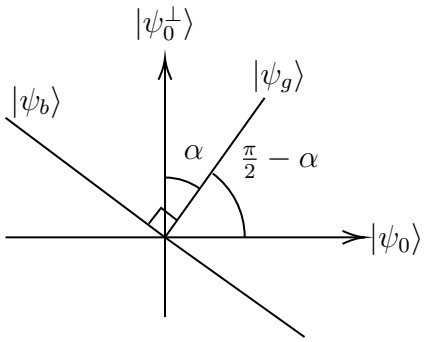
$$I_G \left|x\right\rangle = \begin{cases} -\left|x\right\rangle & x \in \{x_1, \ldots, x_r\} \\ \left|x\right\rangle & \text{otherwise} \end{cases}$$

Then $Q = -H_n I_0 H_n I_G$. We may define

$$\left|\psi_g\right\rangle = \frac{1}{\sqrt{r}} \sum_{i=1}^{r} \left|x_i\right\rangle,\ \left|\psi_b\right\rangle = \frac{1}{\sqrt{N-r}} \sum_{i=r+1}^{N} \left|x_i\right\rangle$$

where $\left|\psi_g\right\rangle$ and $\left|\psi_b\right\rangle$ are good and bad items respectively. Then

$$\left|\psi_0\right\rangle = \frac{1}{\sqrt{N}} \sum_{x \in B_n} \left|x\right\rangle = \frac{\sqrt{r}}{\sqrt{N}} \left|\psi_g\right\rangle + \frac{\sqrt{N-r}}{\sqrt{N}} \left|\psi_b\right\rangle$$



Then $Q = I_{\left|\psi_0^\perp\right\rangle} I_{\left|\psi_g\right\rangle}$, and

$$\cos\alpha = \left\langle\psi_0^\perp|\psi_g\right\rangle,\ \sin\alpha = \left\langle\psi_0|\psi_g\right\rangle = \sqrt{r/N}$$

10 Shor's factoring algorithm

Let N be a given positive integer. The output of the algorithm is

(i) a factor K of N, or

(ii) N if N is prime.

Claim. The algorithm runs in $O(n^3)$ time where n is the number of binary digits expressing N.

Classically the best known algorithm is the *number field sieve algorithm*, with run time

$$\exp\left[O\left(n^{1/3}(\log n)^{2/3}\right)\right]$$

The key idea here is to convert factoring N to preiodicity determination, where we use QFT but some modifications are needed.

10.1 Factoring as a periodicity determination problem

The steps are as follows:

(1) Choose an integer a where $1 < a < N$ uniformly at random.

(2) Use Euclid's algorithm (a poly-time algorithm) to compute $b = \gcd(a, N)$. Then,

  (i) If $b > 1$, then b divides N and the output is b (a factor of N).

  (ii) If $b = 1$ then a, N are coprime and we seek help from Number Theory.

Theorem (Euler's theorem)**.** *If a, N are coprime, then there is a least integer r where $1 < r < N$ such that*

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

Consider the *modular exponential function*

$$\begin{aligned} f : \mathbb{Z} &\to \mathbb{Z}_N \\ k &\mapsto a^k \pmod{N} \end{aligned}$$

Then it immediately follows that $f(k_1 + k_2) = f(k_1)f(k_2)$. From Euler's theorem, we know that there exists r such that

$$f(r) = a^r \equiv 1 \pmod{N}$$

Then for any k,

$$f(k + r) = f(k)f(r) = f(k)$$

i.e. f is periodic with period r. Since r is the least integer satisfying the relation, f must be a one-to-one function in each period.

[Lecture 20 finish]

Suppose we can compute r using QFT, then if r is even,

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

i.e. N divides $a^r - 1$ exactly. But N cannot divide $a^{r/2} - 1$ because r is the *minimal* integer such that $a^r - 1 \equiv 0 \pmod{N}$. Then, if N does not divide $a^{r/2} + 1$, then N partially divides $a^{r/2} \pm 1$. We can use Euclid's algorithm to find $\gcd(a^{r/2} \pm 1, N)$ and get factors of N.

Therefore, we can find factors of N if a, N are coprime provided that

(i) r is even, and

(ii) $a^{r/2} + 1 \neq 0 \pmod{N}$.

Theorem. *If N is odd and $N \neq p^l$ where p is prime and $l \in \mathbb{Z}_+$, then if a where $1 < a < N$ is chosen uniformly at random and a, N are coprime, then the probability that the above two conditions hold is at least $1/2$.*

Hence we can factor with probability at least $1/2$. Given a candidate factor we *check* if in $\text{poly}(n)$ time via test division into N. Repeating the process k times (e.g. $k = 10$) then

$$p_{\text{err}}^{(10)} = (1 - p_{\text{succ}})^{10} \leqslant \left(1 - \frac{1}{2}\right)^{10} = \frac{1}{2^{10}} = \varepsilon$$

We see that if $k = \log(1/\varepsilon)$ then $p_{\text{succ}} \geqslant 1 - \varepsilon$.

Example ($N = 15$)**.** Suppose we choose $a = 7$, then a, N are coprime. Hence $f(k) = a^k \pmod{N} = 7^k \pmod{15}$, with values

$$1, 7, 4, 13, 1, 7, 4, 13, \ldots$$

Hence the period is 4, which is even. Then $a^r - 1 = (7^4 - 1) = (7^2 - 1)(7^2 + 1) = 48 \times 50$, so N does not divide either factor. Then

$$\gcd(48, 15) = 3, \ \gcd(50, 15) = 5$$

Summary.

(i) Is N even? If yes, output 2. If no, go to (ii).

(ii) Is $N = p^l$ for some prime p and $l \in \mathbb{Z}_+$? If yes, compute roots of $N^{1/k}$ where $k = 2, 3, \ldots, \log N$, which can be done in $\text{poly}(n)$ time. If any of these roots is an integer, we get a factor as an output. Otherwise, go to (iii).

(iii) Choose $1 < a < N$ uniformly at random. Compute $b = \gcd(a, N)$, then

- If $b > 1$, then output b.
- If $b = 1$, then a, N are coprime and go to (iv).

(iv) Using the quantum algorithm, find the period r of the function

$$f : \mathbb{Z} \to \mathbb{Z}_N, \ f(k) = a^k \pmod{N}$$

- If r is odd, return to (iii) and repeat.
- If r is even,
  - Compute $a^{r/2} - 1 \pmod{N}$ (or $a^{r/2} + 1 \pmod{N}$), and
  - Compute $t = \gcd(N, a^{r/2} - 1)$.
- If $t > 1$, output t.
- If $t = 1$, then the algorithm has failed and return to (iii).

10.2 Computing period of modular exponential function

Note that $f : \mathbb{Z} \to \mathbb{Z}_N$. To use the quantum period finding algorithm we'll need to restrict the domain of f, say to $\mathbb{Z}_M$ for some $M \in \mathbb{Z}$. However, the restricted function is *not* in general periodic. But...

Claim. For sufficiently large $M(= O(N^2))$ there are enough periods so that the incomplete one has negligible effect on the period finding algorithm.

If we choose $M = 2^m$ where m is the smallest integer such that $2^m > N^2$, then $M = 2^m = Br+b$ where B is the number of complete periods and b is the length of the incomplete one. (Compare this with $N = Ar$ in §8.2.)

Steps for finding r

(i) Construct, as always, a uniform superposition state

$$|\psi_m\rangle = \frac{1}{\sqrt{2^m}} \sum_{x \in \mathbb{Z}_M} |x\rangle$$

noting that it is a m-qubit state.

(ii) Let the output register be $|0\rangle \in \mathcal{B}_N = \{|0\rangle, |1\rangle, \ldots, |N-1\rangle\}$ and act on them with U_f:

$$U_f |\psi_m\rangle |0\rangle = \frac{1}{\sqrt{2^m}} \sum_{x \in \mathbb{Z}_M} |x\rangle |f(x)\rangle$$

(iii) Measure on the second register in the basis $\mathcal{B}_N$. Suppose the output is $y = f(x_0)$ (uniformly at random) where $0 \leqslant x_0 \leqslant r-1$. Because r is the period, $f(x_0) = f(x_0 + jr)$ where

$$j = \begin{cases} 1, 2, \ldots, B-1 & x_0 > b \\ 1, 2, \ldots, B & x_0 \leqslant b \end{cases}$$

In other words, $j = 0, 1, \ldots, A-1$ where

$$A = \begin{cases} B & x_0 > b \\ B+1 & x_0 \leqslant b \end{cases}$$

The probability of getting $y = f(x_0)$ is

$$p(y) = \left\| \frac{1}{\sqrt{2^m}} \sum_{y=0}^{A-1} |x_0 + yr\rangle \right\|^2$$

The post-measurement state of the first register becomes

$$|\text{per}\rangle = \frac{1}{\sqrt{A}} \sum_{j=0}^{A-1} |x_0 + jr\rangle, \; A = \begin{cases} B+1 = \lfloor 2^m/r \rfloor + 1 & x_0 \leqslant b \\ B = \lfloor 2^m/r \rfloor & x_0 > b \end{cases}$$

(iv) Apply QFT_{2^m},

$$\begin{aligned}\mathrm{QFT}_{2^m}|\mathrm{per}\rangle &= \frac{1}{\sqrt{A}}\frac{1}{\sqrt{2^m}}\sum_{c=0}^{2^m-1}\sum_{j=0}^{A-1}\omega^{(x_0+jr)c}|c\rangle \\ &= \sum_{c=0}^{2^m-1} g(c)\,|c\rangle\end{aligned}$$

where $\omega = e^{2\pi i/2^m}$ and

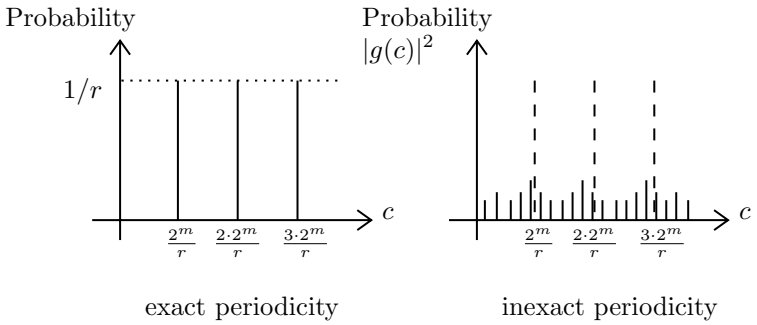
$$g(c) = \frac{1}{\sqrt{A}}\frac{1}{\sqrt{2^m}}\sum_{j=0}^{A-1}\omega^{x_0 c}\omega^{jrc} = \frac{\omega^{x_0 c}}{\sqrt{A}\sqrt{2^m}}\left[\sum_{j=0}^{A-1}(\omega^{rc})^j\right]$$

$$\text{where } \left[\sum_{j=0}^{A-1}(\omega^{rc})^j\right] = \sum_{j=0}^{A-1}\alpha^j = \begin{cases} A & \alpha = 1 \\ \frac{1-\alpha^A}{1-\alpha} = 0 & \text{else}\end{cases}$$

[Lecture 21 finish]

Recall when the periodicity is exact i.e. in §8.2, we had equal discrete probability. Now we have inexact periodicity where $2^m = Br + b$, and the sum above is non-zero only if $\alpha \neq 1$. Hence we get non-zero probability for getting c values even when $\alpha = \omega^{rc} \neq 1$.

Claim. A measurement on $\mathrm{QFT}_{2^m}|\mathrm{per}\rangle$ yields a value of c which is close to a multiple of $2^m/r$ with high probability.



exact periodicity inexact periodicity

Consider $k \cdot 2^m/r$ where $k = 0, 1, \ldots, r-1$. Note that this is no longer an integer.

Claim. Each $k \cdot 2^m/r$ is within a distance of $1/2$ from a *unique nearest integer*.

Proof. Suppose $k \cdot 2^m/r = s + 1/2$ for some $s \in \mathbb{Z}$ i.e. it is in the middle of s and $s+1$, so

$$k \cdot \frac{2^m}{r} = \frac{2s+1}{2} \qquad (\dagger)$$

for some $s \in \mathbb{Z}$. We know that $r < N$ and $2^m > N^2$, so $r < 2^m$. Hence any factor of 2 in r gets cancelled by the factors of 2 in 2^m, and there cannot be any remnant 2 in the denominator. But $(\dagger)$ contradicts this, so we can't have anything in the middle, and the nearest integer would be unique. □

We consider integer values of c (there are r of them) such that

$$\left| c - k \cdot \frac{2^m}{r} \right| < \frac{1}{2}$$

We chose $2^m > N^2$ exactly to get the above relation, to get a unique integer within $1/2$ of $k \cdot 2^m/r$ for each $k \in \{0, 1, \ldots, r-1\}$.

Theorem. *If we do a measurement on*

$$\mathrm{QFT}_{2^m} |\mathrm{per}\rangle = \sum_{c=0}^{2^m} g(c) |c\rangle$$

in the basis $\mathcal{B}_M = \{|0\rangle, |1\rangle, \ldots, |M-1\rangle\}$ *where* $M = 2^m$*, and for all* $k = 0, \ldots, r-1$ *let* $c \equiv c_k$ *be the unique integer such that*

$$\left| c - k \cdot \frac{2^m}{r} \right| < \frac{1}{2}$$

Then

$$p(\text{outcome} = c_k) > \frac{\gamma}{r}, \ \gamma \approx \frac{4}{\pi^2}$$

The proof is non-examinable. We will be interested in those c_k such that k is coprime to r. By the coprimality theorem, the probability of obtaining such a *good* value of c is

$$\Omega\left(1/\log\log r\right) > \Omega\left(1/\log\log N\right)$$

Therefore by $O(\log\log N)$ repetitions we can obtain a good value of c for a desired level of probability. Upon this, we have obtained a value of c such that

$$\left| c - k \cdot \frac{2^m}{r} \right| < \frac{1}{2} \text{ and } k, r \text{ are coprime}$$

The task now is then to get r from such a value of c.

10.3 Getting r from a good value of c

We have

$$\left| \frac{c}{2^m} - \frac{k}{r} \right| < \frac{1}{2^{m+1}}$$

Note that $r < N$ and $2^m > N^2$, and so $2^{m+1} > 2N^2$, hence

$$\left| \frac{c}{2^m} - \frac{k}{r} \right| < \frac{1}{2N^2}, \ r < N \qquad (*)$$

Note that $c/2^m$ is a known fraction.

Claim. There is *at most one* fraction k/r with denominator $r < N$ satisfying $(*)$.

Proof by contradiction. Suppose k'/r' and k''/r'' both satisfy $(*)$ and they are not equal. Then

$$\left|\frac{k'}{r'} - \frac{k''}{r''}\right| = \left|\frac{k'r'' - k''r'}{r'r''}\right| \geqslant \frac{1}{r'r''} > \frac{1}{N^2}$$

where the first inequality follows from our assumption that the fractions are not equal, and that k', r', etc. are all integers, and the second follows from the relation $r', r'' < N$. But by assumption, k'/r' and k''/r'' are both within a distance of $1/2N^2$ of $c/2^m$ by $(*)$, so the distance between them is less than $1/N^2$, and we have a contradiction. □

Hence there is one unique fraction k/r with $r < N$. This uniqueness was the reason we chose $2^m > N^2$, which guarantees that k/r is uniquely determined by $c/2^m$.

Example. Consider $N = 39$, and suppose $a = 7$ (coprime to N). Let r be the period of $f(k) = a^k \pmod N$. Then

$$N^2 = 1521,\ 2^{10} < N^2 < 2^{11}$$

So we choose $m = 11$, and suppose the measurement on $\mathrm{QFT}_{2^m} |\mathrm{per}\rangle$ yields $c = 853$, then our theory tells us that with high probability,

$$\left|c - k \cdot \frac{2^m}{r}\right| < \frac{1}{2}$$

for some $k \in \{0, \ldots, r-1\}$. If this is actually the case, then k/r is the unique fraction with denominator $r < N$ such that

$$\left|\frac{c}{2^m} - \frac{k}{r}\right| < \frac{1}{2^{m+1}}$$

We can use a calculator to check all fractions a/b where $1 \leqslant a < b < N = 39$ to see which one(s) satisfy the

$$\left|\frac{c}{2^m} - \frac{a}{b}\right| = \left|\frac{853}{2048} - \frac{a}{b}\right| < \frac{1}{2^{m+1}}$$

There are $O(N^2)$ of such fractions, but we find that only one fraction $a/b = 5/12$ satisfy it. This is consistent with $(k, r) = (5, 12)$ or $(10, 24)$ or $(15, 36)$ (we don't go further to make sure they're in $\mathbb{Z}_{39}$). For k, r to be coprime, we choose the first pair. We can check that for $f(k) = 7^k \pmod{39}$, we have $f(12) = 1$.

[Lecture 22 finish]

Theorem (Theory of continued fractions)**.** *Any rational number s/t $(s < t)$ can be expressed as a continued fraction*

$$\frac{s}{t} = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\cdots + \frac{1}{a_l}}}}$$

where $a_1, \ldots, a_l$ are positive integers.

Proof. To do this, we frist write

$$\frac{s}{t} = \frac{1}{t/s};\ \frac{t}{s} = a_1 + \frac{s_1}{t_1},\ a_1 \geqslant 1,\ s_1 < t_1 = s$$

We can then repeat this for s_1/t_1 to get $t_1/s_1 = a_2 + s_2/t_2$, where $s_2 < t_2 = s_1$. Then

$$\frac{s}{t} = \frac{1}{a_1 + \frac{1}{a_2 + \frac{s_2}{t_2}}}$$

and we continue to get a sequence of positive integers a_k, s_k, t_k such that $s_k < t_k = s_{k-1}$. Observe that t_k is strictly decreasing and non-negative, so the process must terminate after, say, l steps. We'll write the continued fraction of s/t to be $[a_1, \ldots, a_l]$. □

Example. Take $s/t = 31/64$, then

$$\frac{31}{64} = \frac{1}{64/31} = \frac{1}{2 + 2/31} = \frac{1}{2 + \frac{1}{15+1/2}}$$

and so the continued fraction of s/t is $[2, 15, 2]$.

Note. *For each $k = 1, 2, \ldots, l$ we can truncate the continued fraction at the k^{th} level to get a sequence of rational numebrs.*

$$\frac{p_1}{q_1} = [a_1] = \frac{1}{a_1};\ \frac{p_2}{q_2} = [a_1, a_2] = \frac{1}{a_1 + 1/a_2} = \frac{a_2}{a_1 a_2 + 1};\ \ldots$$

Here p_k/q_k is called the k^{th} convergent continued fraction of s/t.

Here we'll state without proof some properties of continued fraction:

Lemma 4. *Let $a_1, \ldots, a_k$ be any positive numbers (not necessarily integers). Set $p_0 = 0, q_0 = 1$ and $p_1 = 1, q_1 = a_1$, then*

(a) $[a_1, \ldots, a_k] = p_k/q_k$ *where* $p_k = a_k p_{k-1} + p_{q-2},\ q_k = a_k q_{k-1} + q_{k-2},\ k \geqslant 2$. *If a_k's are integers then so are p_k and q_k.*

(b) $a_k p_{k-1} - p_k q_{k-1} = (-1)^k$ *for* $k \geqslant 1$.

(c) $\gcd(p_k, q_k) = 1$ *for all* $k \geqslant 1$.

Theorem 5. *Suppose the continued fraction of s/t is $[a_1, \ldots, a_l]$, and let $p_k/q_k = [a_1, \ldots, a_k]$ be the k^{th} convergent with $k = 1, \ldots, l$. If s, t are m-bit length integers (after s/t is cancelled down to the lowest term), then the length l of the continued fraction is $O(m)$ and the continued fraction and its convergents can be calculated in $O(m^3)$ time.*

Theorem 6. *Let $0 < x < 1$ be a rational number and suppose p/q is a rational number such that*

$$\left| x - \frac{p}{q} \right| < \frac{1}{2q^2} \qquad (\ddagger)$$

Then p, q is a convergent of the continued fraction of x.

Now we go back to our problem, and recall the relation

$$\left|\frac{c}{2^m} - \frac{k}{r}\right| < \frac{1}{2^{m+1}} \underset{\text{(i)}}{\leqslant} \frac{1}{2N^2} \underset{\text{(ii)}}{\leqslant} \frac{1}{2r^2}$$

where in (i) we used $2^m > N^2$ and in (ii) $r < N$. Hence we obtain an inequality of the form (‡), with $x = c/2^m$ where $0 \leqslant c \leqslant 2^m - 1$ and $p/q = k/r$. By Theorem 6, k/r is thus a convergent of the continued fraction of $c/2^m$.

Now c and 2^m are $O(m)$ bit integers, and $2^m = O(N^2)$ where N is n-bit in length, hence c and 2^m are $O(n)$ bit integers. Therefore by Theorem 5, all convergents of $c/2^m$ can be computed in $O(n^3)$ time, with $O(n)$ such convergents.

We saw there is a *unique* fraction such that the above inequalities hold, so we can compute all convergents of $c/2^m$ and check the list of $O(n)$ convergents to find the unique one that satisfies the above relation.

Example (Lecture 22 revisited)**.** Again with $N = 39$ and $a = 7$, we want to find the period r of $f(k) = 7^k \pmod{39}$. Here $m = 11$. Suppose $c = 853$ then

$$\frac{c}{2^m} = \frac{853}{2048} = [2, 2, 2, 42, 4]$$

The convergents are

$$[2] = \frac{1}{2},\ [2, 2] = \frac{2}{5},\ [2, 2, 2] = \frac{5}{12}, \ldots$$

So we only have to find 5 fractions, and in this case only 5/12 works, so $r = 12$.

10.4 Complexity of Shor's algorithm*

Let N be of n-bit length. Let's go through the algorithm again:

(i) The uniform superposition state

$$|\psi_m\rangle = H^{\otimes m} |0\rangle^{\otimes m}$$

where $m = O(n)$, takes $O(n)$ Hadamard gates.

(ii) To compute

$$|f\rangle = U_f |\psi_m\rangle |0\rangle$$

we'll need $O(n^3)$ steps (by repeated squareing, cf. Question 2 of Example Sheet 3) to implement $f(k)$.

(iii) $O(n)$ single qubit measurements on the second (n-qubit) register take .

(iv) QFT on $|\text{per}\rangle$ requires $O(n^2)$ steps.

(v) $O(n)$ single qubit measurements on the first register to get c.

To arrive at c, we've totalled $O(n^3)$ steps. To get a good value of c, we repeat the previous process $O(\log\log N) = O(\log n)$ times. To get r from a good c, we use the continue fraction theory which requires $O(n^3)$ steps, as we discussed above.

Now consider §10.1, if r is even and $t = \gcd(a^{r/2-1}, N) \neq 1$, then the algorithm terminates and we get our desired output. But if r is odd, or if r is even but $t = 1$, then the algorithm has failed and we need to restart.

However, the good case we want occurs with any fixed constant level of probability $1 - \varepsilon$ after $O(\log 1/\varepsilon)$ repetitions which does not depend on n. So the total complexity of Shor's algorithm is $O(n^3)$[1].

[Lecture 23 finish]

Efficient implementation of QFT_N for $N = 2^n$

Claim. The size of the quantum circuit is

$$O(n^2) = O((\log N)^2) \equiv O(\mathrm{poly}(\log N))$$

Remark. For $N \neq 2^n$, we approximate QFT_N by QFT_{2^k} with 2^k close to N. Then the probability of success reduces by just a small amount.

The key idea is to note that

$$\mathrm{QFT}_N |x\rangle = \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} \omega^{xy} |y\rangle$$

is actually a tensor product of single qubit states. Since $N = 2^n$ we can represent $x, y \in \mathbb{Z}_N$ by n-bit strings

$$x = (x_0, \ldots, x_{n-1}),\ y = (y_0, \ldots, y_{n-1})$$

such that

$$\begin{aligned} x &= x_0 + 2x_1 + 2^2 x_2 + \cdots + 2^{n-1} x_{n-1} \\ y &= y_0 + 2y_1 + 2^2 y_2 + \cdots + 2^{n-1} y_{n-1} \end{aligned}$$

Then the phase is equal to

$$\omega^{xy} = \exp\left[2\pi i \cdot \frac{xy}{2^n}\right]$$

In $xy/2^n$ we can discard all the terms which are whole numbers, since if $xy/2^n = m + \alpha$ where $m \in \mathbb{Z}$ and $\alpha < 1$, then

$$\omega^{xy} = e^{2\pi i(m+\alpha)} = e^{2\pi i\alpha}$$

Retaining only the fractional parts, we can write

$$\frac{xy}{2^n} = y_{n-1}(.x_0) + y_{n-2}(.x_1 x_0) + \cdots + y_0(.x_{n-1} x_{n-2} \ldots x_0)$$

[1] Actually, with a good algorithm it is $O(n^2 (\log n)^2 \log\log n)$ instead.

where $(.x_0)$ etc. denote binary expansions:

$$.x_0 = \frac{x_0}{2},\ .x_1x_0 = \frac{x_1}{2} + \frac{x_0}{2^2}$$

To see this, let's write out

$$\frac{xy}{2^n} = \frac{1}{2^n}(x_0 + 2x_1 + 2^2x_2 + \cdots + 2^{n-1}x_{n-1})(y_0 + 2y_1 + 2^2y_2 + \cdots + 2^{n-1}y_{n-1})$$

If we consider the terms involving y_{n-1}, we'll see that the only remaining (fractional) term is obtained from x_0 term in the expansion of x, leading to $x_0/2$. While for y_0, all terms from x can be kept. Then we can write

$$\begin{aligned}
\mathrm{QFT}\,|x\rangle &= \sum_{y_0,\ldots,y_{n-1}\in\{0,1\}} e^{2\pi i xy/2^n}\,|y_{n-1}\rangle\,|y_{n-2}\rangle\ldots|y_0\rangle \\
&= \Big(\sum_{y_{n-1}} \exp\left[2\pi i y_{n-1}(.x_0)\right]|y_{n-1}\rangle\Big)\ldots\Big(\sum_{y_0}\exp\left[2\pi i y_0(.x_{n-1}\ldots x_0)\right]|y_0\rangle\Big) \\
&= \frac{1}{\sqrt{2}}\Big(|0\rangle + \exp\left[2\pi i(.x_0)\right]|1\rangle\Big)\cdots\frac{1}{\sqrt{2}}\Big(|0\rangle + \exp\left[2\pi i(.x_{n-1}\ldots x_0)\,|1\rangle\,\right]\Big)
\end{aligned}$$

Hence it follows that $\mathrm{QFT}_N\,|x\rangle$ is a tensor product of n single qubit states.

The quantum circuit

Recall the Hadamard gate and the Phase gate

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix},\ R_d = \begin{pmatrix}1 & 0\\ 0 & e^{i\pi/2^d}\end{pmatrix}$$

For $x \in \{0,1\}$, it can be easily checked that we have

$$H\,|x\rangle = \frac{1}{\sqrt{2}}\big[\,|0\rangle + e^{2\pi i(.x)}\,|1\rangle\,\big]$$

As for the Phase gate, note that

$$R_d = \begin{pmatrix}1 & 0\\ 0 & \exp\left[i\pi\cdot\frac{1}{2^d}\right]\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & \exp\left[2\pi i(.0\ldots01)\right]\end{pmatrix}$$

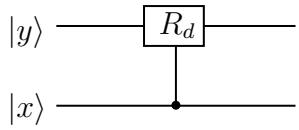
where the binary expansion has a 1 at the $(d+1)^{\text{st}}$ location. Hence it follows that, in the computational basis,

$$R_d\,|0\rangle = |0\rangle\ \ \forall d \geqslant 0 \text{ and } R_d\,|1\rangle = e^{2\pi i(.0\ldots01)}\,|1\rangle$$

In the case $d = 1$,

$$R_1\,|1\rangle = e^{2\pi i(.01)}\,|1\rangle = \exp\left[2\pi i\left(\frac{0}{2} + \frac{1}{2^2}\right)\right]|1\rangle = e^{\pi i/2}\,|1\rangle$$

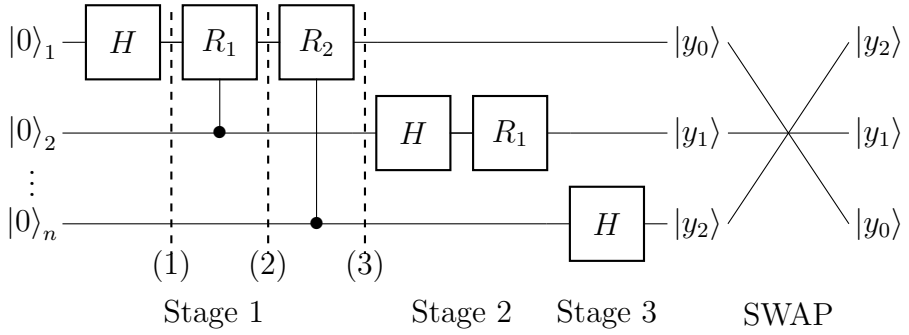
Similarly we can find $R_2\,|1\rangle$. We can use the 2-qubit controlled phase gates



where

$$C - R_d \left|0\right\rangle \left|\psi\right\rangle = \left|0\right\rangle \left|\psi\right\rangle, \; C - R_d \left|1\right\rangle \left|\psi\right\rangle = \left|1\right\rangle R_d \left|\psi\right\rangle$$

Then for $N = 8$, $n = 3$, the quantum circuit for QFT_N would be



We claim that

$$|y_0\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle + e^{2\pi i(.x_2 x_1 x_0)} |1\rangle\right)$$

To see this, consider the state after each gate:

- At (1), we have $\frac{1}{\sqrt{2}}\big[\, |0\rangle + e^{2\pi i(.x_2)} |1\rangle \,\big]$.
- At (2), $\frac{1}{\sqrt{2}}\big[\, |0\rangle + e^{2\pi i(.x_2)} e^{2\pi i(.0x_1)} |1\rangle \,\big]$.
- At (3), $\frac{1}{\sqrt{2}}\big[\, |0\rangle + e^{2\pi i(.x_2)} e^{2\pi i(.0x_1)} e^{2\pi i(.00x_0)} |1\rangle \,\big]$.

For $N = 8$, $n = 3$, we have $n = 3$ Hadamard gates and $2 + 1 = 3$ controlled phase gates. So for $N = 2^n$, n Hadamard gates, $n + (n-1) + \ldots 1 = n(n-1)/2$ controlled phase gates and n 2-qubit SWAP gates. In total, $O(n^2)$ gates.

10.5 Is NP equal to P?

The complexity class **NP** contain non-deterministic computations that are "hard to solve" i.e. there doesn't exist a poly-time algorithm, but are "easy to verify" in poly-time. A language L is in **NP** if it has a poly-time verifier V. A *verifier* V is a computation with inputs w and c such that

(a) If $w \in L$ then for some c, $V(w, c)$ halts with "accept".

(b) If $w \notin L$ then for any c, $V(w, c)$ halts with "reject".

and note that the size of w is poly(n).

We now consider the *satisfiability problem* (SAT): Given $\phi : B_n \to B$, is there an assignment $x_1 = b_1, \ldots, x_n = b_n$ such that $\phi(x_1 \ldots x_n) = 1$? Any such $(b_1 \ldots b_n)$ is called a satisfying assignment. To check this, given a candidate assignment we put it in ϕ to see if the result if one. SAT illustrates fundamental connection between **NP** and search problems. SAT is in **NP**, but we don't know if it is in **P**.

[Lecture 24 finish]