Introduction to Quantum Computing

Script accompanying IN2381 at TUM (WiSe 2020/2021)

Christian B. Mendl

November 6, 2020

Contents

1 Introduction

(see corresponding slides)

2 Introduction to quantum mechanics

2.1 Quantum bits (qubits)

(Nielsen and Chuang 2010, section 1.2)

Classical bit: 0, 1

Quantum bit "qubit": superposition of 0 and 1: a quantum state $|\psi\rangle$ is described as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad \alpha, \beta \in \mathbb{C} \tag{2.1}$$

with $|\alpha|^2 + |\beta|^2 = 1$ (normalization)

Quantum mechanics uses the so-called *ket* notation $|\,\rangle$ (motivation for that will become clear later)

Mathematical description: $|\psi\rangle \in \mathbb{C}^2$ (two-dimensional complex vector), identify $|0\rangle$ and $|1\rangle$ with basis states:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \rightsquigarrow \quad |\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{2.2}$$

Different from classical bits, cannot (in general) directly observe/measure a qubit (the amplitudes α and β). Instead: measurement (with respect to computational basis) will result in

$$\begin{aligned} 0 &\quad \text{with probability } |\alpha|^2, \\ 1 &\quad \text{with probability } |\beta|^2. \end{aligned} \tag{2.3}$$

The measurement also *changes* the qubit ("wavefunction collapse"): If measuring 0, the qubit will be $|\psi\rangle = |0\rangle$ directly after the measurement, and likewise $|\psi\rangle = |1\rangle$ if measuring 1. In practice: can estimate the probabilities $|\alpha|^2$ and $|\beta|^2$ in experiments by repeating the same experiment many times (i.e., via outcome statistics). These repetitions are also called "trials" or "shots".

Circuit notation (with the double line denoting classical information):
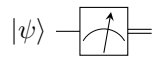
What is a qubit physically? Many possible realizations, e.g., $|0\rangle$ and $|1\rangle$ are:

- two different polarizations of a photon, e.g., left-circular or right-circular
- alignment of a nuclear or electron spin: $\uparrow, \downarrow$
- ground state or excited state of an atom
- clockwise or counterclockwise loop current states in a Josephson function "superconducting qubit"

A useful graphical depiction of a qubit is the *Bloch sphere* representation: If α and β happen to be real-valued, then one can find an angle $\theta \in \mathbb{R}$ such that

$$\alpha = \cos\frac{\theta}{2}, \quad \beta = \sin\frac{\theta}{2} \tag{2.4}$$

($\rightsquigarrow |\alpha|^2 + |\beta|^2 = \cos(\frac{\theta}{2})^2 + \sin(\frac{\theta}{2})^2 = 1$ ✓)

In general: represent

$$\alpha = e^{i\gamma} \cos \frac{\theta}{2},$$

$$\beta = e^{i(\gamma+\varphi)} \sin \frac{\theta}{2}$$

(2.5)

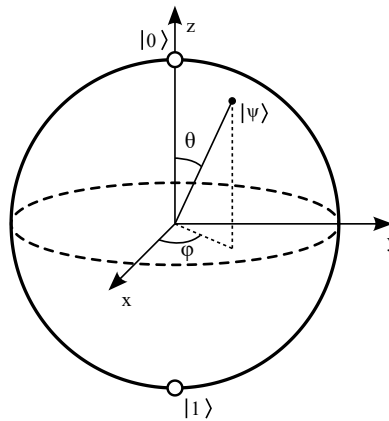using so-called phase angles γ and $\gamma + \varphi$. Then:

$$|\psi\rangle = e^{i\gamma} \cos \frac{\theta}{2}|0\rangle + e^{i(\gamma+\varphi)} \sin \frac{\theta}{2}|1\rangle = e^{i\gamma} \left(\cos \frac{\theta}{2}|0\rangle + e^{i\varphi} \sin \frac{\theta}{2}|1\rangle \right).$$

(2.6)

The overall phase factor $e^{i\gamma}$ can be ignored here for our purposes.

Thus $|\psi\rangle$ is characterized by two angles θ and φ; these specify the point defined as

$$\vec{r} = \begin{pmatrix} \cos(\varphi)\sin(\theta) & \sin(\varphi)\sin(\theta) & \cos(\theta) \end{pmatrix}$$

(2.7)

on the surface of a sphere: **Bloch sphere** (Felix Bloch)

Source: https://commons.wikimedia.org/wiki/File:Bloch_sphere.svg

As an exercise, explain why $|0\rangle$ and $|1\rangle$ correspond to the north and south pole, respectively.

2.2 Single qubit gates

(Nielsen and Chuang 2010, sections 1.3.1, 2.1.8, 4.2)

How to manipulate quantum states (besides measurements)?

Principle of *time evolution*: the quantum state $|\psi\rangle$ at current time point t transitions to a quantum state $|\psi'\rangle$ at a later time point $t' > t$. This transition is described by a complex, unitary matrix U:

$$|\psi'\rangle = U|\psi\rangle$$

(2.8)

$U|\psi\rangle$ is the usual matrix-vector product, and *unitary* means that $U^\dagger U = I$ (identity matrix), with $U^\dagger$ the conjugate-transpose matrix.

Explicitly: denoting

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad |\psi'\rangle = \begin{pmatrix} \alpha' \\ \beta' \end{pmatrix}, \qquad U = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

(2.9)

then

$$|\psi'\rangle = U|\psi\rangle = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} a\alpha + b\beta \\ c\alpha + d\beta \end{pmatrix}.$$

(2.10)

Circuit notation:

3

quantum gate

$$|\psi\rangle - \boxed{U} - |\psi'\rangle$$

time flow
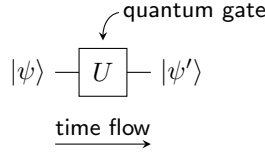
Note:

- Circuit is read from left to right, but matrix times vector ($U|\psi\rangle$) works from right to left.

- U preserves normalization since U is unitary, i.e., $\|U|\psi\rangle\| = \||\psi\rangle\|$ for all $|\psi\rangle \in \mathbb{C}^2$.

Examples:

- quantum analogue of the classical NOT-gate ($0 \leftrightarrow 1$): flip $|0\rangle \leftrightarrow |1\rangle$; this is achieved by the so-called Pauli-X gate:

$$X \equiv \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2.11}$$

  (check: $X|0\rangle = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = |1\rangle$ and $X|1\rangle = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = |0\rangle$ ✓)

- Pauli-Y gate:

$$Y \equiv \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{2.12}$$

- Pauli-Z gate:

$$Z \equiv \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.13}$$

  Z gate leaves $|0\rangle$ unchanged, but flips the sign of the coefficient of $|1\rangle$

  Recall the above Bloch sphere representation:

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle \tag{2.14}$$

  Then (using $e^{i\pi} = -1$)

$$Z|\psi\rangle = \cos\frac{\theta}{2}|0\rangle - e^{i\varphi}\sin\frac{\theta}{2}|1\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i(\varphi+\pi)}\sin\frac{\theta}{2}|1\rangle \tag{2.15}$$

  $\rightsquigarrow$ new Bloch sphere angles: $\theta' = \theta$, $\varphi' = \varphi + \pi$ (rotation by $\pi \,\widehat{=}\, 180°$ around z-axis)

X, Y, Z gates (also denoted σ_1, σ_2, σ_3) are the so-called **Pauli matrices**.

The **Pauli vector** $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is a vector of 2×2 matrices.

- Hadamard gate:

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{2.16}$$
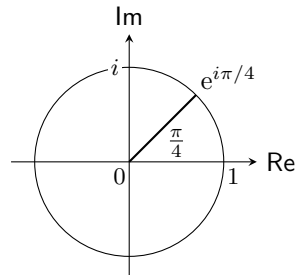
$$\alpha|0\rangle + \beta|1\rangle - \boxed{H} - \alpha\frac{|0\rangle+|1\rangle}{\sqrt{2}} + \beta\frac{|0\rangle-|1\rangle}{\sqrt{2}}$$

- Phase gate:

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \tag{2.17}$$

- T gate:

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} \tag{2.18}$$

4

General definition of matrix exponential (via power series expansion), see also Tutorial 1:

$$\mathrm{e}^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k, \quad A \in \mathbb{C}^{n \times n} \tag{2.19}$$

with $A^k = A \cdot A \cdots A$ (k times). For $A^2 = I$, $x \in \mathbb{R}$ a real number:

$$\mathrm{e}^{iAx} = \cos(x)I + i \sin(x)A \tag{2.20}$$

(generalizes Euler's formula $\mathrm{e}^{ix} = \cos(x) + i\sin(x)$). This can be used to define the following **rotation operators** via the Pauli matrices: for $\theta \in \mathbb{R}$:

$$R_x(\theta) := \mathrm{e}^{-i\theta X/2} = \cos(\theta/2)I - i\sin(\theta/2)X = \begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}, \tag{2.21}$$

$$R_y(\theta) := \mathrm{e}^{-i\theta Y/2} = \cos(\theta/2)I - i\sin(\theta/2)Y = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}, \tag{2.22}$$

$$R_z(\theta) := \mathrm{e}^{-i\theta Z/2} = \cos(\theta/2)I - i\sin(\theta/2)Z = \begin{pmatrix} \mathrm{e}^{-i\theta/2} & 0 \\ 0 & \mathrm{e}^{i\theta/2} \end{pmatrix}. \tag{2.23}$$

General case: rotation about an axis $\vec{v} \in \mathbb{R}^3$ (normalized such that with $\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2} = 1$); using the notation (in what follows)

$$\vec{v} \cdot \vec{\sigma} \equiv v_1\sigma_1 + v_2\sigma_2 + v_3\sigma_3 = \begin{pmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 \end{pmatrix} \in \mathbb{C}^{2 \times 2} \tag{2.24}$$

It holds that $(\vec{v} \cdot \vec{\sigma})^2 = I$ (see Tutorial 1). With that, we define

$$R_{\vec{v}}(\theta) := \mathrm{e}^{-i\theta(\vec{v}\cdot\vec{\sigma})/2} = \cos(\theta/2)I - i\sin(\theta/2)(\vec{v} \cdot \vec{\sigma}). \tag{2.25}$$

Note: R_x, R_y, R_z are special cases, corresponding to $\vec{v} = (1,0,0)$, $\vec{v} = (0,1,0)$, $\vec{v} = (0,0,1)$, respectively.

Can derive that Bloch sphere representation of $R_{\vec{v}}(\theta)$ is a "conventional" rotation (in three dimensions) by angle θ about axis $\vec{v}$! To explain this statement, let $|\psi\rangle$ be an arbitrary single qubit state, and denote the corresponding Bloch sphere point by $\vec{r}$. Let

$$|\psi'\rangle = R_{\vec{v}}(\theta)|\psi\rangle, \tag{2.26}$$

with corresponding point on the Bloch sphere $\vec{r}'$. Then $\vec{r}'$ results from $\vec{r}$ by a rotation around axis $\vec{v}$ by angle θ according to the right hand rule.
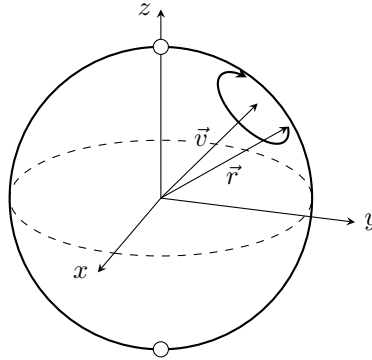
Z-Y decomposition of an arbitrary 2×2 unitary matrix: For any unitary matrix $U \in \mathbb{C}^{2\times 2}$ there exit real numbers $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ such that

$$U = \mathrm{e}^{i\alpha} \underbrace{\begin{pmatrix} \mathrm{e}^{-i\beta/2} & 0 \\ 0 & \mathrm{e}^{i\beta/2} \end{pmatrix}}_{R_z(\beta)} \underbrace{\begin{pmatrix} \cos\frac{\gamma}{2} & -\sin\frac{\gamma}{2} \\ \sin\frac{\gamma}{2} & \cos\frac{\gamma}{2} \end{pmatrix}}_{R_y(\gamma)} \underbrace{\begin{pmatrix} \mathrm{e}^{-i\delta/2} & 0 \\ 0 & \mathrm{e}^{i\delta/2} \end{pmatrix}}_{R_z(\delta)}. \tag{2.27}$$

2.3 Multiple qubits

(Nielsen and Chuang 2010, sections 1.2.1, 2.1.7)

So far: single qubit, superposition of basis states $|0\rangle$ and $|1\rangle$

For two qubits, this generalizes to

$$\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\} \tag{2.28}$$

as computational basis states: all combinations (bitstrings) of 0s and 1s

General two-qubit quantum state:

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle \tag{2.29}$$

with amplitudes $a_{ij} \in \mathbb{C}$, such that $|\alpha_{00}|^2 + |\alpha_{01}|^2 + |\alpha_{10}|^2 + |\alpha_{11}|^2 = 1$ (normalization)

Can identify the basis states with unit vectors:

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \tag{2.30}$$

thus:

$$|\psi\rangle = \begin{pmatrix} \alpha_{00} \\ \alpha_{01} \\ \alpha_{10} \\ \alpha_{11} \end{pmatrix} \in \mathbb{C}^4 \tag{2.31}$$

What happens if we measure only *one* qubit of a two qubit state?
Let's say we measure the first qubit: obtain result

$$\begin{aligned} 0 \quad &\text{with probability } |\alpha_{00}|^2 + |\alpha_{01}|^2, \\ 1 \quad &\text{with probability } |\alpha_{10}|^2 + |\alpha_{11}|^2. \end{aligned} \tag{2.32}$$

Wavefunction directly after measurement:

$$\begin{aligned} \text{if measured 0:} \quad |\psi'\rangle &= \frac{\alpha_{00}|00\rangle + \alpha_{01}|01\rangle}{\sqrt{|\alpha_{00}|^2 + |\alpha_{01}|^2}}, \\ \text{if measured 1:} \quad |\psi'\rangle &= \frac{\alpha_{10}|10\rangle + \alpha_{11}|11\rangle}{\sqrt{|\alpha_{10}|^2 + |\alpha_{11}|^2}}, \end{aligned} \tag{2.33}$$

where the factors in the denominators ensure proper normalization.

Mathematical formalism for constructing two qubit states: *tensor products* of vector spaces

Can combine two (arbitrary) vector spaces V and W to form the **tensor product** $V \otimes W$. The elements of $V \otimes W$ are linear combinations of "tensor products" $|v\rangle \otimes |w\rangle$ consisting of elements $|v\rangle \in V$ and $|w\rangle \in W$.

Example: let $V = \mathbb{C}^2$, $W = \mathbb{C}^2$ be the single qubit spaces with basis $\{|0\rangle, |1\rangle\}$, then

$$\frac{1}{2} \underbrace{|0\rangle \otimes |0\rangle}_{=|00\rangle} + \frac{5i}{7} \underbrace{|1\rangle \otimes |0\rangle}_{=|10\rangle} \tag{2.34}$$

is an element of $V \otimes W$.

Let $\{|i\rangle_V : i = 1, \ldots, m\}$ be a basis of V, and $\{|j\rangle_W : j = 1, \ldots, n\}$ be a basis of W, then

$$\{|i\rangle_V \otimes |j\rangle_W : i = 1, \ldots, m, j = 1, \ldots, n\} \tag{2.35}$$

is a basis of $V \otimes W$. In particular:

$$\dim(V \otimes W) = \dim(V) \cdot \dim(W). \tag{2.36}$$

Note: $|i\rangle_V \otimes |j\rangle_W$ is also written as $|i\rangle|j\rangle$ or $|ij\rangle$.

Example: the basis construction for $V = \mathbb{C}^2$, $W = \mathbb{C}^2$ leads to the above Eq. (2.28).

Basic properties of tensor products:

- for all $|v\rangle \in V$, $|w\rangle \in W$ and $\alpha \in \mathbb{C}$:

$$\alpha(|v\rangle \otimes |w\rangle) = (\alpha|v\rangle) \otimes |w\rangle = |v\rangle \otimes (\alpha|w\rangle) \tag{2.37}$$

- for all $|v_1\rangle, |v_2\rangle \in V$ and $|w\rangle \in W$:

$$(|v_1\rangle + |v_2\rangle) \otimes |w\rangle = |v_1\rangle \otimes |w\rangle + |v_2\rangle \otimes |w\rangle \tag{2.38}$$

- for all $|v\rangle \in V$ and $|w_1\rangle, |w_2\rangle \in W$:

$$|v\rangle \otimes (|w_1\rangle + |w_2\rangle) = |v\rangle \otimes |w_1\rangle + |v\rangle \otimes |w_2\rangle \tag{2.39}$$

Vector notation (using standard basis), e.g., $|v\rangle = v_1|0\rangle + v_2|1\rangle = \left(\begin{smallmatrix} v_1 \\ v_2 \end{smallmatrix}\right)$, $|w\rangle = w_1|0\rangle + w_2|1\rangle = \left(\begin{smallmatrix} w_1 \\ w_2 \end{smallmatrix}\right)$:

$$\begin{aligned} |v\rangle \otimes |w\rangle &= (v_1|0\rangle + v_2|1\rangle) \otimes (w_1|0\rangle + w_2|1\rangle) \\ &= v_1 w_1|00\rangle + v_1 w_2|01\rangle + v_2 w_1|10\rangle + v_2 w_2|11\rangle \end{aligned} \tag{2.40}$$

Thus:

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \otimes \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} v_1 w_1 \\ v_1 w_2 \\ v_2 w_1 \\ v_2 w_2 \end{pmatrix} \tag{2.41}$$

Example: $\begin{pmatrix} 2 \\ 3 \end{pmatrix} \otimes \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 10 \\ 14 \\ 15 \\ 21 \end{pmatrix}$

Note: not every element of $V \otimes W$ can be written in the form $|v\rangle \otimes |w\rangle$ with $|v\rangle \in V$ and $|w\rangle \in W$, for example the so-called Bell state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{2.42}$$

(see also Exercise 1.2(d)).

Assuming that V and W are equipped with an inner (scalar) product $\langle \cdot | \cdot \rangle$ (see also Primer 1.2), define inner product on $V \otimes W$ by:

$$\left\langle \sum_j \alpha_j |v_j\rangle |w_j\rangle \,\Big|\, \sum_k \beta_k |\tilde{v}_k\rangle |\tilde{w}_k\rangle \right\rangle := \sum_j \sum_k \alpha_j^* \beta_k \langle v_j | \tilde{v}_k \rangle \langle w_j | \tilde{w}_k \rangle. \tag{2.43}$$

Generalization to n qubits: 2^n computational basis states

$$\{|00\cdots0\rangle, |00\cdots1\rangle, |0\cdots10\rangle, \ldots, |11\cdots1\rangle\} \tag{2.44}$$

(all binary strings of length n)

Thus: general n-qubit quantum state, also denoted **quantum register**, given by:

$$|\psi\rangle = \sum_{x_0=0}^{1} \sum_{x_1=0}^{1} \cdots \sum_{x_{n-1}=0}^{1} \alpha_{x_{n-1},\ldots,x_1,x_0} |x_{n-1},\ldots,x_1,x_0\rangle = \sum_{x=0}^{2^n-1} \alpha_x |x\rangle \qquad (2.45)$$

with $\alpha_x \in \mathbb{C}$ for all $x \in \{0,1,\ldots,2^n-1\}$, such that $\sum_{x=0}^{2^n-1} |\alpha_x|^2 = 1$ (normalization).

⤳ exponentially many amplitudes α_x, in general "hard" to simulate on classical computers (for large n) due to this "curse of dimensionality".

2.4 Multiple qubit gates

(Nielsen and Chuang 2010, sections 1.3.2, 1.3.4, 2.1.7)

As for single qubits, an operation on multiple qubits is described be a unitary matrix U.

For n qubits: $U \in \mathbb{C}^{2^n \times 2^n}$

Example: **controlled-NOT** gate (also denoted **CNOT**): two qubits: control and target; target qubit gets flipped if control is 1:

$$\underset{\substack{\nearrow \quad \nwarrow \\ \text{control} \qquad \text{target}}}{|00\rangle \mapsto |00\rangle,} \quad |01\rangle \mapsto |01\rangle, \quad \underbrace{|10\rangle \mapsto |11\rangle, \quad |11\rangle \mapsto |10\rangle}_{\text{target gets flipped}} \qquad (2.46)$$
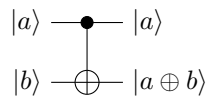
Can be expressed as:

$$|a,b\rangle \mapsto |a, a \oplus b\rangle \qquad (2.47)$$

(where $\oplus$ is addition modulo 2, same as XOR)
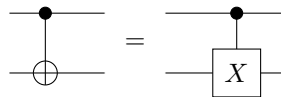
Circuit notation:

$$
\begin{array}{c}
|a\rangle \; \text{———}\bullet\text{———} \; |a\rangle \\
\quad\quad\quad | \\
|b\rangle \; \text{———}\oplus\text{———} \; |a \oplus b\rangle
\end{array}
$$

Matrix representation (with respect to computational basis in Eq. (2.28)):

$$U_{\mathsf{CNOT}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \boxed{0 & 1} \\ 0 & 0 & \boxed{1 & 0} \end{pmatrix} \qquad (2.48)$$
$$\underset{\text{Pauli-}X}{}$$

(quick check: U_{CNOT} is unitary ✓)

By identifying the lower-right block as Pauli-X matrix, we arrive at the following alternative circuit representation:

$$
\begin{array}{ccc}
\bullet & & \bullet \\
| & = & | \\
\oplus & & \boxed{X}
\end{array}
$$

Can generalize Pauli-X to any unitary operator U acting on target qubit ⤳ controlled-U gate:

$$\underset{\substack{\nearrow \quad \nwarrow \\ \text{control} \qquad \text{target}}}{|00\rangle \mapsto |00\rangle,} \quad |01\rangle \mapsto |01\rangle, \quad \underbrace{|10\rangle \mapsto |1\rangle \otimes (U|0\rangle), \quad |11\rangle \mapsto |1\rangle \otimes (U|1\rangle)}_{U \text{ is applied to target}} \qquad (2.49)$$

Circuit notation and corresponding matrix representation:

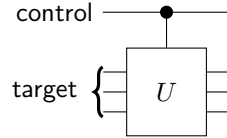$$\begin{array}{c}\text{circuit}\end{array} \;\widehat{=}\; \begin{pmatrix} 1 & & \\ & 1 & \\ & & \boxed{U} \end{pmatrix} \tag{2.50}$$

Example: controlled-Z:

$$\begin{array}{c}\text{circuit}\end{array} \;\widehat{=}\; \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} \tag{2.51}$$

As an exercise, convince yourself that controlled-Z is invariant when flipping the roles of the control and target qubit.

The controlled-U construction also works for multiple target qubits:

Note: single qubit and CNOT gates are *universal*: they can be used to implement an arbitrary unitary operation on n qubits. (Quantum analogue of universality of classical NAND gate; proof in Nielsen and Chuang 2010, section 4.5)

Example of a circuit consisting only of single qubit gates and CNOTs:
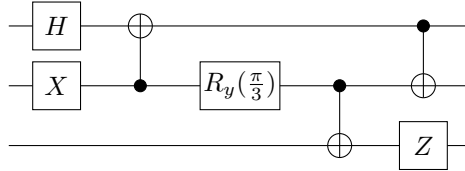
Matrix Kronecker products   How can we compute the matrix representation of single qubit gates acting in parallel, as shown by the following circuit? Here A and B are unitary 2×2 matrices.
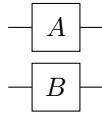
Operation on basis states (for $a, b \in \{0, 1\}$):

$$|a, b\rangle \mapsto (A|a\rangle) \otimes (B|b\rangle) \tag{2.52}$$

Example: $A = I$ (identity matrix), $B = Y$ (Pauli-Y, see Eq. (2.12)):

$$
\begin{aligned}
|00\rangle &\mapsto |0\rangle \otimes \underbrace{(Y|0\rangle)}_{=i|1\rangle} = |0\rangle \otimes (i|1\rangle) = i|01\rangle \\
|01\rangle &\mapsto |0\rangle \otimes \underbrace{(Y|1\rangle)}_{=-i|0\rangle} = -i|00\rangle \\
|10\rangle &\mapsto |1\rangle \otimes (Y|0\rangle) = i|11\rangle \\
|11\rangle &\mapsto |1\rangle \otimes (Y|1\rangle) = -i|10\rangle
\end{aligned}
\tag{2.53}
$$

Matrix representation:

$$
\boxed{I} \atop \boxed{Y} \quad \widehat{=} \quad
\begin{pmatrix}
0 & -i & 0 & 0 \\
i & 0 & 0 & 0 \\
0 & 0 & 0 & -i \\
0 & 0 & i & 0
\end{pmatrix}
= \begin{pmatrix} Y & 0 \\ 0 & Y \end{pmatrix} = I \otimes Y
\tag{2.54}
$$

with Y labelling the dashed blocks, and the 2×2 zero matrix indicated.

General formula: **Kronecker product** (see also exercise 1.2): matrix representation of tensor products of operators:
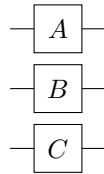
$$
A \otimes B =
\begin{pmatrix}
a_{11}B & a_{12}B & \cdots & a_{1n}B \\
a_{21}B & a_{22}B & \cdots & a_{2n}B \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1}B & a_{m2}B & \cdots & a_{mn}B
\end{pmatrix}
\in \mathbb{C}^{mp \times nq}
\tag{2.55}
$$

for $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{p \times q}$ (arbitrary dimensions)

Another example:

$$
\boxed{Y} \atop \boxed{I} \quad \widehat{=} \quad Y \otimes I =
\begin{pmatrix} 0 \cdot I & -i \cdot I \\ i \cdot I & 0 \cdot I \end{pmatrix}
=
\begin{pmatrix}
0 & 0 & -i & 0 \\
0 & 0 & 0 & -i \\
i & 0 & 0 & 0 \\
0 & i & 0 & 0
\end{pmatrix}
\tag{2.56}
$$

The Kronecker product can be generalized to an arbitrary number of tensor factors, for example $A \otimes B \otimes C$:

$$
\boxed{A} \atop \boxed{B} \atop \boxed{C}
$$

Basic properties of matrix Kronecker products:

(a) $(A \otimes B)^* = A^* \otimes B^*$ (element-wise complex conjugation)

(b) $(A \otimes B)^T = A^T \otimes B^T$ (transposition)

(c) $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$ (conjugate-transpose)

(d) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$ (associative property)

(e) $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ for matrices A, B, C, D with compatible dimensions

(f) Kronecker product of two Hermitian matrices is Hermitian; expressed in formulas: if $A^\dagger = A$ and $B^\dagger = B$, then $(A \otimes B)^\dagger = A \otimes B$

(g) Kronecker product of two unitary matrices is unitary

As a short exercise, prove (g) using (c) and (e).

2.5 Quantum measurement

(Nielsen and Chuang 2010, sections 1.3.3, 2.2.3, 2.2.5)

We have already discussed measurement of a single qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ with respect to the computational basis $\{|0\rangle, |1\rangle\}$

classical data
(measurement outcome 0 or 1)
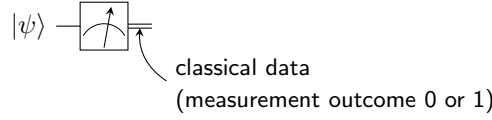
Linear algebra: can switch to a different (orthonormal) basis to represent a qubit, for example

$$|+\rangle := \frac{|0\rangle + |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{2.57a}$$

$$|-\rangle := \frac{|0\rangle - |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{2.57b}$$

Representation of $|\psi\rangle$ with respect to $\{|+\rangle, |-\rangle\}$ basis:

$$\alpha|0\rangle + \beta|1\rangle = \alpha\frac{|+\rangle + |-\rangle}{\sqrt{2}} + \beta\frac{|+\rangle - |-\rangle}{\sqrt{2}} = \frac{\alpha+\beta}{\sqrt{2}}|+\rangle + \frac{\alpha-\beta}{\sqrt{2}}|-\rangle \tag{2.58}$$

Can perform measurement with respect to orthonormal basis $\{|+\rangle, |-\rangle\}$ (instead of $\{|0\rangle, |1\rangle\}$), will obtain result

$$\begin{aligned} + &\quad \text{with probability } \frac{|\alpha + \beta|^2}{2}, \\ - &\quad \text{with probability } \frac{|\alpha - \beta|^2}{2}. \end{aligned} \tag{2.59}$$

As before, the probabilities are the squared absolute values of the coefficients in front of the basis states.

Wavefunction collapse: immediately after measurement, qubit will be in state $|+\rangle$ if measured "+" and likewise in state $|-\rangle$ if measured "−".

In general, given an *orthonormal* basis $\{|u_1\rangle, |u_2\rangle\}$, one can represent a qubit as $|\psi\rangle = \alpha_1|u_1\rangle + \alpha_2|u_2\rangle$ and measure with respect to this orthonormal basis; will obtain measurement result "u_1" or "u_2" with respective probabilities $|\alpha_1|^2$ and $|\alpha_2|^2$.

Recall that *orthonormal* means that $|u_1\rangle, |u_2\rangle$ are normalized (have length 1) and are orthogonal to each other, which can be succinctly expressed as

$$\langle u_i | u_j \rangle = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{2.60}$$

It turns out that quantum measurements as discussed so far are special cases of the following (even more general) definition:

Quantum measurements are described by a collection $\{M_m\}$ of **measurement operators** acting on the quantum system, with the index m labeling possible measurement outcomes. Denoting the state of the quantum system immediately before the measurement by $|\psi\rangle$, result m occurs with probability

$$p(m) = \langle \psi | M_m^\dagger M_m | \psi \rangle = \|M_m|\psi\rangle\|^2, \tag{2.61}$$

and the state after the measurement is

$$\frac{M_m|\psi\rangle}{\|M_m|\psi\rangle\|}. \tag{2.62}$$

The measurement operators satisfy the completeness relation

$$\sum_m M_m^\dagger M_m = I, \tag{2.63}$$

such that the probabilities sum to 1:

$$\sum_m p(m) = \sum_m \langle\psi|M_m^\dagger M_m|\psi\rangle = \langle\psi|\underbrace{\sum_m M_m^\dagger M_m}_{=I}|\psi\rangle = \langle\psi|\psi\rangle = 1. \tag{2.64}$$

Example: measurement of a qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ with respect to computational basis $\{|0\rangle, |1\rangle\}$ is a special case: set

$$M_0 := |0\rangle\langle0| = \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{2\times1} \underbrace{\begin{pmatrix} 1 & 0 \end{pmatrix}}_{1\times2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{2.65a}$$

$$M_1 := |1\rangle\langle1| = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.65b}$$

(Note that for this example, $M_m^\dagger = M_m$ and $M_m^2 = M_m$)

Check completeness relation: $M_0^\dagger M_0 + M_1^\dagger M_1 = M_0 + M_1 = I$ ✓

Measurement probabilities:

$$p(0) = \langle\psi|M_0^\dagger M_0|\psi\rangle = \langle\psi|M_0|\psi\rangle = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha^*\alpha = |\alpha|^2, \tag{2.66a}$$

$$p(1) = \langle\psi|M_1^\dagger M_1|\psi\rangle = \langle\psi|M_1|\psi\rangle = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \beta^*\beta = |\beta|^2, \tag{2.66b}$$

agrees with hitherto definition.

Special case: *projective measurements* (together with unitary transformations equivalent to general measurement framework).

Before defining projective measurements below, we take a brief mathematical excursion:

Definition. *A square matrix $P \in \mathbb{C}^{n\times n}$ is called an* **orthogonal projection matrix** *if it is Hermitian ($P^\dagger = P$) and satisfies $P^2 = P$.*

Let V be a k-dimensional subspace of $\mathbb{C}^n$, and $\{|u_1\rangle, \dots, |u_k\rangle\}$ an orthonormal ($\langle u_i|u_j\rangle = \delta_{ij}$) basis of V (which always exists). Then

$$P = \sum_{j=1}^k |u_j\rangle\langle u_j| \tag{2.67}$$

(with $\langle u_j| = |u_j\rangle^\dagger$) is the **projector** onto V, acting on a vector $|w\rangle \in \mathbb{C}^n$ as

$$P|w\rangle = \sum_{j=1}^k |u_j\rangle\langle u_j|w\rangle. \tag{2.68}$$

It is straightforward to show that P is an orthogonal projection matrix, that P is independent of the chosen basis, and that $P|v\rangle = |v\rangle$ for all $|v\rangle \in V$.

Relation to spectral decomposition: Every normal matrix $A \in \mathbb{C}^{n \times n}$ (i.e., $[A, A^\dagger] = A\,A^\dagger - A^\dagger\,A = 0$) is diagonalizable by an orthonormal basis, that is, there exist a unitary $U \in \mathbb{C}^{n \times n}$ and eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that

$$A = U \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} U^\dagger = \sum_{j=1}^{n} \lambda_j \,|u_j\rangle\langle u_j| \tag{2.69}$$

where the $|u_j\rangle$ are the column vectors of U: $U = \big(u_1|u_2|\cdots|u_n\big)$.

Some eigenvalues can coincide, for example $\lambda_1 = \lambda_2$; then

$$\lambda_1 \,|u_1\rangle\langle u_1| + \lambda_2 \,|u_2\rangle\langle u_2| = \lambda_1 \, P, \tag{2.70}$$

with $P = |u_1\rangle\langle u_1| + |u_2\rangle\langle u_2|$ the projector onto the (two-dimensional) eigenspace corresponding to eigenvalue λ_1. In general, we can write the spectral decomposition as

$$A = \sum_m \lambda_m \, P_m \tag{2.71}$$

where the sum is over all *distinct* eigenvalues and the P_m are the projectors onto the corresponding eigenspaces.

A **projective measurement** is described by an **observable** M, a Hermitian operator acting on the quantum system. Write the spectral decomposition of the observable as

$$M = \sum_m \lambda_m P_m, \tag{2.72}$$

where P_m is the projector onto the eigenspace of M with eigenvalue λ_m. The possible outcomes of the measurement correspond to the eigenvalues λ_m. Upon measuring a quantum state $|\psi\rangle$, the probability of getting result λ_m is given by

$$p(\lambda_m) = \langle\psi|P_m|\psi\rangle. \tag{2.73}$$

In this case, the state of the quantum system directly after the measurement is

$$\frac{P_m|\psi\rangle}{\|P_m|\psi\rangle\|} = \frac{P_m|\psi\rangle}{\sqrt{p(m)}}. \tag{2.74}$$

Projective measurements are special cases of the general measurement framework, namely, if the measurement operators are projectors.

Average value of a projective measurement:

$$\mathbb{E}[M] = \sum_m \lambda_m \, p(\lambda_m) = \sum_m \lambda_m \langle\psi|P_m|\psi\rangle = \langle\psi| \left(\sum_m \lambda_m P_m \right) |\psi\rangle = \langle\psi|M|\psi\rangle \tag{2.75}$$

If the state $|\psi\rangle$ is obvious from the context, one also uses the notation $\langle M\rangle = \langle\psi|M|\psi\rangle$. Define corresponding standard deviation by

$$\Delta(M) = \sqrt{\langle M^2\rangle - \langle M\rangle^2} \tag{2.76}$$

Examples:

- Measuring a qubit with respect to computational basis $\{|0\rangle, |1\rangle\}$ (see above Eq. (2.65)) is actually a projective measurement

- In general: measurement w.r.t. orthonormal basis $\{|u_1\rangle, |u_2\rangle\}$ is a projective measurement: set

$$P_m = |u_m\rangle\langle u_m| \quad \text{for } m = 1, 2 \tag{2.77}$$

(check: $\sum_{m=1}^{2} P_m = I$ ✓), define M by

$$M = \sum_{m=1}^{2} \lambda_m P_m \quad \text{with arbitrary } \lambda_1, \lambda_2 \in \mathbb{R} \text{ such that } \lambda_1 \neq \lambda_2 \tag{2.78}$$

Then

$$p(\lambda_m) = \langle\psi|P_m|\psi\rangle = \langle\psi|u_m\rangle\langle u_m|\psi\rangle = |\langle u_m|\psi\rangle|^2 \quad \text{for } m = 1, 2 \tag{2.79}$$

Note:

$$|\psi\rangle = \sum_{m=1}^{2} P_m|\psi\rangle = \sum_{m=1}^{2} |u_m\rangle\langle u_m|\psi\rangle, \tag{2.80}$$

i.e., the numbers $\langle u_m|\psi\rangle$ are the coefficients of $|\psi\rangle$ with respect to $\{|u_1\rangle, |u_2\rangle\}$ basis

- Measuring Pauli-Z:

$$Z = 1 \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{P_1} + (-1) \cdot \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{P_2} \tag{2.81}$$

agrees with measurement w.r.t. computational basis $\{|0\rangle, |1\rangle\}$

3 Entanglement and its applications

A n-qubit state $|\psi\rangle$ ($n \geq 2$) is called **entangled** if it cannot be written as tensor product of single qubit states, i.e.,

$$|\psi\rangle \neq |\varphi_{n-1}\rangle \otimes \cdots \otimes |\varphi_0\rangle \quad \text{for any } |\varphi_0\rangle, \ldots |\varphi_{n-1}\rangle \in \mathbb{C}^2 \tag{3.1}$$

Example: Bell states, also denoted EPR states (Einstein, Podolsky, Rosen):
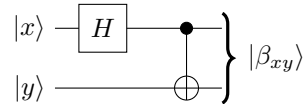
$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{3.2a}$$

$$|\beta_{01}\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \tag{3.2b}$$

$$|\beta_{10}\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \tag{3.2c}$$

$$|\beta_{11}\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \tag{3.2d}$$

Quantum circuit to create Bell states (see also Exercise 2.1(b)):

14

3.1 Quantum teleportation

(Nielsen and Chuang 2010, section 1.3.7)

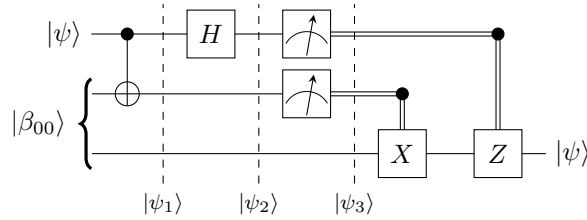Scenario: two (experimental physicists) Alice and Bob live far away from each other:

Alice Bob

When visiting each other a long time ago, they generated the EPR pair $|\beta_{00}\rangle$, each keeping one qubit of the pair.

Now Alice's task is to send another (unknown) qubit $|\psi\rangle$ to Bob.

Note: measurement is not an option since this does not reveal the unknown amplitudes α and β of $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$

Quantum circuit for teleporting $|\psi\rangle$ (first two qubit lines belong to Alice, last line to Bob):

Input:

$$|\psi\rangle|\beta_{00}\rangle \equiv |\psi\rangle \otimes |\beta_{00}\rangle = \frac{1}{\sqrt{2}}\Big(\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|00\rangle + |11\rangle)\Big) \tag{3.3}$$

After Alice sends her two qubits through CNOT:

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}\Big(\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|10\rangle + |01\rangle)\Big) \tag{3.4}$$

After Hadamard:

$$\begin{aligned}|\psi_2\rangle &= \frac{1}{2}\Big(\alpha(|0\rangle + |1\rangle)(|00\rangle + |11\rangle) + \beta(|0\rangle - |1\rangle)(|10\rangle + |01\rangle)\Big) \\ &= \frac{1}{2}\Big(|00\rangle(\alpha|0\rangle + \beta|1\rangle) + |01\rangle(\alpha|1\rangle + \beta|0\rangle) + |10\rangle(\alpha|0\rangle - \beta|1\rangle) + |11\rangle(\alpha|1\rangle - \beta|0\rangle)\Big)\end{aligned} \tag{3.5}$$

(second line follows from regrouping)

Alice then measures her qubits w.r.t. computational basis, corresponds to projective measurement with $P_1 = |00\rangle\langle00| \otimes I$, $P_2 = |01\rangle\langle01| \otimes I$, $P_3 = |10\rangle\langle10| \otimes I$, $P_4 = |11\rangle\langle11| \otimes I$. Thus, if Alice measures 00, then $|\psi_2\rangle$ will "collapse" to

$$|00\rangle(\alpha|0\rangle + \beta|1\rangle) = |00\rangle|\psi\rangle, \tag{3.6}$$

i.e., Bob's qubit is now $|\psi\rangle$!

Similarly, directly from second line in Eq. (3.5):

$$00 \rightarrow \alpha|0\rangle + \beta|1\rangle \tag{3.7a}$$
$$01 \rightarrow \alpha|1\rangle + \beta|0\rangle \tag{3.7b}$$
$$10 \rightarrow \alpha|0\rangle - \beta|1\rangle \tag{3.7c}$$
$$11 \rightarrow \alpha|1\rangle - \beta|0\rangle \tag{3.7d}$$

Alice transmits her measurement result (as classical information) to Bob; Bob then applies Pauli-X and/or Pauli-Z gates if necessary to recover $|\psi\rangle$.

Even though wavefunction collapse is instantaneous, no faster-than-light information transfer is possible due to the required classical communication.

3.2 EPR and the Bell inequality

(Nielsen and Chuang 2010, section 2.6)

EPR: **E**instein, **P**odolsky, **R**osen

Famous EPR paper: *Can quantum-mechanical description of physical reality be considered complete?* (Einstein, Podolsky, and Rosen 1935)

The authors argue that quantum mechanics is incomplete since it lacks certain "elements of reality". Sufficient to be considered "element of reality": property (e.g., of a qubit) can be predicted with certainty.

Consider the following scenario: Alice and Bob are far from each other, but share the entangled two-qubit state

$$|\beta_{11}\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \tag{3.8}$$

(also denoted spin-singlet). The first qubit belongs to Alice and the second qubit to Bob. Alice and Bob measure the observable $\vec{v}\cdot\vec{\sigma} = v_1 X + v_2 Y + v_3 Z$ (with $\vec{v} \in \mathbb{R}^3$, $\|\vec{v}\| = 1$) on their respective qubit. (Recall that $\vec{v}\cdot\vec{\sigma}$ is Hermitian and unitary, and has eigenvalues ± 1, see Eq. (2.24) and Tutorial 1(d).) Alice performs her measurement immediately before Bob. Examples:

- $\vec{v} = (0,0,1)$, observable $Z = 1 \cdot |0\rangle\langle 0| + (-1) \cdot |1\rangle\langle 1|$, i.e., standard measurement w.r.t. computational basis (see above Eq. (2.81)). If Alice measures eigenvalue

$$\begin{aligned} +1, &\quad \text{the wavefunction collapses to } |01\rangle, \\ -1, &\quad \text{the wavefunction collapses to } |10\rangle, \end{aligned} \tag{3.9}$$

  such that Bob will always obtain the opposite measurement result (since his qubit is $|1\rangle$ if Alice's qubit collapses to $|0\rangle$, and other way around)!

  (Note: "measuring eigenvalue $+1$" is just another name for "obtaining result 0" in reference to $|0\rangle$, and likewise "measuring eigenvalue -1" is equivalent to saying "obtaining result 1".)

- $\vec{v} = (1,0,0)$, observable X: eigenstates $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ with corresponding eigenvalues ± 1; in other words: Alice and Bob perform measurement w.r.t. basis $\{|+\rangle, |-\rangle\}$. Can represent the wavefunction as

$$|\beta_{11}\rangle = -\frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle), \tag{3.10}$$

  namely:

$$-\frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle) = -\frac{1}{\sqrt{2}}\left(\frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) - \frac{1}{2}(|0\rangle - |1\rangle)(|0\rangle + |1\rangle)\right)$$
$$= -\frac{1}{2\sqrt{2}}(|00\rangle - |01\rangle + |10\rangle - |11\rangle - |00\rangle - |01\rangle + |10\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) = |\beta_{11}\rangle \tag{3.11}$$

  If Alice measures eigenvalue $+1$, wavefunction will collapse to $|+\rangle$, i.e., Bob's qubit is in state $|-\rangle$ and he will certainly measure eigenvalue -1. Conversely if Alice measures eigenvalue -1, Bob will measure eigenvalue $+1$.

- general $\vec{v}$, observable $\vec{v} \cdot \vec{\sigma}$: denote its orthonormal eigenstates by $|a\rangle, |b\rangle$, then there exist complex numbers $\alpha, \beta, \gamma, \delta$ such that

$$|0\rangle = \alpha|a\rangle + \beta|b\rangle, \tag{3.12a}$$
$$|1\rangle = \gamma|a\rangle + \delta|b\rangle. \tag{3.12b}$$

Inserted into $|\beta_{11}\rangle$ (cf. Exercise 5.1(a)):

$$\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) = (\alpha\delta - \beta\gamma)\frac{1}{\sqrt{2}}(|ab\rangle - |ba\rangle) \tag{3.13}$$

We observe that

$$\alpha\delta - \beta\gamma = \det(U) \quad \text{with} \quad U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}; \tag{3.14}$$

U is the base change matrix between the orthonormal $\{|0\rangle, |1\rangle\}$ and $\{|a\rangle, |b\rangle\}$ bases, and hence unitary. Any unitary matrix U satisfies $|\det(U)| = 1$ since

$$|\det(U)|^2 = \det(U)^* \det(U) = \det(U^\dagger) \det(U) = \det(U^\dagger U) = \det(I) = 1. \tag{3.15}$$

Here we have used that in general $\det(A^T) = \det(A)$, $\det(A)^* = \det(A^*)$, and $\det(A)\det(B) = \det(AB)$. Since $|\det(U)| = 1$, we can represent $\det(U) = e^{i\theta}$ for some angle $\theta \in \mathbb{R}$. In summary:

$$\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) = e^{i\theta}\frac{1}{\sqrt{2}}(|ab\rangle - |ba\rangle). \tag{3.16}$$

⤳ as before: Bob will obtain opposite measurement result of Alice (in case Alice measures "a", Bob will measure "b" and other way around; global phase factor $e^{i\theta}$ not relevant for measurements). Therefore Alice can predict Bob's measurement result. On the other hand, there is no possibility that Alice could somehow influence Bob's measurement after performing her measurement since they are far apart (speed of light too slow to travel from Alice to Bob).

EPR argument: "property" $\vec{v} \cdot \vec{\sigma}$ of a qubit is thus an "element of reality"; however, quantum mechanics does not a priori specify this property for all possible $\vec{v}$ (but only probabilities) and is thus an incomplete description of reality. Instead: "hidden variable theory": there must be additional variables "hidden" in a qubit which determine Bob's measurement of $\vec{v} \cdot \vec{\sigma}$ for all possible $\vec{v}$ (but this idea turned out to be wrong).

Bell's inequality: experimental test which can invalidate *local* hidden variable theories (Bell 1964).
Here **local** means that no faster-than-light communication is possible (which is almost universally accepted, otherwise one could send information backwards in time according to special relativity).

John Bell

Experimental schematic for Bell's inequality: many repetitions (to collect statistics) of the following setup: Charlie experimentally prepares two particles and sends one to Alice and one to Bob, who perform measurements on the received particle:

Alice:
Q or R

Bob:
S or T

Charlie
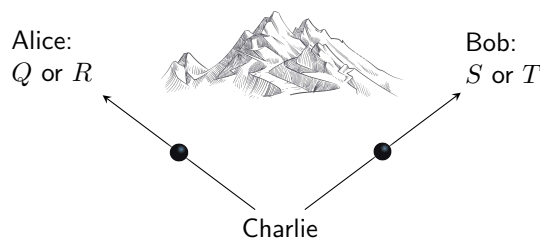
We assume that the particles have some (abstract) properties, each of which can assume two possible values ± 1. Alice measures a property denoted P_Q of her particle, or (another) property P_R, randomly deciding which one to measure; corresponding measurement values $Q \in \{\pm 1\}$ or $R \in \{\pm 1\}$. Analogously for Bob with $S, T \in \{\pm 1\}$.

Alice and Bob perform their measurements (almost) simultaneously, such that no information about the result can be transmitted in between (assuming no faster-than-light communication).

After completing the experiment, Alice and Bob meet to analyze their measurement data.

Consider the quantity

$$QS + RS + RT - QT = (Q + R)S + (R - Q)T \tag{3.17}$$

Since the properties only assume values ± 1, one concludes that $Q + R = 0$ and $R - Q = \pm 2$, or $Q + R = \pm 2$ and $R - Q = 0$; therefore $QS + RS + RT - QT = \pm 2$. Denote by $p(q, r, s, t)$ the probability that the two-particle system before measurement is in state $Q = q$, $R = r$, $S = s$, $T = t$, then the average (expectation value) of the quantity in Eq. (3.17) is

$$\mathbb{E}[QS + RS + RT - QT] = \sum_{q,r,s,t} p(q,r,s,t) \underbrace{(qs + rs + rt - qt)}_{=\pm 2} \leq \sum_{q,r,s,t} p(q,r,s,t) \cdot 2 = 2. \tag{3.18}$$

The last equality follows from the fact that probabilities sum to 1. Since $\mathbb{E}$ is linear, one arrives at the following **Bell inequality**:

$$\mathbb{E}[QS] + \mathbb{E}[RS] + \mathbb{E}[RT] - \mathbb{E}[QT] \leq 2. \tag{3.19}$$

Each term can be experimentally evaluated; e.g., for $\mathbb{E}[QS]$, Alice and Bob average over the cases when Alice measured P_Q and Bob measured P_S.

Now consider the following realization of the setup, and the predictions by quantum mechanics: Charlie prepares the two qubit (singlet) state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), \tag{3.20}$$

and sends the first qubit to Alice and the second qubit to Bob.

Observables for measurements (subscripts 1 and 2 indicate which qubit the observable acts on):

$$Q = Z_1, \qquad S = \frac{-Z_2 - X_2}{\sqrt{2}}, \tag{3.21a}$$

$$R = X_1, \qquad T = \frac{Z_2 - X_2}{\sqrt{2}}. \tag{3.21b}$$

Measurement averages (cf. Exercise 5.1(b)):

$$\langle QS \rangle = \frac{1}{\sqrt{2}}, \qquad \langle RS \rangle = \frac{1}{\sqrt{2}}, \tag{3.22a}$$

$$\langle RT \rangle = \frac{1}{\sqrt{2}}, \qquad \langle QT \rangle = -\frac{1}{\sqrt{2}}. \tag{3.22b}$$

(Recall that $\langle M \rangle \equiv \langle \psi | M | \psi \rangle$ for any observable M.) Inserted into Eq. (3.19):

$$\langle QS \rangle + \langle RS \rangle + \langle RT \rangle - \langle QT \rangle = 2\sqrt{2} \nleq 2 \tag{3.23}$$

violates Bell's inequality!

Actual laboratory experiments (using photons) agree with prediction by quantum mechanics, thus not all (implicit) assumptions leading to the Bell inequality can be satisfied:

- "Realism": physical properties P_Q, P_R, P_S, P_T have definite values independent of observation (that is, measurement)

- Locality: Alice performing her measurement cannot influence Bob's measurement and vice versa

In summary, nature is not "locally realistic". (Most common point of view: realism does not hold.)

Practical lesson: use entanglement as resource.

4 The density operator

So far: state vector $|\psi\rangle$ describing a quantum system

Convenient alternative formulation for quantum systems about which we only have partial information: *density operator* (also denoted *density matrix*)

4.1 Ensembles of quantum states

(Nielsen and Chuang 2010, section 2.4.1)

Consider a quantum system which is in one of several states $|\psi_i\rangle$ with probability p_i: **ensemble** of quantum states $\{p_i, |\psi_i\rangle\}$

The **density operator** ρ of the ensemble $\{p_i, |\psi_i\rangle\}$ is defined as

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|. \tag{4.1}$$

Quantum mechanics in terms of density operators:

- Unitary operations: a unitary transformation U maps $|\psi_i\rangle \mapsto U|\psi_i\rangle$ and the ensemble to $\{p_i, U|\psi_i\rangle\}$; thus the density operator is transformed as

$$\rho \xmapsto{U} \sum_i p_i U|\psi_i\rangle\langle\psi_i|U^\dagger = U\rho U^\dagger. \tag{4.2}$$

- Measurements: measurement operators $\{M_m\}$, if system is in state $|\psi_i\rangle$, then probability for result m, given state i, is

$$p(m|i) = \langle\psi_i|M_m^\dagger M_m|\psi_i\rangle = \mathrm{tr}\big[M_m^\dagger M_m|\psi_i\rangle\langle\psi_i|\big] \tag{4.3}$$

according to Eq. (2.61), where tr is the matrix trace.

(Note: $\langle v|A|w\rangle = \mathrm{tr}[A(|w\rangle\langle v|)]$ holds in general for any square matrix A and vectors $|v\rangle$, $|w\rangle$.)

Thus the overall probability for measurement result m equals

$$\begin{aligned}
p(m) &= \sum_i p(m|i)p_i = \sum_i \mathrm{tr}\big[M_m^\dagger M_m|\psi_i\rangle\langle\psi_i|\big]\,p_i \\
&= \mathrm{tr}\Big[M_m^\dagger M_m \sum_i p_i|\psi_i\rangle\langle\psi_i|\Big] = \mathrm{tr}[M_m^\dagger M_m\rho].
\end{aligned} \tag{4.4}$$

Density operator ρ_m after obtaining result m? State i collapses to

$$|\psi_i\rangle \mapsto \frac{M_m|\psi_i\rangle}{\|M_m|\psi_i\rangle\|} =: |\psi_i^m\rangle. \tag{4.5}$$

Thus

$$\rho_m = \sum_i p(i|m)\, |\psi_i^m\rangle\langle\psi_i^m| = \sum_i p(i|m)\frac{M_m|\psi_i\rangle\langle\psi_i|M_m^\dagger}{\|M_m|\psi_i\rangle\|^2}$$
$$= \sum_i p_i \frac{M_m|\psi_i\rangle\langle\psi_i|M_m^\dagger}{p(m)} = \frac{M_m\rho M_m^\dagger}{\mathrm{tr}[M_m^\dagger M_m\rho]}. \tag{4.6}$$

Here we have used that $\|M_m|\psi_i\rangle\|^2 = p(m|i)$, and

$$\frac{p(i|m)}{p(m|i)} = \frac{p_i}{p(m)} \tag{4.7}$$

according to Bayes' theorem. Note that ρ_m is now expressed solely in terms of ρ and the measurement operators, without explicit reference to the ensemble $\{p_i, |\psi_i\rangle\}$.

4.2   General properties of the density operator

(Nielsen and Chuang 2010, section 2.4.2)

Characterization of density operators: An operator ρ is the density operator associated to some ensemble $\{p_i, |\psi_i\rangle\}$ if and only if

1. $\mathrm{tr}[\rho] = 1$ (trace condition)

2. ρ is a positive operator (positivity condition)

Remark: ρ is called a **positive operator** if it is Hermitian and all its eigenvalues are ≥ 0, equivalently, if $\langle\varphi|\rho|\varphi\rangle \geq 0$ for all vectors $|\varphi\rangle$.

Proof.

"$\Rightarrow$" Suppose $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$, then

$$\mathrm{tr}[\rho] = \sum_i p_i\, \mathrm{tr}[|\psi_i\rangle\langle\psi_i|] = \sum_i p_i \underbrace{\langle\psi_i|\psi_i\rangle}_{=1} = 1, \tag{4.8}$$

and for any state $|\varphi\rangle$:

$$\langle\varphi|\rho|\varphi\rangle = \sum_i p_i\langle\varphi|\psi_i\rangle\langle\psi_i|\varphi\rangle = \sum_i p_i\, |\langle\varphi|\psi_i\rangle|^2 \geq 0. \tag{4.9}$$

"$\Leftarrow$" ρ is an operator (i.e., a Hermitian matrix), and thus, according to the spectral theorem, there exist real eigenvalues λ_j and corresponding orthonormal eigenvectors $|\varphi_j\rangle$ such that

$$\rho = \sum_j \lambda_j|\varphi_j\rangle\langle\varphi_j|. \tag{4.10}$$

Since ρ satisfies the trace condition,

$$1 = \mathrm{tr}[\rho] = \sum_j \lambda_j\, \mathrm{tr}[|\varphi_j\rangle\langle\varphi_j|] = \sum_j \lambda_j, \tag{4.11}$$

and due to the positivity condition, $\lambda_j \geq 0$ for all j. Thus we can interpret the eigenvalues λ_j as probabilities $\rightsquigarrow \{\lambda_j, |\varphi_j\rangle\}$ is an ensemble and gives rise to ρ.

$\square$

From now on, we *define* a **density operator** as positive operator ρ with $\text{tr}[\rho] = 1$.

Language regarding density operators:

"pure state"	versus	"mixed state"

"pure state"

Quantum system in a state $|\psi\rangle$, corresponding density operator $\rho = |\psi\rangle\langle\psi|$, such that

$$\text{tr}[\rho^2] = \text{tr}[|\psi\rangle\langle\psi|\psi\rangle\langle\psi|] = \langle\psi|\psi\rangle^2 = 1.$$

"mixed state"

Density operator ρ describing quantum system cannot be written as $\rho = |\psi\rangle\langle\psi|$; intuition: in the ensemble representation $\{p_i, |\psi_i\rangle\}$ of ρ, all probabilities p_i are strictly smaller than 1. Then

$$\text{tr}[\rho^2] = \sum_i p_i^2 < 1.$$

In general: let ρ be a density operator. Then $\text{tr}[\rho^2] \le 1$, and $\text{tr}[\rho^2] = 1$ if and only if ρ describes a pure quantum state.

Proof. Denote the eigenvalues of ρ by $\{\lambda_i\}$, then $0 \le \lambda_i \le 1$ since ρ is positive and $1 = \text{tr}[\rho] = \sum_i \lambda_i$. Moreover,

$$\text{tr}[\rho^2] = \sum_i \lambda_i^2 \le 1, \tag{4.12}$$

with "$= 1$" precisely if one eigenvalue is 1 and the others are 0. $\qquad\square$

Ensemble representation is not unique! Example:

$$\rho = \frac{3}{4}|0\rangle\langle0| + \frac{1}{4}|1\rangle\langle1| = \frac{1}{2}|a\rangle\langle a| + \frac{1}{2}|b\rangle\langle b| \tag{4.13}$$

with

$$|a\rangle = \sqrt{\frac{3}{4}}|0\rangle + \sqrt{\frac{1}{4}}|1\rangle \tag{4.14a}$$

$$|b\rangle = \sqrt{\frac{3}{4}}|0\rangle - \sqrt{\frac{1}{4}}|1\rangle \tag{4.14b}$$

(But note that $|0\rangle, |1\rangle$ are the unique (up to phase factors) eigenvectors of ρ in this example, and that $\langle a|b\rangle \neq 0$.)

For the following: given an ensemble $\{p_i, |\psi_i\rangle\}$, set $|\tilde{\psi}_i\rangle = \sqrt{p_i}|\psi_i\rangle$ such that $\rho = \sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|$. We say that the ensemble $\{|\tilde{\psi}_i\rangle\}$ *generates* the density operator ρ. (The vectors $|\tilde{\psi}_i\rangle$ are not normalized in general.)

To relate an ensemble $\{|\tilde{\psi}_i\rangle\}_{i=1,\dots,m}$ to another $\{|\tilde{\varphi}_j\rangle\}_{j=1,\dots,n}$ in case $m \neq n$, we "pad" one of the ensembles with zero vectors, such that without loss of generality $m = n$.

Unitary freedom in the ensemble for density matrices: The sets $\{|\tilde{\psi}_i\rangle\}$ and $\{|\tilde{\varphi}_j\rangle\}$ generate the same density matrix if and only if

$$|\tilde{\psi}_i\rangle = \sum_j u_{ij}|\tilde{\varphi}_j\rangle \tag{4.15}$$

for some unitary matrix (u_{ij}).

Sketch of proof.

"$\Leftarrow$" Insert definitions.

"⇒" Use the spectral decomposition of the density matrix, $\rho = \sum_k \lambda_k |\chi_k\rangle\langle\chi_k|$ with $\langle\chi_k|\chi_\ell\rangle = \delta_{k\ell}$, set $|\tilde{\chi}_k\rangle = \sqrt{\lambda_k}|\chi_k\rangle$, express $|\tilde{\psi}_i\rangle = \sum_k v_{ik}|\tilde{\chi}_k\rangle$ for some complex coefficients v_{ik}. Then

$$\sum_k |\tilde{\chi}_k\rangle\langle\tilde{\chi}_k| = \rho = \sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| = \sum_{k,\ell} \left(\sum_i v_{ik}v_{i\ell}^*\right) |\tilde{\chi}_k\rangle\langle\tilde{\chi}_\ell|. \tag{4.16}$$

This equation can only be satisfied (since the vectors $|\tilde{\chi}_k\rangle$ are orthogonal and thus $|\tilde{\chi}_k\rangle\langle\tilde{\chi}_\ell|$ linearly independent) if

$$\sum_i v_{ik}v_{i\ell}^* = \delta_{k\ell}, \tag{4.17}$$

in other words, if (v_{ik}) is a unitary matrix. By the same arguments, $|\tilde{\varphi}_j\rangle = \sum_k w_{jk}|\tilde{\chi}_k\rangle$ for a unitary matrix (w_{jk}). Thus

$$|\tilde{\psi}_i\rangle = \sum_k v_{ik}|\tilde{\chi}_k\rangle = \sum_{k,j} v_{ik}w_{jk}^*|\tilde{\varphi}_j\rangle = \sum_j \left(vw^\dagger\right)_{ij}|\tilde{\varphi}_j\rangle, \tag{4.18}$$

and $vw^\dagger$ is (as product of two unitary matrices) again unitary.

$\square$

The Bloch sphere picture for qubits can be generalized to mixed states by the representation

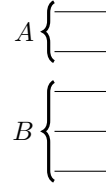$$\rho = \frac{I + \vec{r}\cdot\vec{\sigma}}{2}, \tag{4.19}$$

with $\vec{r} \in \mathbb{R}^3$ the Bloch vector of ρ (see Exercise 6.2).

4.3 The reduced density operator

(Nielsen and Chuang 2010, section 2.4.3)

Consider a composite quantum system consisting of subsystems A and B, for example: A: m qubits, B: n qubits.
Let the quantum system be described by a density operator denoted ρ^{AB}.

Define the **reduced density operator** for system A by

$$\rho^A = \mathrm{tr}_B[\rho^{AB}], \tag{4.20}$$

where tr_B is the *partial trace* over system B (see below), and analogously

$$\rho^B = \mathrm{tr}_A[\rho^{AB}]. \tag{4.21}$$

Examples:

- For any quantum states $|a_1\rangle, |a_2\rangle \in A$ and $|b_1\rangle, |b_2\rangle \in B$:

$$\mathrm{tr}_B[|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|] = |a_1\rangle\langle a_2| \cdot \mathrm{tr}[|b_1\rangle\langle b_2|] = |a_1\rangle\langle a_2| \cdot \langle b_2|b_1\rangle, \tag{4.22}$$

which can actually be used as definition of the partial trace, together with requiring linearity. Note that in this equation, $\langle b_2|b_1\rangle$ is a complex number, whereas $|a_1\rangle\langle a_2|$ is a matrix (column vector times a row vector). Also note that $|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2| = |a_1 b_1\rangle\langle a_2 b_2|$.

- Given a density matrix ρ for subsystem A and a density matrix σ for subsystem B, suppose that the overall density matrix is the Kronecker product of ρ and σ:

$$\rho^{AB} = \rho \otimes \sigma. \tag{4.23}$$

Then

$$\operatorname{tr}_B[\rho \otimes \sigma] = \rho \cdot \underbrace{\operatorname{tr}[\sigma]}_{=1} = \rho, \qquad \operatorname{tr}_A[\rho \otimes \sigma] = \underbrace{\operatorname{tr}[\rho]}_{=1} \cdot \sigma = \sigma. \tag{4.24}$$

- $\rho^{AB} = |\psi\rangle\langle\psi|$ with $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = |\beta_{00}\rangle$ (Bell state), and A: first qubit, B: second qubit. Explicitly expanding ρ^{AB} leads to

$$\rho^{AB} = \frac{1}{2}\big(|00\rangle + |11\rangle\big)\big(\langle 00| + \langle 11|\big) = \frac{1}{2}\big(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|\big). \tag{4.25}$$

Now Eq. (4.22) allows to compute the partial trace over subsystem B:

$$
\begin{aligned}
\rho^A = \operatorname{tr}_B[\rho^{AB}] &= \frac{1}{2}\big(\operatorname{tr}_B[|00\rangle\langle 00|] + \operatorname{tr}_B[|00\rangle\langle 11|] + \operatorname{tr}_B[|11\rangle\langle 00|] + \operatorname{tr}_B[|11\rangle\langle 11|]\big) \\
&= \frac{1}{2}\big(|0\rangle\langle 0| \underbrace{\langle 0|0\rangle}_{=1} + |0\rangle\langle 1| \underbrace{\langle 1|0\rangle}_{=0} + |1\rangle\langle 0| \underbrace{\langle 0|1\rangle}_{=0} + |1\rangle\langle 1| \underbrace{\langle 1|1\rangle}_{=1}\big) \\
&= \frac{1}{2}\big(|0\rangle\langle 0| + |1\rangle\langle 1|\big) = \frac{I}{2}.
\end{aligned}
\tag{4.26}
$$

Note that the composite system is in the "pure state" $|\psi\rangle$, whereas the subsystem is described by the "mixed state" $\frac{I}{2}$. (This is indeed a mixed state since $\operatorname{tr}[(\frac{I}{2})^2] = \frac{1}{4}\operatorname{tr}[I] = \frac{1}{2} < 1$.)

Motivation/justification for partial trace: Let M be any observable on subsystem A, then we want that ρ^A yields the same statistics for measuring M as ρ^{AB} for measuring $M \otimes I$, where I here denotes the identity matrix on subsystem B. In particular:

$$\langle M \rangle = \operatorname{tr}[M\rho^A] \overset{!}{=} \operatorname{tr}[(M \otimes I)\rho^{AB}] = \langle M \otimes I \rangle \tag{4.27}$$
$$\underset{\text{on } A}{\nwarrow} \qquad \underset{\text{on } AB}{\nwarrow}$$

for all density operators ρ^{AB}. The partial trace operation for computing ρ^A from ρ^{AB} is the unique operation with this property (Nielsen and Chuang 2010, Box 2.6).

Application to quantum teleportation: Why does quantum teleportation not allow for faster-than-light communication via the instantaneous wavefunction collapse?

Recall the corresponding quantum circuit (see Sect. 3.1):
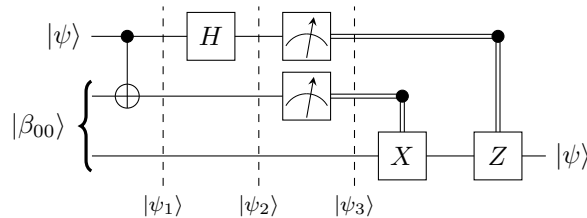
At $|\psi_3\rangle$, Alice has completed her measurements (her qubits have "collapsed"), but Bob does not know her measurements results yet.

Intermediate state $|\psi_2\rangle$ (see Eq. (3.5) above):

$$|\psi_2\rangle = \frac{1}{2}\Big(|00\rangle(\alpha|0\rangle + \beta|1\rangle) + |01\rangle(\alpha|1\rangle + \beta|0\rangle) + |10\rangle(\alpha|0\rangle - \beta|1\rangle) + |11\rangle(\alpha|1\rangle - \beta|0\rangle)\Big). \tag{4.28}$$

Thus, directly after Alice's measurements, system is in state (from Bob's perspective, who does not know the measurement results yet):

$$|\varphi_1\rangle = |00\rangle(\alpha|0\rangle + \beta|1\rangle) \quad \text{with probability } \tfrac{1}{4}, \tag{4.29a}$$

$$|\varphi_2\rangle = |01\rangle(\alpha|1\rangle + \beta|0\rangle) \quad \text{with probability } \tfrac{1}{4}, \tag{4.29b}$$

$$|\varphi_3\rangle = |10\rangle(\alpha|0\rangle - \beta|1\rangle) \quad \text{with probability } \tfrac{1}{4}, \tag{4.29c}$$

$$|\varphi_4\rangle = |11\rangle(\alpha|1\rangle - \beta|0\rangle) \quad \text{with probability } \tfrac{1}{4}. \tag{4.29d}$$

Corresponding density matrix of ensemble $\{\tfrac{1}{4}, |\varphi_j\rangle\}_{j=1,\dots,4}$:

$$
\begin{aligned}
\rho^{AB} &= \frac{1}{4}\sum_{j=1}^{4} |\varphi_j\rangle\langle\varphi_j| \\
&= \frac{1}{4}\big(|00\rangle\langle00| \otimes (\alpha|0\rangle + \beta|1\rangle)(\alpha^*\langle0| + \beta^*\langle1|) + |01\rangle\langle01| \otimes (\alpha|1\rangle + \beta|0\rangle)(\alpha^*\langle1| + \beta^*\langle0|) \\
&\quad + |10\rangle\langle10| \otimes (\alpha|0\rangle - \beta|1\rangle)(\alpha^*\langle0| - \beta^*\langle1|) + |11\rangle\langle11| \otimes (\alpha|1\rangle - \beta|0\rangle)(\alpha^*\langle1| - \beta^*\langle0|)\big).
\end{aligned}
\tag{4.30}
$$

In this expression, the terms corresponding to the first two (Alice's) qubits are of the form $|a_1 a_2\rangle\langle a_1 a_2|$ with $a_1, a_2 \in \{0,1\}$, and tracing them out gives $\operatorname{tr}_A[|a_1 a_2\rangle\langle a_1 a_2|] = \langle a_1 a_2 | a_1 a_2\rangle = \langle a_1 | a_1\rangle\langle a_2 | a_2\rangle = 1$. Thus the reduced density operator describing Bob's qubit is:

$$
\begin{aligned}
\rho^B &= \operatorname{tr}_A[\rho^{AB}] \\
&= \frac{1}{4}\big((\alpha|0\rangle + \beta|1\rangle)(\alpha^*\langle0| + \beta^*\langle1|) + (\alpha|1\rangle + \beta|0\rangle)(\alpha^*\langle1| + \beta^*\langle0|) \\
&\quad + (\alpha|0\rangle - \beta|1\rangle)(\alpha^*\langle0| - \beta^*\langle1|) + (\alpha|1\rangle - \beta|0\rangle)(\alpha^*\langle1| - \beta^*\langle0|)\big) \\
&= \frac{1}{4}\big(2(|\alpha|^2 + |\beta|^2)|0\rangle\langle0| + 2(|\alpha|^2 + |\beta|^2)|1\rangle\langle1|\big) \\
&= \frac{1}{2}\big(|0\rangle\langle0| + |1\rangle\langle1|\big) = \frac{I}{2},
\end{aligned}
\tag{4.31}
$$

independent of $|\psi\rangle$! (For the third equal sign we have used that the terms proportional to $|0\rangle\langle1|$ and $|1\rangle\langle0|$ cancel, and for the fourth equal sign that $|\alpha|^2 + |\beta|^2 = 1$.) Since $\rho^B = \frac{I}{2}$, any measurement by Bob cannot reveal any information about $|\psi\rangle$, i.e., Alice cannot transmit information (encoded in α, β) via the instantaneous wavefunction collapse to Bob.

5 The quantum Fourier transform and its applications

Most famous application: Shor's algorithm for factoring a n-bit integer in $\mathcal{O}(n^2 \log n \log \log n)$; for comparison: best known classical algorithm: number field sieve, $\exp(\Theta(n^{1/3} \log(n)^{2/3}))$ (with Θ: asymptotically the same).

5.1 The quantum Fourier transform

(Nielsen and Chuang 2010, section 5.1)

Discrete Fourier transform defined by

$$\mathcal{F}_N : \mathbb{C}^N \to \mathbb{C}^N, \quad \mathcal{F}_N(x) = y \quad \text{with} \quad y_k = \frac{1}{\sqrt{N}}\sum_{j=0}^{N-1} x_j\, e^{2\pi i j k/N} \quad \forall k = 0,\dots,N-1. \tag{5.1}$$

Note: $\mathcal{F}_N$ is a unitary transformation, i.e., $\langle \mathcal{F}_N(x') | \mathcal{F}_N(x)\rangle = \langle x' | x\rangle$ for all $x, x' \in \mathbb{C}^N$; in particular, the matrix representation of $\mathcal{F}_N$ is a unitary $\mathbb{C}^{N\times N}$ matrix.

Definition of **quantum Fourier transform (QFT)** acting on orthonormal basis $\{|0\rangle, |1\rangle, \ldots |N-1\rangle\}$:

$$|j\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{2\pi i jk/N} |k\rangle. \tag{5.2}$$

Thus, for a superposition of basis states as input:

$$\sum_{j=0}^{N-1} x_j |j\rangle \stackrel{\text{QFT}}{\mapsto} \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} e^{2\pi i jk/N} x_j |k\rangle = \sum_{k=0}^{N-1} \mathcal{F}_N(x)_k |k\rangle. \tag{5.3}$$

From now on: assume that N is a power of 2, i.e., $N = 2^n$ for some $n \in \mathbb{N}$; thus basis $\{|0\rangle, |1\rangle, \ldots, |2^n - 1\rangle\}$ can be regarded as computational basis of a n-qubit quantum register.

For $j \in \{0, 1, \ldots, 2^n - 1\}$, use binary representation $j = j_{n-1} j_{n-2} \cdots j_1 j_0$ to describe input:

$$|j_{n-1}\rangle \;\text{———}$$

$$|j_{n-2}\rangle \;\text{———}$$

$$\vdots \qquad\qquad \vdots$$

$$|j_1\rangle \;\text{———}$$

$$|j_0\rangle \;\text{———}$$

For the following: extend binary representation to floating-point, i.e.,

$$0.abcd\cdots = \frac{a}{2} + \frac{b}{4} + \frac{c}{8} + \frac{d}{16} + \ldots \tag{5.4}$$

With this notation, the quantum Fourier transform can be represented in product form:

$$|j_{n-1} \ldots j_1 j_0\rangle \stackrel{\text{QFT}}{\mapsto} \frac{1}{2^{n/2}} \left(|0\rangle + e^{2\pi i 0.j_0} |1\rangle\right) \left(|0\rangle + e^{2\pi i 0.j_1 j_0} |1\rangle\right) \cdots \left(|0\rangle + e^{2\pi i 0.j_{n-1} \ldots j_1 j_0} |1\rangle\right) \tag{5.5}$$

Check:

$$|j\rangle \stackrel{\text{QFT}}{\mapsto} \frac{1}{2^{n/2}} \sum_{k=0}^{2^n - 1} e^{2\pi i jk/2^n} |k\rangle$$

$$= \frac{1}{2^{n/2}} \sum_{k_{n-1}=0}^{1} \cdots \sum_{k_0=0}^{1} e^{2\pi i j \left(\sum_{\ell=1}^{n} k_{n-\ell} 2^{-\ell}\right)} |k_{n-1} \ldots k_1 k_0\rangle$$

$$= \frac{1}{2^{n/2}} \bigotimes_{\ell=1}^{n} \left(\sum_{k_{n-\ell}=0}^{1} e^{2\pi i j k_{n-\ell} 2^{-\ell}} |k_{n-\ell}\rangle \right) \tag{5.6}$$

$$= \frac{1}{2^{n/2}} \bigotimes_{\ell=1}^{n} \left(|0\rangle + e^{2\pi i j \, 2^{-\ell}} |1\rangle\right)$$

$$= \frac{1}{2^{n/2}} \left(|0\rangle + e^{2\pi i \, 0.j_0} |1\rangle\right) \otimes \left(|0\rangle + e^{2\pi i \, 0.j_1 j_0} |1\rangle\right) \otimes \cdots \otimes \left(|0\rangle + e^{2\pi i \, 0.j_{n-1} \ldots j_0} |1\rangle\right).$$

For the last equal sign we have used that $e^{2\pi i q} = 1$ for all $q \in \mathbb{Z}$, thus, e.g., $e^{2\pi i \, ab.cd} = e^{2\pi i \, 0.cd}$.

Eq. (5.5) can be realized by the following quantum circuit (up to reversing the order of the qubits at the end), with

$$R_k := \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i / 2^k} \end{pmatrix} : \tag{5.7}$$
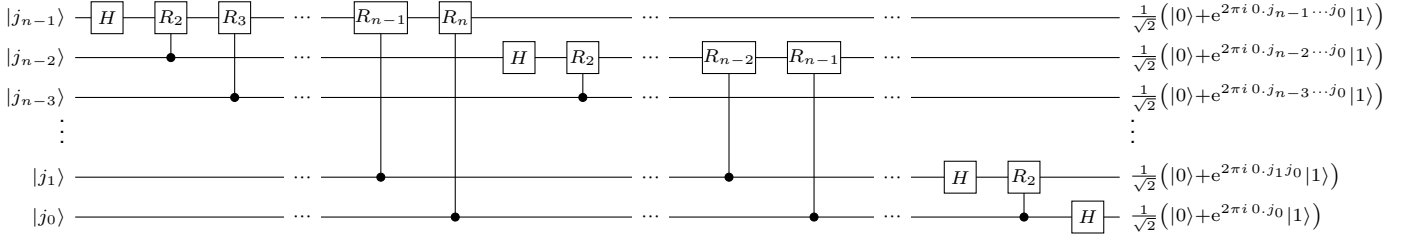
To verify that this circuit works as intended:

- Consider first (top) qubit with input $|j_{n-1}\rangle$: action of Hadamard gate (2.16) can be represented as:

$$H|x\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + (-1)^x|1\rangle\big) \quad \text{for} \quad x \in \{0,1\}. \tag{5.8}$$

Note that $(-1)^x = e^{2\pi i\, 0.x}$ since $e^0 = 1$, $e^{2\pi i\, 0.5} = e^{\pi i} = -1$. Thus the quantum state after the first Hadamard gate equals

$$\frac{1}{\sqrt{2}}\big(|0\rangle + e^{2\pi i\, 0.j_{n-1}}|1\rangle\big)|j_{n-2}\ldots j_1 j_0\rangle. \tag{5.9}$$

- State after first controlled-R_2 gate (acts on first qubit, but only if $j_{n-2} = 1$), using the representation $e^{2\pi i/4} = e^{2\pi i\, 0.01}$ for R_2 and that $e^{2\pi i\, 0.j_{n-1}} e^{2\pi i\, 0.0j_{n-2}} = e^{2\pi i\, 0.j_{n-1}j_{n-2}}$:

$$\frac{1}{\sqrt{2}}\big(|0\rangle + e^{2\pi i\, 0.j_{n-1}j_{n-2}}|1\rangle\big)|j_{n-2}\ldots j_1 j_0\rangle. \tag{5.10}$$

- State after all controlled-R_k gates have acted on first qubit:

$$\frac{1}{\sqrt{2}}\big(|0\rangle + e^{2\pi i\, 0.j_{n-1}\ldots j_1 j_0}|1\rangle\big)|j_{n-2}\ldots j_1 j_0\rangle. \tag{5.11}$$
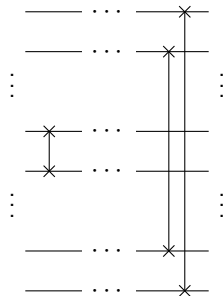
- State after Hadamard gate acting on second qubit:

$$\frac{1}{\sqrt{2}}\big(|0\rangle + e^{2\pi i\, 0.j_{n-1}\ldots j_1 j_0}|1\rangle\big)\frac{1}{\sqrt{2}}\big(|0\rangle + e^{2\pi i\, 0.j_{n-2}}|1\rangle\big)|j_{n-3}\ldots j_1 j_0\rangle. \tag{5.12}$$

$\vdots$

- Output:

$$\frac{1}{2^{n/2}}\big(|0\rangle + e^{2\pi i\, 0.j_{n-1}\ldots j_1 j_0}|1\rangle\big)\big(|0\rangle + e^{2\pi i\, 0.j_{n-2}\ldots j_1 j_0}|1\rangle\big)\cdots\big(|0\rangle + e^{2\pi i\, 0.j_0}|1\rangle\big). \tag{5.13}$$

Still required: reverse order of qubits at the end, can be implemented by swap gates:

Remark: quantum circuit provides another confirmation that quantum Fourier transform is unitary, since each gate is unitary.

Number of gates:

# gates	description	
n	Hadamard and controlled-R_k gates acting on first qubit	
$n-1$		second qubit
$\vdots$	$\vdots$	
1		last qubit
$\frac{1}{2}n(n+1)$		

Final reversing of qubit order: $\frac{n}{2}$ (n even) or $\frac{n-1}{2}$ (n odd) swap gates
In summary: quantum Fourier transform requires $\mathcal{O}(n^2)$ gates.

For comparison: classical FFT: $\mathcal{O}(n\, 2^n)$ operations (exponential in n)

Practical difficulty in using quantum Fourier transform: cannot measure output amplitudes directly.

Excursion: short review of Fourier transformation   Recall definition (see Eq. (5.1) above):

$$\mathcal{F}_N : \mathbb{C}^N \to \mathbb{C}^N, \quad \mathcal{F}_N(x) = y \quad \text{with} \quad y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j\, e^{2\pi ijk/N} \quad \forall k = 0, \ldots, N-1. \quad (5.14)$$

Intuition: vector x is discretization of a periodic function $f : [0,1] \to \mathbb{C}$ (for example, sampling of an audio signal):
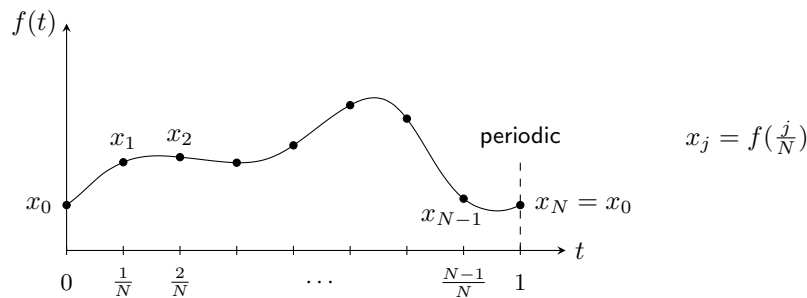
Since f is periodic, Fourier analysis states that f can be decomposed into a sum of oscillations with different frequencies (Fourier series):
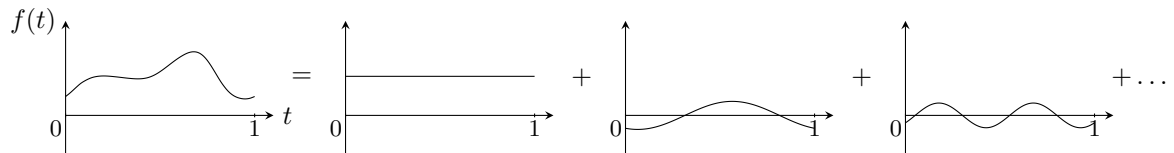
Mathematically:

$$f(t) = \sum_{k=-\infty}^{\infty} y_k\, e^{-2\pi ikt}, \quad (5.15)$$

with $e^{-2\pi ikt}$ an "oscillator" with frequency $2\pi k$, and y_k the corresponding amplitude.

If we only consider f evaluated at discrete (uniformly spaced) points $0, \frac{1}{N}, \ldots \frac{N-1}{N}$, then a finite sum $\sum_{k=0}^{N-1} y_k\, e^{-2\pi ikt}$ is sufficient to recover f at these points.

Discrete Fourier transform $\mathcal{F}_N$ finds coefficients y_k given the discretization x of f.

$\mathcal{F}_N$ is a unitary transformation, i.e., its matrix representation is unitary (see also Tutorial 8); thus $\mathcal{F}_N^{-1} = \mathcal{F}^\dagger$. Intuition: "Fourier modes" $(\frac{1}{\sqrt{N}} e^{-2\pi ijk/N})_{j=0,\dots,N-1}$ are orthonormal:

$$\left\langle \begin{pmatrix} \frac{1}{\sqrt{N}} e^{-2\pi i\, 0\cdot k'/N} \\ \frac{1}{\sqrt{N}} e^{-2\pi i\, 1\cdot k'/N} \\ \vdots \\ \frac{1}{\sqrt{N}} e^{-2\pi i(N-1)k'/N} \end{pmatrix} \middle| \begin{pmatrix} \frac{1}{\sqrt{N}} e^{-2\pi i\, 0\cdot k/N} \\ \frac{1}{\sqrt{N}} e^{-2\pi i\, 1\cdot k/N} \\ \vdots \\ \frac{1}{\sqrt{N}} e^{-2\pi i(N-1)k/N} \end{pmatrix} \right\rangle = \frac{1}{N} \sum_{j=0}^{N-1} e^{2\pi ij(k'-k)/N} = \delta_{k,k'} \qquad (5.16)$$

for any $k, k' \in \{0, \dots, N-1\}$.

Since $\mathcal{F}_N^{-1} = \mathcal{F}^\dagger$, we can identify the inverse discrete Fourier transform (compared to (5.15), only a finite sum is required here):

$$x_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} y_k\, e^{-2\pi ijk/N} \qquad (5.17)$$

$$\underset{\substack{\| \\ f(\frac{j}{N})}}{}$$

In summary:

$$x \xrightleftharpoons[\mathcal{F}_N^{-1}]{\mathcal{F}_N} y$$

5.2 Phase estimation

(Nielsen and Chuang 2010, section 5.2)

"Auxiliary step" for many quantum algorithms

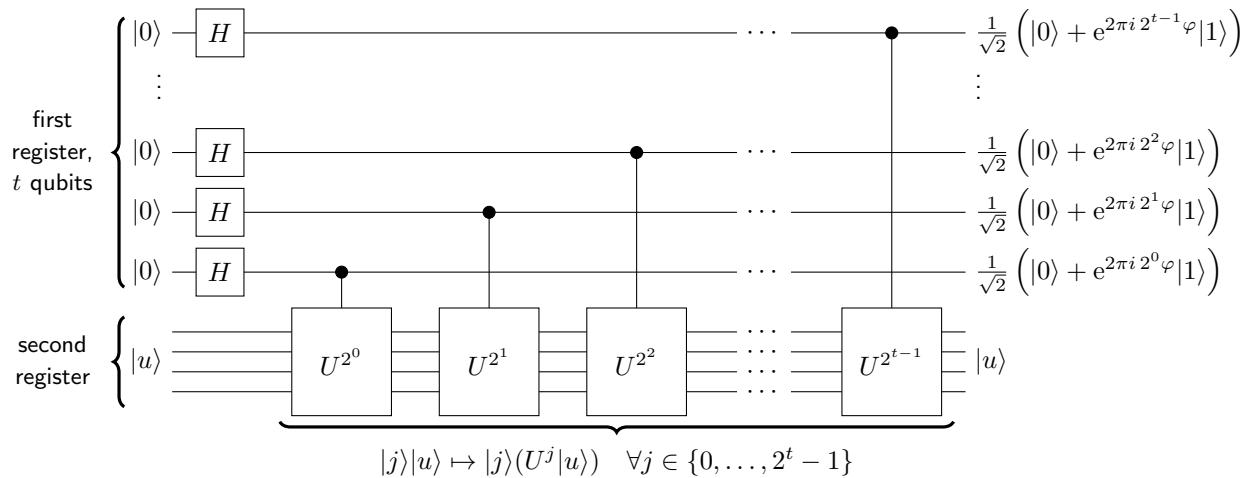Setup: unitary operator U with eigenvector $|u\rangle$ and corresponding (unknown) eigenvalue $e^{2\pi i\varphi}$:

$$U|u\rangle = e^{2\pi i\varphi}|u\rangle \qquad (5.18)$$

(without loss of generality $0 \le \varphi \le 1$)

Goal: compute φ, assuming that there is an *oracle* (black box) available which can prepare $|u\rangle$ and perform controlled-U^{2^j} operations for any $j \in \mathbb{N}$

(Note: $U^{2^j}|u\rangle = \underbrace{U \cdots U}_{2^j \text{ times}}|u\rangle = e^{2\pi i\, 2^j \varphi}|u\rangle$)

First stage:

$$|j\rangle|u\rangle \mapsto |j\rangle(U^j|u\rangle) \quad \forall j \in \{0, \dots, 2^t - 1\}$$

Second register always remains in eigenstate $|u\rangle$.

Output of first register:

$$|\chi_1\rangle = \frac{1}{2^{t/2}}\left(|0\rangle + \mathrm{e}^{2\pi i\,2^{t-1}\varphi}|1\rangle\right) \otimes \cdots \otimes \left(|0\rangle + \mathrm{e}^{2\pi i\,2^1\varphi}|1\rangle\right) \otimes \left(|0\rangle + \mathrm{e}^{2\pi i\,2^0\varphi}|1\rangle\right) = \frac{1}{2^{t/2}}\sum_{k=0}^{2^t-1}\mathrm{e}^{2\pi i\varphi k}|k\rangle \tag{5.19}$$

Second stage: Inverse Fourier transform applied to first register (e.g., by reversing the circuit for the forward Fourier transform and taking the adjoint of all gates)

Recall definition of quantum Fourier transform (see (5.2) above):

$$|j\rangle \mapsto \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}\mathrm{e}^{2\pi ijk/N}|k\rangle, \quad \text{here: } N = 2^t \tag{5.20}$$

By comparing with the output of the first stage in Eq. (5.19): if φ can be written exactly as $\varphi = \frac{j}{2^t}$ for some $j \in \{0, 1, \ldots, 2^t - 1\}$ (equivalently: if φ can be exactly represented using t bits: $\varphi = 0.\varphi_{t-1}\ldots\varphi_1\varphi_0$), then the inverse Fourier transform results in state $|\varphi_{t-1}\ldots\varphi_1\varphi_0\rangle \rightsquigarrow$ measurement returns φ exactly!

Analysis for general φ: final quantum state in first register after inverse Fourier transform (applied to output of first stage, see Eq. (5.19)):

$$|\chi_2\rangle := \mathcal{F}_{2^t}^{-1}|\chi_1\rangle = \frac{1}{2^t}\sum_{k,\ell=0}^{2^t-1}\mathrm{e}^{-2\pi ik\ell/2^t}\mathrm{e}^{2\pi i\varphi k}|\ell\rangle = \sum_{\ell=0}^{2^t-1}\underbrace{\frac{1}{2^t}\sum_{k=0}^{2^t-1}\mathrm{e}^{2\pi i(\varphi-\ell/2^t)k}}_{=:a_\ell}|\ell\rangle = \sum_{\ell=0}^{2^t-1}a_\ell|\ell\rangle \tag{5.21}$$

Example for $t = 3$, $\varphi = 0.65$:
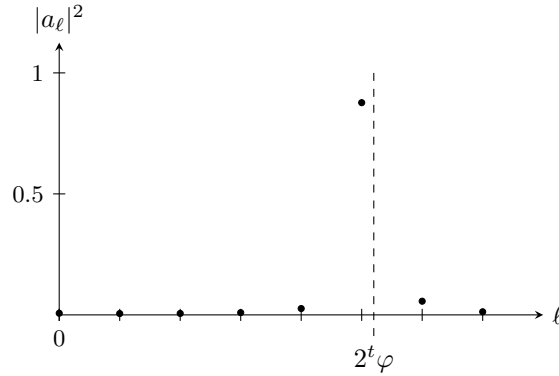
Let $b \in \{0, 1, \ldots, 2^t - 1\}$ be the largest integer such that

$$\frac{b}{2^t} = 0.b_{t-1}\ldots b_1 b_0 \leq \varphi, \tag{5.22}$$

i.e., $\delta := \varphi - \frac{b}{2^t}$ satisfies $0 \leq \delta \leq 2^{-t}$, then the following result holds (see Nielsen and Chuang 2010, section 5.2.1 for a derivation). Denote the final measurement outcome by m, and a given "deviation" tolerance by $e_{\mathsf{tol}} \in \mathbb{N}$, then the probability for a deviation larger than the tolerance is bounded by

$$\mathbb{P}\big(|m - b| > e_{\mathsf{tol}}\big) \leq \frac{1}{2\left(e_{\mathsf{tol}} - 1\right)}. \tag{5.23}$$

(In other words: the larger the deviation, the more unlikely it is to occur.)
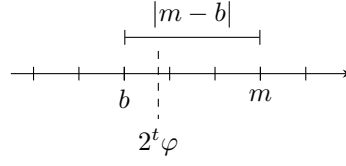
For a desired accuracy 2^{-n} (n bits), we have to choose $e_{\text{tol}} = 2^{t-n} - 1$; for $t = n + p$ qubits in the first register, this happens with probability
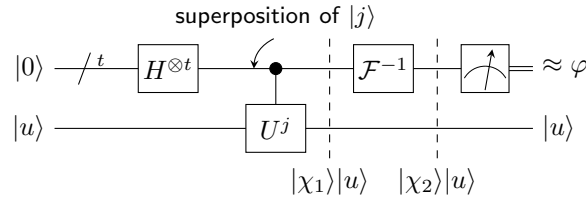
$$\mathbb{P}\big(|m-b| \leq e_{\text{tol}}\big) = 1 - \mathbb{P}\big(|m-b| > e_{\text{tol}}\big) \overset{(5.23)}{\geq} 1 - \frac{1}{2\,(e_{\text{tol}} - 1)} = 1 - \underbrace{\frac{1}{2\,(2^p - 2)}}_{=:\epsilon} = 1 - \epsilon. \quad (5.24)$$

Solving for p leads to $2\epsilon = \frac{1}{2^p - 2} \Leftrightarrow 2^p = 2 + \frac{1}{2\epsilon} \Leftrightarrow p = \log_2(2 + \frac{1}{2\epsilon})$. Thus: to obtain φ accurate to n bits with probability of at least $1 - \epsilon$, choose

$$t = n + \left\lceil \log_2\left(2 + \frac{1}{2\epsilon}\right) \right\rceil. \quad (5.25)$$

($\lceil \cdot \rceil$: round "upwards" to next integer)

Schematic of overall algorithm:

Algorithm: quantum phase estimation

Inputs:

(a) black box for performing controlled-U^j operations for integer j

(b) eigenstate $|u\rangle$ of U with (unknown) eigenvalue $e^{2\pi i\varphi}$

(c) $t = n + \lceil \log_2(2 + \frac{1}{2\epsilon}) \rceil$ qubits initialized to $|0\rangle$, given a desired accuracy 2^{-n} and success probability $1 - \epsilon$

Output: n-bit approximation $\tilde{\varphi}$ to φ

1:	$|0\rangle|u\rangle$	initial state
2:	$\mapsto \dfrac{1}{\sqrt{2^t}} \displaystyle\sum_{j=0}^{2^t - 1} |j\rangle|u\rangle$	create superposition (Hadamard gates)
3:	$\mapsto \dfrac{1}{\sqrt{2^t}} \displaystyle\sum_{j=0}^{2^t - 1} |j\rangle(U^j|u\rangle)$	apply black box
	$= \dfrac{1}{\sqrt{2^t}} \displaystyle\sum_{j=0}^{2^t - 1} e^{2\pi ij\varphi}|j\rangle|u\rangle$	
4:	$\mapsto |\chi_2\rangle|u\rangle$	inverse Fourier transform
5:	$\mapsto \tilde{\varphi}$	measure first register

Remark: Algorithm can also be run even if one cannot prepare an eigenstate $|u\rangle$ of U: consider a general state $|\psi\rangle$ as input of the second register; $|\psi\rangle$ can be represented as linear combination of eigenstates $(|u_k\rangle)_k$ of U:

$$|\psi\rangle = \sum_k c_k |u_k\rangle \tag{5.26}$$

Denote eigenvalue corresponding to eigenstate $|u_k\rangle$ by $\mathrm{e}^{2\pi i\varphi_k}$;
$\rightsquigarrow$ output of algorithm:

$$\sum_k c_k |\chi_{2,k}\rangle |u_k\rangle \tag{5.27}$$

with $|\chi_{2,k}\rangle$ an approximation of φ_k ($|\chi_{2,k}\rangle$ puts largest weight on computational basis states corresponding to digital approximation of φ_k)
$\rightsquigarrow$ measurement results in φ_k accurate to n bits with probability of at least $|c_k|^2(1-\epsilon)$

5.3 Applications: order-finding and factoring

(Nielsen and Chuang 2010, section 5.3)

Integer factoring, e.g., $15 = 3 \cdot 5$

Order-finding problem and factoring problem are equivalent in the sense that solving one problem efficiently allows to solve the other efficiently, too.

5.3.1 Order-finding

Given positive integers x and N with $x < N$ and no common factors, that is, $\gcd(x, N) = 1$, the **order** of x modulo N is the least positive integer r such that

$$x^r = 1 \mod N. \tag{5.28}$$

Note: the order is the period of the function $f : a \mapsto x^a \mod N$, i.e., $f(a+r) = f(a)$ for all a.

In the following: denote the number of bits required to specify N by L, i.e., $L = \lceil \log_2(N) \rceil$

No known classical algorithm for order-finding exists with runtime polynomial in L!

Quantum algorithm for order-finding: apply phase estimation to the following unitary operator U acting on L qubits:

$$U|y\rangle = \begin{cases} |x \cdot y \mod N\rangle, & 0 \le y < N \quad \xleftarrow{\text{only this case relevant here}} \\ |y\rangle, & N \le y < 2^L \end{cases} \tag{5.29}$$

Example: $N = 12$ (thus $L = 4$), $x = 7$:

$$U|5\rangle = |7 \cdot 5 \mod 12\rangle = |35 \mod 12\rangle = |11\rangle \tag{5.30}$$

(For mapping a state like $|5\rangle$ to a qubit register, we identify the number with its binary representation, i.e., $|5\rangle = |0101_2\rangle \equiv |0\rangle|1\rangle|0\rangle|1\rangle$)

U is indeed unitary (permutation of computational basis states); requires that x and N have no common factors; thus, in particular, x has an inverse modulo N: there exists an integer $\tilde{x}$ such that $x\tilde{x} = 1 \mod N$.

Remark: controlled-U^j operation can be realized by "modular exponentiation" (Nielsen and Chuang 2010, Box 5.2), requiring $\mathcal{O}(L^3)$ gates.

Let r be the order of x modulo N, define

$$|u_s\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} \mathrm{e}^{-2\pi i sk/r} |x^k \mod N\rangle \quad \text{for } s = 0, 1, \dots, r-1 \tag{5.31}$$

Each $|u_s\rangle$ is an eigenstate of U with eigenvalue $\mathrm{e}^{2\pi i s/r}$:

$$
\begin{aligned}
U|u_s\rangle &= \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} \mathrm{e}^{-2\pi i s k/r} |x^{k+1} \mod N\rangle \\
&= \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} \mathrm{e}^{2\pi i s/r} \mathrm{e}^{-2\pi i s(k+1)/r} |x^{k+1} \mod N\rangle = \mathrm{e}^{2\pi i s/r} |u_s\rangle
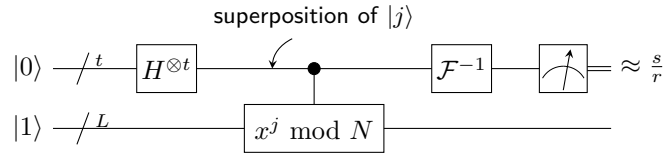\end{aligned}
\tag{5.32}
$$

Phase estimation would allow us to (approximately) obtain $\frac{s}{r}$; however, preparing $|u_s\rangle$ would require knowledge of r, which we are trying to compute!

Instead: prepare as initial state the superposition (cf. Exercise 9.2(a))

$$
\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} |u_s\rangle = |1\rangle
\tag{5.33}
$$

Applying U^j gives: $U^j|1\rangle = U^{j-1}|x \cdot 1 \mod N\rangle = \cdots = |x^j \mod N\rangle$

Schematic overview:

Set $t = n + \lceil \log_2(2 + \frac{1}{2\epsilon}) \rceil$ as in Eq. (5.25) above. For $s \in \{0, 1, \ldots, r-1\}$, the phase estimation algorithm returns an estimate $\tilde{\varphi}$ of $\frac{s}{r}$ accurate to n bits with probability at least $(1-\epsilon)/r$ (division by r is due to the equal superposition of the states $|u_s\rangle$, see (5.33).)

Finally, compute s and r given $\tilde{\varphi}$ via *continued fractions algorithm* (Nielsen and Chuang 2010, Box 5.3), requires accuracy $n = 2L + 1$ bits.

In case s and r have a common factor (which can be canceled from $\frac{s}{r}$), the algorithm returns s' and r' with $\frac{s'}{r'} = \frac{s}{r}$ but $s' < s$ and $r' < r$. In particular, r' is not the sought order of x modulo N. One solution to this problem: repeat algorithm twice to obtain (s'_1, r'_1) and (s'_2, r'_2), then the order r is (with some probability) the least common multiple of r'_1 and r'_2.

5.3.2 Factoring

Definition **prime factorization**: given a composite integer N, what are the prime numbers which – when multiplied together – equal N?

In the following: try to find a non-trivial factor of N.

Strategy: reduction of factoring to order-finding based on two observations:

(a) Can compute a non-trivial factor of N if we can find a non-trivial ($x \neq \pm 1 \mod N$) solution x to the equation $x^2 = 1 \mod N$

(b) Can show that a randomly chosen y which is co-prime to N (i.e., y and N do not have a common factor, $\gcd(y, N) = 1$) is likely to have an order r which is even, and be such that $y^{r/2} \neq \pm 1 \mod N$, thus $x = y^{r/2} \mod N$ is a solution of (a).

(Note that final algorithm is thus probabilistic.)

Observation (a) is based on the following theorem:

Theorem 1. *Suppose N is an L bit composite number, and x is a non-trivial solution to the equation $x^2 = 1 \mod N$ in the range $1 \leq x \leq N$, that is, neither $x = 1 \mod N$ nor $x = N - 1 = -1 \mod N$. Then at least one of $\gcd(x-1, N)$ and $\gcd(x+1, N)$ is a non-trivial factor of N that can be computed using $\mathcal{O}(L^3)$ operations.*

Proof. Since $x^2 = 1 \mod N$, N must divide $x^2 - 1 = (x+1)(x-1)$, thus N must have a common factor with one or the other of $(x + 1)$ and $(x - 1)$. By assumption, $x + 1 < N$ and thus also $x - 1 < N$, therefore the common factor cannot be N itself. Using Euclid's algorithm, one can compute $\gcd(x - 1, N)$ and $\gcd(x + 1, N)$ and thus obtain a non-trivial factor of N using $\mathcal{O}(L^3)$ operations. $\qquad\square$

Observation (b) is based on the following theorem:

Theorem 2. *Suppose $N = p_1^{\alpha_1} \cdots p_m^{\alpha_m}$ is the prime factorization of an odd composite integer. Let x be an integer chosen uniformly at random, subject to the requirements that $1 \leq x \leq N - 1$ and x is co-prime to N. Let r be the order of x modulo N. Then*

$$\mathbb{P}\big(r \text{ is even and } x^{r/2} \neq -1 \mod N\big) \geq 1 - \frac{1}{2^m}. \tag{5.34}$$

For a proof see Nielsen and Chuang 2010, appendix A4.3.

Combining observations (a) and (b) leads to the following algorithm:

Algorithm: reduction of factoring to order-finding

Input: composite integer N

Output: non-trivial factor of N

Runtime: $\mathcal{O}(\log_2(N)^3)$ operations; succeeds with probability $\mathcal{O}(1)$

1: If N is even, return the factor 2.

2: Determine whether $N = a^b$ for integers $a \geq 1$ and $b \geq 2$ (using a classical algorithm), and if so return the factor a.

3: Randomly choose x in the range $1 \leq x \leq N - 1$. If $\gcd(x, N) > 1$, then return the factor $\gcd(x, N)$.

4: Use the order-finding subroutine to find the order r of x modulo N.

5: If r is even and $x^{r/2} \neq -1 \mod N$, then compute $\gcd(x^{r/2} - 1, N)$ and $\gcd(x^{r/2} + 1, N)$, one of which must be a non-trivial factor, and return this factor; otherwise the algorithm fails.

Remarks:

- If the algorithm "fails", simply restart at step 3.

- The "fast return" in case $\gcd(x, N) > 1$ in step 3 ensures that the condition "x co-prime to N" in Theorem 2 is satisfied when proceeding to steps 4 and 5.

- Reaching step 5 implies that the algorithm did not return at step 2, and thus $m \geq 2$ in Theorem 2.

- Only the order-finding subroutine is "difficult", i.e., there is no known classical algorithm with runtime polynomial in $\log_2(N)$.

6 Quantum error-correction

(Nielsen and Chuang 2010, section 10)

6.1 Introduction

Principle of classical error-correction: encode data by adding redundant information, e.g., using the "repetition code"

$$0 \mapsto 0\,0\,0$$
$$1 \mapsto 1\,1\,1$$

Quantum version?

Difficulties:

- No-cloning theorem (see Tutorial 2) forbids to copy an arbitrary quantum state.
- Errors are continuous: amplitudes of a quantum state, e.g., $a, b \in \mathbb{C}$ for a single qubit $|\psi\rangle = a|0\rangle + b|1\rangle$, form "analog" information.
- Measurements potentially destroy quantum information (wavefunction collapse).

6.1.1 Three qubit bit flip code

Error model: flip $|0\rangle \leftrightarrow |1\rangle$ occurs with probability p, i.e., qubit $|\psi\rangle$ gets mapped to $X|\psi\rangle$ with probability p. This is also known as **bit flip channel**:

$$\rho' = (1-p)|\psi\rangle\langle\psi| + p\,X|\psi\rangle\langle\psi|X, \tag{6.1}$$
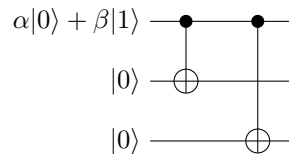
where ρ' denotes the output density matrix (cf. quantum ensembles in Sect. 4.1)

Idea: encode $\alpha|0\rangle + \beta|1\rangle$ using 3 qubits as $\alpha|000\rangle + \beta|111\rangle \rightsquigarrow$ logical qubits:

$$|0_L\rangle = |000\rangle \tag{6.2a}$$
$$|1_L\rangle = |111\rangle \tag{6.2b}$$

Corresponding circuit:

Now each qubit is sent through the bit flip channel; suppose at most a single flip occurred $\rightsquigarrow$ can restore original state as follows:

Error correction procedure

1. *Error detection* or *syndrome diagnosis*: perform measurement (*without destroying the quantum information!*) using the projection operators:

$$
\begin{aligned}
P_0 &= |000\rangle\langle000| + |111\rangle\langle111| & \text{no error occurred} & \tag{6.3a}\\
P_1 &= |100\rangle\langle100| + |011\rangle\langle011| & \text{bit flip on first qubit} & \tag{6.3b}\\
P_2 &= |010\rangle\langle010| + |101\rangle\langle101| & \text{bit flip on second qubit} & \tag{6.3c}\\
P_3 &= |001\rangle\langle001| + |110\rangle\langle110| & \text{bit flip on third qubit} & \tag{6.3d}
\end{aligned}
$$

Example: bit flip occurred on first qubit:

$$\alpha|000\rangle + \beta|111\rangle \xrightarrow{X_1} \alpha|100\rangle + \beta|011\rangle =: |\psi'\rangle \qquad (6.4)$$

⤳ measurement probabilities (cf. (2.73)):

$$
\begin{aligned}
P_0 : \quad & \langle\psi'|P_0|\psi'\rangle = 0 \quad (\text{since } \langle100|000\rangle = 0 \text{ and } \langle011|111\rangle = 0) && (6.5\text{a}) \\
P_1 : \quad & \langle\psi'|P_1|\psi'\rangle = \big(\alpha^*\langle100| + b^*\langle011|\big)P_1\big(\alpha|100\rangle + \beta|011\rangle\big) \\
& \qquad = |\alpha|^2\langle100|100\rangle\langle100|100\rangle + |\beta|^2\langle011|011\rangle\langle011|011\rangle \\
& \qquad = |\alpha|^2 + |\beta|^2 = 1 && (6.5\text{b}) \\
P_2 : \quad & \langle\psi'|P_2|\psi'\rangle = \cdots = 0 && (6.5\text{c}) \\
P_3 : \quad & \langle\psi'|P_3|\psi'\rangle = \cdots = 0 && (6.5\text{d})
\end{aligned}
$$

⤳ measurement reveals that bit flip occurred with certainty on first qubit, does not destroy quantum superposition since

$$P_1|\psi'\rangle = |\psi'\rangle \qquad (6.6)$$

2. *Recovery*: if flip occurred on qubit j, can restore original state by flipping this qubit again, i.e., applying Pauli-X

Circuit notation for the above example:
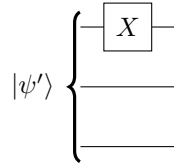
Alternative viewpoint of the syndrome measurements: perform measurements of the observables $Z_1 Z_2 \equiv Z \otimes Z \otimes I$ and $Z_2 Z_3 \equiv I \otimes Z \otimes Z$ (using the notation Z_j: acting on qubit j)

$$(6.7)$$

Eigenvalues of $Z_1 Z_2$? Computational basis states are eigenbasis since, for all $a, b, c \in \{0, 1\}$,

$$Z_1 Z_2 |a, b, c\rangle = (Z|a\rangle) \otimes (Z|b\rangle) \otimes |c\rangle = ((-1)^a|a\rangle) \otimes ((-1)^b|b\rangle) \otimes |c\rangle = (-1)^{a+b}|a, b, c\rangle; \quad (6.8)$$

corresponding eigenvalues are thus ± 1

Intuition: $Z_1 Z_2$ compares first two qubits, eigenvalue $1 \leftrightarrow a = b$, eigenvalue $-1 \leftrightarrow a \neq b$.

Similarly for $Z_2 Z_3$: compare second and third qubit

In summary: obtain 2 bits of classical information, equivalent to measuring the projection operators P_0, P_1, P_2, P_3

6.1.2   Three qubit phase flip code

Qubits can be affected by various types of errors (additional to bit flip $|0\rangle \leftrightarrow |1\rangle$)

phase flip error:
$$\alpha|0\rangle + \beta|1\rangle \mapsto \alpha|0\rangle - \beta|1\rangle \quad \text{with probability } p, \tag{6.9}$$

that is, Pauli-Z is applied with probability p

Note: phase flip error does not have classical equivalent

Idea for error correction: map phase flip error model to bit flip error model by changing to qubit basis $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$

Observation:
$$Z|\pm\rangle = \tfrac{1}{\sqrt{2}}(Z|0\rangle \pm Z|1\rangle) = \tfrac{1}{\sqrt{2}}(|0\rangle \mp |1\rangle) = |\mp\rangle \tag{6.10}$$

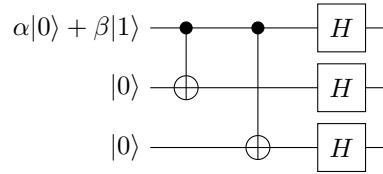i.e., Z acts as $|+\rangle \leftrightarrow |-\rangle$, analogous to X: $|0\rangle \leftrightarrow |1\rangle$

$\rightsquigarrow$ logical qubits:

$$|0_L\rangle = |+++\rangle \tag{6.11a}$$
$$|1_L\rangle = |---\rangle \tag{6.11b}$$

Use Hadamard gates to change between $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$ basis
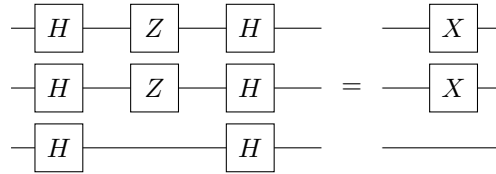
Corresponding circuit for encoding:

1. *Error detection*: measurement using projection operators
   $$P'_j = H^{\otimes 3} P_j H^{\otimes 3} \quad \text{for } j = 0, \dots, 3, \tag{6.12}$$

   equivalently: using observables

   $$\begin{aligned}
   H^{\otimes 3}(Z_1 Z_2) H^{\otimes 3} &= (H \otimes H \otimes H)(Z \otimes Z \otimes I)(H \otimes H \otimes H) \\
   &= (HZH) \otimes (HZH) \otimes (HIH) = X \otimes X \otimes I = X_1 X_2
   \end{aligned} \tag{6.13a}$$

   (here we have used that $HZH = X$) and

   $$H^{\otimes 3}(Z_2 Z_3) H^{\otimes 3} = \cdots = X_2 X_3 \tag{6.13b}$$

   corresponding eigenvalues ± 1 allow to compare qubits, analogous to bit flip errors

2. *Recovery*: apply Hadamard-conjugated bit flip recovery operators, for example: phase flip occurred on first qubit $\rightsquigarrow$ recovery using $HX_1 H = Z_1$

6.2   The Shor code

So far: considered two error models separately

It turns out: can protect against an arbitrary error on a single qubit!

Shor code: concatenate bit and phase flip error procedures:

first stage (phase flip model)

$$|0\rangle \mapsto |+++\rangle \tag{6.14a}$$
$$|1\rangle \mapsto |---\rangle \tag{6.14b}$$

second stage (bit flip model):

$$|+\rangle \mapsto \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \tag{6.15a}$$

$$|-\rangle \mapsto \frac{1}{\sqrt{2}}(|000\rangle - |111\rangle) \tag{6.15b}$$

⤳ concatenation gives nine qubit code with logical qubits:

$$|0_L\rangle = \frac{1}{2\sqrt{2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \tag{6.16a}$$

$$|1_L\rangle = \frac{1}{2\sqrt{2}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) \tag{6.16b}$$

Error correction procedure analogous to before (now both bit and phase flip errors)

Examples:

- Bit flip occurred on first qubit: measuring $Z_1 Z_2$ reveals that first or second qubit has flipped, $Z_2 Z_3$ that second and third qubit are the same ⤳ first qubit must have flipped, correct error by applying X_1

- Bit flip on fifth qubit: measure $Z_4 Z_5$ and $Z_5 Z_6$, correct error by applying X_5 (similar to first example)

- Phase flip occurred on first qubit ⤳ first three-qubit block:

$$|000\rangle \overset{Z_1}{\mapsto} |000\rangle, \tag{6.17a}$$

$$|111\rangle \overset{Z_1}{\mapsto} -|111\rangle \tag{6.17b}$$

(Z_2 and Z_3 have same effect), thus

$$|0_L\rangle \overset{Z_1}{\mapsto} \frac{1}{2\sqrt{2}}(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle), \tag{6.18a}$$

$$|1_L\rangle \overset{Z_1}{\mapsto} \frac{1}{2\sqrt{2}}(|000\rangle + |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) \tag{6.18b}$$

Syndrome measurement for detecting phase flip errors: observables $X_1 X_2 X_3 X_4 X_5 X_6$ and $X_4 X_5 X_6 X_7 X_8 X_9$

Recovery from phase flip error on first block: apply $Z_1 Z_2 Z_3$

Does the procedure still work if both a bit and phase flip error occurred (on the same qubit)? ⤳ yes!

Consider again error on first qubit (works analogously for the other qubits):

$$|0_L\rangle, \text{first block:} \quad \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \overset{Z_1}{\mapsto} \frac{1}{\sqrt{2}}(|000\rangle - |111\rangle) \overset{X_1}{\mapsto} \frac{1}{\sqrt{2}}(|100\rangle - |011\rangle), \quad (6.19a)$$

$$|1_L\rangle, \text{first block:} \quad \frac{1}{\sqrt{2}}(|000\rangle - |111\rangle) \overset{Z_1}{\mapsto} \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \overset{X_1}{\mapsto} \frac{1}{\sqrt{2}}(|100\rangle + |011\rangle) \quad (6.19b)$$

Syndrome measurements reveal that both bit and phase flip errors occurred, error can be corrected by applying $Z_1 X_1$ (note: apply X_1 first, then Z_1)

Procedure actually works for an *arbitrary* error, as long as it affects only a single qubit (even replacing the qubit by a completely random state)

Effect of noise in general described by a quantum operation (see also Exercise 10.2 and Nielsen and Chuang 2010, section 8.2):

A **quantum operation** describes in general terms how a quantum system evolves, typically governed by the unitary time evolution, when applying quantum gates or performing measurements, or also under the influence of noise processes. Abstractly,

$$\rho' = \mathcal{E}(\rho), \quad (6.20)$$

where $\mathcal{E}$ is the quantum operation and ρ the density matrix of the initial quantum system. For example, a unitary transformation is written as $\mathcal{E}(\rho) = U\rho U^\dagger$ (see (4.2)), and a measurement with outcome m as $\mathcal{E}_m(\rho) = M_m \rho M_m^\dagger$ (possibly up to a normalization factor, see (4.6)). The following **operator-sum representation**

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger \quad (6.21)$$

with E_k complex matrices satisfying $\sum_k E_k^\dagger E_k \leq I$, captures in greatest generality any quantum operation compatible with the laws of quantum mechanics. (The sum is finite, and we write $A \leq B$ for matrices A and B if $B - A$ is positive semidefinite.) The quantum operation is **trace-preserving**, i.e., $\mathrm{tr}[\mathcal{E}(\rho)] = \mathrm{tr}[\rho]$ for any ρ, precisely if $\sum_k E_k^\dagger E_k = I$.

Note that for a pure initial state $\rho = |\psi\rangle\langle\psi|$,

$$\mathcal{E}(|\psi\rangle\langle\psi|) = \sum_k E_k|\psi\rangle\langle\psi|E_k^\dagger \quad (6.22)$$

In the present context of quantum error correction, with an error on a single qubit: all E_k act (non-trivially) only on this qubit, say the first one; we represent

$$E_k = e_{k,0}I + e_{k,1}X_1 + e_{k,2}Z_1 + e_{k,3}\underbrace{X_1 Z_1}_{=-iY_1} \quad (6.23)$$

with $e_{k,j} \in \mathbb{C}$ for $j = 0, \ldots, 3$. (This representation always exists and is unique since the identity and Pauli matrices form a basis of 2×2 matrices.) Thus

$$\begin{aligned} E_k|0_L\rangle = \; & e_{k,0}\frac{1}{2\sqrt{2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \\ & + e_{k,1}\frac{1}{2\sqrt{2}}(|100\rangle + |011\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \\ & + e_{k,2}\frac{1}{2\sqrt{2}}(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \\ & + e_{k,3}\frac{1}{2\sqrt{2}}(|100\rangle - |011\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle), \end{aligned} \quad (6.24)$$

and similarly for $E_k|1_L\rangle$

Effect of syndrome diagnosis: project onto eigenspaces of syndrome measurement operators; the ones relevant for the first qubit are $Z_1 Z_2$ and $X_1 X_2 X_3 X_4 X_5 X_6$. Writing the initial state (before the error) as $|\psi\rangle = \alpha|0_L\rangle + \beta|1_L\rangle$, and assuming, for example, that both measurements result in eigenvalue -1, then for each k:

$$E_k|\psi\rangle = E_k(\alpha|0_L\rangle + \beta|1_L\rangle) \xrightarrow{\substack{\text{syndrome} \\ \text{measurement}}} e_{k,3}\frac{\alpha}{2\sqrt{2}}(|100\rangle - |011\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)$$

$$e_{k,3}\frac{\beta}{2\sqrt{2}}(|100\rangle + |011\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle);$$

$$(6.25)$$

in other words: the syndrome measurement collapses $E_k|\psi\rangle$ to a state proportional to $X_1 Z_1|\psi\rangle$.

Another example: assuming the measurement of $Z_1 Z_2$ results in eigenvalue 1, and of $X_1 X_2 X_3 X_4 X_5 X_6$ in eigenvalue -1, then for each k:

$$E_k|\psi\rangle = E_k(\alpha|0_L\rangle + \beta|1_L\rangle) \xrightarrow{\substack{\text{syndrome} \\ \text{measurement}}} e_{k,2}\frac{\alpha}{2\sqrt{2}}(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)$$

$$e_{k,2}\frac{\beta}{2\sqrt{2}}(|000\rangle + |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle);$$

$$(6.26)$$

in other words: the syndrome measurement collapses $E_k|\psi\rangle$ to a state proportional to $Z_1|\psi\rangle$.

In general: denoting the projection operator onto the simultaneous eigenspace of the operators $Z_1 Z_2$ and $X_1 X_2 X_3 X_4 X_5 X_6$ corresponding to the syndrome measurement outcome by P, then the state $\tilde{\rho}$ after the syndrome measurement will be proportional to

$$\tilde{\rho} \propto P\mathcal{E}(|\psi\rangle\langle\psi|)P = \sum_k PE_k|\psi\rangle\langle\psi|E_k^\dagger P \qquad (6.27)$$

$\tilde{\rho}$ can thus be regarded as ensemble of the (un-normalized) quantum states $PE_k|\psi\rangle$; as exemplified in Eqs. (6.25) and (6.26), these states are actually all the same (independent of k), except for (irrelevant) prefactors! In (6.25), $PE_k|\psi\rangle \propto X_1 Z_1|\psi\rangle$, and in (6.26), $PE_k|\psi\rangle \propto Z_1|\psi\rangle$. The corresponding recovery operation then restores the initial $|\psi\rangle$; e.g., for (6.25), applying $Z_1 X_1$.

Remarks:

- There exist other quantum error correction schemes (besides the Shor code) for protecting against single qubit errors

- Important overarching framework to understand quantum error correction: stabilizer formalism (see below)

6.3 Theory of quantum error-correction

Quantum states encoded in a subspace C of some larger Hilbert space, e.g., for three qubit bit flip code: $C = \mathrm{span}\{|0_L\rangle, |1_L\rangle\} = \mathrm{span}\{|000\rangle, |111\rangle\}$

Denote projector onto code space C by P, e.g., for the three qubit bit flip code: $P = |000\rangle\langle 000| + |111\rangle\langle 111|$

Theorem 3 (Quantum error-correction conditions). *Let C be a quantum code, and let P be the projector onto C. Suppose $\mathcal{E}$ is a quantum operation with operation elements $\{E_k\}$. A necessary and sufficient condition for the existence of an error-correction operation $\mathcal{R}$ correcting $\mathcal{E}$ on C is that*

$$PE_k^\dagger E_\ell P = \alpha_{k\ell} P \tag{6.28}$$

for some Hermitian matrix $(\alpha_{k\ell})$ of complex numbers.

The operation elements $\{E_k\}$ are denoted errors, and we call them a "correctable set of errors" if such an $\mathcal{R}$ exists. Mathematically:

$$(\mathcal{R} \circ \mathcal{E})(\rho) \propto \rho \tag{6.29}$$

with $\circ$ the function composition symbol and $\propto$ meaning "proportional to" (to account for measurements, for example)

For a proof of Theorem 3, see Nielsen and Chuang 2010, section 10.3.

6.4 Stabilizer codes

(Nielsen and Chuang 2010, section 10.5)

6.4.1 The stabilizer formalism

Example: Bell state $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ satisfies

$$Z_1 Z_2 |\psi\rangle = |\psi\rangle \quad \text{and} \quad X_1 X_2 |\psi\rangle = |\psi\rangle; \tag{6.30}$$

$\rightsquigarrow |\psi\rangle$ is "stabilized" by these operators

It turns out: $|\psi\rangle$ is the unique quantum state (up to an overall phase factor) which is stabilized by $Z_1 Z_2$ and $X_1 X_2$, i.e., this is a method for characterizing $|\psi\rangle$

Note: $Z_1 Z_2, X_1 X_2 \in G_2$ (Pauli group)

Definition of the **Pauli group** G_1:

$$G_1 = \{\pm I, \pm iI, \pm X, \pm iX, \ldots, \pm iZ\} \tag{6.31}$$

with matrix multiplication as group operation (cf. Tutorial 11)

In general, the **Pauli group** G_n is defined as the set of all n-fold tensor products of identity and Pauli matrices, with prefactors ± 1 and $\pm i$. For example, $-iX_1 Y_3 X_4 Z_5 \equiv -iX \otimes I \otimes Y \otimes X \otimes Z \in G_5$. G_n is closed under matrix multiplication (as required), since matrix multiplication factorizes with respect to the tensor product structure, i.e.,

$$(A_1 \otimes A_2 \otimes \cdots \otimes A_n)(B_1 \otimes B_2 \otimes \cdots \otimes B_n) = (A_1 B_1) \otimes (A_2 B_2) \otimes \cdots \otimes (A_n B_n) \tag{6.32}$$

for matrices A_j, B_j of compatible dimensions, and the product of two Pauli matrices is again a Pauli matrix or the identity (up to these prefactors); e.g., $XZ = -iY$ or $Y^2 = I$. In general, for $\sigma_1 = X$, $\sigma_2 = Y$, $\sigma_3 = Z$ and $\alpha, \beta \in \{1, 2, 3\}$:

$$\sigma_\alpha \sigma_\beta = \begin{cases} I & \text{if } \alpha = \beta \\ \sum_{\gamma=1}^3 i\epsilon_{\alpha\beta\gamma} \sigma_\gamma & \text{if } \alpha \neq \beta \end{cases}, \tag{6.33}$$

where the "Levi-Civita symbol" $\epsilon_{\alpha\beta\gamma} = 1$ if α, β, γ is an even permutation of $1, 2, 3$ (like $3, 1, 2$), $\epsilon_{\alpha\beta\gamma} = -1$ if α, β, γ is an odd permutation of $1, 2, 3$ (like $2, 1, 3$) and $\epsilon_{\alpha\beta\gamma} = 0$ otherwise. Note that, in particular, the sum over γ in (6.33) only contains a single non-zero term.

General principle of the stabilizer formalism: S subgroup of G_n, define the n-qubit subspace V_S as the vector space **stabilized** by S:

$$V_S = \left\{ |\psi\rangle \in \mathbb{C}^{2^n} : g|\psi\rangle = |\psi\rangle \text{ for all } g \in S \right\} \tag{6.34}$$

V_S is indeed a vector space: linear combinations are also in V_S. The elements of S are called **stabilizer operators**.

Example: $S = \{I, Z_1Z_2, Z_2Z_3, Z_1Z_3\}$ is a subgroup of G_3, corresponding vector space stabilized by S: $V_S = \text{span}\{|000\rangle, |111\rangle\}$ (three qubit bit flip code)

In the following, we use a compact description of S via generators:

Definition. *A set of elements* $g_1, \ldots, g_\ell$ *in a group* G *is said to* generate *the group if every element of* G *can be written as product of elements from* $\{g_1, \ldots, g_\ell\}$. *Notation:*

$$G = \langle g_1, \ldots, g_\ell \rangle \tag{6.35}$$

Example: $\{I, Z_1Z_2, Z_2Z_3, Z_1Z_3\} = \langle Z_1Z_2, Z_2Z_3 \rangle$ since $Z_1Z_3 = (Z_1Z_2)(Z_2Z_3)$ and $I = (Z_1Z_2)^2$

Note: V_S can possibly be the trivial vector space: $V_S = \{0\}$ (zero vector, not the same as $|0\rangle$!); e.g., if $-I \in S$: then the condition $-I|\psi\rangle = |\psi\rangle$ has the only solution $|\psi\rangle = 0$

Condition on S such that S stabilizes a non-trivial vector space?

  (a) The elements of S must commute, i.e., $[g, h] = gh - hg = 0$ for all $g, h \in S$

  (b) $-I \notin S$

To justify (a): any $g, h \in S \subseteq G_n$ either commute or anti-commute (see Exercise 11.2). Assume that they anti-commute ($gh = -hg$), then for any $|\psi\rangle \in V_S$:

$$|\psi\rangle = gh|\psi\rangle = -hg|\psi\rangle = -|\psi\rangle, \tag{6.36}$$

thus $|\psi\rangle = 0$, and we conclude that $V_S = \{0\}$.

It turns out: conditions (a) and (b) are also sufficient. Assume from now on that they are satisfied.

Check matrix representation   (see also Exercise 11.2)

Given an element $g \in G_n$, we define a row vector of length $2n$, denoted $r(g)$, as follows (ignoring the allowed prefactor ± 1 or $\pm i$ of g):

matrix in g acting on qubit j	jth entry of $r(g)$	$(n+j)$th entry of $r(g)$
I	0	0
X	1	0
Y	1	1
Z	0	1

For example, $n = 5$ and

$$g = iX_2Y_4Z_5 \quad \leadsto \quad r(g) = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \tag{6.37}$$

Useful properties:

  • For all $g, g' \in G_n$:
$$r(gg') = r(g) + r(g') \tag{6.38}$$
  with gg' the usual group operation (matrix-matrix product), and addition bitwise modulo 2

- $g, g' \in G_n$ commute ($[g, g'] = 0$) if and only if

$$r(g)\Lambda r(g')^T = 0 \mod 2 \quad \text{with} \quad \Lambda := \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \tag{6.39}$$

For a list of generators $g_1, \ldots, g_\ell$, define the **check matrix** C as the $\ell \times 2n$ matrix with rows $r(g_1), \ldots, r(g_\ell)$:

$$C = \begin{pmatrix} r(g_1) \\ \vdots \\ r(g_\ell) \end{pmatrix} \tag{6.40}$$

Thus, according to Eq. (6.39), $g_1, \ldots, g_\ell$ pairwise commute if and only if

$$C \Lambda C^T = 0 \quad \text{(zero matrix, modulo 2)} \tag{6.41}$$

Want generators to be independent, such that none can be omitted:

$$\langle g_1, \ldots g_{j-1}, g_{j+1}, \ldots, g_\ell \rangle \subsetneq \langle g_1, \ldots, g_\ell \rangle \quad \text{for any } j \in \{1, \ldots, \ell\} \tag{6.42}$$

Can probe whether generators are indeed independent via the following

Proposition. *Let $S = \langle g_1, \ldots, g_\ell \rangle$ be such that $-I \notin S$. The generators $g_1, \ldots, g_\ell$ are independent if and only if the rows of the corresponding check matrix are linearly independent.*
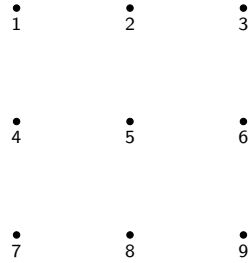
(Nielsen and Chuang 2010, Proposition 10.3; see also there for an accompanying proof)

6.4.2 Surface codes

Promising route towards quantum error correction via surface codes

Setup for a "planar code": qubits arranged on a $d \times d$ grid, with d an odd integer
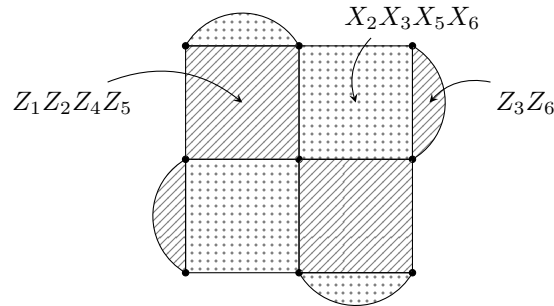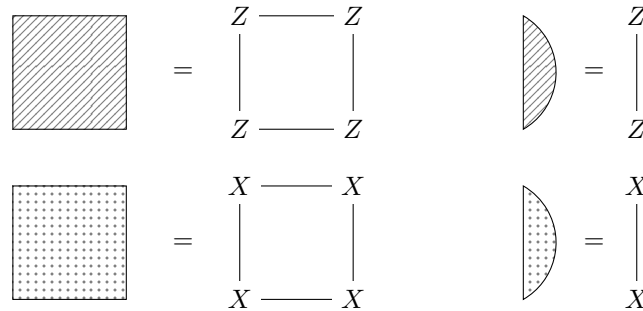
Example for $d = 3$:

$\leadsto$ quantum Hilbert space dimension is 2^{d^2}

Square of four qubits denoted *plaquette*:

Define generators $g_1, \ldots g_{d^2-1}$ by X or Z matrices acting on qubits of a plaquette or boundary pairs:


42

with the pictorial definitions

$$
\begin{array}{ccc}
\boxed{\text{(hatched square)}} & = & \begin{array}{ccc} Z & \!\!\!\!\!\rule[0.5ex]{1.5em}{0.4pt}\!\!\!\! & Z \\ | & & | \\ Z & \!\!\!\!\!\rule[0.5ex]{1.5em}{0.4pt}\!\!\!\! & Z \end{array}
\end{array}
\qquad
\begin{array}{ccc}
\boxed{\text{(half disk)}} & = & \begin{array}{c} Z \\ | \\ Z \end{array}
\end{array}
$$

$$
\begin{array}{ccc}
\boxed{\text{(dotted square)}} & = & \begin{array}{ccc} X & \!\!\!\!\!\rule[0.5ex]{1.5em}{0.4pt}\!\!\!\! & X \\ | & & | \\ X & \!\!\!\!\!\rule[0.5ex]{1.5em}{0.4pt}\!\!\!\! & X \end{array}
\end{array}
\qquad
\begin{array}{ccc}
\boxed{\text{(dotted half disk)}} & = & \begin{array}{c} X \\ | \\ X \end{array}
\end{array}
$$

Logical code space is simultaneous $+1$ eigenspace of all stabilizer generators. Can derive for the present setup that each additional generator divides the logical code space dimension by 2; thus the actual code space dimension is

$$
\frac{2^{d^2}}{2^{\sharp \text{generators}}} = \frac{2^{d^2}}{2^{d^2-1}} = 2, \tag{6.43}
$$

i.e., a single logical qubit!

Logical operations (Pauli gates) acting on this qubit? Define X_L and Z_L as strings of X or Z operators connecting the top and bottom or left and right boundaries of the grid:
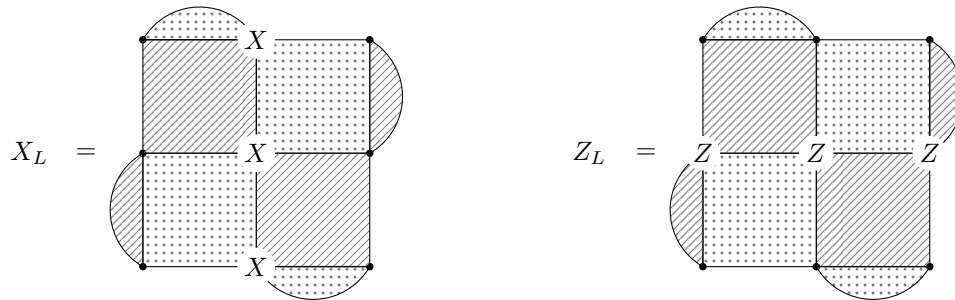
Both X_L and Z_L commute with all stabilizers.

Advantage (especially for larger grid dimensions d): can recover from local errors (noise affecting a localized region on the grid)

Remark: there exist several variants of surface codes, like the *Toric code* (see, e.g., Dennis, Kitaev, Landahl, and Preskill 2002)

References

Bell, J. S. (1964). "On the Einstein Podolsky Rosen paradox". In: *Physics Physique Fizika* 1, pp. 195–200. DOI: 10.1103/PhysicsPhysiqueFizika.1.195.

Dennis, E., A. Kitaev, A. Landahl, and J. Preskill (2002). "Topological quantum memory". In: *J. Math. Phys.* 43, pp. 4452–4505. DOI: 10.1063/1.1499754. URL: https://arxiv.org/abs/quant-ph/0110143.

Einstein, A., B. Podolsky, and N. Rosen (1935). "Can quantum-mechanical description of physical reality be considered complete?" In: *Phys. Rev.* 47, pp. 777–780. DOI: 10.1103/PhysRev.47.777.

Nielsen, M. and I. L. Chuang (2010). *Quantum Computation and Quantum Information*. Cambridge University Press.

Index