

Algebraic foundations of quantum computing¹

January 26, 2021

¹If you are a teacher/lecturer/professor, and would like to use these notes for any course you are teaching, please go ahead! I would appreciate you letting me know at michael.mcgettrick@nuigalway.ie, then I can even send you PDF without the URL at the bottom, or source L^AT_EX

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

1

Qubits, Bloch sphere, qudits, unitary matrices

Definition: Qubit

A qubit is 2 component vector

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \quad (1)$$

with $z_1, z_2 \in \mathbb{C}$ and $|z_1|^2 + |z_2|^2 = 1$.

Writing $z_1 = r_1 e^{i\phi_1}$ and $z_2 = r_2 e^{i\phi_2}$, the equation $|z_1|^2 + |z_2|^2 = 1$ becomes $r_1^2 + r_2^2 = 1$. So we can set $r_1 = \cos(\theta/2)$, $r_2 = \sin(\theta/2)$, with $0 \leq \theta \leq \pi$.

Instead of the four parameters $\{r_1, \phi_1, r_2, \phi_2\}$ we now have three parameters $\{\theta, \phi_1, \phi_2\}$.

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

3

Outline

1. Qubits, Bloch sphere, qudits, unitary matrices	3
2. Two Qubits, Tensor Products, No Cloning Theorem	29
3. Entanglement	38
4. Density operators	53
5. (Von Neumann) entropy	87
6. CPTP maps, Kraus operators and quantum channels	101
7. Positive Operator-Valued Measures (POVMs)	117
8. Deutsch-Jozsa algorithm	125
9. Quantum circuits and gates	135
10. Quantum Teleportation	142
11. (Computational) Complexity	151
12. Grover's algorithm	158
13. Quantum Fourier Transform and Shor's algorithm	174

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

2

Qubits, Bloch sphere, qudits, unitary matrices

Definition: Ket vector

The Ket vector (notation) is the column vector

$$|z\rangle = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \quad (2)$$

Definition: Bra vector

The Bra vector (notation) is the row vector

$$\langle z| = (z_1^* \quad z_2^*) \quad (3)$$

So, $\langle z|$ is the complex conjugate transpose of $|z\rangle$.

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

4

Definition: Inner Product $\langle | \rangle$

We can multiply the $\langle z|$ and $|z\rangle$ vectors in the obvious way, as follows.

$$\langle z|z\rangle = (z_1^* \quad z_2^*) \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = |z_1|^2 + |z_2|^2 = 1 \quad (4)$$

Exercise 1

Which of the following vectors are qubits (and, if not a qubit, explain why not):

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} \frac{e^{i\pi/7}}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}, \begin{pmatrix} 1/3 + i/3 \\ 2/3 - i/\sqrt{3} \end{pmatrix}, \quad (5)$$

Observables

A qubit gives us information about the possible outcomes of measuring a certain observable (property) of the quantum system/particle in question, as follows:

1. An observable O described by a qubit has exactly two possible outcomes (lets say, generically, o_1 and o_2). (As an example, which is not realistic but nonetheless gives the idea, imagine “measuring” the color (by looking at something) in a world where everything was blue or red.)
2. r_1^2 (or $|z_1|^2$) is the probability of observing outcome o_1 , while r_2^2 (or $|z_2|^2$) is the probability of observing outcome o_2 . Of course, these two probabilities sum to one, as there are only two possible outcomes for a qubit.

We traditionally represent the basis vectors as

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (6)$$

so that we can write

$$|z\rangle = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = z_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + z_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = z_1 |0\rangle + z_2 |1\rangle \quad (7)$$

Note:

1. In Physics it is more traditional to use α and β instead of z_1 and z_2 , we will do this at times (α and β are complex scalars whose modulus is not greater than one).
2. Other notations that are common in books, notes, online, etc. are $|\text{up}\rangle$ for $|0\rangle$ and $|\text{down}\rangle$ for $|1\rangle$. The “up” and “down” refer to spin properties of fundamental particles.

After measuring observable O ,

- ▶ if you actually obtained outcome o_1 , then the subsequent state of the system is $|0\rangle$
- ▶ if you actually obtained outcome o_2 , then the subsequent state of the system is $|1\rangle$

Nota bene: Measuring a system in principle “interferes” with the system: The state after measurement is not in general the same as the state before measurement. (Think of our previous example - where the observable is having color red or blue: To check this, we have to physically shine a light on the “object”, which in principle at least could interfere with / change the object.) We will see that, in fact measurement corresponds to projecting on to a vector in the orthonormal basis of the Hilbert Space in which the qubit lives.

Orthonormal bases

- ▶ $\{|0\rangle, |1\rangle\}$ form an orthonormal basis, i.e.
 $\langle 0|0\rangle = 1, \langle 0|1\rangle = 0, \langle 1|0\rangle = 0, \langle 1|1\rangle = 1, \dots$
- ▶ Another orthonormal basis we will use is written as
 $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}, \quad |-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. We leave it as an exercise to the reader to check that $\langle +|+\rangle = \langle -|-\rangle = 1$ and $\langle +|-\rangle = \langle -|+\rangle = 0$

Exercise 2

Show that

$$\frac{1}{\sqrt{7}}|0\rangle + i\frac{\sqrt{6}}{\sqrt{7}}|1\rangle \quad \text{and} \quad \frac{-\sqrt{6}}{\sqrt{7}}|0\rangle + i\frac{1}{\sqrt{7}}|1\rangle \quad (8)$$

form an orthonormal basis.

The overall (global) phase factor $e^{i\phi_1}$ makes no difference to any physical measurements we can make (it does not affect the probabilities of any outcomes). In other words, $|\psi\rangle$ and $e^{i\delta}|\psi\rangle$ are indistinguishable physically. We simply drop this phase factor and write

$$|\psi\rangle = \cos(\theta/2)|0\rangle + \sin(\theta/2)e^{i\phi}|1\rangle \quad (11)$$

By choosing $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$, we represent uniquely **physically, but not mathematically** any qubit. Choosing two angles like this may remind you of “longitude” and “latitude” in geography: Precisely this analogy leads to a picture of any qubit on a sphere: The Bloch sphere (also variously known as Poincaré or Riemann sphere).

A general state of a qubit can be written as

$$|\psi\rangle = \cos(\theta/2)e^{i\phi_1}|0\rangle + \sin(\theta/2)e^{i\phi_2}|1\rangle. \quad (9)$$

Exercise 3

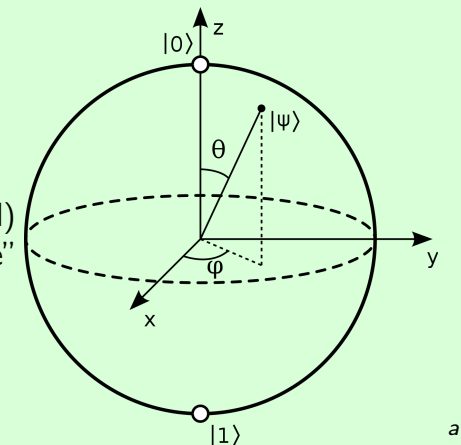
Determine the parameters θ, ϕ_1 and ϕ_2 that correspond to the states in Equation 8.

The multiplicative factors $e^{i\phi_1}$ and $e^{i\phi_2}$ in Equation 9 are called **phases**. We can write

$$|\psi\rangle = e^{i\phi_1} \left(\cos(\theta/2)|0\rangle + \sin(\theta/2)e^{i(\phi_2-\phi_1)}|1\rangle \right) \quad (10)$$

Bloch sphere

- ▶ θ is “latitude” while ϕ is “longitude”.
- ▶ $|0\rangle$ is the “North Pole” ($\theta = 0$, and ϕ is immaterial) while $|1\rangle$ is the “South Pole” ($\theta = \pi$).
- ▶ $|+\rangle$ and $|-\rangle$ are states on the Equator ($\theta = \pi/2$ with $\phi = 0, \pi$, respectively).



^aSource: Wikipedia

One must be careful with interpreting the Bloch sphere picture. One cannot “add” vectors pictorially, as one would do in Cartesian space \mathbb{R}^3 . In particular, two “vectors” of equal length, pointing in opposite directions in the Bloch sphere do not cancel (e.g. $|0\rangle + |1\rangle \neq 0$)! Indeed, for any vector on the sphere, the one pointing in opposite direction to it is orthogonal!

Definition: Hadamard matrix

Because of its ubiquity, we introduce the Hadamard matrix

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad (12)$$

You should check that $H|0\rangle = |+\rangle$ and $H|1\rangle = |-\rangle$. Furthermore, since $H^2 = I$, we also have $H|+\rangle = |0\rangle$ and $H|-\rangle = |1\rangle$.

Definition: Qudit

A qudit is a d component vector

$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_d \end{pmatrix} \quad (13)$$

with $z_j \in \mathbb{C}$ and $\sum_j |z_j|^2 = 1$.

Exercise 4

Calculate the matrix M that sends the basis $\{|0\rangle, |1\rangle\}$ to the basis defined in Equation 8. Is $(M^{\top*})M = I$?

Definition: Superposition

A qubit with $0 < r_1 < 1$ is said to be in superposition. Since neither r_1 nor r_2 are zero, there is a non-zero probability of measuring both $|0\rangle$ and $|1\rangle$: It is “in two states at the same time”^a.

^a“God does not play dice with the universe”, *Albert Einstein*

Definition: Interference

Constructive/destructive interference simply refers to the addition/subtraction of components of the qubit vector.

How does “classical” computing compare with “quantum” computing, mathematically?

Classical Computing (as Boolean functions)

- **“States”:** Bit (Binary Digit) chosen from $B = \{0, 1\}$
- **Operations:** Logic gates, AKA Boolean functions $f : B^n \rightarrow B$.

Example: The AND gate is a function $f_{\text{AND}} : B^2 \rightarrow B$ with $f_{\text{AND}}((0, 0)) = 0$, $f_{\text{AND}}((0, 1)) = 0$, $f_{\text{AND}}((1, 0)) = 0$, $f_{\text{AND}}((1, 1)) = 1$.

Classical Computing (using vectors and matrices)

- **“States”:**

$$0 \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad 1 \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (14)$$

Classical Computing (using vectors and matrices) ... continued

- **Operations:** The four possible unary operators can be represented using 2x2 matrices. For example,

$$\text{NOT} \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (15)$$

	Classical	Classical (take 2)	Quantum
Fundamental objects ("states")	bits {0, 1}	vectors	qubits
Operations	Boolean functions	2x2 matrices	Unitary matrices
Number of distinct states	2	2	∞

Exercise 5

Which of the following matrices are Unitary?

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

$$\begin{pmatrix} 3+i & -i-1 \\ 2i+1 & i-3 \end{pmatrix} \quad \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Unitary matrices have the very important property that **they preserve norm**, i.e. our inner product $\langle | \rangle$. If U is Unitary, then $\langle U\psi_1 | U\psi_2 \rangle = \langle \psi_1 | U^\dagger U \psi_2 \rangle = \langle \psi_1 | I \psi_2 \rangle = \langle \psi_1 | \psi_2 \rangle$. (If we wished to avoid the bra-ket notation, we could just write

$$(U\psi_1)^\dagger U\psi_2 = \psi_1^\dagger U^\dagger U \psi_2 = \psi_1^\dagger I \psi_2.$$

In particular when $\psi_1 = \psi_2$ we have $\langle U\psi | U\psi \rangle = \langle \psi | \psi \rangle = 1$.

Definition: Unitary matrices

A matrix U with entries in \mathbb{C} is said to be **Unitary** if and only if $U^\dagger U = ((U^\top)^*)U = I$

Example: We will show that

$$U = \frac{1}{2} \begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix} \quad (16)$$

is Unitary. We have

$$((U^\top)^*)U = \frac{1}{4} \begin{pmatrix} 1-i & 1+i \\ 1+i & 1-i \end{pmatrix} \begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (17)$$

- The set of all $n \times n$ Unitary matrices form a group $U(n)$ (we have already seen each Unitary matrix necessarily has an inverse).
- If a Unitary matrix has in fact only Real entries, it is an Orthogonal matrix.
- The eigenvalues of a Unitary matrix are all "unimodular", i.e. they have modulus equal to one, i.e. the eigenvalues are $\lambda_j = e^{i\theta_j}$
- A **subgroup** of $U(n)$ is $SU(n)$, the matrices that have determinant equal to one. (In general, the determinant of a Unitary matrix is complex unimodular.)

- To date, we have been discussing qubits in what is termed the **standard basis** (or computational basis), $\{|0\rangle, |1\rangle\}$. We may write any qubit state in this basis, or in any other orthonormal basis. What is more significant though is that we can **measure** any qubit state in any orthonormal basis we chose. **Why is this important?** We commented earlier on phases, and on global versus relative phase. We said for example, if you measure $\psi_1 = (|0\rangle - |1\rangle)/\sqrt{2}$ in the standard basis, you will get 0 or 1 (or more correctly observable o_1 or o_2) with equal probabilities. Measuring in the standard basis could not distinguish ψ_1 from $\psi_2 = (|0\rangle + |1\rangle)/\sqrt{2}$. But, if we were to measure in the $\{|+\rangle, |-\rangle\}$ basis, we could distinguish ψ_1 from ψ_2 . **The construction/development/invention of an algorithm in quantum computing involves deciding both what sequence of Unitary operations to carry out on a state(s) and what basis to measure the final state(s) in.**

How is this possible? The steps are

1. P presents the coin in state $|0\rangle$ (heads up)
2. Q operates with the Hadamard matrix: $H|0\rangle = |+\rangle$
3. P can either flip the coin, or not: In both cases, the state remains unchanged,
 $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2} \rightarrow (|1\rangle + |0\rangle)/\sqrt{2} = |+\rangle$.
4. Q operates with the Hadamard matrix: $H|+\rangle = |0\rangle$

and so the final coin is always “heads up”. Note that, of course, this is an “unfair” game in that Q is allowed quantum moves, (and has 2 moves) while P is restricted to classical operations. Nonetheless, it demonstrates how in this game a quantum player can always beat a classical player.

Diversion: Let the quantum games begin!

PQ PENNY FLIP: “The starship Enterprise is facing some immanent—and apparently inescapable—calamity when Q appears on the bridge and offers to help, provided Captain Picard can beat him at penny flipping: Picard is to place a penny head up in a box, whereupon they will take turns (Q, then Picard, then Q) flipping the penny (or not), without being able to see it. Q wins if the penny is head up when they open the box.”^a

^ataken verbatim from Meyer, <https://arxiv.org/abs/quant-ph/9804010>

So: Picard agrees. He loses the first game. “Best of three!” he declares. He loses all 3. Figuring his luck must turn [50/50 chances to win each game, right?], he keeps playing. 40 games later, he has still lost every game..... 😞

The representation of any operator A as a 2×2 matrix in an orthonormal basis $E = \{e_1, e_2\}$ is

$$\begin{pmatrix} \langle e_1 | A | e_1 \rangle & \langle e_1 | A | e_2 \rangle \\ \langle e_2 | A | e_1 \rangle & \langle e_2 | A | e_2 \rangle \end{pmatrix} \quad (18)$$

Example: Basis change

Question: What is the representation of the Hadamard matrix in the basis $\{f_1, f_2\} = \{\frac{\sqrt{3}}{2}|0\rangle + \frac{1}{2}|1\rangle, \frac{1}{2}|0\rangle - \frac{\sqrt{3}}{2}|1\rangle\}$?

Answer: We calculate

$$\begin{pmatrix} \langle f_1 | H | f_1 \rangle & \langle f_1 | H | f_2 \rangle \\ \langle f_2 | H | f_1 \rangle & \langle f_2 | H | f_2 \rangle \end{pmatrix} \quad (19)$$

We have

$$\begin{aligned} \blacktriangleright H|f_1\rangle &= \frac{\sqrt{3}}{2}H|0\rangle + \frac{1}{2}H|1\rangle = \frac{\sqrt{3}}{2}|+\rangle + \frac{1}{2}|-\rangle = \\ &= \frac{\sqrt{3}+1}{2}|0\rangle + \frac{\sqrt{3}-1}{2}|1\rangle \end{aligned}$$

Example: Basis change

... continued

$$\blacktriangleright H|f_2\rangle = \frac{1}{2}H|0\rangle - \frac{\sqrt{3}}{2}H|1\rangle = \frac{1}{2}|+\rangle - \frac{\sqrt{3}}{2}|-\rangle = \frac{1-\sqrt{3}}{2}|0\rangle + \frac{1+\sqrt{3}}{2}|1\rangle$$

Now Expression (19) becomes

$$\left(\begin{array}{cc} (\frac{\sqrt{3}}{2}\langle 0| + \frac{1}{2}\langle 1|)(\frac{\sqrt{3}+1}{2}|0\rangle + \frac{\sqrt{3}-1}{2}|1\rangle) & (\frac{\sqrt{3}}{2}\langle 0| + \frac{1}{2}\langle 1|)(\frac{1-\sqrt{3}}{2}|0\rangle + \frac{1+\sqrt{3}}{2}|1\rangle) \\ (\frac{1}{2}\langle 0| - \frac{\sqrt{3}}{2}\langle 1|)(\frac{\sqrt{3}+1}{2}|0\rangle + \frac{\sqrt{3}-1}{2}|1\rangle) & (\frac{1}{2}\langle 0| - \frac{\sqrt{3}}{2}\langle 1|)(\frac{1-\sqrt{3}}{2}|0\rangle + \frac{1+\sqrt{3}}{2}|1\rangle) \end{array} \right)$$

$$= \begin{pmatrix} \frac{\sqrt{3}}{2}(\frac{\sqrt{3}+1}{2}) + \frac{1}{2}(\frac{\sqrt{3}-1}{2}) & \frac{\sqrt{3}}{2}(\frac{1-\sqrt{3}}{2}) + \frac{1}{2}(\frac{1+\sqrt{3}}{2}) \\ \frac{1}{2}(\frac{\sqrt{3}+1}{2}) + \frac{\sqrt{3}}{2}(\frac{1-\sqrt{3}}{2}) & \frac{1}{2}(\frac{1-\sqrt{3}}{2}) - \frac{\sqrt{3}}{2}(\frac{1+\sqrt{3}}{2}) \end{pmatrix}$$

$$= \left(\frac{1}{2}\right) \begin{pmatrix} \sqrt{3}+1 & \sqrt{3}-1 \\ \sqrt{3}-1 & -\sqrt{3}-1 \end{pmatrix}$$

Definition: Outer product

The **outer product** of a state $|\psi\rangle$ is given by

$$|\psi\rangle\langle\psi| \quad (20)$$

Example (Outer product)

The outer product of

$$|\psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (21)$$

is

$$|\psi\rangle\langle\psi| = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)^{\top*} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (22)$$

Definition: Projection

A projection operator P projects an element of a vector space on to a subspace, and has the property $P^2 = P$

For any qubit/qudit $|\psi\rangle$, $|\psi\rangle\langle\psi|$ is a projection operator. This is easy to see since $P = |\psi\rangle\langle\psi| \implies P^2 = (|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = P$ (since $\langle\psi|\psi\rangle = 1$ because of normalization).

For any qubit $|\psi\rangle$, $|\psi\rangle\langle\psi|$ is a (2×2) matrix with rank 1. *These matrices are not invertible (since their rank is one but dimension is 2).* (Reminder: **rank** is the number of rows/columns of the matrix that are linearly independent)

Definition: Measurement

For any chosen orthonormal basis $|e_i\rangle$, measurement (of a qubit state $|\psi\rangle$) corresponds to projection using the operators $|e_i\rangle\langle e_i|$. The state then “collapses” to $|e_i\rangle$ with probability $\langle e_i|\psi\rangle$.

Exercise 6

Calculate the projection matrices $|e_i\rangle\langle e_i|$ in the basis $|+\rangle, |-\rangle$.



Definition: Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (23)$$

Any 2×2 complex matrix can be written as a linear combination of the identity and the three Pauli matrices, i.e.

$$\alpha I + \beta \sigma_x + \gamma \sigma_y + \delta \sigma_z.$$

Since σ_z is diagonal, the computational (standard) basis vectors $|0\rangle$ and $|1\rangle$ are eigenvectors of σ_z . By contrast, the standard computational basis vectors are *not* eigenvectors of σ_x . Instead, since $\sigma_x |+\rangle = |+\rangle$ while $\sigma_x |-\rangle = -|-\rangle$, σ_x is diagonal in the $\{|+\rangle, |-\rangle\}$ basis.

Exercise 6

Check that σ_x behaves like a NOT operation in the standard computational basis, while σ_z behaves like a NOT in the $\{|+\rangle, |-\rangle\}$ basis.

We now study the possible states of multiple qubit systems, beginning with 2 qubits. We may index these qubits at times using A and B (in books/literature, *Alice* and *Bob*). For out (Bra)ket notation, we will use

$$|\psi\rangle \otimes |\phi\rangle \text{ or } |\psi\rangle |\phi\rangle \text{ or } |\psi\phi\rangle \quad (24)$$

The notation \otimes is called a *tensor product*:

Definition: Tensor Product

The tensor product of two vectors

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \text{ and } w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix} \quad (25)$$

is the vector

$$v \otimes w = \begin{pmatrix} v_1 w \\ v_2 w \\ \vdots \\ v_n w \end{pmatrix} \quad (26)$$

which has nm components.

Example: Tensor Product

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \otimes \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \begin{pmatrix} 3 \\ 4 \end{pmatrix} \\ 2 \begin{pmatrix} 3 \\ 4 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 6 \\ 8 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \otimes \begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 8 \\ 10 \\ 12 \\ 15 \end{pmatrix} \quad (27)$$

Note that in general the tensor product is *not* commutative:
 $v \otimes w \neq w \otimes v$.

In the standard basis we have

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (28)$$

The most general state of a 2-qubit system is

$$\alpha_1 |00\rangle + \alpha_2 |01\rangle + \alpha_3 |10\rangle + \alpha_4 |11\rangle \quad (29)$$

with $\alpha_j \in \mathbb{C}$ and

$$\sum_j |\alpha_j|^2 = 1 \quad (30)$$

Exercise 7

Given the 2 qubit state

$$-\frac{1}{2} |00\rangle + i\frac{1}{3} |01\rangle + -i\frac{1}{4} |10\rangle + \frac{\sqrt{83}}{12} |11\rangle \quad (31)$$

- ▶ If Alice measures her qubit in state $|0\rangle$, what is the probability that Bob will measure his qubit in state $|0\rangle$?
- ▶ If Bob measures his qubit in state $|0\rangle$, what is the probability that Alice will measure her qubit in state $|0\rangle$?

Note: This exercise is a problem in conditional probability (check your favorite probability or statistics course!) The relevant equation is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (32)$$

We may interpret

- ▶ $|\alpha_1|^2$ is the probability, on measurement, of finding Alice's qubit in state $|0\rangle$ and Bob's qubit in state $|0\rangle$
- ▶ $|\alpha_2|^2$ is the probability, on measurement, of finding Alice's qubit in state $|0\rangle$ and Bob's qubit in state $|1\rangle$
- ▶ $|\alpha_3|^2$ is the probability, on measurement, of finding Alice's qubit in state $|1\rangle$ and Bob's qubit in state $|0\rangle$
- ▶ $|\alpha_4|^2$ is the probability, on measurement, of finding Alice's qubit in state $|1\rangle$ and Bob's qubit in state $|1\rangle$

The no-cloning theorem (or "What do you mean I can't copy my USB stick 😊")

It is impossible to copy a quantum state.^a

Theorem: For an (arbitrary, unknown) quantum state $|\psi\rangle$ (the one we want to copy), there is no Unitary operation that will map $|\psi\rangle |\phi\rangle$ to $|\psi\rangle |\psi\rangle$.

Proof: We prove by contradiction. Suppose we had such a Unitary operation U , with $U(|\psi\rangle |\phi\rangle) = |\psi\rangle |\psi\rangle$. Set $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$. Since U is a linear operator, copying any qubit state, we have

$$U((\alpha |0\rangle + \beta |1\rangle) |\phi\rangle) = \alpha U(|0\rangle |\phi\rangle) + \beta U(|1\rangle |\phi\rangle) = \alpha |00\rangle + \beta |11\rangle$$

But we also must have

$$\begin{aligned} U((\alpha |0\rangle + \beta |1\rangle) |\phi\rangle) &= ((\alpha |0\rangle + \beta |1\rangle)(\alpha |0\rangle + \beta |1\rangle)) \\ &= \alpha^2 |00\rangle + \alpha\beta |01\rangle + \alpha\beta |10\rangle + \beta^2 |11\rangle \end{aligned}$$

^aIn Quantum Computing, you can't "copy and paste". See also footnote on page 143.

The no-cloning theorem ... continued

Equating these two expressions would force either

1. $\alpha = 0$, $\beta = \beta^2$, or
2. $\beta = 0$, $\alpha = \alpha^2$

Either of these would then contradict our original assumption that U would copy an arbitrary/general qubit state. Therefore, there is no such U \square

Definition: Separability

A 2-qubit state

$$\alpha_1 |00\rangle + \alpha_2 |01\rangle + \alpha_3 |10\rangle + \alpha_4 |11\rangle \quad (33)$$

is said to be **separable** if and only if (iff) it can be written as $|\psi_1\rangle \otimes |\psi_2\rangle$ where $|\psi_1\rangle$ is the first qubit state ($= \gamma_1 |0\rangle + \delta_1 |1\rangle$ say) and $|\psi_2\rangle$ is the second qubit state ($= \gamma_2 |0\rangle + \delta_2 |1\rangle$ say).

Definition: Entangled

A 2-qubit state is said to be **entangled** iff it is not separable.

Examples:

- ▶ The state $(|00\rangle + |01\rangle)/\sqrt{2}$ is separable. We can write it as $|\psi_1\rangle |\psi_2\rangle$ where $|\psi_1\rangle = |0\rangle$ and $|\psi_2\rangle = (|0\rangle + |1\rangle)/\sqrt{2} = |+\rangle$. (Remember: We are abusing the tensor product notation $|\psi_1\psi_2\rangle \equiv |\psi_1\rangle |\psi_2\rangle \equiv |\psi_1\rangle \otimes |\psi_2\rangle$)
- ▶ The state $|\psi\rangle = (|00\rangle + |01\rangle + |10\rangle)/\sqrt{3}$ is entangled. (We will not prove this yet. But if you think carefully you will realize any 2-qubit state which has exactly three terms from the four possible ones $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ is entangled.)
- ▶ The state $(|00\rangle + |01\rangle + |10\rangle + |11\rangle)/2$ is separable. It can be written as $|++\rangle$.
- ▶ The state $(|00\rangle + |11\rangle)/\sqrt{2}$ is entangled. We will see soon that it is in some sense “more entangled” than the entangled state $|\psi\rangle$ above.

- ▶ The state $\sqrt{2}[|00\rangle - |01\rangle + \sqrt{3}|10\rangle - \sqrt{3}|11\rangle]/4$ is separable. It can be written as $[(|0\rangle + \sqrt{3}|1\rangle)/2] |-\rangle$.
 - ▶ The state $(|00\rangle + 2|01\rangle + \sqrt{5}|10\rangle + \sqrt{6}|11\rangle)/4$ is entangled.
- Let us denote for the moment the set $S = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ of basis 2-qubit states. Closer examination shows (why?) that
1. if $|\psi\rangle$ has only one term from S it is separable;
 2. if $|\psi\rangle$ has exactly 3 terms from S it is entangled;
 3. if $|\psi\rangle$ has 2 or 4 terms from S , it may be entangled or separable.

Definition: Bell pair (Bell state)

A Bell^a pair is the two qubit state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \quad (34)$$

^aJohn Stewart Bell (1928-1990) was a quantum physicist from Belfast.

Note that while $|\psi\rangle$ above is a Bell pair - in fact there are four different Bell pairs in the standard definition, the others being $(|00\rangle - |11\rangle)/\sqrt{2}$, $(|01\rangle + |10\rangle)/\sqrt{2}$, $(|01\rangle - |10\rangle)/\sqrt{2}$. The Bell pair state is maximally entangled, in that, according to measures we will discuss soon, it is “more entangled” than any other 2-qubit state.

For the moment, we present more of a plausibility argument / qualitative discussion of “how much” entanglement is in a state. We ask the two questions:

1. If Alice were to actually measure her qubit, how much subsequent uncertainty would there be in the measurement outcome of Bob's qubit?
2. If Bob were to actually measure his qubit, how much subsequent uncertainty would there be in the measurement outcome of Alice's qubit?

$(|00\rangle + |01\rangle + |10\rangle)/\sqrt{3}$ If Alice measures first if she gets 1, she will know for certain Bob's qubit is in state 0; if she gets 0, she has no new information about what state Bob's qubit is in (it could be 0 or 1 with equal probability). If Bob measures first if he gets 1, he will know for certain Alice's qubit is in state 0, but if he gets 0, he has no new information about Alice's qubit (which is equiprobable to be in 0 or 1). So, on both sides, measurement can give us extra information about the second party, but in general will not give us complete information, pointing to partial entanglement.

$(|00\rangle + |01\rangle)/\sqrt{2}$ If Alice measures first she will get zero for the result, and no further information about Bob's qubit. If Bob measures first he may get zero or one, but no new information about Alice's qubit.

$(|00\rangle + |11\rangle)/\sqrt{2}$ If Alice measures first if she gets zero, she knows for sure Bob's qubit is in state zero; if she gets 1, she knows for sure Bob's qubit is in state 1. If Bob measures first if he gets zero, he knows for sure Alice's qubit is in state zero; if he gets 1, he knows for sure Alice's qubit is in state 1. In both cases, whichever party measures first determines completely the state of both qubits, which is the property of a maximally entangled state.

We may re-write the definition of (2-qubit) entanglement in terms of properties of the 4-vector. Let the 2-qubit state be given by components $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)^T$ with $\sum_j |\alpha_j|^2 = 1$. Let the associated 1-qubit states be $(\beta_1, \beta_2)^T$ and $(\gamma_1, \gamma_2)^T$ with $\sum_j |\beta_j|^2 = 1 = \sum_j |\gamma_j|^2$. Then the system is entangled iff there is no solution to the set of equations

$$\alpha_1 = \beta_1 \gamma_1 \quad \alpha_2 = \beta_1 \gamma_2 \quad \alpha_3 = \beta_2 \gamma_1 \quad \alpha_4 = \beta_2 \gamma_2 \quad (35)$$

Example: Maximally entangled state

Considering the Bell pair $(|00\rangle + |11\rangle)/\sqrt{2}$, we have $\alpha_1 = \alpha_4 = 1/\sqrt{2}$ and $\alpha_2 = \alpha_3 = 0$. So, for example, from Equation (35), $\alpha_2 = 0 \implies$ either $\beta_1 = 0$ or $\gamma_2 = 0$. But

- ▶ $\beta_1 = 0 \implies \alpha_1 = 0$ (contradiction)
- ▶ $\gamma_2 = 0 \implies \alpha_4 = 0$ (contradiction)

and so Equation (35) has no solution.

Our previous discussion of entanglement between **two** qubits is often referred to as **bipartite entanglement**. We now consider the three qubit tensor product

$|\psi_1\rangle \otimes |\psi_2\rangle \otimes |\psi_3\rangle = |\psi_1\rangle |\psi_2\rangle |\psi_3\rangle = |\psi_1\psi_2\psi_3\rangle$. This can of course be written as $\alpha_1|000\rangle + \alpha_2|001\rangle + \alpha_3|010\rangle + \alpha_4|011\rangle + \alpha_5|100\rangle + \alpha_6|101\rangle + \alpha_7|110\rangle + \alpha_8|111\rangle$ with $\sum_j |\alpha_j|^2 = 1$.

Discussion:

GHZ state: When Alice measures her qubit, she obtains 0 or 1 with equal probability. The subsequent state (for Bob and Charlie) is completely separable, either $|00\rangle$ or $|11\rangle$.

W state: When Alice measures her qubit, she obtains 0 with probability 2/3 and 1 with probability 1/3.

- ▶ If she measures zero, the subsequent Bob/Charlie state is a maximally entangled Bell pair $(|01\rangle + |10\rangle)/\sqrt{2}$.
- ▶ If she measures 1, the subsequent Bob/Charlie state is the separable state $|00\rangle$.

(Check here - there is nothing special about Alice measuring first - analogous statements apply if Bob or Charlie measured first.)

Two important states exhibiting **tripartite entanglement** are the following:

Definition: GHZ state

The GHZ^a state is

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$$

^aGreenberger-Horne-Zeilinger

Definition: W state

The W state is

$$|W\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle)$$

Clearly for tripartite entanglement, we have a number of possibilities.....

1. Alice could be entangled with Bob, but not with Charlie
2. Bob could be entangled with Charlie, but not with Alice
3. Charlie could be entangled with Alice, but not with Bob
4. Nobody is entangled with anyone else
5. Everyone is (partially) entangled with everyone else (GHZ, W)
6. ?

and for multipartite entanglement, even more possibilities 😊

Definition: Mixed states

A **mixed state** is a probability distribution (sometimes called a *statistical ensemble*) over **pure states** (i.e. a probability distribution over ket vectors $\{p_i, |\psi_i\rangle\}$ with $0 \leq p_i \leq 1$ and $\sum_j p_j = 1$). (A pure state can in reverse be defined as a mixed state where exactly one of the p_j is one, and all the other p_j are zero.)

Example 1:

The mixed state

$$\left\{ \left\{ \frac{1}{8}, |0\rangle \right\}, \left\{ \frac{7}{8}, |1\rangle \right\} \right\}$$

corresponds to a state which is either in state $|0\rangle$ (with probability 0.125) or in state $|1\rangle$ (with probability 0.875).

Neither of the previous examples are maximally mixed states (can you say why?), but $\{\{0.5, |0\rangle\}, \{0.5, |1\rangle\}\}$ and $\{\{0.5, |+\rangle\}, \{0.5, |-\rangle\}\}$ are maximally mixed.

It may seem that a pure qubit state $A = |+\rangle$ and a mixed state $B = \{\{0.5, |0\rangle\}, \{0.5, |1\rangle\}\}$ correspond to one another: In each case,

- S1: the probability to measure $|0\rangle$ or $|1\rangle$ is the same, right?
- S2: the subsequent qubit state after measurement is the same, right?

These statements S1 and S2 are correct, but it does not mean that A and B correspond, as we now show.

The mixed state

$$\left\{ \left\{ \frac{1}{10}, |0\rangle \right\}, \left\{ \frac{3}{10}, |+\rangle \right\}, \left\{ \frac{6}{10}, |-\rangle \right\} \right\}$$

corresponds to a state which is

- $|0\rangle$ with probability 0.1
- $|+\rangle$ with probability 0.3
- $|-\rangle$ with probability 0.6

(Note, in this example, the states are not all mutually orthogonal.)

Definition: Maximally mixed state

A maximally mixed state $\{p_i, |\psi_i\rangle\}$ occurs precisely when $p_1 = p_2 = p_3 = \dots = p_n = 1/n$ (each state has an equal probability of occurring) and the states $\{|\psi_i\rangle\}$ are orthonormal. Such a uniform ensemble is often denoted by π .

1. INITIAL (PURE)

$$\text{STATE: } A = |+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$$

2. Operate with

$$\text{Hadamard: } |+\rangle \rightarrow H|+\rangle = |0\rangle$$

3. Measure in

computational basis $\{|0\rangle, |1\rangle\}$: We obtain $|0\rangle$ with certainty (probability 1).

1. INITIAL (MIXED) STATE:

$$B = \{\{0.5, |0\rangle\}, \{0.5, |1\rangle\}\}$$

2. Operate with Hadamard: New

$$\text{ensemble is } \{\{0.5, H|0\rangle\}, \{0.5, H|1\rangle\}\} = \{\{0.5, |+\rangle\}, \{0.5, |-\rangle\}\}$$

3. Measure in computational basis

$$\begin{aligned} \{|0\rangle, |1\rangle\}: \text{ We obtain } |0\rangle \text{ with probability } & (0.5)P(|0\rangle | |+\rangle) + (0.5)P(|0\rangle | |-\rangle) = \\ & 0.5(|\langle 0 | H | + \rangle|^2 + |\langle 0 | H | - \rangle|^2) = \\ & (0.5)(0.5) + (0.5)(0.5) = 0.5 \end{aligned}$$

The **density matrix** formalism consists in representing a quantum state (ket vector) using instead a matrix as follows:

Definition: Density Matrix

The density matrix ρ corresponding to a (pure) quantum state $|\psi\rangle$ is the matrix (outer product)

$$\rho = |\psi\rangle \langle \psi| \quad (36)$$

$$|\psi\rangle = |+\rangle$$

In this case,

$$\begin{aligned} \rho &= |\psi\rangle \langle \psi| = |\psi\rangle (|\psi\rangle)^{\top*} = \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} \right) \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \end{aligned} \quad (37)$$

THEOREM: The density matrix ρ is a Hermitian, positive semi-definite matrix with trace one

PROOF:

- **Hermitian** We must show that $\rho = \rho^{\top*}$ where $\rho = |\psi\rangle \langle \psi|$. We have $\rho^{\top*} = (|\psi\rangle \langle \psi|)^{\top*} = (|\psi\rangle^* \langle \psi|^{\top})^{\top} = \langle \psi|^{\top*} |\psi\rangle^{\top*} = |\psi\rangle \langle \psi| = \rho$ (we have used the property of the transpose, that for any 2 matrices A and B , $(AB)^{\top} = B^{\top} A^{\top}$) \square
- **Positive semi definite** Writing x as the (column) vector $|x\rangle$ we have $x^{\top*} \rho x = |x\rangle^{\top*} |\psi\rangle \langle \psi| x = \langle x|\psi\rangle \langle \psi|x\rangle = (\langle x|\psi\rangle)(\langle x|\psi\rangle)^* = |\langle x|\psi\rangle|^2 \geq 0$ \square
- **Trace one** If $|\psi\rangle = (v_1, v_2, \dots, v_n)^{\top}$, normalization implies $\sum_j |v_j|^2 = 1$. But, the diagonal entries of $\rho = |\psi\rangle \langle \psi|$ are $\rho_{ii} = v_i^* v_i = |v_i|^2$. So, the trace is $\text{Tr}(\rho) = \sum_j |v_j|^2 = 1$. \square

$$|\psi\rangle = \sqrt{3}|0\rangle/2 - i|1\rangle/2 \quad \text{We have now}$$

$$\begin{aligned} \rho &= |\psi\rangle \langle \psi| = \left(\frac{1}{2} \begin{pmatrix} \sqrt{3} \\ -i \end{pmatrix} \right) \left(\frac{1}{2} \begin{pmatrix} \sqrt{3} & i \end{pmatrix} \right) \\ &= \frac{1}{4} \begin{pmatrix} 3 & \sqrt{-3} \\ -\sqrt{-3} & 1 \end{pmatrix} \end{aligned} \quad (38)$$

Definition: Hermitian

A complex matrix M is **Hermitian** iff $M^{\top*} = M$

Definition: Positive semi-definite

A complex $n \times n$ matrix M is **positive semi-definite** iff, for all $x \in \mathbb{C}^n$, $x^{\top*} M x \geq 0$

Exercise 8

For a mixed state $\{p_j, |\psi_j\rangle\}$, check that the definition

$$\rho = \sum_k p_k |\psi_k\rangle \langle \psi_k| \quad (39)$$

satisfies the three conditions (Hermitian, positive semi-definite, trace one) to make it a density operator.

The density matrix provides a unique (basis independent) representation of any (pure or mixed) state.

We return to our comparison of the pure state $A = |+\rangle$ and the mixed state $B = \{0.5, |0\rangle\}, \{0.5, |1\rangle\}$ to show they have different density matrices. We saw in Equation (37) that

$$\rho_A = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

For the mixed state B we have

$$\begin{aligned}\rho_B &= \sum_j p_j |\psi_j\rangle \langle \psi_j| = p_1 |\psi_1\rangle \langle \psi_1| + p_2 |\psi_2\rangle \langle \psi_2| \\ &= 0.5 |0\rangle \langle 0| + 0.5 |1\rangle \langle 1| = 0.5 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 0.5 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= 0.5 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 0.5 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.\end{aligned}$$

Since $\rho_A \neq \rho_B$, these represent different quantum states.

Example: Density matrix from Bloch sphere representation

We will calculate the density matrix corresponding to the (pure) state defined by the Bloch sphere parameters $\theta = 2\pi/3$, $\phi = 7\pi/6$. The Bloch sphere representation gives

$$|\psi\rangle = \cos(\theta/2) |0\rangle + \sin(\theta/2) e^{i\phi} |1\rangle \quad (40)$$

We have

$$\begin{aligned}\cos(\theta/2) &= \cos(\pi/3) = 1/2 & \sin(\theta/2) &= \sin(\pi/3) = \sqrt{3}/2 \\ e^{i\phi} &= \cos \phi + i \sin \phi \\ \cos \phi &= \cos(7\pi/6) = \cos(\pi + \pi/6) \\ &= \cos(\pi) \cos(\pi/6) - \sin(\pi) \sin(\pi/6) \\ &= -\cos(\pi/6) = -\sqrt{3}/2\end{aligned} \quad (41)$$

Note that the definition of the density matrix for a mixed state, Equation (39), can also be written as $\rho = \sum_i p_i \rho_i$, where ρ_i are the individual density matrices of each pure state in the ensemble.

Example: Density matrix for a mixed state

Let us calculate the density matrix for the mixed state $\{\{1/3, |1\rangle\}, \{2/3, |+\rangle\}\}$.

The individual pure state density matrices are

$$\rho_{|1\rangle} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

$$\rho_{|+\rangle} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

$$\Rightarrow \rho = \frac{1}{3} \rho_{|1\rangle} + \frac{2}{3} \rho_{|+\rangle} = \frac{1}{3} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \frac{2}{3} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

... continued

$$\begin{aligned}\sin(\phi) &= \sin(7\pi/6) = \sin(\pi + \pi/6) \\ &= \sin(\pi) \cos(\pi/6) + \cos(\pi) \sin(\pi/6) \\ &= -\sin(\pi/6) = -1/2 \\ \Rightarrow e^{i\phi} &= -\frac{\sqrt{3} + i}{2} \Rightarrow |\psi\rangle = \frac{1}{2} |0\rangle - \frac{\sqrt{3}}{2} \left(\frac{\sqrt{3} + i}{2} \right) |1\rangle \\ &= \frac{1}{2} |0\rangle - \left(\frac{3 + i\sqrt{3}}{4} \right) |1\rangle \\ \Rightarrow \rho_{|\psi\rangle} &= \left(\frac{1}{2} |0\rangle - \left(\frac{3 + i\sqrt{3}}{4} \right) |1\rangle \right) \left(\frac{1}{2} \langle 0| - \left(\frac{3 - i\sqrt{3}}{4} \right) \langle 1| \right)\end{aligned}$$

... continued

$$= \frac{1}{2} \begin{pmatrix} 1 & \\ -(3 + i\sqrt{3})/2 & \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & (i\sqrt{3} - 3)/2 \\ & \end{pmatrix}$$

$$= \frac{1}{4} \begin{pmatrix} 1 & (i\sqrt{3} - 3)/2 \\ -(3 + i\sqrt{3})/2 & 3 \end{pmatrix}$$

Aside: We have used some trigonometric identities, which you probably should memorize (if you don't know them already):

- ▶ $\sin^2 A + \cos^2 A = 1$
- ▶ $\sin(A + B) = \sin A \cos B + \cos A \sin B$
- ▶ $\cos(A + B) = \cos A \cos B - \sin A \sin B$
- ▶ $\sin(-A) = -\sin(A)$
- ▶ $\cos(-A) = \cos(A)$

We saw above that

$$\rho_{|1\rangle} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \rho_{|+\rangle} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Both these matrices are rank one (for $\rho_{|1\rangle}$, the first row is zero times the second; for $\rho_{|+\rangle}$, the two rows are the same). It is easy also to check that $\rho^2 = \rho$ in each case. By contrast, the mixed state we analyzed, $\{\{1/3, |1\rangle\}, \{2/3, |+\rangle\}\}$ has

$$\rho = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \quad (42)$$

which has rank two, with $\rho^2 \neq \rho$.

The following are also used often (but can be derived quickly from $\sin(A + B)$ and $\cos(A + B)$ formulae above):

- ▶ $\sin(2A) = 2 \sin(A) \cos(A)$
- ▶ $\cos(2A) = \cos^2(A) - \sin^2(A)$

You may have pondered some further questions about density matrices, such as

Is it possible/easy to check if a given density matrix is that of a pure state or of a mixed state?

The answer here is, for a *pure* state, $\rho^2 = \rho$, or, alternatively, $\text{rank}(\rho) = 1$. (Reminder: The *rank* of a matrix is the number of linearly independent rows in the matrix. For a qubit, this is very easy to check - just see if one row is a multiple of the other one!)

Given a density matrix for a pure state, how would we recover the (a ?) corresponding ket vector?

We carry out an eigenvalue/eigenvector calculation on ρ . The (a ?) eigenvector is the ket vector in the standard basis. Consider again Equation (38). The density matrix is

$$\frac{1}{4} \begin{pmatrix} 3 & \sqrt{-3} \\ -\sqrt{-3} & 1 \end{pmatrix} \quad (43)$$

We calculate the eigenvalues²

$$\begin{vmatrix} 3/4 - \lambda & \sqrt{-3}/4 \\ -\sqrt{-3}/4 & 1/4 - \lambda \end{vmatrix} = 0 = \left(\frac{3}{4} - \lambda\right) \left(\frac{1}{4} - \lambda\right) - \left(\frac{\sqrt{-3}}{4}\right) \left(\frac{-\sqrt{-3}}{4}\right)$$

$$= \lambda^2 - \lambda = \lambda(\lambda - 1) \implies \lambda = 0 \text{ or } \lambda = 1. \quad (44)$$

²This is instructive, but in fact, we waste our time here: For a pure state the eigenvalues are always zero and one

Choosing $\lambda = 1$ gives as eigenvector

$$\begin{pmatrix} 3/4 - 1 & \sqrt{-3}/4 \\ -\sqrt{-3}/4 & 1/4 - 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0 \implies \frac{-1}{4}a + \frac{\sqrt{-3}}{4}b = 0$$

$$\implies a = \sqrt{-3}b \implies \begin{pmatrix} a \\ b \end{pmatrix} = a \begin{pmatrix} 1 \\ \frac{1}{\sqrt{-3}} \end{pmatrix} \quad (45)$$

Ignoring global phases (so a is a positive real number), normalization gives us $a^2(1 + 1/3) = 1 = 4a^2/3 \implies a = \sqrt{3}/2$, so we get the ket

$$|\psi\rangle = \begin{pmatrix} \sqrt{3}/2 \\ -i/2 \end{pmatrix} \quad (46)$$

as in Equation (38).

Note that

- The *Trace* of a matrix A is by definition $\text{Trace}(A) = \sum_i A_{ii}$, i.e. the sum of the diagonal entries in the matrix.

►

$$\text{Trace}(AB) = \text{Trace}(BA) \quad (49)$$

(even when matrices A and B do not commute).

- $\text{Trace}(ABC) = \text{Trace}(BCA) = \text{Trace}(CAB)$ (Cyclic permutation does not change the trace.)
- In particular, we will often use the fact that $\text{Trace}(|\psi_1\rangle\langle\psi_2|) = \langle\psi_2|\psi_1\rangle$

If we operate on a ket $|\psi\rangle$ with a unitary matrix, $|\psi\rangle \rightarrow U|\psi\rangle$, what does this correspond to for the corresponding density matrix?

$$\rho = |\psi\rangle\langle\psi| \rightarrow (U|\psi\rangle)(\langle\psi|U^\dagger) = U\rho U^\dagger \quad (47)$$

How do we calculate the expectation value of measuring an observable O using the density matrix formalism?

$$\begin{aligned} \langle\psi|O|\psi\rangle &= \sum_j \sum_k (|\psi\rangle^\dagger)_j O_{jk} (|\psi\rangle)_k = \sum_j \sum_k O_{jk} (|\psi\rangle)_k (|\psi\rangle^\dagger)_j \\ &= \sum_j \sum_k O_{jk} (|\psi\rangle)_k (\langle\psi|)_j = \sum_j \sum_k O_{jk} (|\psi\rangle\langle\psi|)_{kj} = \text{Trace}(O\rho) \end{aligned} \quad (48)$$

We now explore how the density matrix formalism extends to two (or more) qubits. Recall that in the standard basis a general two qubit state can be written as in Equation (33):

$$|\psi\rangle = \alpha_1 |00\rangle + \alpha_2 |01\rangle + \alpha_3 |10\rangle + \alpha_4 |11\rangle \quad (50)$$

We can use the same definition, Equation (36), $\rho = |\psi\rangle\langle\psi|$, which gives

$$\rho = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} \begin{pmatrix} \alpha_1^* & \alpha_2^* & \alpha_3^* & \alpha_4^* \end{pmatrix} = \begin{pmatrix} |\alpha_1|^2 & \alpha_1\alpha_2^* & \alpha_1\alpha_3^* & \alpha_1\alpha_4^* \\ \alpha_2\alpha_1^* & |\alpha_2|^2 & \alpha_2\alpha_3^* & \alpha_2\alpha_4^* \\ \alpha_3\alpha_1^* & \alpha_3\alpha_2^* & |\alpha_3|^2 & \alpha_3\alpha_4^* \\ \alpha_4\alpha_1^* & \alpha_4\alpha_2^* & \alpha_4\alpha_3^* & |\alpha_4|^2 \end{pmatrix}$$

Let us look at some examples:

$|\psi\rangle = |00\rangle$ For this separable state, we have two equivalent ways to look at the resulting density matrix:

- Writing $|\psi\rangle$ as a 4-vector $|\psi\rangle = (1 \ 0 \ 0 \ 0)^T$ we have

$$\rho = |\psi\rangle\langle\psi| = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} (1 \ 0 \ 0 \ 0) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- Alternatively, since $|\psi\rangle$ is separable, we write the tensor product $|\psi\rangle = |00\rangle = |0\rangle \otimes |0\rangle$. This gives

$$\begin{aligned} \rho &= (|0\rangle \otimes |0\rangle)(\langle 0| \otimes \langle 0|) = (|0\rangle\langle 0|) \otimes (|0\rangle\langle 0|) \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned} \quad (51)$$

$$= \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ For the maximally entangled Bell state,

$$\begin{aligned} \rho &= \frac{1}{2} (|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)(\langle 0| \otimes \langle 0| + \langle 1| \otimes \langle 1|) \\ &= \frac{1}{2} (|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 1| \otimes |0\rangle\langle 1| + |1\rangle\langle 0| \otimes |1\rangle\langle 0| \\ &\quad + |1\rangle\langle 1| \otimes |1\rangle\langle 1|) \\ &= \frac{1}{2} \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right. \\ &\quad \left. + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] \end{aligned}$$

$|\psi\rangle = (|00\rangle + |01\rangle)/\sqrt{2}$ For this separable state,

$$\begin{aligned} \rho &= \frac{1}{2} (|0\rangle \otimes |0\rangle + |0\rangle \otimes |1\rangle)(\langle 0| \otimes \langle 0| + \langle 0| \otimes \langle 1|) \\ &= \frac{1}{2} (|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 0| \otimes |0\rangle\langle 1| + |0\rangle\langle 0| \otimes |1\rangle\langle 0| \\ &\quad + |0\rangle\langle 0| \otimes |1\rangle\langle 1|) \\ &= \frac{1}{2} \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right. \\ &\quad \left. + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] \\ &= \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned} \quad (52)$$

$$\begin{aligned} &= \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

$|\psi\rangle = (|00\rangle + |01\rangle + |10\rangle)/\sqrt{3}$ This is a (non-maximally) entangled state. With three terms, the density matrix will have $3 \times 3 = 9$ terms. If we follow our previous examples, we get nine 4×4 matrices (with mostly zeroes in them).

It is considerably shorter to write

$$|\psi\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \Rightarrow \rho = \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 0 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (53)$$

How do we interpret the *overall* (global) state of the two qubits in these examples? We may have three “viewpoints”:

1. global (“referee”)
2. local 1 (Alice, first qubit)
3. local 2 (Bob, second qubit)



- We have of course a corresponding definition for $\rho_B = \text{Tr}_A \rho_{AB}$.
- Writing ρ_{AB} in the standard basis

$$\rho_{AB} = \sum_{ijkl} \alpha_{ijkl} |i\rangle \langle j| \otimes |k\rangle \langle l|$$

the reduced density matrices are

$$\rho_A = \sum_{ijk} \alpha_{ijkk} |i\rangle \langle j| \otimes \langle k|k\rangle \quad \rho_B = \sum_{ikl} \alpha_{iikl} \langle i|i\rangle \otimes |k\rangle \langle l|$$

where we have used the cyclic property of the trace to interchange $\sum_j \langle j|j\rangle$ and $\sum_j |j\rangle \langle j|$

Suppose, having prepared their two qubits in a (possibly entangled) state, Bob stays on planet earth with his qubit, while Alice leaves Wonderland and brings her qubit to Alpha Centauri (4.3 light years away). We can be sure there is no further communication between them, and each can only see their own qubit. Given a description ρ_{AB} of the overall system, can we determine what Alice and Bob “see” locally, i.e. a ρ_A for Alice’s qubit, and ρ_B for Bob’s qubit? The matrix to be calculated for Alice (ρ_A) or Bob (ρ_B) is termed a *reduced density matrix*.

Definition: Reduced density matrix

For a bipartite system given by density matrix ρ_{AB} , the **reduced density matrix** for the first party (Alice) is

$$\rho_A = \text{Tr}_B \rho_{AB} \quad \text{or in more detail} \quad \rho_A = \sum_i (I \otimes \langle b_i|) \rho_{AB} (I \otimes |b_i\rangle)$$

where $|b_i\rangle$ is any orthonormal basis for system B .

Let’s calculate some reduced density matrices from our recent examples.

$|\psi\rangle = |00\rangle$ We calculated in Equation (51) that this separable state has

$$\rho_{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$



$$\rho_A = \text{Tr}_B \rho_{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \text{Tr} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes 1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$



$$\rho_B = \text{Tr}_A \rho_{AB} = \text{Tr} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 1 \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$|\psi\rangle = (|00\rangle + |01\rangle)/\sqrt{2}$ From Equation (52)

$$\rho_{AB} = \frac{1}{2} \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right. \\ \left. + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right]$$



$$\rho_A = \text{Tr}_B \rho_{AB} = \frac{1}{2} \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \text{Tr} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \text{Tr} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right. \\ \left. + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \text{Tr} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \text{Tr} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] \\ = \frac{1}{2} \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes 1 + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes 1 \right] = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

As expected, Alice “sees” her qubit in the $|0\rangle$ state.

We note that $\text{Tr}\{|i\rangle\langle j|\} = \text{Tr}\{\langle j|i\rangle\} = \delta_{ij}$, where δ is the Kronecker delta function defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (54)$$



$$\rho_A = \text{Tr}_B \rho_{AB} = \frac{1}{2} (|0\rangle\langle 0| \otimes \text{Tr}\{|0\rangle\langle 0|\} + |0\rangle\langle 0| \otimes \text{Tr}\{|0\rangle\langle 1|\} \\ + |0\rangle\langle 0| \otimes \text{Tr}\{|1\rangle\langle 0|\} + |0\rangle\langle 0| \otimes \text{Tr}\{|1\rangle\langle 1|\}) \\ = \frac{1}{2} (|0\rangle\langle 0| \otimes 1 + |0\rangle\langle 0| \otimes 0 + |0\rangle\langle 0| \otimes 0 + |0\rangle\langle 0| \otimes 1) \\ = |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$



$$\rho_B = \text{Tr}_A \rho_{AB} = \frac{1}{2} \left[\text{Tr} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \text{Tr} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right. \\ \left. + \text{Tr} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \text{Tr} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] \\ = \frac{1}{2} \left[1 \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 1 \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 1 \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 1 \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] \\ = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

As expected, Bob “sees” his qubit in the $|+\rangle$ state. To avoid having to write down so many matrices, let's redo the calculation using the BraKet notation from the second line of Equation (52).

$$\rho \equiv \rho_{AB} = \frac{1}{2} (|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 0| \otimes |0\rangle\langle 1| + |0\rangle\langle 0| \otimes |1\rangle\langle 0| \\ + |0\rangle\langle 0| \otimes |1\rangle\langle 1|)$$



$$\rho_B = \text{Tr}_A \rho_{AB} = \frac{1}{2} (\text{Tr}\{|0\rangle\langle 0|\} \otimes |0\rangle\langle 0| + \text{Tr}\{|0\rangle\langle 0|\} \otimes |0\rangle\langle 1| \\ + \text{Tr}\{|0\rangle\langle 0|\} \otimes |1\rangle\langle 0| + \text{Tr}\{|0\rangle\langle 0|\} \otimes |1\rangle\langle 1|) \\ = \frac{1}{2} (|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$


$|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ For the maximally entangled Bell state,

$$\rho \equiv \rho_{AB} = \frac{1}{2} (|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 1| \otimes |0\rangle\langle 1| + |1\rangle\langle 0| \otimes |1\rangle\langle 0| \\ + |1\rangle\langle 1| \otimes |1\rangle\langle 1|)$$



$$\rho_A = \text{Tr}_B \rho_{AB} = \frac{1}{2} (|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$


which is a mixed state.



$$\rho_B = \text{Tr}_A \rho_{AB} = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \rho_A \quad (55)$$


$$|\psi\rangle = (|00\rangle + |01\rangle + |10\rangle)/\sqrt{3}$$

$$\rho \equiv \rho_{AB} = \frac{1}{3}(|00\rangle + |01\rangle + |10\rangle)(\langle 00| + \langle 01| + \langle 10|)$$



$$\begin{aligned} \rho_A &= \text{Tr}_B \rho_{AB} = \frac{1}{3}(|0\rangle\langle 0|\delta_{00} + |0\rangle\langle 0|\delta_{01} + |0\rangle\langle 1|\delta_{00} + |0\rangle\langle 0|\delta_{10} \\ &\quad + |0\rangle\langle 0|\delta_{11} + |0\rangle\langle 1|\delta_{10} + |1\rangle\langle 0|\delta_{00} + |1\rangle\langle 0|\delta_{01} + |1\rangle\langle 1|\delta_{00}) \\ &= \frac{1}{3}(2|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{3} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \end{aligned}$$

which is a mixed state.




$$\begin{aligned} \rho_B &= \text{Tr}_A \rho_{AB} = \frac{1}{3}(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1| + |0\rangle\langle 0|) \\ &= \frac{1}{3} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \rho_A \end{aligned} \quad (56)$$

While we have seen how to get the 4×4 density matrices from tensor products of two 2×2 individual qubit density matrices, what about the other way around?


We now describe a “trick” for the quick calculation of reduced density matrices ρ_A and ρ_B from a given 4×4 bipartite state density matrix

$$\rho_{AB} = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix}.$$



Add the numbers in the color coded areas to get the 4 numbers in ρ_A

$$\begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix} \xrightarrow{\text{Tr}_B} \begin{pmatrix} * & * \\ * & * \end{pmatrix}$$




Add the numbers in the color coded areas to get the 4 numbers in ρ_B

$$\begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix} \xrightarrow{\text{Tr}_A} \begin{pmatrix} * & * \\ * & * \end{pmatrix}$$

Example: Reduced density matrix

We calculated in Equation (52) that the density matrix for the state $|\psi\rangle = (|00\rangle + |01\rangle)/\sqrt{2}$ is


$$\rho_{|\psi\rangle} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$



$$\rho_A = \text{Tr}_B \rho_{|\psi\rangle} = \text{Tr}_B \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Example: Reduced density matrix

...continued



$$\rho_B = \text{Tr}_A \rho_{|\psi\rangle} = \text{Tr}_A \begin{pmatrix} \begin{matrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{matrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Definition: Shannon entropy

The **Shannon entropy** of a probability distribution $p_1, p_2, p_3, \dots, p_n$ with $0 \leq p_i \leq 1$, $\sum_i p_i = 1$, is

$$S = - \sum_{i=1}^n p_i \log p_i$$

Note:

- ▶ We will take log as the log to base two, \log_2 . Other definitions can be used (and are used), but logs to different bases just differ by a scaling factor: $a = \log_2 x$, $b = \log_B(x) \implies B^b = x \implies \log_2(B^b) = \log_2 x = b \log_2 B = a$. This means to convert back and forth from \log_2 to \log_B for any base B we multiply or divide by $\log_2 B$.
- ▶ You may think S "looks" negative with the minus sign in front. In fact it is positive (or zero). p_i is positive, but $\log p_i$ is negative (or zero). (Can you reason why?)



- ▶ We may interpret the entropy as the amount of uncertainty associated with an event (or a probability distribution): Zero entropy means complete certainty, while the larger S is, the more random is the outcome.
- ▶ In information theory terms, the larger the entropy, the less information is transmitted.

Entropy of uniform distribution (max entropy)

$$p_i = 1/n \implies S = - \sum p_i \log p_i = - \sum (1/n) \log(1/n) = - \log(1/n) = - \log(n^{-1}) = \log n$$

Zero entropy

For some particular $1 \leq j \leq n$, $p_j = 1$, and $p_i = 0$ for all $i \neq j$. Then $S = - \sum p_i \log p_i = -0 \log 0 - 0 \log 0 - 0 \log 0 \dots - 1 \log 1 \dots - 0 \log 0 = 0$ (since $\log 1 = 0$).

Definition: Von Neumann entropy

The **Von Neumann entropy** associated with a density matrix ρ is

$$S(\rho) = -\text{Tr}(\rho \log \rho)$$

If you're wondering, for $\log \rho$, what is the "log of a matrix", consider $e^X = \rho \implies X = \log \rho$, where

$$e^X = I + X^2/2 + X^3/3! + X^4/4! + \dots$$

Operationally, if $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$ are the eigenvalues of the $n \times n$ density matrix ρ , then

$$S(\rho) = -\sum_{i=1}^n \lambda_i \log \lambda_i.$$

The **Von Neumann entropy** of a density matrix is the **Shannon entropy** of its eigenvalues.

Von Neumann entropy of a pure state

A pure state corresponds to a rank one density matrix, one of whose eigenvalues is one, with all other eigenvalues zero. The resulting Von Neumann entropy is therefore zero. We give some explicit eigenvalue calculations.

$$\begin{aligned} \rho &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \implies \begin{vmatrix} 1/2 - \lambda & 1/2 \\ 1/2 & 1/2 - \lambda \end{vmatrix} = 0 \\ &= \left(\frac{1}{2} - \lambda\right)^2 - \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \lambda^2 - \lambda = \lambda(\lambda - 1) \\ &\implies \lambda = 0 \text{ or } \lambda = 1 \implies S(\rho) = -1 \log 1 - 0 \log 0 = 0. \end{aligned}$$

Von Neumann entropy of a pure state

... continued

► From Equation (53),

$$\begin{aligned} \rho &= \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ \implies |\rho - \lambda I| &= \begin{vmatrix} \frac{1}{3} - \lambda & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} - \lambda & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} - \lambda & 0 \\ 0 & 0 & 0 & -\lambda \end{vmatrix} = 0 \\ &= \lambda \left(\left(\frac{1}{3} - \lambda\right) \left[\left(\frac{1}{3} - \lambda\right)^2 - \frac{1}{9} \right] + \frac{1}{3} \left[\frac{1}{9} - \frac{1}{3} \left(\frac{1}{3} - \lambda\right) \right] \right) \\ &\quad + \frac{1}{3} \left[\frac{1}{9} - \frac{1}{3} \left(\frac{1}{3} - \lambda\right) \right] = \lambda^3(\lambda - 1) \end{aligned}$$

Von Neumann entropy of a pure state

... continued

which means the matrix has the four eigenvalues 0, 0, 0, 1 (or rather eigenvalue 1 with multiplicity 1 and eigenvalue 0 with multiplicity 3). The Von Neumann entropy is now $S(\rho) = \sum_i \lambda_i \log \lambda_i = 0 \log 0 + 0 \log 0 + 0 \log 0 + 1 \log 1 = 0$. Note that this is the entropy of a bipartite state

Some important properties of the Von Neumann entropy are

1. **Positivity:** $S(\rho) \geq 0$ (and $S(\rho) = 0$ iff ρ is a pure state).
2. the maximum value for $S(\rho)$ is $\log(n)$ and occurs for density matrix $\rho = \frac{1}{n}I$ (the maximally mixed state).
3. **Concavity:** $S(\sum_j p_j \rho_j) \geq \sum_j p_j S(\rho_j)$, for non-negative p_j summing to one.
4. **Subadditivity:** $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$, with equality precisely when ρ_{AB} is a separable state.

³if and only if. For statements A and B, when we write A iff B, we mean both $A \implies B$ and $B \implies A$.

Von Neumann entropy of a mixed state

We exhibit the subadditivity property, using the example of the pure bipartite state just considered, $|\psi\rangle = (|00\rangle + |01\rangle + |10\rangle)/\sqrt{3}$. We found that the reduced density matrices for both parties are equal and correspond to the mixed states

$$\rho_A = \rho_B = \frac{1}{3} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

The eigenvalue calculation gives

$$\begin{aligned} \left| \begin{matrix} \frac{2}{3} - \lambda & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} - \lambda \end{matrix} \right| &= 0 = \left(\frac{2}{3} - \lambda \right) \left(\frac{1}{3} - \lambda \right) - \left(\frac{1}{3} \right) \left(\frac{1}{3} \right) \\ &= \lambda^2 - \lambda + \frac{1}{9} \implies \lambda = \frac{1 \pm \sqrt{1^2 - 4(1)(1/9)}}{2(1)} = \frac{1}{2} \left(1 \pm \frac{\sqrt{5}}{3} \right) \\ &\approx 0.87 \text{ or } 0.127. \quad S(\rho_A) \approx -0.87 \log 0.87 - 0.127 \log 0.127 \approx 0.552. \end{aligned}$$

Definition: Purity

The **purity** of a quantum state with density matrix ρ is $P(\rho) = \text{Tr}(\rho^2)$.

- ▶ A pure state has purity equal to one.
- ▶ For a maximally mixed state, $P(\rho) = 1/2$ (for qubits). In the general case, $P(\rho) = 1/n$.
- ▶ We may also describe the purity in terms of the eigenvalues of ρ : $P(\rho) = \sum_j \lambda_j^2$.
 - ▶ PURE: $P(\rho) = 1^2 + 0^2 + \dots + 0^2 = 1$
 - ▶ MAXIMALLY MIXED: $P(\rho) = 1/n^2 + 1/n^2 + \dots + 1/n^2 = n/(n^2) = 1/n$

We return to our discussion of entanglement, which we described qualitatively but not quantitatively. For a **pure bipartite** state, the entanglement measure is defined as follows.

Definition: Entanglement entropy

For a pure bipartite state with density matrix ρ_{AB} , the Entanglement (or Entanglement entropy) is the Von Neumann entropy of the reduced density matrix.

Where it says in this definition “the” reduced density matrix, you may wonder “which one?” It turns out not to matter, we can calculate either $-\text{Tr} \rho_A \log(\rho_A)$ or $-\text{Tr} \rho_B \log(\rho_B)$ since $\text{Tr}_A(\rho_{AB}) = \text{Tr}_B(\rho_{AB})$ for pure bipartite systems (see Equations (55, 56)).

Example: State purity

- ▶ Calculate the purity of the state represented by the density matrix

$$\frac{1}{3} \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 2 \end{pmatrix}.$$

We have

$$\rho^2 = \frac{1}{9} \begin{pmatrix} 3 & 3\sqrt{2} \\ 3\sqrt{2} & 6 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 2 \end{pmatrix} \implies \text{Tr}\{\rho^2\} = 1.$$

- ▶ Calculate the purity of the mixed state $\{\{1/3, |1\rangle\}, \{2/3, |+\rangle\}\}$.
We saw in Equation (42) that the corresponding density matrix is

$$\rho = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \implies \rho^2 = \frac{1}{9} \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \implies \text{Tr}\{\rho^2\} = \frac{7}{9}.$$

Question: If the purity of the mixed state

$$\left\{ \left\{ \frac{1}{2}, \frac{1}{\sqrt{p}} |0\rangle + \sqrt{\frac{p-1}{p}} |1\rangle \right\}, \left\{ \frac{1}{2}, |1\rangle \right\} \right\}$$

is given as 5/9, calculate the probability of measuring $|1\rangle$.

Answer:

We calculate the density matrix:

$$\begin{aligned} \rho &= \frac{1}{2} \begin{pmatrix} \frac{1}{\sqrt{p}} \\ \frac{\sqrt{p-1}}{\sqrt{p}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{p}} & \frac{\sqrt{p-1}}{\sqrt{p}} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} \\ &= \frac{1}{2p} \begin{pmatrix} 1 & \sqrt{p-1} \\ \sqrt{p-1} & p-1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \frac{1}{p} & \frac{\sqrt{p-1}}{\sqrt{p}} \\ \frac{\sqrt{p-1}}{\sqrt{p}} & 2 - \frac{1}{p} \end{pmatrix} \end{aligned}$$

$$\Rightarrow \rho^2 = \frac{1}{4} \begin{pmatrix} \frac{1}{p^2} + \frac{p-1}{p^2} & * \\ * & 2 - \frac{1}{p} + \frac{p-1}{p^2} \end{pmatrix} \Rightarrow \text{Tr}\{\rho^2\} = \frac{2p^2 + p - 1}{4p^2}.$$

We write * since we don't bother calculating these terms, they have no effect on the trace. Setting $(2p^2 + p - 1)/(4p^2) = 5/9$ gives $p = 3$. We then calculate

$$\text{Prob}(|1\rangle) = \frac{1}{2} \left(\sqrt{\frac{p-1}{p}} \right)^2 + \frac{1}{2} = \left(\frac{1}{2} \right) \left(\frac{2}{3} \right) + \frac{1}{2} = \frac{5}{6}$$

Theorem: Unitary operations do not change purity

Proof: We have seen (Equation (47)) that under a Unitary operation U ,

$$\begin{aligned} \rho &= |\psi\rangle \langle\psi| \rightarrow \rho' = (U|\psi\rangle)(\langle\psi|U^\dagger) = U\rho U^\dagger \\ \Rightarrow (\rho')^2 &= U\rho U^\dagger U\rho U^\dagger = U\rho(I)\rho U^\dagger = U\rho^2 U^\dagger \\ \Rightarrow \text{Tr}\{(\rho')^2\} &= \text{Tr}\{U\rho^2 U^\dagger\} = \text{Tr}\{\rho^2 U^\dagger U\} = \text{Tr}\{\rho^2\} \end{aligned}$$

where we have used $UU^\dagger = I$ and the cyclic property of the trace ($\text{Tr}\{ABC\} = \text{Tr}\{BCA\}$, etc.)

If the 2×2 density matrix

$$\rho = \begin{pmatrix} a & b \\ b^* & c \end{pmatrix}$$

represents a pure quantum state, then we must have

$$ac = |b|^2.$$

- **Proof 1:** This matrix will have rank one (and hence represent a pure state) if (for example) the first row is a multiple of the second: $a = kb^*$ and $b = kc$. But this implies $ac = (kb^*)(b/k) = bb^* = |b|^2$.
- **Proof 2:** Doing the eigenvalue calculation gives characteristic polynomial $\lambda^2 - \lambda + ac - |b|^2$. (We have used the fact that $\text{Tr}\{\rho\} = 1 = a + c$.) From this, we obtain the polynomial $\lambda^2 - \lambda$ (which is factorizable to give eigenvalues 1 and 0) only if $ac - |b|^2 = 0$.

CPTP stands for **C**ompletely **P**ositive **T**race **P**reserving. These are *superoperators* that map linear operators (e.g. density operators) to other linear operators: $\rho \longrightarrow \rho' = \mathcal{L}(\rho)$.

Recall first that density matrices

- ▶ have non-negative eigenvalues between zero and one (**physics interpretation: these are probabilities**). This is the **positive** part.
- ▶ have eigenvalues that sum to one (**physics interpretation: probabilities sum to one**). This is the **trace one** part.

So, a CPTP map will send a positive, trace one density matrix to another positive, trace one density matrix.

This would be fine if we were working with a single quantum state (on one dimension), but in bipartite, tripartite, multipartite situations, it is not sufficient. For simplicity, we consider the bipartite case, with Hilbert spaces $\mathcal{H}_A \otimes \mathcal{H}_B$.

Example: Positive but not completely positive ... continued

We now consider $\mathcal{L} \otimes I$ acting on the bipartite state $|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. We have

$$\begin{aligned}\rho_{|\psi\rangle} &= \frac{1}{2}(|00\rangle + |11\rangle) \otimes (\langle 00| + \langle 11|) \\ &= \frac{1}{2}(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 1| \otimes |1\rangle\langle 0| \\ &\quad + |1\rangle\langle 0| \otimes |1\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|)\end{aligned}$$

Then

$$\begin{aligned}(\mathcal{L} \otimes I)\rho_{|\psi\rangle} &= \frac{1}{2}(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 0| \otimes |0\rangle\langle 1| \\ &\quad + |0\rangle\langle 1| \otimes |1\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|)\end{aligned}$$

where we have shaded in corresponding colors the terms changed by the transpose operation.

Definition: Completely positive

The superoperator $\mathcal{L}: \mathcal{H}_A \rightarrow \mathcal{H}_A, \rho \mapsto \rho'$ is **completely positive** iff it is positive on \mathcal{H}_A and also $\mathcal{L} \otimes I$ is positive on $\mathcal{H}_A \otimes \mathcal{H}_B$, for any \mathcal{H}_B .

The intuitive interpretation here is that, not only must the superoperator preserve positivity (*probabilities*) on \mathcal{H}_A , it must also preserve probabilities on larger spaces ($\mathcal{H}_A \otimes \mathcal{H}_B$) into which \mathcal{H}_A is embedded.

We give an example of a superoperator which is positive but not completely positive.

Example: Positive but not completely positive

$\mathcal{L}(\rho) = \rho^\top$. Clearly this is trace preserving (taking the transpose does not change diagonal elements). The characteristic polynomial of ρ^\top equals that of ρ , meaning they have identical eigenvalues.

Example: Positive but not completely positive ... continued

$$\rho' = (\mathcal{L} \otimes I)\rho_{|\psi\rangle} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

This is positive iff all eigenvalues are greater than or equal to zero. We calculate

$$|\rho'| = \begin{vmatrix} \frac{1}{2} - \lambda & 0 & 0 & 0 \\ 0 & -\lambda & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\lambda & 0 \\ 0 & 0 & 0 & \frac{1}{2} - \lambda \end{vmatrix} = \left(\frac{1}{2} - \lambda\right) \begin{vmatrix} -\lambda & \frac{1}{2} & 0 \\ \frac{1}{2} & -\lambda & 0 \\ 0 & 0 & \frac{1}{2} - \lambda \end{vmatrix}$$

$= \left(\frac{1}{2} - \lambda\right) [(-\lambda)(-\lambda)\left(\frac{1}{2} - \lambda\right) - \frac{1}{2}\left(\frac{1}{2}\right)\left(\frac{1}{2} - \lambda\right)]$
 $= \left(\frac{1}{2} - \lambda\right)^2 \left(\lambda^2 - \frac{1}{4}\right) \implies \lambda = \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}$. Since a negative eigenvalue appears, we conclude $\mathcal{L} \otimes I$ is not positive on $\mathcal{H}_A \otimes \mathcal{H}_B$.

Let us write out explicitly the action of the transpose superoperator $\mathcal{L}(\rho) = \rho^\top$. We consider the simplest case of two qubits, so ρ_A and ρ_B are 2×2 matrices, and the overall system has a 4×4 density matrix ρ_{AB} . Then

$$(\mathcal{L} \otimes I)\rho_{AB} = (\mathcal{L} \otimes I) \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix} = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix}$$

while

$$(I \otimes \mathcal{L})\rho_{AB} = (I \otimes \mathcal{L}) \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix} = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix}$$

We now give some examples of quantum channels and their associated Kraus operators....and the effect the channel has on the quantum state. In what follows sometimes we will use notation like $|\psi_S\rangle \otimes |\psi_E\rangle$ where S denotes our system (think *qubit*) and E denotes the environment. You may think interchangeably of “Alice and Bob” or “System and Environment”.

Example: The amplitude damping channel

For simplicity $|\psi_S\rangle$ and $|\psi_E\rangle$ are both qubit states, and we are interested in what the channel does to $|\psi_S\rangle$. The behaviour of the amplitude damping channel on the overall system is given by a Unitary matrix U that acts as follows (with $0 \leq p \leq 1$):

$$\begin{aligned} U(|00\rangle) &= |00\rangle & U(|11\rangle) &= |11\rangle \\ U(|01\rangle) &= \sqrt{1-p}|01\rangle + i\sqrt{p}|10\rangle \\ U(|10\rangle) &= i\sqrt{p}|01\rangle + \sqrt{1-p}|10\rangle \end{aligned}$$

Definition: Quantum Channel

A **Quantum Channel** is a CPTP map

Quantum Channels (and so CPTP maps) can be described by giving their **Kraus operators** which we now define.

Kraus Representation Theorem

The Kraus representation theorem^a states that a CPTP map \mathcal{L} can be represented as

$$\mathcal{L}(\rho) = \sum_k M_k \rho M_k^\dagger \quad \text{with} \quad \sum_k M_k^\dagger M_k = I \quad (57)$$

where M_k are known as the *Kraus operators*.

^aThe proof of this theorem uses something called the “Choi-Jamiołkowski isomorphism”, which is beyond the scope of this course.

Example: The amplitude damping channel ... continued

As can be seen, the channel “mixes” the $|01\rangle$ and $|10\rangle$ states. In the standard basis we can write

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-p} & i\sqrt{p} & 0 \\ 0 & i\sqrt{p} & \sqrt{1-p} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(You should check that U is indeed Unitary, i.e. that $UU^\dagger = I$, the 4×4 identity matrix, so it is a permissible quantum operation on the overall system+environment.) We write U in bra-ket notation as

$$\begin{aligned} U &= |00\rangle\langle 00| + \sqrt{1-p}|01\rangle\langle 01| + i\sqrt{p}|01\rangle\langle 10| \\ &\quad + i\sqrt{p}|10\rangle\langle 01| + \sqrt{1-p}|10\rangle\langle 10| + |11\rangle\langle 11| \end{aligned} \quad (58)$$

where we color the system in blue and the environment in magenta.

Example: The amplitude damping channel ... continued

You should think of the Kraus operators as the action of U restricted to the system. They are given by

$$M_0 = \langle 0 | U | 0 \rangle = |0\rangle \langle 0| + \sqrt{1-p} |1\rangle \langle 1| \equiv \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}$$

$$M_1 = \langle 1 | U | 0 \rangle = i\sqrt{p} |0\rangle \langle 1| \equiv \begin{pmatrix} 0 & i\sqrt{p} \\ 0 & 0 \end{pmatrix}$$

where the first and fifth terms in Equation (58) contribute to M_0 , while only the third term contributes to M_1 . We check now that Equation (57) holds:

$$\begin{aligned} \sum_k M_k^\dagger M_k &= M_0^\dagger M_0 + M_1^\dagger M_1 \\ &= \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}^2 + \begin{pmatrix} 0 & 0 \\ -i\sqrt{p} & 0 \end{pmatrix} \begin{pmatrix} 0 & i\sqrt{p} \\ 0 & 0 \end{pmatrix} \end{aligned}$$

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

109

CPTP maps, Kraus operators and quantum channels

Note:

1. In the new density matrix,
 - ▶ the entry in the upper left corner has increased
 - ▶ the off-diagonal entries have decreased by a factor of $\sqrt{1-p}$
 - ▶ the entry in the bottom right corner has decreased by a factor of $(1-p)$.
2. Repeated application of this channel (with equal or varying values of p) will “push” the system state density matrix to

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

corresponding to $|\psi\rangle = |0\rangle$

Exercise

Can you construct a quantum channel that will “push” the system state towards $|\psi\rangle = |1\rangle$?

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

111

Example: The amplitude damping channel ... continued

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1-p \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & p \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Finally, the channel changes the density matrix as follows:

$$\begin{aligned} \mathcal{L}(\rho) &= \sum_k M_k \rho M_k^\dagger = M_0 \rho M_0^\dagger + M_1 \rho M_1^\dagger \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1-p \end{pmatrix} \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1-p \end{pmatrix} \\ &+ p \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \rho_{11} + p\rho_{12} & \sqrt{1-p}\rho_{21} \\ \sqrt{1-p}\rho_{22} & \rho_{11} - p\rho_{11} \end{pmatrix} \end{aligned}$$

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

110

CPTP maps, Kraus operators and quantum channels

Example: the phase damping channel

The Unitary defining this channel acts as follows:

$$U(|10\rangle) = |10\rangle \quad U(|11\rangle) = |11\rangle$$

$$U(|00\rangle) = \sqrt{1-p} |00\rangle + i\sqrt{p} |01\rangle$$

$$U(|01\rangle) = i\sqrt{p} |00\rangle + \sqrt{1-p} |01\rangle$$

This “mixes” $|00\rangle$ and $|01\rangle$. The matrix and bra-ket representations are

$$U = \begin{pmatrix} \sqrt{1-p} & i\sqrt{p} & 0 & 0 \\ i\sqrt{p} & \sqrt{1-p} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{aligned} U &= |10\rangle \langle 10| + \sqrt{1-p} |00\rangle \langle 00| + i\sqrt{p} |01\rangle \langle 00| \\ &+ i\sqrt{p} |00\rangle \langle 01| + \sqrt{1-p} |01\rangle \langle 01| + |11\rangle \langle 11| \end{aligned} \quad (59)$$

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

112

Example: the phase damping channel ... continued

We calculate the Kraus operators

$$M_0 = \langle 0 | U | 0 \rangle = |1\rangle \langle 1| + \sqrt{1-p} |0\rangle \langle 0| \equiv \begin{pmatrix} \sqrt{1-p} & 0 \\ 0 & 1 \end{pmatrix}$$

$$M_1 = \langle 1 | U | 0 \rangle = i\sqrt{p} |0\rangle \langle 0| \equiv \begin{pmatrix} i\sqrt{p} & 0 \\ 0 & 0 \end{pmatrix}$$

$$\sum_k M_k^\dagger M_k = M_0^\dagger M_0 + M_1^\dagger M_1 = \begin{pmatrix} 1-p & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} p & 0 \\ 0 & 0 \end{pmatrix} = I$$

$$\begin{aligned} \mathcal{L}(\rho) &= \sum_k M_k \rho M_k^\dagger = \begin{pmatrix} \sqrt{1-p} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \begin{pmatrix} \sqrt{1-p} & 0 \\ 0 & 1 \end{pmatrix} \\ &\quad + \begin{pmatrix} i\sqrt{p} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \begin{pmatrix} -i\sqrt{p} & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \rho_{11} & \sqrt{1-p}\rho_{12} \\ \sqrt{1-p}\rho_{21} & \rho_{22} \end{pmatrix} \end{aligned}$$

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

113

CPTP maps, Kraus operators and quantum channels

The Kraus representation (Equation (57)) is not unique. Lets show this for the phase damping channel. Using the Bra-Ket expression (Equation(59)) for U , we calculate (different) Kraus operators with regard to the $|+\rangle, |-\rangle$ basis:

$$\begin{aligned} M_+ &= \langle + | U | 0 \rangle = \frac{1}{\sqrt{2}} |1\rangle \langle 1| + \sqrt{\frac{1-p}{2}} |0\rangle \langle 0| + i\sqrt{\frac{p}{2}} |0\rangle \langle 0| \\ &\equiv \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{1-p} + i\sqrt{p} & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} M_- &= \langle - | U | 0 \rangle = \frac{1}{\sqrt{2}} |1\rangle \langle 1| + \sqrt{\frac{1-p}{2}} |0\rangle \langle 0| - i\sqrt{\frac{p}{2}} |0\rangle \langle 0| \\ &\equiv \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{1-p} - i\sqrt{p} & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Since $(\sqrt{1-p} + i\sqrt{p})(\sqrt{1-p} - i\sqrt{p}) = 1$,
 $M^\dagger M = \frac{1}{2} I \implies M_+^\dagger M_+ + M_-^\dagger M_- = I$.

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

115

Example: the phase damping channel ... continued

Note:

1. the phase damping channel reduces (by a factor of $\sqrt{1-p}$) the off-diagonal entries in the density matrix
2. the system density matrix is pushed towards a diagonal density matrix - i.e. a mixed state
3. the phase damping channel moves pure states towards mixed states: This is an example of what is known as **decoherence**
4. a maximally mixed state is obtained if we start with $|+\rangle$ (i.e. an equal superposition of $|0\rangle$ and $|1\rangle$)

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

114

CPTP maps, Kraus operators and quantum channels

Finally

$$\begin{aligned} \mathcal{L}(\rho) &= M_+ \rho M_+^\dagger + M_- \rho M_-^\dagger \\ &= \frac{1}{2} \begin{pmatrix} \sqrt{1-p} + i\sqrt{p} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \begin{pmatrix} \sqrt{1-p} - i\sqrt{p} & 0 \\ 0 & 1 \end{pmatrix} \\ &\quad + \frac{1}{2} \begin{pmatrix} \sqrt{1-p} - i\sqrt{p} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \begin{pmatrix} \sqrt{1-p} + i\sqrt{p} & 0 \\ 0 & 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} (\sqrt{1-p} + i\sqrt{p})\rho_{11} & (\sqrt{1-p} + i\sqrt{p})\rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \begin{pmatrix} \sqrt{1-p} - i\sqrt{p} & 0 \\ 0 & 1 \end{pmatrix} \\ &\quad + \frac{1}{2} \begin{pmatrix} (\sqrt{1-p} - i\sqrt{p})\rho_{11} & (\sqrt{1-p} - i\sqrt{p})\rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \begin{pmatrix} \sqrt{1-p} + i\sqrt{p} & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \rho_{11} & \sqrt{1-p}\rho_{12} \\ \sqrt{1-p}\rho_{21} & \rho_{22} \end{pmatrix} \end{aligned}$$

which is identical to what we previously calculated using Kraus operators M_0 and M_1 .

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

116

Definition: POVM

A **Positive Operator Valued Measure** is a set of operators $\{E_i\}$ that are

- ▶ Hermitian: $E_i = E_i^\dagger$
- ▶ Positive semi-definite: $\langle \psi | E_i | \psi \rangle \geq 0$ for any $|\psi\rangle$
- ▶ Complete: $\sum_i E_i = I$

Note that

- ▶ The number of POVMs in the set $\{E_i\}$ is not restricted - it can be greater than the dimension of the Hilbert space (so for qubits, we can have 3 or more E_i).
- ▶ The probability of obtaining outcome i is $\text{Tr}\{\rho E_i\}$, for a quantum state with density matrix ρ .
- ▶ Projective measurements are POVMs with the extra property that each E_i is a projector, i.e. $E_i^2 = E_i$.

POVM Example

... continued

Consider the state $|\psi\rangle = \frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle$ with

$$\rho_{|\psi\rangle} = \frac{1}{25} \begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix}$$

Measuring this state using this POVM would give

- ▶ Outcome 1 with probability

$$\text{Tr}\left\{\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix} \frac{1}{25} \begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix}\right\} = \frac{9}{50}$$

- ▶ Outcome 2 with probability

$$\text{Tr}\left\{\frac{1}{4} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \frac{1}{25} \begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix}\right\} = \frac{1}{100} \text{Tr} \begin{pmatrix} -3 & -4 \\ 3 & 4 \end{pmatrix} = \frac{1}{100}$$

POVM Example

$$E_1 = \frac{1}{2}|0\rangle\langle 0|; \quad E_2 = \frac{1}{2}|-\rangle\langle -|; \quad E_3 = 1 - E_1 - E_2.$$

or equivalently

$$E_1 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \quad E_2 = \frac{1}{4} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}; \quad E_3 = \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}.$$

It is easy to check these are not projectors: For example

$$E_1^2 = \frac{1}{4} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2} E_1 \neq E_1$$

The completeness property is easily seen from the definition of $E_3 = 1 - E_1 - E_2$.

POVM Example

... continued

- ▶ Outcome 3 with probability

$$\text{Tr}\left\{\frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \frac{1}{25} \begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix}\right\} = \frac{1}{100} \text{Tr} \begin{pmatrix} 21 & 28 \\ 45 & 60 \end{pmatrix} = \frac{81}{100}$$

We now consider a different example with five operators,

$$E_k = \frac{2}{5} \begin{pmatrix} \cos^2 \theta_k & \cos \theta_k \sin \theta_k \\ \cos \theta_k \sin \theta_k & \sin^2 \theta_k \end{pmatrix}$$

with $k \in \{0, 1, 2, 3, 4\}$ and $\theta_k = 2\pi k/5$. Let's check these are not projectors:

$$\begin{aligned} E_k^2 &= \frac{4}{25} \begin{pmatrix} \cos^4 \theta_k + \cos^2 \theta_k \sin^2 \theta_k & \cos^3 \theta_k \sin \theta_k + \cos \theta_k \sin^3 \theta_k \\ \cos^3 \theta_k \sin \theta_k + \cos \theta_k \sin^3 \theta_k & \cos^2 \theta_k \sin^2 \theta_k + \sin^4 \theta_k \end{pmatrix} \\ &= \frac{4}{25} \begin{pmatrix} \cos^2 \theta_k (\cos^2 \theta_k + \sin^2 \theta_k) & \cos \theta_k \sin \theta_k (\cos^2 \theta_k + \sin^2 \theta_k) \\ \cos \theta_k \sin \theta_k (\cos^2 \theta_k + \sin^2 \theta_k) & \sin^2 \theta_k (\cos^2 \theta_k + \sin^2 \theta_k) \end{pmatrix} \end{aligned}$$

$$= \frac{4}{25} \begin{pmatrix} \cos^2 \theta_k & \cos \theta_k \sin \theta_k \\ \cos \theta_k \sin \theta_k & \sin^2 \theta_k \end{pmatrix} = \frac{2}{5} E_k \neq E_k.$$

The spectral decomposition (calculation of eigenvalues and eigenvectors) follows:

$$0 = |E_k - \lambda I| = \begin{vmatrix} 0.4 \cos^2 \theta_k - \lambda & 0.4 \cos \theta_k \sin \theta_k \\ 0.4 \cos \theta_k \sin \theta_k & 0.4 \sin^2 \theta_k - \lambda \end{vmatrix}$$

$$= \lambda^2 - 0.4\lambda(\sin^2 \theta_k + \cos^2 \theta_k) = \lambda(\lambda - 0.4) \implies \lambda = 0 \text{ or } \lambda = 0.4.$$

Eigenvector associated with eigenvalue $\lambda = 0.4$ (In the following we abbreviate $\cos \theta_k$ by c_k , $\sin \theta_k$ by s_k etc.):

$$\begin{pmatrix} 0.4c_k^2 - 0.4 & 0.4c_k s_k \\ 0.4c_k s_k & 0.4s_k^2 - 0.4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies (c_k^2 - 1)a + c_k s_k b = 0$$

$$\implies c_k s_k b = (1 - c_k^2)a = s_k^2 a \implies a = \frac{c_k}{s_k} b$$

Normalization: $a^2 + b^2 = 1 = b^2 \left(\frac{c_k^2}{s_k^2} + 1 \right) = \frac{b^2}{s_k^2} \implies b = s_k$

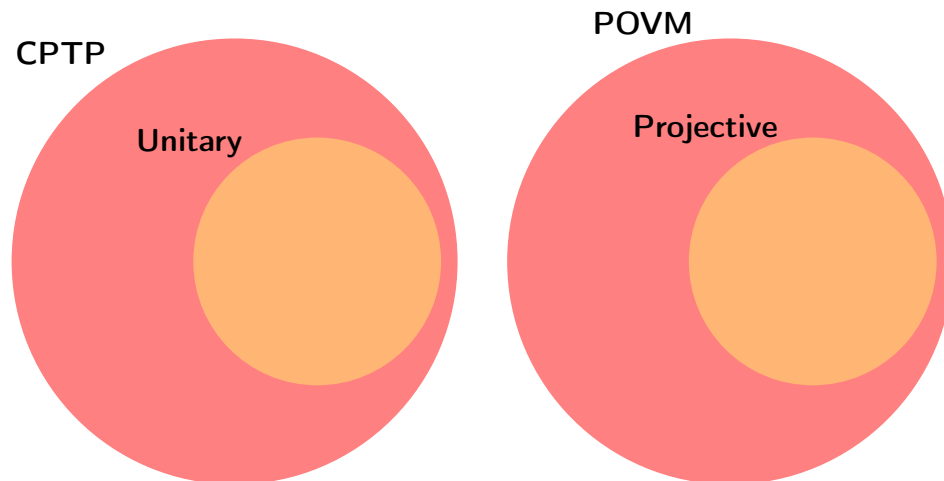


Figure 1: Venn Diagram illustrations of the relation between CPTP, POVM, Unitary and Projective Operations: Unitary \subset CPTP and Projective \subset POVM.

We have the eigenvalue equation $E_k v_k = \lambda v_k$ with

$$\lambda = 0.4 = \frac{2}{5} \quad v_k = \begin{pmatrix} \cos \theta_k \\ \sin \theta_k \end{pmatrix}$$

and we can write

$$E_k = \frac{2}{5} |v_k\rangle \langle v_k|$$

Definition: Kraus operator decomposition of a POVM

A Kraus operator decomposition of the POVM $\{E_k\}$ is a (non-unique) family of matrices $\{M_k\}$ with $E_k = M_k^\dagger M_k$.

Example (Kraus operator decomposition of a POVM)

From the preceding example, set $M_k = \sqrt{0.4} |v_k\rangle \langle v_k|$. Then $M_k^\dagger M_k = M_k^2 = 0.4 |v_k\rangle \langle v_k| v_k\rangle \langle v_k| = 0.4 |v_k\rangle \langle v_k| = E_k$

Returning momentarily to a single pure qubit state $\alpha |0\rangle + \beta |1\rangle$, let us represent $|\alpha|^2 = p$ and $|\beta|^2 = 1 - p$, so that p and $1 - p$ will represent the probabilities of measuring $|0\rangle$ or $|1\rangle$. We show in schematic form what the associated density matrices “look like”:

$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} .9 & .3 \\ .3 & .1 \end{pmatrix}$	$\begin{pmatrix} .8 & .4 \\ .4 & .2 \end{pmatrix}$	$\begin{pmatrix} .7 & .46 \\ .46 & .3 \end{pmatrix}$	$\begin{pmatrix} .6 & .5 \\ .5 & .4 \end{pmatrix}$
$p = 1$	$p = 0.9$	$p = 0.8$	$p = 0.7$	$p = 0.6$

$\begin{pmatrix} .5 & .5 \\ .5 & .5 \end{pmatrix}$	$\begin{pmatrix} .4 & .5 \\ .5 & .6 \end{pmatrix}$	$\begin{pmatrix} .3 & .46 \\ .46 & .7 \end{pmatrix}$	$\begin{pmatrix} .2 & .4 \\ .4 & .8 \end{pmatrix}$	$\begin{pmatrix} .1 & .3 \\ .3 & .9 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$
$p = 0.5$	$p = 0.4$	$p = 0.3$	$p = 0.2$	$p = 0.1$	$p = 0$

Figure 2: The red shading represents the (positive) magnitude of the corresponding entry in the matrix, ranging from zero (white) to one (red).

The **Deutsch-Jozsa algorithm** is the first quantum algorithm (dating from 1992) to “beat” any known classical algorithm. As with many breakthroughs, it concerns itself with an “unusual” problem/thought experiment:

QUESTION:

Suppose I give you a function $f(x)$, defined on a finite subset of the integers. The range of the function is **true/false** or 0/1. I **promise** you (pinky promise ☺) that the function f has exactly one of the following properties:

- ▶ **property 1:** $f(x)$ has the same value for all x , (**constant**), or
- ▶ **property 2:** $f(x)$ has value 0 for half of the values of x , and value 1 for the other half (**balanced**).

The question is: How will we go about determining which property holds?

- ▶ $f : \{0, 1, 2, 3, 4, 5\} \rightarrow \{0, 1\}$ with

x	0	1	2	3	4	5
f(x)	0	0	1	0	1	1

Again, evaluating from the left,

we must check $f(0)$, then $f(1)$, then $f(2)$, and at that point we conclude $f(x)$ must be **balanced**.

- ▶ $f : \{0, 1, 2, 3, 4, 5\} \rightarrow \{0, 1\}$ with

x	0	1	2	3	4	5
f(x)	1	1	1	0	0	0

Evaluating from the left, we

need four comparisons to determine that f is in fact **balanced** rather than constant.

- ▶ $f : \{0, 1, 2, 3, 4, 5\} \rightarrow \{0, 1\}$ with

x	0	1	2	3	4	5
f(x)	0	0	0	0	0	0

Evaluating from the left, we

need four comparisons to determine that f is in fact **constant** rather than balanced.

ANSWER: Let's think at first about the classical “best response” we can give to this problem. We can evaluate $f(x)$ for a few values of x . If we by chance find different results for $f(x)$ among the few we test, we conclude **property 2** holds and we are finished. On the other hand, if amongst the few values of $f(x)$ we evaluate, we find the same result, we can conclude nothing. We must continue until we have evaluated $f(x)$ for at least half the values in the domain of f . Examples:

- ▶ $f : \{0, 1, 2, 3, 4, 5\} \rightarrow \{0, 1\}$ with

x	0	1	2	3	4	5
f(x)	0	1	1	0	0	1

If we evaluate “from the left”,

after two comparisons, we determine the function must be **balanced** (since $f(0) \neq f(1)$).

We are now considering questions in computer science, in particular in the computational complexity of a particular algorithm, i.e. how many “steps” does it take to determine the result. For this particular problem, it should be clear that the classical computational complexity (the number of operations we have to make as a function of the size of the input) is:

Best case scenario: 2 comparisons

Worst case scenario: $N/2$ comparisons

This means that the average case scenario is linear in N .

Deutsch-Jozsa Algorithm

We know that the Hadamard matrix has the behaviour $H|0\rangle = |+\rangle$ and $H|1\rangle = |-\rangle$. We write this as

$$H|0\rangle = \frac{1}{\sqrt{2}} \sum_{y \in \{0,1\}} (-1)^{0 \cdot y} |y\rangle \quad H|1\rangle = \frac{1}{\sqrt{2}} \sum_{y \in \{0,1\}} (-1)^{1 \cdot y} |y\rangle \quad (61)$$

Deutsch-Jozsa Algorithm

... continued

We can package the two Equations (61) into one equation:

$$H|x\rangle = \frac{1}{\sqrt{2}} \sum_y (-1)^{x \cdot y} |y\rangle$$

with $x, y \in \{0, 1\}$. Extending this to 2 qubits, we get

$$\begin{aligned} (H|x_1\rangle) \otimes (H|x_2\rangle) &= \frac{1}{2} \left(\sum_{y_1} (-1)^{x_1 \cdot y_1} |y_1\rangle \right) \otimes \left(\sum_{y_2} (-1)^{x_2 \cdot y_2} |y_2\rangle \right) \\ &= \frac{1}{2} \sum_{y_1, y_2} (-1)^{x_1 \cdot y_1 + x_2 \cdot y_2} (|y_1\rangle \otimes |y_2\rangle) = \frac{1}{2} \sum_{y_1, y_2} (-1)^{x \cdot y} |y_1 y_2\rangle. \end{aligned}$$

For n qubits, writing x for bit string $x_1 x_2 \dots x_n$ etc.,

$$H^{\otimes n} |x\rangle = \frac{1}{\sqrt{2^n}} \sum_y (-1)^{x \cdot y} |y\rangle. \quad (62)$$

Deutsch-Jozsa Algorithm

... continued

Let $f : \{0, 1, \dots, 2^n - 1\} \rightarrow \{0, 1\}$ be the function in question (with $N = 2^n$). We construct a Unitary operation

$$U_f : |x\rangle |y_0\rangle \rightarrow |x\rangle |y_0 \oplus f(x)\rangle \quad (63)$$

with $x \in \{0, 1\}^n$ and $y_0 \in \{0, 1\}$ (\oplus here is addition modulo 2).

For example:

$n = 1, f(0) = f(1) = 1$ Then $U_f(|00\rangle) = |01\rangle$, $U_f(|01\rangle) = |00\rangle$, $U_f(|10\rangle) = |11\rangle$, $U_f(|11\rangle) = |10\rangle$. This is a constant function.

$n = 2, f(0) = f(2) = 1, f(1) = f(3) = 0$

$ \psi\rangle$	$ 00\rangle$	$ 10\rangle$	$ 20\rangle$	$ 30\rangle$	$ 01\rangle$	$ 11\rangle$	$ 21\rangle$	$ 31\rangle$
$U_f \psi\rangle$	$ 01\rangle$	$ 10\rangle$	$ 21\rangle$	$ 30\rangle$	$ 00\rangle$	$ 11\rangle$	$ 20\rangle$	$ 31\rangle$

(In this table - the red digits are integers representing 2-digit binary strings, for example 3 is 11, 1 is 01, etc.) It should be clear from these examples that U_f is a permutation matrix.

Deutsch-Jozsa Algorithm

... continued

A final ingredient we need is the property that

$$U_f(|x\rangle |-\rangle) = (-1)^{f(x)} |x\rangle |-\rangle$$

which we can show by examining the two cases:

- $f(x) = 0$: $U_f(|x\rangle |-\rangle) = |x\rangle |-\rangle = (-1)^0 |x\rangle |-\rangle$
- $f(x) = 1$: $U_f(|x\rangle |-\rangle) = |x\rangle (|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle) = |x\rangle (|1\rangle - |0\rangle) = -|x\rangle |-\rangle = (-1)^1 |x\rangle |-\rangle$

The algorithm is

- **Step 1** Input the state $|0\rangle^{\otimes n} |1\rangle$
- **Step 2** Apply H to all qubits to get

$$\frac{1}{\sqrt{2^n}} \sum_x |x\rangle |-\rangle$$

Deutsch-Jozsa Algorithm

... continued

- **Step 3** Apply our Unitary U_f to get

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

- **Step 4** Apply $H^{\otimes n} \otimes I$ (and use Equation (62)) to get

$$\begin{aligned} &\frac{1}{\sqrt{2^n}} \sum_x (-1)^{f(x)} \left(\frac{1}{\sqrt{2^n}} \sum_y (-1)^{x \cdot y} |y\rangle \right) |-\rangle \\ &= \frac{1}{2^n} \sum_{x,y} (-1)^{f(x) + x \cdot y} |y\rangle |-\rangle \end{aligned}$$

Deutsch-Jozsa Algorithm

... continued

- **Step 5** Measure. The probability of getting the $|0\rangle|-\rangle$ state is

$$\left| \frac{1}{2^n} \sum_x (-1)^{f(x)} \right|^2 = \begin{cases} 1 & \text{if } f(x) \text{ is constant} \\ 0 & \text{if } f(x) \text{ is balanced} \end{cases}$$

This algorithm gives exponential speedup over the best known classical algorithm. Lets compare the number of operations in the worst case scenario for $n = 20$ (i.e. $N = 2^{20} = 1048576$):

- **Classical:** Worst case, we have to check 524289 values of $f(x)$ (i.e. $1048576/2 + 1$).
- **Quantum:** Go through the steps to count the number of operations we applied: • 20 Hadamard operations • One application of U_f • 20 more Hadamard operations • measurement: A total of about 42 operations.

In the circuit model, we represent schematically quantum algorithms as follows:

- Time flows from left to right.
- At the very left, we have the input (qubit state(s)).
- At the very right, we have the output (projective/POVM measurements).
- In the middle we have Unitary/CPTP operations (represented mostly by boxes or rectangles).
- Each “wire” running from left to right represents a single qubit.
- For each Unitary operation, the number of input wires should equal the number of output wires, to ensure reversibility.
- Unitary operations (U) act on a single or on multiple qubits. If U has n input wires from the left (and therefore also n output wires), it is represented by a $2^n \times 2^n$ matrix.

Some further observations. . .

1. Applying the Hadamard operation to every qubit gives us an equal “mixture” of all the numbers from 0 to $2^n - 1$. This allows us in some sense to apply $f(x)$, via U_f , to all possible inputs simultaneously (hence references, in popular media / books / online, to “massive parallelism” in quantum computing).
2. In this algorithm and others in quantum computing, the speedup is obtained in a problem where we ask some **global** question about a function (not a local question). (A local question would be something like “What is the value of $f(5)$?”). The quantum algorithm does not actually obtain the answer to any such local question.) The quest for new, efficient algorithms in quantum computing focuses on which such global questions can we ask that may have useful applications.
3. If we want to actually ask “What is the value of $f(5)$?”, there is no quantum algorithm (currently) that will improve on the classical algorithm

Lets look at some sample circuits:

$$-|0\rangle + |1\rangle = -|-\rangle = \sigma_X(H(|1\rangle)):$$

$$|1\rangle \xrightarrow{H} \xrightarrow{X} -|0\rangle + |1\rangle$$

Note here we write X instead of the Pauli matrix σ_X .

$$\frac{1}{2}(|00\rangle - |01\rangle - |10\rangle + |11\rangle) = (I \otimes H)(H \otimes \sigma_X)|10\rangle:$$

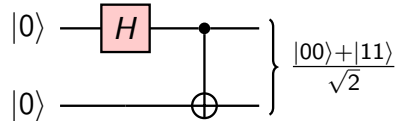
$$\begin{array}{c} |1\rangle \xrightarrow{H} \\ |0\rangle \xrightarrow{X} \end{array} \xrightarrow{H} \left. \begin{array}{c} \text{---} \\ \text{---} \end{array} \right\} \frac{|00\rangle - |01\rangle - |10\rangle + |11\rangle}{2}$$

Circuit to create a Bell State:

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = (CNOT)(H \otimes I)|00\rangle:$$

CNOT is the Control-NOT gate, defined by

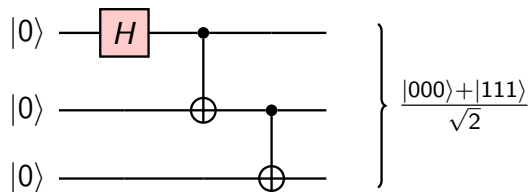
$ \psi\rangle$	$ 00\rangle$	$ 01\rangle$	$ 10\rangle$	$ 11\rangle$
CNOT $ \psi\rangle$	$ 00\rangle$	$ 01\rangle$	$ 11\rangle$	$ 10\rangle$



For CNOT, the first qubit controls the second qubit: The first qubit does not change, and if the first qubit is zero, it does nothing to the second qubit, while if it is one, it flips the second qubit. The CNOT Unitary matrix is a permutation matrix. On its own, the NOT gate is of course just Pauli X, σ_x . Because of its wide use, the CNOT has its own special symbol in quantum circuits, with a bullet representing the control and a larger circle for the target.

Circuit to generate the tripartite GHZ state:

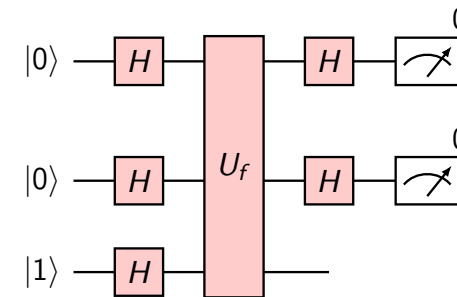
$$|GHZ\rangle = (CNOT(2,3))(CNOT(1,2))(H \otimes I \otimes I)|000\rangle:$$



Observe that the last CNOT gate could alternatively have its control on the first qubit. You should check explicitly the matrices corresponding to these previous circuit examples. We show here just the matrices for the Bell State circuit:

Circuit for Deutsch-Jozsa algorithm (for a function with $n = 2$ whose domain is $\{0, 1, 2, 3\}$):

$$(|0\rangle\langle 0| \otimes |0\rangle\langle 0| \otimes I)(H \otimes H \otimes I)(U_f)(H \otimes H \otimes H)|001\rangle:$$



Note that we show generalized Unitary operations on a number of qubits as a rectangle extending across the corresponding qubit wires.

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = (CNOT)(H \otimes I)|00\rangle:$$

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \right) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Control gates

- The CNOT can be (and is) extended to CCNOT, CCCNOT etc.: In all cases, the target qubit is flipped (Pauli X) only if all the control qubits are $|1\rangle$.
- CNOT is also generalized to C_U , the Control-Unitary gate, where a single qubit Unitary operator U is applied to the target qubit only if the control qubit is $|1\rangle$.
- If you put on your programming / computer science hat for a minute - the Control gate family looks very much like the `if` statement in any programming language.

We can define

$$CNOT(|a\rangle|b\rangle) = \begin{cases} |a\rangle|b\rangle & \text{if } a = 0 \\ |a\rangle|b \oplus 1\rangle & \text{if } a = 1 \end{cases}$$

$$CCNOT(|a\rangle|b\rangle|c\rangle) = \begin{cases} |a\rangle|b\rangle|c \oplus 1\rangle & \text{if } a = b = 1 \\ |a\rangle|b\rangle|c\rangle & \text{otherwise} \end{cases}$$

$$C_U(|a\rangle \otimes |b\rangle) = \begin{cases} |a\rangle \otimes |b\rangle & \text{if } a = 0 \\ |a\rangle \otimes U|b\rangle & \text{if } a = 1 \end{cases}$$

$$CC_U(|a\rangle \otimes |b\rangle \otimes |c\rangle) = \begin{cases} |a\rangle \otimes |b\rangle \otimes U|c\rangle & \text{if } a = b = 1 \\ |a\rangle \otimes |b\rangle \otimes |c\rangle & \text{otherwise} \end{cases}$$

etcetera

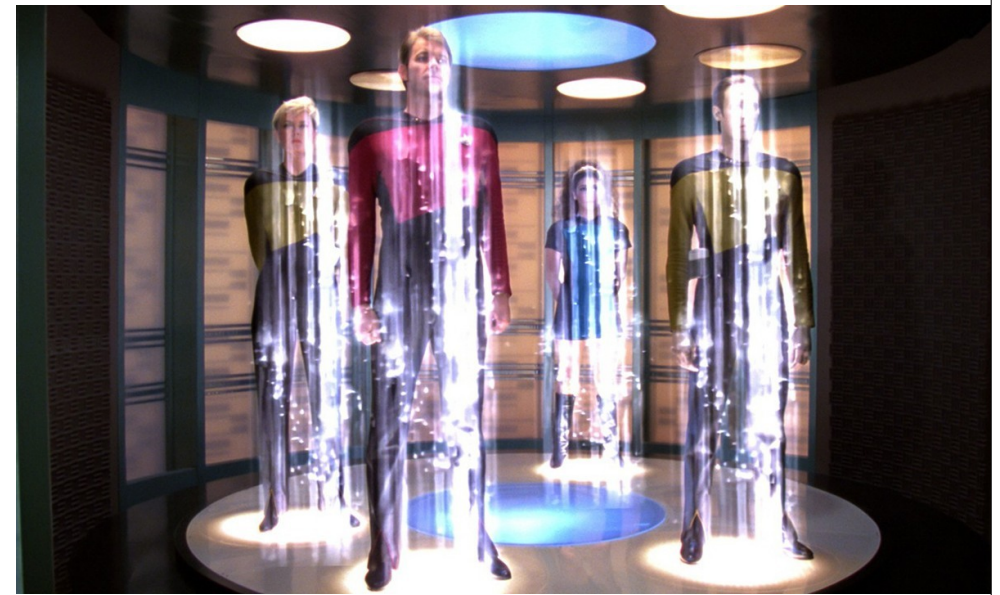


Figure 3: Teleportation, as imagined onboard the USS Enterprise in the science fiction series “Star Trek”

Quantum Teleportation

We now examine the phenomenon of **quantum teleportation**⁴, which allows quantum information to be sent using classical communication. At the cost of

- ▶ destroying one shared entangled Bell state
- ▶ transmitting two classical bits of information (or, a number from the set $\{0, 1, 2, 3\}$)

this allows Alice to send a qubit state to Bob. The steps are:

- ▶ **Step 0** Alice and Bob prepare a Bell pair $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ in their lab in Shanghai. Bob takes his qubit to Tokyo.
- ▶ **Step 1** In the Shanghai lab, Alice is given an (unknown) qubit $|\psi\rangle$ by her colleague Charlie. She must transmit the qubit state exactly to Bob, using classical communication, and without travelling to Tokyo. The overall state of the system is $|\psi\rangle \otimes |\Phi^+\rangle = |\psi\rangle \otimes (\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle))$ where we color blue/red the qubits in Shanghai/Tokyo respectively.

⁴In Quantum Computing, you can “cut and paste”. See also footnote on page 36.

Quantum Teleportation

Setting $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ our system is in the state

$$\begin{aligned} & (\alpha|0\rangle + \beta|1\rangle) \otimes \left(\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \right) \\ &= \frac{1}{\sqrt{2}} (\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle) \end{aligned}$$

- ▶ **Step 2** Alice performs a CNOT operation, followed by a Hadamard on her first qubit:

$$\begin{aligned} & \frac{1}{\sqrt{2}} (\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle) \\ & \xrightarrow{CNOT \otimes I} \frac{1}{\sqrt{2}} (\alpha|000\rangle + \alpha|011\rangle + \beta|110\rangle + \beta|101\rangle) \\ & \xrightarrow{H \otimes I \otimes I} \frac{1}{2} (\alpha|000\rangle + \alpha|100\rangle + \alpha|011\rangle + \alpha|111\rangle \\ & \quad + \beta|010\rangle - \beta|110\rangle + \beta|001\rangle - \beta|101\rangle) \end{aligned}$$

Note that this state can be re-written as

$$\frac{1}{2} (|00\rangle (\alpha|0\rangle + \beta|1\rangle) + |01\rangle (\alpha|1\rangle + \beta|0\rangle) + |10\rangle (\alpha|0\rangle - \beta|1\rangle) + |11\rangle (\alpha|1\rangle - \beta|0\rangle))$$

- **Step 3** Alice measures her two qubits in Shanghai: With equal probability, she obtains $|00\rangle$ or $|01\rangle$ or $|10\rangle$ or $|11\rangle$. She phones Bob in Tokyo to tell him her result.
- **Step 4** Finally, Bob operates on his single qubit as follows:

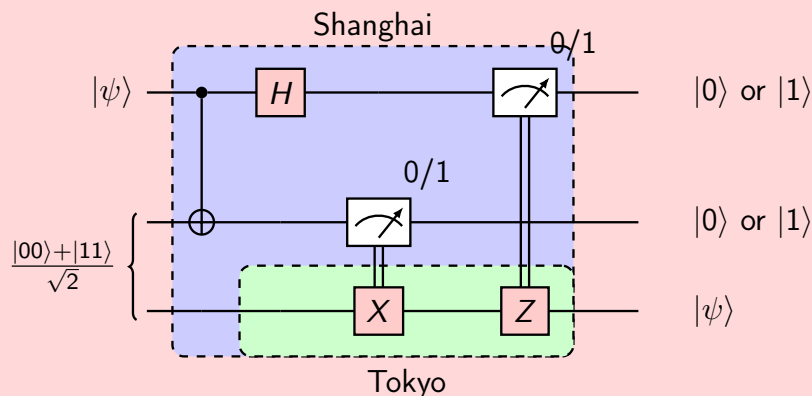
Alice's result:	00	01	10	11
Bob's operation:	I	σ_X	σ_Z	σ_X first, then σ_Z

Bob's final qubit state is now $\alpha|0\rangle + \beta|1\rangle$.

In quantum teleportation,

1. **No matter, energy or information is transmitted instantaneously.** Bob and Alice both make local measurements/operations, and information is transmitted classically (at a speed below the speed of light, so as not to break laws of general relativity).
2. Alice's qubit (or rather the one she was given by Charlie in the example) is not physically transmitted from the Shanghai to the Tokyo lab. Rather, the state of that qubit is destroyed in the Shanghai lab, and that state "appears" in the Tokyo lab **at a later time**.
3. Neither Alice nor Bob need to know anything about the qubit state they are teleporting to successfully execute the protocol.
4. In 2017, teleportation was carried out over a distance of nearly 1400 kilometers in China.

Quantum Teleportation Circuit



Quantum teleportation preserves entanglement!

We now show that if Alice (in Shanghai) shares an entangled state with Charlie (in Moscow), when she quantum teleports her state to Bob (in Tokyo), Bob's new state is entangled with Charlie.

- Let the Charlie-Alice (partially) entangled state be $|\psi\rangle = \alpha|00\rangle + \beta|11\rangle$ (This can vary from the unentangled case when $|\alpha| = 0$ or $|\alpha| = 1$ to the maximally entangled case when $|\alpha|\sqrt{2} = 1$.)
- As before, Alice and Bob prepare a Bell pair $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ in their lab in Shanghai, and Bob flies to Tokyo with his qubit in his suitcase.
- The overall system is now in the state

$$|\psi\rangle \otimes |\Phi^+\rangle = (\alpha|00\rangle + \beta|11\rangle) \otimes \left(\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \right)$$

Quantum teleportation preserves entanglement! ...continued

$$|\psi\rangle \otimes |\Phi^+\rangle = \frac{1}{\sqrt{2}} (\alpha |000\rangle + \alpha |001\rangle + \beta |110\rangle + \beta |111\rangle)$$

- 2 qubits in Alice's lab in Shanghai
- 1 qubits in Bob's lab in Tokyo
- 1 qubits in Charlie's lab in Moscow

- Alice operates with CNOT to get

$$\frac{1}{\sqrt{2}} (\alpha |000\rangle + \alpha |001\rangle + \beta |110\rangle + \beta |110\rangle)$$

- Alice operates with Hadamard to get

$$\frac{1}{2} (\alpha |000\rangle + \alpha |010\rangle + \alpha |001\rangle + \alpha |011\rangle + \beta |101\rangle - \beta |110\rangle + \beta |100\rangle - \beta |110\rangle)$$

Computational Complexity theory is the branch of computer science that concerns itself with (loosely speaking) the amount of *time* or *space* an algorithm will take, and the dependence of that time/space on the size of the input.

- By *time* we mean the time it will take some computer program that implements the algorithm to run on some computer. By *space* we mean the “amount” (in bits/qubits) of storage space the algorithm will require. In a very practical sense, one can understand why these quantities are important: **When you buy a computer at your local store, among the top things you look for are a fast CPU (“time”) and lots of RAM / hard disk (“space”).**
- It's clear the time for an algorithm will differ on different CPU's, in different operating systems, in different programming languages, etc.: These dependencies are important of course, but are not part of complexity theory.

Quantum teleportation preserves entanglement! ...continued

Alice as before measures her two (blue) qubits. The possible outcomes are:

Probability	Alice's outcome	Charlie-Bob state
0.25	00	$\alpha 00\rangle + \beta 11\rangle$
0.25	01	$\alpha 01\rangle + \beta 10\rangle$
0.25	10	$\alpha 00\rangle - \beta 11\rangle$
0.25	11	$\alpha 01\rangle - \beta 10\rangle$

- Alice transmits classically her result to Bob, who carries out the corresponding σ_X and/or σ_Z operations (if any) to obtain the final Charlie-Bob state $\alpha |00\rangle + \beta |11\rangle$.

- We will consider (as in most of classical computer science) mainly time complexity.
- Nota Bene: The word “complexity” in this context has nothing to do with how complicated or simple an algorithm may be; furthermore, it certainly has nothing to do with the Complex Numbers \mathbb{C} .

We use so-called **Big-Oh** notation in complexity theory.

Definition: Big-Oh notation

We say “ f is Big Oh of g ” and we write $f(n) = \mathcal{O}(g(n))$ iff there exists a positive integer m and a positive constant (real number) K such that

$$f(n) \leq K g(n) \quad (64)$$

for all $n > m$.

Let's prove that $n^3 = \mathcal{O}(2^n)$: We are only interested in positive integer values of n . We tabulate a few values:

n	1	2	3	4	5	6	7	8	9	10	11
$f(n) = n^3$	1	8	27	64	125	216	343	512	729	1000	1331
$g(n) = 2^n$	2	4	8	16	32	64	128	256	512	1024	2048

Note that, between $n = 9$ and $n = 10$, one function "passes out" the other function, in that

$$\begin{cases} f(n) > g(n) & \text{for } n=9 \\ f(n) < g(n) & \text{for } n=10 \end{cases}$$

We prove that $n^3 \leq 2^n$ for $n > m = 10$ (and we can take $K = 1$ in Equation (64)). We use induction:

Base Case: $n = m = 10 : 10^3 \leq 2^{10}$

Inductive Step: Assume true for p , i.e. $p^3 \leq 2^p$.

Some common complexity classes (and functions) follow:

constant	logarithmic	polynomial				exponential			factorial
		linear	quadratic	cubic	...	2^n	3^n	...	
1	$\log n$	n	n^2	n^3	...	2^n	3^n	...	$n!$

In this table, reading from left to right, if f is to the left of g in the list, then $f(n) = \mathcal{O}(g(n))$.

Example: Sequential (Linear) search

1	6	5	7	2	4	3	8
---	---	---	---	---	---	---	---

Suppose we want to search for a particular number in the list, and our (best!) algorithm is to look step by step from the left.

We must use this assumption (along with the base case) to prove the statement is true for $p + 1$, i.e. $(p + 1)^3 \leq 2^{(p+1)}$. But

$$(p + 1)^3 = p^3 + 3p^2 + 3p + 1, \text{ and } 2^{(p+1)} = 2^p + 2^p.$$

From our assumption, $p^3 \leq 2^p$, so we are left to prove that $3p^2 + 3p + 1 \leq 2^p$. But

$$\begin{aligned} 3p^2 + 3p + 1 &\leq 3p^2 + 3p^2 + 3p^2 = 9p^2 \\ &\leq p^3 \text{ (since } p \geq 10) \\ &\leq 2^p \text{ (from our assumption)} \end{aligned}$$

So we have proved that $(p + 1)^3 \leq 2^{(p+1)}$ and we are done. ■

Example: Sequential (Linear) search

... continued

- Search for 5 would necessitate 3 comparisons
- Search for 3 would necessitate 7 comparisons
- Search for 2 would necessitate 5 comparisons

etcetera. It should be clear, on average about 4 comparisons are needed. For a list of length n (and we only care about very large n), about $n/2$ comparisons are needed. We say this algorithm is linear in n (which is the size of the input), or $\mathcal{O}(n)$.

Example: Shortest path in a (complete) graph

Suppose we have an idealized road network with $n(n - 1)/2$ direct roads connecting n cities. This is a graph with an edge between every two nodes. Given a number attached to each edge, representing the distance, find the shortest path between two given nodes/cities.

Example: Shortest path in a (complete) graph ... continued

The **Brute Force** algorithm for solving this problem considers all possible paths between the two vertices, calculates the total length for each path, and then finds the smallest one. In this problem, the number of cities is n , the number of edges is $n(n-1)/2$, but the number of different paths (of length 1, or 2, or ... $(n-1)$) is proportional to $n!$. So, the Brute Force algorithmic solution is $\mathcal{O}(n!)$.^a To put in to context how bad factorial complexity is, if n was just 20, $n!$ is more than $10^{17} = 100,000,000,000,000,000$.

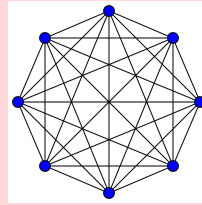


Figure 4: K_8 , the complete graph on 8 vertices.

^athere are other algorithms for this problem with better algorithmic complexity

Grover's algorithm

The **Grover Algorithm** presents a solution to the unstructured search problem that is quadratically faster ($\mathcal{O}(\sqrt{N})$) than any classical algorithm ($\mathcal{O}(N)$). For example, to search a list with a million items in it, we expect a classical algorithm to average half a million operations, while the Grover algorithm will necessitate about a thousand operations.

We assume for simplicity we are looking for an item that occurs exactly once in a list of length $2^n = N$. Let our list be $[x_0, x_1, x_2, \dots, x_{N-1}]$ and, for example we are looking for a particular letter (e.g. $x_5 = a$). We can represent this as a function

$$f : \{0, 1\}^n \rightarrow \{0, 1\}$$

with $f(x_5) = 1$ (for example), and $f(x_i) = 0$ for $i \neq 5$ (i.e., we label the item we are looking for with a one).

Notation: $\{0, 1\}^n$ is the set of all binary strings of length n . For example

$0010 \in \{0, 1\}^4$, $00111 \in \{0, 1\}^5$, $\{0, 1\}^2 = \{00, 01, 10, 11\}$.



Grover's algorithm

As in the Deutsch-Jozsa algorithm (see Equation (63)), we define the Unitary $U_f : |x\rangle |y_0\rangle \rightarrow |x\rangle |y_0 \oplus f(x)\rangle$ so that we have as before

$$U_f(|x\rangle |-\rangle) = (-1)^{f(x)} |x\rangle |-\rangle.$$

We represent the list of length N using $\log N$ qubits, and define

$$\begin{aligned} |a\rangle & \dots \text{the marked element we are searching for} \\ |\psi\rangle &= \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle \dots \text{an equal superposition of all states} \end{aligned}$$

Note that the inner product satisfies $\sqrt{2^n} \langle a | \psi \rangle = 1$: These two states are "nearly" orthogonal, since we are interested in large n , so $\frac{1}{\sqrt{2^n}}$ is "close to" zero. From our N dimensional space, we consider the 2-dimensional subspace generated by $|a\rangle$ and an equal superposition of all the *other* states,

$$|\phi\rangle = \frac{1}{\sqrt{2^n - 1}} \sum_{x \in \{0,1\}^n, x \neq a} |x\rangle.$$

Note that

$$|\psi\rangle = \frac{\sqrt{2^n - 1}|\phi\rangle + |a\rangle}{\sqrt{2^n}} \quad \text{and} \quad \langle\psi|\phi\rangle = \sqrt{1 - \frac{1}{2^n}}$$

We have

$$\begin{aligned} U_f(|a\rangle|-\rangle) &= (-1)^{f(a)}|a\rangle|-\rangle = -|a\rangle|-\rangle \\ U_f(|\phi\rangle|-\rangle) &= |\phi\rangle|-\rangle \end{aligned} \quad (65)$$

Define $U_\psi = 2|\psi\rangle\langle\psi| - I$, so that we have

$$\begin{aligned} U_\psi|a\rangle &= (2|\psi\rangle\langle\psi| - I)|a\rangle = \frac{2}{\sqrt{2^n}}|\psi\rangle - |a\rangle \\ &= \frac{\sqrt{2^n - 1}}{2^{n-1}}|\phi\rangle + \left(\frac{1}{2^{n-1}} - 1\right)|a\rangle \\ U_\psi|\phi\rangle &= (2|\psi\rangle\langle\psi| - I)|\phi\rangle = 2\sqrt{1 - \frac{1}{2^n}}|\psi\rangle - |\phi\rangle \\ &= \left(\frac{2^n - 1}{2^{n-1}} - 1\right)|\phi\rangle + \frac{\sqrt{2^n - 1}}{2^{n-1}}|a\rangle = \left(1 - \frac{1}{2^{n-1}}\right)|\phi\rangle + \frac{\sqrt{2^n - 1}}{2^{n-1}}|a\rangle \end{aligned}$$

We can re-write this in the concise format

$$U_\psi \begin{pmatrix} |a\rangle \\ |\phi\rangle \end{pmatrix} = \begin{pmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} |a\rangle \\ |\phi\rangle \end{pmatrix}$$

with

$$\cos\theta = \left(1 - \frac{1}{2^{n-1}}\right).$$

(You should check that

$$\cos^2\theta + \sin^2\theta = \left(1 - \frac{1}{2^{n-1}}\right)^2 + \left(\frac{\sqrt{2^n - 1}}{2^{n-1}}\right)^2 = 1.)$$

We re-write Equation (65)

$$U_f \begin{pmatrix} |a\rangle|-\rangle \\ |\phi\rangle|-\rangle \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} |a\rangle|-\rangle \\ |\phi\rangle|-\rangle \end{pmatrix},$$

by abuse of notation, as

$$U_f \begin{pmatrix} |a\rangle \\ |\phi\rangle \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} |a\rangle \\ |\phi\rangle \end{pmatrix}.$$

We define the **Grover iterate**

$$G(\theta) = U_\psi U_f = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

Note that $G^k(\theta) = G(k\theta)$. Since θ is small, this rotation matrix **G** increases the amplitude of $|a\rangle$ while decreasing the amplitude of $|\phi\rangle$, so-called **amplitude amplification**.

The Grover algorithm swaps the (large) amplitude attached to $|\phi\rangle$ (i.e. $\sqrt{(2^n - 1)/2^n}$)⁵ with the (small) amplitude attached to $|a\rangle$ (i.e. $2^{-n/2}$). Since $|\phi\rangle$ and $|a\rangle$ are orthogonal, we should carry out k Grover iterates to obtain a rotation through $\pi/2$, i.e. $k\theta = \pi/2$. For small θ we approximate $\theta \approx \sin\theta$ so that

$$k\theta \approx k\sin\theta = k\frac{\sqrt{2^n - 1}}{2^{n-1}} = \frac{\pi}{2} \implies k \approx \frac{\pi}{4} \frac{2^n}{\sqrt{2^n - 1}} \approx \frac{\pi}{4} \sqrt{2^n} = \frac{\pi}{4} \sqrt{N}$$

⁵Be very careful when writing (by hand) 2^{n-1} or $2^n - 1$. **They are different!** The -1 is in the exponent in one case, but not in the other. To be safe it is wise to use brackets $2^{(n-1)}$

Grover's Algorithm

- ▶ **Initial State:** $|0\rangle^{\otimes n}|1\rangle$
- ▶ **Step 1:** Apply $H^{\otimes(n+1)}$
- ▶ **Step 2:** Apply the Grover iterate $(U_\psi \otimes I)U_f$ about $\frac{\pi}{4}\sqrt{2^n}$ times
- ▶ **Measure**

Example: Find **s** in $[y, w, c, t, d, q, \mathbf{s}, b]$

$N = 8, n = \log N = \log 8 = 3$, and **s** is represented by the state $|110\rangle$. Our function f is

x	y	w	c	t	d	q	s	b
x (binary)	000	001	010	011	100	101	110	111
f(x)	0	0	0	0	0	0	1	0

Example: Find s in $[y, w, c, t, d, q, s, b]$... continued

At the end of Step 1, we have the state

$$|\psi\rangle = \frac{1}{\sqrt{8}} \sum_{x \in \{0,1\}^3} |x\rangle \otimes |-\rangle.$$

(In what follows, we stop writing explicitly the ancilla qubit $|-\rangle$.)

The Grover iterate is $G = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$

$$= \frac{1}{2^{n-1}} \begin{pmatrix} 2^{n-1} - 1 & \sqrt{2^n - 1} \\ -\sqrt{2^n - 1} & 2^{n-1} - 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 3 & \sqrt{7} \\ -\sqrt{7} & 3 \end{pmatrix}$$

Furthermore, $|a\rangle = |110\rangle$ and $|\phi\rangle = \frac{1}{\sqrt{7}} \sum_{x \in \{0,1\}^3, x \neq a} |x\rangle$

$$\Rightarrow |\psi\rangle = \frac{|a\rangle + \sqrt{7}|\phi\rangle}{\sqrt{8}}.$$

Example: Find s in $[y, w, c, t, d, q, s, b]$... continued

Our component vector in the 2D space spanned by $\{|a\rangle, |\phi\rangle\}$ is

$$\frac{1}{\sqrt{8}} \begin{pmatrix} 1 \\ \sqrt{7} \end{pmatrix}.$$

The first application of the Grover iterate G gives

$$G|\psi\rangle = \frac{1}{4} \begin{pmatrix} 3 & \sqrt{7} \\ -\sqrt{7} & 3 \end{pmatrix} \frac{1}{\sqrt{8}} \begin{pmatrix} 1 \\ \sqrt{7} \end{pmatrix} = \frac{1}{2\sqrt{8}} \begin{pmatrix} 5 \\ \sqrt{7} \end{pmatrix}.$$

The second application of G gives

$$G^2|\psi\rangle = \frac{1}{4} \begin{pmatrix} 3 & \sqrt{7} \\ -\sqrt{7} & 3 \end{pmatrix} \frac{1}{2\sqrt{8}} \begin{pmatrix} 5 \\ \sqrt{7} \end{pmatrix} = \frac{1}{4\sqrt{8}} \begin{pmatrix} 11 \\ -\sqrt{7} \end{pmatrix},$$

and the third gives $G^3|\psi\rangle = \frac{1}{8\sqrt{8}} \begin{pmatrix} 13 \\ -7\sqrt{7} \end{pmatrix}.$

Example: Find s in $[y, w, c, t, d, q, s, b]$... continued

If we measure the state (and hence stop the algorithm) after j Grover iterations, what is the probability of finding s ?

Grover iterate	0	1	2	3
Probability of finding s	$1/8$ $= 0.125$	$25/32$ ≈ 0.78	$121/128$ ≈ 0.945	$169/512$ ≈ 0.33

Remark that

- ▶ Even after one Grover iteration, we have 78% chance of finding s if we were to measure.
- ▶ The maximum probability occurs after 2 iterations, with about 94.5% success probability.

Example: Find s in $[y, w, c, t, d, q, s, b]$... continued

- ▶ If we run the algorithm too long - on to the third iteration, the success probability **decreases** to about 33%.
- ▶ The number of iterations we **should** carry out is $\pi\sqrt{N}/4 = \pi/\sqrt{2} \approx 2.22$.

Exercise (Grover Algorithm)

Check that the matrix G in the previous $N = 8$ example is Unitary

You may be uneasy with the fact that, in this example, optimum behaviour of the algorithm still only gives a probability (however high) of obtaining the correct answer. Variations of the Grover algorithm give certainty of obtaining the correct answer (this is beyond the scope of this course).

We describe here though a simple small case where the algorithm is 100% successful.

Example: Grover search of a list of length four

Our 2-qubit basis is $S = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ with

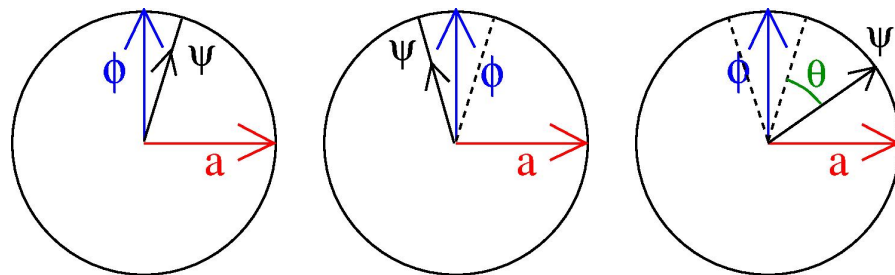
$$|\phi\rangle = \frac{1}{\sqrt{3}} \sum_{b \neq a, b \in S} |b\rangle \quad (66)$$

$$|\psi\rangle = \frac{1}{4} (|00\rangle + |01\rangle + |10\rangle + |11\rangle) = \frac{|a\rangle + \sqrt{3}|\phi\rangle}{2}$$

(Remember: $|a\rangle$ is the vector we are looking for, $|\phi\rangle$ is an equal superposition of all the (3) other vectors, hence the factor of $\sqrt{3}$.)

$$G = \frac{1}{2} \begin{pmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix} \Rightarrow G|\psi\rangle = \frac{1}{4} \begin{pmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

So, after a single iteration, we measure the correct result with probability 1.



(a) Time $t = 0$. $|\phi\rangle$ is orthogonal to $|a\rangle$.

(b) Time $t = 1$. Using U_f , $|\psi\rangle$ is reflected in (the one dimensional line generated by) $|\phi\rangle$.

(c) Time $t = 2$. Using U_ψ , $|\psi\rangle$ is reflected in (the one dimensional line generated by) the original (at time $t = 0$) $|\psi\rangle$.

Figure 6: Geometrical view of one Grover iteration. All vectors are drawn on the unit circle. The net effect of the two reflections is rotation through angle θ . The overall effect is to move the overall state of the system, $|\psi\rangle$, towards the desired solution, $|a\rangle$.

Geometrical view of Grover iterations

It is well known that the composition of two successive reflections in lines through the origin is a rotation about the origin through an angle which is twice the angle between the two lines. The combined effect of the reflections U_f and U_ψ is to produce a rotation. Note firstly that $U_\psi = 2|\psi\rangle\langle\psi| - I$ is a reflection in (the line generated by) the vector $|\psi\rangle$, since

- U_ψ operating on $|\psi\rangle$ has no effect.
- If $|\psi'\rangle$ is orthogonal to $|\psi\rangle$, then U_ψ changes $|\psi'\rangle$ to $-|\psi'\rangle$

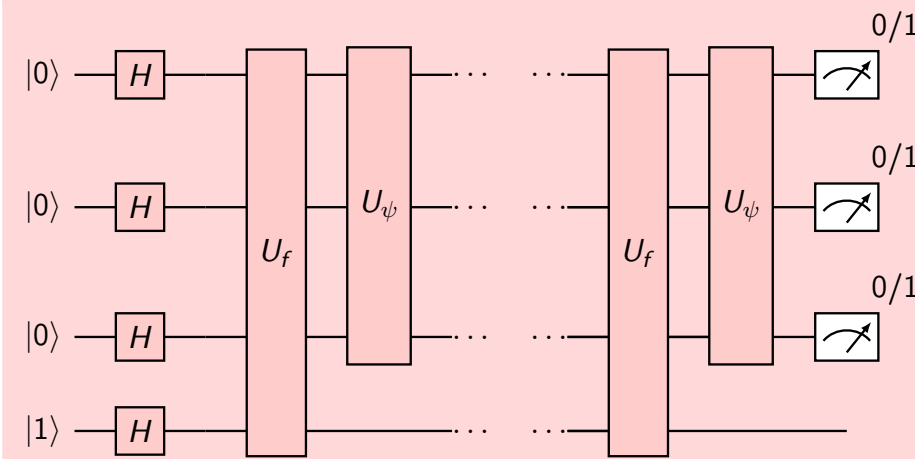
For a general vector $|\psi'\rangle$, we break in to components parallel and perpendicular to $|\psi\rangle$.

Secondly, U_f can be written as $U_f = 2|\phi\rangle\langle\phi| - I$ (see Equation (65)), and so can be interpreted as a reflection in (the line generated by) $|\phi\rangle$.

Following our last example of searching a list of length four, in that case, $n = 2$ and (using Equation (66)) the inner product of the two unit vectors $|\psi\rangle$ and $|\phi\rangle$ is

$$\langle\psi|\phi\rangle = ||\psi\rangle|||\phi\rangle| \cos \delta = \cos \delta = \frac{\sqrt{3}}{2}$$

which means δ is 30 degrees ($\pi/6$). From Figure 6, θ is twice the angle between $|\psi\rangle$ and $|\phi\rangle$, so $\theta = \pi/3$, which is the angle between $|\psi\rangle$ and $|a\rangle$. Hence a single rotation is sufficient.

Grover algorithm circuit ($n = 3$)

The Quantum Fourier Transform (sometimes abbreviated QFT) is the discrete Fourier transform as applied to amplitudes of a quantum state. We study it both

- ▶ as a generalization of the ubiquitous Hadamard matrix (which, it turns out, is the QFT in 2D)
- ▶ for its role in quantum algorithms

Definition: Discrete Fourier Transform (DFT)

The DFT of a set of N complex numbers x_0, x_1, \dots, x_{N-1} is the set of N complex numbers y_0, y_1, \dots, y_{N-1} given by

$$y_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \omega^{-jk} x_k \quad (67)$$

with $\omega = e^{2\pi i/N} = \cos(2\pi/N) + i \sin(2\pi/N)$. ω is called an N th root of unity, since $\omega^N = 1$.

Example: DFT of $[1, 1, 1, 1]$

From Equation (67) we get

$$\begin{aligned} y_j &= \frac{1}{2} \sum_{k=0}^3 \omega^{-jk} = \frac{1}{2} (1 + \omega^{-j} + (\omega^{-j})^2 + (\omega^{-j})^3) \\ &= \frac{1}{2} (1 + (-i)^j + (-1)^j + i^j) \quad (\text{since } \omega = i) \end{aligned}$$

which gives

$$\begin{aligned} y_0 &= 2; & y_1 &= \frac{1}{2} (1 - i - 1 + i) = 0 \\ y_2 &= \frac{1}{2} (1 - 1 + 1 - 1) = 0; & y_3 &= \frac{1}{2} (1 + i - 1 - i) = 0 \end{aligned}$$

Example: DFT of $[1, 2, 1, 2]$

$$y_j = \frac{1}{2} (1 + 2\omega^{-j} + \omega^{-2j} + 2\omega^{-3j}) = \frac{1}{2} + (-i)^j + \frac{(-1)^j}{2} + i^j$$

which gives

$$\begin{aligned} y_0 &= 3; & y_1 &= \frac{1}{2} - i - \frac{1}{2} + i = 0 \\ y_2 &= \frac{1}{2} - 1 + \frac{1}{2} - 1 = -1; & y_3 &= \frac{1}{2} + i - \frac{1}{2} - i = 0 \end{aligned}$$

Note here in our very small example, the vector is periodic, and the DFT has peaks at 0 and at 2 — a small sign of its use in evaluating the periodicity of a function.

Definition: Inverse (Discrete) Fourier Transform (DFT)

$$x_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \omega^{jk} y_k$$

Lets check this is in fact the inverse:

$$\begin{aligned} x_j &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \omega^{jk} y_k = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \omega^{jk} \left(\frac{1}{\sqrt{N}} \sum_{p=0}^{N-1} \omega^{-kp} x_p \right) \\ &= \frac{1}{N} \sum_{k,p} \omega^{k(j-p)} x_p \\ &= \frac{1}{N} \sum_k \left(\omega^{kj} x_0 + \omega^{k(j-1)} x_1 + \dots + x_j + \dots + \omega^{k(j-N+1)} x_{N-1} \right) \end{aligned}$$

But $\sum_k \omega^{kj} = \sum_k (\omega^j)^k$ is a geometric series that sums to $(1 - (\omega^j)^N)/(1 - \omega^j)$ which is zero since $\omega^N = 1$

Definition: Quantum Fourier Transform (QFT)

This is the $N \times N$ matrix $F_N = \frac{1}{\sqrt{N}} \sum_{j,k=0}^{N-1} \omega^{jk} |j\rangle \langle k|$

$$= \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(N-1)} \\ 1 & \omega^3 & \omega^6 & \dots & \omega^{3(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \dots & \omega^{(N-1)(N-1)} \end{pmatrix}$$

with $\omega = e^{2\pi i/N}$

For example

$$F_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = H$$

So, we finally get

$$x_j = \frac{1}{N} \sum_{k=0}^{N-1} x_j = \frac{1}{N} (Nx_j) = x_j \quad (68)$$

We have the following equivalent definitions:

Definition: Quantum Fourier Transform (QFT)

This is the mapping (in the standard basis)

$$|j\rangle \longrightarrow \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \omega^{jk} |k\rangle \quad (69)$$

Note that for qubits ($N = 2$), $\omega = -1$ so Equation (69) is just

$$|0\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle); \quad |1\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

$$F_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix} \quad F_4 = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix} \quad (70)$$

F_N has the following properties:

- ▶ It is Hermitian: $F_N^{\top*} = F_N$
- ▶ The magnitude of each complex number in the matrix is identical (and equal to $1/\sqrt{N}$).
- ▶ The columns form an orthonormal set of vectors (as do the rows).

Example: QFT of $|\psi\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$

We have

$$F_4 |\psi\rangle = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Example: QFT of $|\psi\rangle = |00\rangle$

$$F_4 |\psi\rangle = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Shor's algorithm is an algorithm for factoring (large) numbers in polynomial time.

Motivation:

Suppose I give you a choice of two problems to solve. Further, suppose I offer a monetary prize for the right answer, which is inversely proportional to the amount of time it takes you to get the answer.

Problem 1 Multiply the numbers 13, 19 and 29.

Problem 2 Factorize (into prime factors) the number 7163.

Which would you pick?

Probably you will (and you should!) pick the first problem. You know the procedure for multiplying numbers, and you can estimate how many operations you should do. But to factorize 7163? Well, you may (and should) know, you can start trying to divide by other primes, starting with 3, 5, 7, etc....but you can probably guess this will take more time than the multiplication in problem 1.

From these examples, we note a general pattern in the behaviour of both classical and quantum Fourier transforms: **If the original data / vector components / quantum amplitudes is fairly uniformly distributed, the Fourier transform will be sharply peaked, and vice versa.**

A third example gives us further insight:

Example: QFT of $|\psi\rangle = |01\rangle$

$$F_4 |\psi\rangle = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ -i \\ -1 \\ i \end{pmatrix}$$

So, **shifting a vector** (in this case, $(1, 0, 0, 0) \rightarrow (0, 1, 0, 0)$) **produces phases in the components of the Fourier transformed vector.**

It turns out that $7163 = (13)(19)(29)$.

- ▶ The best current "classical" algorithm for factoring large integers is approximately exponential. Shor's algorithm is polynomial.
- ▶ Factoring large numbers may seem like an obscure mathematical task - but it's presumed difficulty is at the heart of RSA encryption, and hence cybersecurity.
- ▶ Using a naïve brute force approach, trying to crack the currently recommended 2048-bit RSA encryption key would require testing 2^{2048} keys, which is a number with well over 500 digits.
- ▶ It is important to point out - there is no rigorous proof that it is impossible to come up with a polynomial-time classical algorithm

Shor's Algorithm

Shor's Algorithm uses the QFT (Quantum Fourier Transform) to find the periodicity of a particular function, and that periodicity is then used (in a classical algorithm) to find a factor for N . In that sense, the factoring process is a hybrid of classical and quantum operations. The quantum part of the algorithm (period finding) is where the improvement is made (from exponential to polynomial time).

We assume we want to find factors of a (large) positive integer N , which is not prime and is not even (so in fact our smallest example would be $N = 9$).

- **Step A** Find x with $x^2 = 1 \pmod{N}$. We can re-write this equation as $x^2 - 1 = 0 \pmod{N} = (x - 1)(x + 1)$. This means $(x - 1)(x + 1) = Nm$ for some positive integer m , so every non-trivial factor p of N must also divide into $x - 1$ or $x + 1$.

Shor's Algorithm

... continued

Finally:

- **Step A2a** Consider the function $f(x) = a^x \pmod{N}$. To determine r , we find the period of this function

x	1	2	3	...	r	$r+1$	$r+2$
$f(x)$	a	a^2	a^3	...	1	a	a^2

The QFT applied to $f(x)$ will have peaks at $r, 2r, 3r$, etc.

Example: Modular periodicities

$$f(x) = 2^x \pmod{9}$$

x	1	2	3	4	5	6	7	8	9
2^x	2	4	8	16	32	64	128	256	512
$2^x \pmod{9}$	2	4	8	7	5	1	2	4	8

The period is $r = 6$.

Shor's Algorithm

... continued

- **Step B** (Using the Euclidean algorithm) calculate $\text{GCD}(x - 1, N)$ and $\text{GCD}(x + 1, N)$ to obtain a factor p .
- **Step C** Reassign $N' = N/p$, and proceed again from Step A with N' to obtain further factors.

We further specify:

- **Step A1** To find x with $x^2 = 1 \pmod{N}$, first pick a (random) a which is coprime to N (i.e. $\text{GCD}(a, N) = 1$).
- **Step A2** Find the order of a , i.e. the least positive integer r with $a^r = 1 \pmod{N}$.
- **Step A3** Assign $x = a^{r/2}$. If r is odd, repeat from step A1 for a different a .

Example: Modular periodicities

... continued

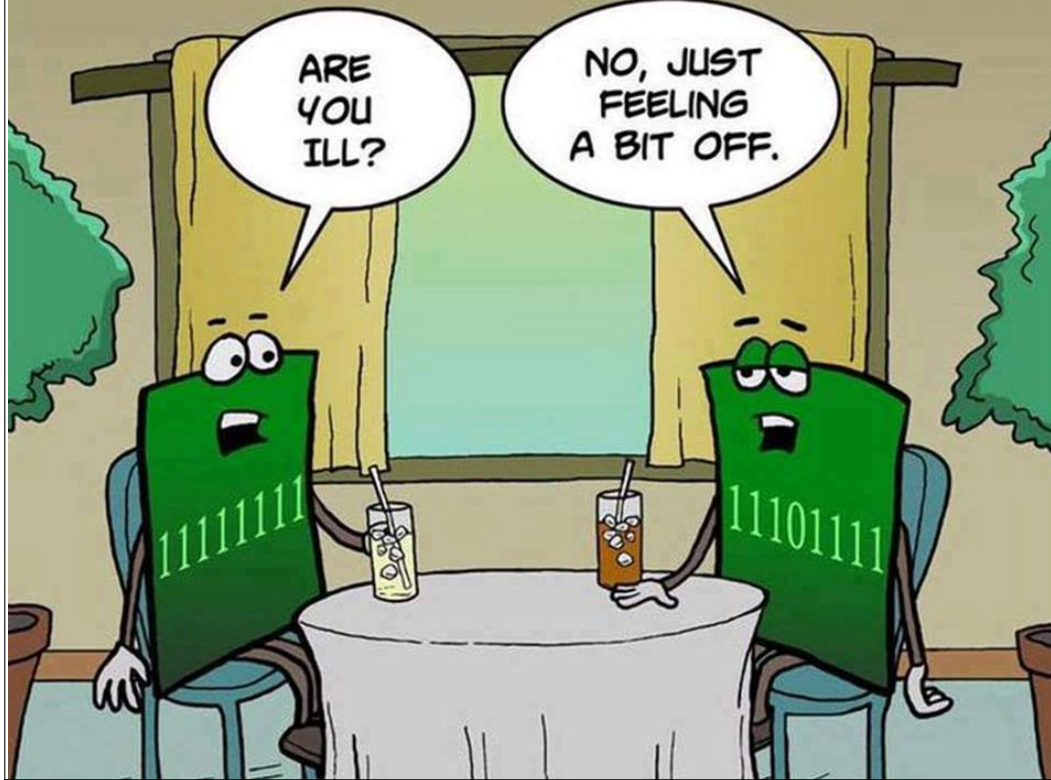
$$f(x) = 3^x \pmod{10}$$

x	1	2	3	4	5	6	7
3^x	3	9	27	81	243	729	2187
$3^x \pmod{10}$	3	9	7	1	3	9	7

The period is $r = 4$.

The computational complexities of the various Fourier transforms are

- $\mathcal{O}(N^2)$ for (Classical) Fourier transform
- $\mathcal{O}(N \log N)$ for (Classical) FFT (Fast Fourier Transform)
- $\mathcal{O}(n^2) = \mathcal{O}((\log N)^2)$ for QFT



Quantum Fourier Transform and Shor's algorithm

To describe our quantum circuit for F_4 , we need to define some more gates.

Definition: Phase gate

The Phase gate, normally denoted by R_ϕ , is a single qubit gate given by

$$R_\phi = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix}.$$

Definition: Control Unitary gate

For any single qubit gate / Unitary U , we can construct a 2 qubit Control Unitary gate, denoted by CU or sometimes $C - U$. Its action is

$$CU|\alpha\beta\rangle = \begin{cases} |\alpha\beta\rangle & \text{if } \alpha = 0 \\ |\alpha\rangle U|\beta\rangle & \text{if } \alpha = 1 \end{cases}$$

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

190

Quantum Fourier Transform and Shor's algorithm

Definition: Control Unitary gate

... continued

The generic format for the 4x4 matrix of a CU gate is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & u_{1,1} & u_{1,2} \\ 0 & 0 & u_{2,1} & u_{2,2} \end{pmatrix} \quad \text{where } U = \begin{pmatrix} u_{1,1} & u_{1,2} \\ u_{2,1} & u_{2,2} \end{pmatrix}$$

Definition: SWAP gate

This swaps the two qubits, $SWAP|\alpha\beta\rangle = |\beta\alpha\rangle$. It is a permutation matrix, given by

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

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

191

QFT Circuit for F_4

We consider again F_4 , which has action

$$F_4(|00\rangle) = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) = \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle + |1\rangle)$$

$$F_4(|01\rangle) = \frac{1}{2}(|00\rangle + i|01\rangle - |10\rangle - i|11\rangle) = \frac{1}{2}(|0\rangle - |1\rangle)(|0\rangle + i|1\rangle)$$

$$F_4(|10\rangle) = \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle) = \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle)$$

$$F_4(|11\rangle) = \frac{1}{2}(|00\rangle - i|01\rangle - |10\rangle + i|11\rangle) = \frac{1}{2}(|0\rangle - |1\rangle)(|0\rangle - i|1\rangle)$$

The resulting 2-qubit states are separable (in fact, this is true for F_N)

<http://www.maths.nuigalway.ie/~gettrick/teach/ma437/>

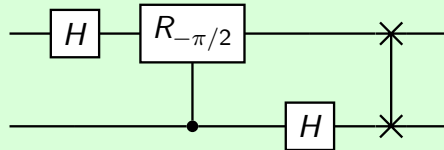
192

QFT Circuit for F_4 ... continued

We can implement this with 4 gates, using

$$F_4 |\alpha\beta\rangle = (SWAP)(I \otimes H)(CR_{\pi/2}(2,1))(H \otimes I) |\alpha\beta\rangle$$

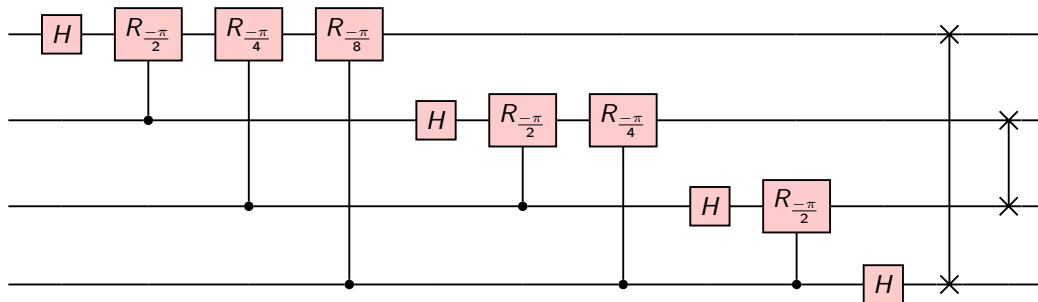
where $SWAP$ swaps the two qubits, while $CR_\phi(j, k)$ is a CONTROL-phase gate, with control j and target k and phase $e^{i\phi}$.



The explicit 4x4 matrices for these gates are

$$H \otimes I = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \quad CR_{\pi/2}(2,1) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -i \end{pmatrix}$$

We show further the circuit for F_{16} .



We can see

- The first wire has a Hadamard and 3 (controlled) phases.
- The second wire has a Hadamard and 2 (controlled) phases.
- The third wire has a Hadamard and 1 (controlled) phase.
- etcetera... at the end we have two swaps.

QFT Circuit for F_4 ... continued

$$I \otimes H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \quad SWAP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(You should check that the product of these four matrices in the order given is indeed the F_4 matrix in Equation (70).)

If we need 4 gates (and therefore 4 unitary operations) to implement F_4 , how many do we need for F_N ? It turns out we can construct a quantum circuit for F_N using a number of gates that is a polynomial function (specifically, quadratic) of the number of qubits ($\log N$).

For n quantum wires, we would have about $n/2$ SWAP gates at the end, and $(n + (n-1) + (n-2) + \dots + 3 + 2 + 1)$ other gates, i.e.

$$\frac{n}{2} + \frac{n}{2}(n+1) = \frac{n}{2}(n+2) = \mathcal{O}(n^2)$$

which shows quadratic complexity in $\log N$.