Unit 1: a brief history of Q

Tzu-Chieh Wei
C. N. Yang Institute for Theoretical Physics and Department of Physics and Astronomy,
State University of New York at Stony Brook, Stony Brook, NY 11794-3840, USA
(Dated: August 28, 2025)

In this unit, we will give a brief review of the necessary Math that you should have known. Then we introduce the very but important basics of quantum principles, and you are ready to study the concepts of quantum bits (qubits) and quantum gates. We will also have the first taste of quantum algorithms and the potential of quantum computation but also possible limitation.

Learning outcomes: (1) You'll be ready to embark on a QIS journey. (2) You'll understand why quantum computing has a lot of potential and why it may also not be all-powerful.

I. WHAT MATHEMATICS DO I NEED?

Mostly linear algebra (vectors and matrices) is needed to get started on quantum information science. We will review the very basics here; if you know all of them, you are really set to learn a lot. Dr. Martin Laforest (University of Waterloo) has written a little book on "The Mathematics of Quantum Mechanics"; it is free to download at `http://dl.icdst.org/pdfs/files3/8950b72535591b7bf7217e2bb5f650a1.pdf` and please browse through it to see if there is any math that you need to refresh yourself of.

Other mathematics such as group theory, representation of groups, and bosonic and fermionic annihilation and creation operators, etc. will be taught and can be learned along the way. We will try to use mathematics as tools and will not dwell too much on its rigor; e.g., we will not define the abstract Hilbert space as in more math oriented quantum mechanics textbook.

Complex numbers. The symbol $i = \sqrt{-1}$ represents the square root of -1 and has the properties that $i^2 = -1$ and $|i| = |\sqrt{-1}| = 1$. We also define the complex conjugate of i: $\bar{i} = i^* = -i$. With the definition of i, we can extend the real numbers to complex numbers and define $z \equiv a + b\,i$ for $a\&b$ being real. It has the property that $\overline{z} = z^* = a - b\,i$ and the absolute value of $|z| = \sqrt{z \cdot \overline{z}}$. From this definition we have $|z| = |a + b\,i|^2 = a^2 + b^2$. The expression z encodes a two-dimensional coordinate, and many mathematical and physical objects in 2D can be conveniently expressed in terms of such complex numbers. Even 2D conformal field theory [1] can be treated this way, using z and $\bar{z}$.

Why do we need complex numbers? In quantum mechanics, the wave function $\psi(x)$ that describes a quantum particle on a line is in general complex. It represents the probability amplitude of finding the particle at location x. The probability density of finding it at x is given by $p(x) = |\psi(x)|^2$ and the distribution satisfies the normalization $\int_{-\infty}^{\infty} p(x)dx = 1$. The i is responsible for interesting interference phenomena in physics. See also the article in Physics Today by Johanna Miller on "Does quantum mechanics need imaginary numbers?" [2] who discussed a proposal and experiment that "Quantum theory based on real numbers can be experimentally falsified" [3], where the setup is a three-player entanglement swapping scenario (which will be discussed in Unit 2). However, it is interesting to point out that it is possible to perform universal quantum computation using real-value gates so that there are only real-valued amplitudes during the computation [4, 5], where the imaginary part can be encoded using an additional qubit.

An important formula relates an angle from the origin on the 2D complex plane and its x- and y-components is the Euler formula: $e^{i\theta} = \cos\theta + i\sin\theta$. Note that $i = e^{i\pi/2}$ and $-1 = e^{i\pi}$.

Vectors in linear algebra and (pure) quantum states. A particle on a line contains infinite many locations and is the standard example in quantum mechanics, e.g., a particle in a box or in a harmonic potential [6]; a much simpler quantum mechanical systems is the two-level system [7]. In general, we can have finite-dimensional or even infinite-dimensional systems. It is useful to use the Dirac notation to represent a quantum state; at this stage you can regard it just as a useful notation with 'bra-kets'. For example, we can label the two basis states (which are orthonormal) of a two-level system by $|0\rangle$ and $|1\rangle$, and they can be understood resepctively as the following two vectors,

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

Another example is the three-level system, whose basis vectors are

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} =: |`1'\rangle,\ \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} =: |`2'\rangle,\ \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} =: |`3'\rangle. \tag{2}$$

Note that $\vec{v}'s$ form an orthogonal basis: $\vec{v}_i^{*T} \cdot \vec{v}_j = \vec{v}_i^{\dagger}\vec{v}_j = \delta_{ij}$. The complex conjugation on $\vec{v}_j$ is needed as the vector space is complex in general. In terms of the Dirac's bra-ket notation, we write$\langle `i'|`j'\rangle = \delta_{ij}$. The notation δ_{ij} is the Kronecker delta function and equals one if $i = j$, but zero otherwise.

Matrices in linear algebra and operations on quantum states. In quantum mechanics, the time evolution is governed by the Schrödinger's equation and you may have seen the evolution operator of the form $e^{-itH/\hbar}$, which is unitary. A unitary matrix U satisfies $U^\dagger U = UU^\dagger = I$, for example, the following 3×3 matrices will act on a three-level system,

$$U = \begin{pmatrix} \cos\theta & i\sin\theta & 0 \\ i\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; U^\dagger = \begin{pmatrix} \cos\theta & -i\sin\theta & 0 \\ -i\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3}$$

If U acts on $|`1'\rangle$, it is done as follows,

$$U|`1'\rangle = \begin{pmatrix} \cos\theta & i\sin\theta & 0 \\ i\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \vec{v}_1 = \begin{pmatrix} \cos\theta & i\sin\theta & 0 \\ i\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ i\sin\theta \\ 0 \end{pmatrix}. \tag{4}$$

There is a bra-ket notation for U: $U = \sum_{ij} U_{ij}|`i'\rangle\langle `j'|$, thus $\langle `i'|U|`j'\rangle = U_{ij}$. Moreover, $(U\vec{v})_i = \sum_{ij} U_{ij}v_j = \langle `i'|U|v\rangle$. An $n \times n$ identity matrix can be written as: $I = \sum_{i=1}^n |`i'\rangle\langle `i'|$.

Projectors and tensor (or Kronecker) product. We have seen, e.g., $|`1'\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ as a vector. We can define a projector $P_1 \equiv \vec{v}_1\vec{v}_1{}^\dagger = |`1'\rangle\langle `1'| = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ that projects to the subspace spanned by the vector $\vec{v}_1$. That is $P_1\vec{v} \sim \vec{v}_1$. As an example, $P_1(a|`1'\rangle + b|`2'\rangle) = a|`1'\rangle$. Note that P_1 is hermitian (or self-adjoint, i.e., $P_1^\dagger \equiv (P_1^*)^T = P_1$) and it squires to itself $P_1^2 = P_1$.

Next we discuss an important operation to combine different vector spaces, i.e., tensor (or Kronecker) product. For example, the tensor product of two vectors,

$$\vec{v}_1 \otimes \vec{v}_2 = \begin{pmatrix} 1 \times \vec{v}_2 \\ 0 \times \vec{v}_2 \\ 0 \times \vec{v}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = |`1'\rangle \otimes |`2'\rangle. \tag{5}$$

The meaning of this can be understood as describing a combined system of two qutrits (quantum trits), one in state $|`1'\rangle$ and the other in state $|`2'\rangle$. You can also see the advantage of using the Dirac notation, rather than writing out all components of the state vector.

We also naturally have the tensor product of two matrices, for example for the two 3×3 matrices A and B (note that they do not need to have the same dimensions),

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & a_{13}B \\ a_{21}B & a_{22}B & a_{23}B \\ a_{31}B & a_{32}B & a_{33}B \end{pmatrix}. \tag{6}$$

Note that it is easy to convince yourself that

$$A \otimes B(|`1'\rangle \otimes |`2'\rangle) = (A|`1'\rangle) \otimes (B|`2'\rangle). \tag{7}$$

Eigenvalue equation for a Hermitian matrix. Assume that H is an $n \times n$ Hermitian matrix (i.e. $H^\dagger = H$), the eigenvalue equation for H is

$$H\vec{v} = \lambda\vec{v}, \tag{8}$$

or in the bra-ket notation: $H|v\rangle = \lambda|v\rangle$. If H is the so-called Hamiltonian of a system, λ's correspond to eigen-energies of the system.

There are n indepdent solutions $\vec{v}_i$ (eigenvectors) and λ_i (eigenvalues); we have that λ_i's are real and eigenvectors can be made orthonormal: $\vec{v}_i^\dagger \vec{v}_j = \delta_{ij}$ (or in bra-ket notation: $\langle v_i|v_j\rangle = \delta_{ij}$). If some λ_i's are the same, they are *degenerate*. For example, the following three Pauli matrices each have two eigenvalues and eigenvectors (what are they?)

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

EPR (Einstein-Podolsky-Rosen) 1935	"Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" [8]
Bell 1964	Inequality to compare classical theory and quantum mechanics [9]
Clauser-Horne-Shimony-Holt (CHSH) 1969	Another inequality easier to test experimentally [10]
Feynman 1981 and 1985	Quantum Computation and Quantum Simulations [11, 12]
Aspect, Granger and Roger 1982	Experimental violation of CHSH inequality [13] (see 2022 Physics Nobel Prize)
Bennett and Brassard 1984	Quantum Key Distribution using non-orthogonal states [14]
Benioff 1990	Turing Machine using Quantum Mechanics [15]
Manin 1990	Idea of Quantum Computation [16]
Ekert 1991	QKD using singlet pairs [17]
Bennett et al. 1993	Quantum teleportation [18]
Shor 1994	Quantum Factoring algorithm [19]
Shor 1995	Quantum Error Correction [20]
Grover 1996	Quantum Search algorithm [21–23]
...	...
Google 2019	Quantum Supremacy 'Demonstration' [24]
...	...
IBM 2023	Quantum Utility 'Demonstration' [25]
...	...

TABLE I. Some incomplete but important early milestones in quantum information.

II. THE NOTION OF QUANTUM BITS

In Table I, we list a few early important developments that spurred the field of quantum information science. The field has evolved from a small community to one that is mutli-disciplinary, and there are a lot of experimental progresses as well as theoretical advancement. We include the Google's experiment on quantum supremacy in the table, although there have been advancement in numerical simulations that could simulate the experiment classically. We also listed IBM's quantum utlity experiment, despite subsequent numerical works that could simulate the experiment in the non-solvable regime. These examples represent healthy competitions between quantum experiments and numerical simulations before one can definitely claim the achievement of quantum advantage.

Quantum mechanics provides the starting point for discussions; however, one needs not go through a whole semester of quantum physics or quantum mechanics to appreciate the concepts of QIS. We will first review some basic quantum mechanical principles and explain some of the early ideas of quantum information processing protocols. There are many textbooks and, among them, I recommend the book by Susskind [7].

Quantum bits and quantum gates. A quantum bit (qubit) is a two-level system, which can be described by a complex vector ψ (usually normalized to have unit length), which lives in a 'Hilbert' space (denoted by $\mathcal{C}^2$), but let's not worry about its rigorous mathematical definition). Usually we write it as

$$|\psi\rangle = \vec{v} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

States in a Hilbert space are defined by rays of vectors, and, for convenience and as a convention, we will normalize the complex vector ψ to have a unit norm. (There may be states that cannot be normalized, like a plane wave; but we can place time in a box by hand.) The meaning is the total probability by adding the distributions $|\alpha|^2$ over $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|\beta|^2$ over $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is unit,

$$\langle\psi| \cdot |\psi\rangle = \langle\psi|\psi\rangle = \vec{v}^* \cdot \vec{v} = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = |\alpha|^2 + |\beta|^2 = 1.$$

(See also the Born rule below regarding the probability of obtaining a specific measurement outcome.) Since it is a two-component vector, it has two basis vectors, and we have chosen the following,

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

Thus we can also use the Dirac's notation,

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle = \alpha|0\rangle + \beta|1\rangle.$$

Quantum gates or quantum operators act on quantum states (thus their dimensions should match), so they behave like a matrix, e.g. the NOT or also called X gate:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

which flips up to down,

$$X|\uparrow\rangle = X\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |\downarrow\rangle.$$

Another example gate is the so-called Hadamard gate, which unfortunately shares the same symbol H as the Hamilonian used by physicists,

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

which takes $|0\rangle$ to $|+\rangle \equiv (|0\rangle + |1\rangle)/\sqrt{2}$, and $|1\rangle$ to $|-\rangle \equiv (|0\rangle - |1\rangle)/\sqrt{2}$. In fact, $H^2 = I$, so the reverse direction also holds.

Note that we use a 'ket' notation $|\psi\rangle$ for ψ, whose 'dual row vector' $\psi^\dagger$ is denoted by a 'bra' notation $\langle\psi|$,

$$\langle\psi| = (|\psi\rangle)^\dagger = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix}.$$

Inner product, outer product, density matrix and its trace. The inner product between a bra and a ket results in a number; in particular, for the same ψ, we have

$$\langle\psi| \cdot |\psi\rangle = \langle\psi|\psi\rangle = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = |\alpha|^2 + |\beta|^2 = 1.$$

The outer product results in a matrix (also called 'density matrix' and denoted by ρ whose subscript ψ reminds you its origin from ψ), also an operator; in this case, it is a projector:

$$\rho_\psi \equiv |\psi\rangle\langle\psi| = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}\begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix}.$$

The trace of this density matrix is actually the norm square,

$$\mathrm{Tr}(\rho_\psi) \equiv \mathrm{Tr}(|\psi\rangle\langle\psi|) = \mathrm{Tr}\left[\begin{pmatrix} \alpha \\ \beta \end{pmatrix}\begin{pmatrix} \alpha^* & \beta^* \end{pmatrix}\right] = \mathrm{Tr}\begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix} = |\alpha|^2 + |\beta|^2 = 1.$$

Its normalization is unity if we have normalized the ket similarly. Interestingly, using the cyclic property of the trace, we have (i.e. trace of outer product = inner product):

$$\mathrm{Tr}(|\psi\rangle\langle\psi|) = \mathrm{Tr}(\langle\psi|\cdot|\psi\rangle) = \langle\psi|\psi\rangle = 1.$$

Bloch sphere . (Note in optics, the corresponding one is the Poincaré sphere.) Given the normalization $|\alpha|^2+|\beta|^2=1$ and the irrelevance of the overall phase (in computing, e.g., observables or in forming the density matrix), we can choose to parameterize α and β by

$$\alpha = \cos(\theta/2),\ \ \beta = e^{i\phi}\sin(\theta/2).$$

To describe a pure qubit requires two real variables; a qubit seems to have more information content than a classical bit, which can only be either 0 or 1. (However, to read out information stored in the qubit, one needs to perform measurement, which does not reveal all information.)

Let us calculate the density matrix,

$$\rho_\psi \equiv |\psi\rangle\langle\psi| = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix} = \begin{pmatrix} \cos^2(\theta/2) = (1+\cos\theta)/2 & \sin(\theta/2)\cos(\theta/2)e^{-i\phi} = \sin\theta e^{-i\phi}/2 \\ \sin(\theta/2)\cos(\theta/2)e^{i\phi} & \sin^2(\theta/2) = (1-\cos\theta)/2 \end{pmatrix}.$$

If we define $r_x = \sin\theta\cos\phi, r_y = \sin\theta\sin\phi$, and $r_z = \cos\theta$, then we have

$$\rho_\psi = \frac{1}{2}\begin{pmatrix} 1+r_z & r_x - ir_y \\ r_x + ir_y & 1-r_z \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{r_x}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \frac{r_y}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \frac{r_z}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We have thus arrived at a combination of the idenity matrix I and the three Pauli matrices $\sigma_x =: X$, $\sigma_y =: Y$, and $\sigma_z =: Z$ (we use both notations). We can write compactly

$$\rho_\psi = (I + \vec{r}\cdot\vec{\sigma})/2,\ \ \vec{\sigma} \equiv (X, Y, Z),$$

where $|\vec{r}| = 1$ in this case of a 'pure' state and the exact forms of the Pauli matrices are

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

Pure states lie on the surface of the Bloch sphere, see Fig. 1, and we will see below that there are states inside the sphere, which are called 'mixed' states.

Of course, density matrices also exist for quantum systems with more levels than two. But the physical picture of the corresponding 'sphere' is harder to obtain as a single constraint of the length $r \le 1$ is not sufficient [26]. For a qutrit (3-level system), see Ref. [27]. As another comment, we point out that even though a general density matrix ρ can contain complex numbers, it can be simulated by another real density matrix in an extended Hilbert space:

$$\tilde{\rho} = \mathrm{Re}(\rho)\otimes\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \mathrm{Im}(\rho)\otimes\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \rho\otimes|i\rangle\langle i| + \rho^*|-i\rangle\langle -i|. \tag{9}$$

Properties of Pauli matrices. These matrices square to identity, anticommute and are cyclic in their commutators:

$$X^2 = Y^2 = Z^2 = I,$$

$$\{X, Y\} \equiv XY + YX = 0 = \{Y, Z\} = \{Z, X\},$$

$$[X, Y] \equiv XY - YX = 2iZ,\ [Y, Z] = 2iX,\ [Z, X] = 2iY.$$

They are related to spin-1/2 angular momentum operators $\hat{S}_\alpha = \hbar\sigma_\alpha/2$ (where $\hbar$ is the reduced Planck constant) and they generate rotation of the qubit around respective axes,

$$R_x(\varphi) \equiv e^{-i\varphi\frac{X}{2}},\ R_y(\varphi) \equiv e^{-i\varphi\frac{Y}{2}},\ R_z(\varphi) \equiv e^{-i\varphi\frac{Z}{2}}. \tag{10}$$

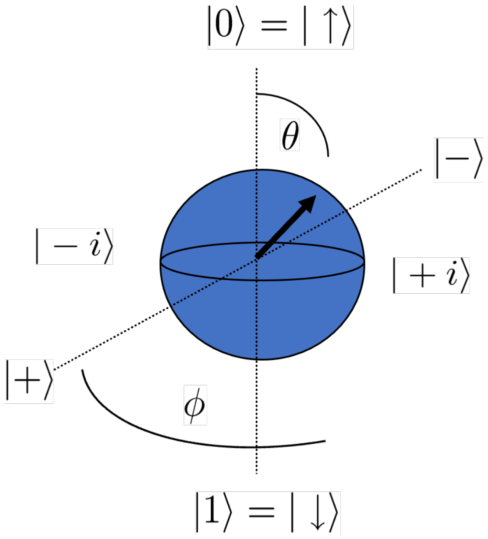

$|0\rangle = |\uparrow\rangle$
θ
$|-\rangle$
$|-i\rangle$
$|+i\rangle$
$|+\rangle$
ϕ
$|1\rangle = |\downarrow\rangle$

FIG. 1. Bloch sphere and some example states: $|0\rangle$, $|1\rangle$, $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$, and $|\pm i\rangle = (|0\rangle \pm i|1\rangle)/\sqrt{2}$.

A general rotation by an angle φ with respect to an axis denoted by a unit vector $\hat{n}$ can be conveniently written as

$$R_{\hat{n}}(\varphi) \equiv e^{-i\frac{\varphi}{2}\hat{n}\cdot\vec{\sigma}} = \cos(\varphi/2)I - i\sin(\varphi/2)\hat{n}\cdot\vec{\sigma}.$$

Note that X, Y, and Z are themselves rotation by 180 degrees (e.g. X flips up to down); Note they are Hermitian, e.g., $X^\dagger = X$ and traceless, e.g. $\mathrm{Tr}(X)$=$\mathrm{Tr}(Y)$=$\mathrm{Tr}(Z)$=0.

As we have seen that the density matrix corresponding to a general pure state is a rank-1 projector (i.e., having only one nonzero eigenvalue), written as $\rho_\psi \equiv |\psi\rangle\langle\psi|$. The associated vector $\vec{r}$ has a constraint of having a unit length $|\vec{r}| = 1$. Using a different expression that $\rho_\psi^2 = |\psi\rangle\langle\psi|\cdot|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \rho_\psi$. We find that $\mathrm{Tr}(\rho_\psi^2) = \mathrm{Tr}(\rho_\psi) = 1$. The trace of the square of a density matrix is sometimes referred to as purity, and a pure state has a unit parity. For a qubit, the unit purity implies $|\vec{r}| = 1$. We can deduce this result by directly squaring ρ_ψ, whose detail is left as an exercise,

$$\rho^2 = \frac{1}{4}(I + \vec{r}\cdot\vec{\sigma})^2 = \frac{1}{4}(I + 2\vec{r}\cdot\vec{\sigma} + |\vec{r}|^2 I) \Rightarrow \mathrm{Tr}(\rho^2) = (1 + |\vec{r}|^2)/2 \le 1.$$

Mixed states. If $|\vec{r}| < 1$, ρ does not represent the density matrix of a pure state, it is a mixed state! In other words, both eigenvalues of a qubit density matrix ρ are nonzero and less than one (rank-two in contrast to rank-one for a pure state). Suppose we diagonalize ρ and obtain two eigenvalues p_1 and p_2 and the associated eigenvectors (eigenstates) $|\psi_1\rangle$ and $|\psi_2\rangle$, then

$$\rho = p_1|\psi_1\rangle\langle\psi_1| + p_2|\psi_2\rangle\langle\psi_2|, \ \text{with}\ p_1 + p_2 = 1, p_i \ge 0, \ \text{and}\, \rho|\psi_i\rangle = p_i|\psi_i\rangle.$$

This shows that a mixed state can come from a statistical mixture of pure states; we can imagine a source randomly emit states $|\psi_i\rangle$ with probability p_i. In this example, we have the 'spectral' decomposition for ρ and $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthonormal eigenstates $\langle\psi_i|\psi_j\rangle = \delta_{ij}$. But, in general, there are *infinitely* many decompositions (into non-orthogonal states).

Exercise: could you give a method to construct such infinitely many decompositions? Hint. Consider a transformation $|\phi_j\rangle = \sum_{k=1}^2 U_{jk}\sqrt{p_k}|\psi_k\rangle$, where $j = 1, ..., n$ any dimension n and U_{jk} is a part of a larger $n \times n$ matrix. Note that $|\phi_j\rangle$'s will not be normalized.

We can see that the minimum number of pure-state components $|\psi_i\rangle$'s in the decompositon of a mixed state ρ is the rank of the matrix ρ. However, from the viewpoint of a source randomly emitting states $|\psi_i\rangle$ with probability p_i, the number of $|\psi_i\rangle$ can be arbitrary and that $|\psi_i\rangle$'s need not be orthogonal. The spectral decomposition is just a special decomposition. Moreover, a source can also emit a mixed state. In general, there are infinite number of ways in decomposing a mixed state (with more than two components), thus we have a statistical ensemble:

$$\rho = \sum_{i=1}^{n} q_j\rho_j, \quad \text{with} \ \sum_j q_j = 1, \quad \rho_j \ge 0 \,\&\, \mathrm{Tr}(\rho_j) = 1.$$

Note that when we write $\rho_j \ge 0$ for a matrix ρ_j, we mean that all eigenvalues are non-negative. The mathematical jargon for this is 'positive semi-definite'.

III. BASIC QUANTUM MECHANICAL RULES

Three essential ingredients of quantum mechanics will be introduced; they are (i) superposition principle (and entanglement), (ii) unitary evolution, and (iii) measurement. Note that in (i) we added entanglement to superposition, as it is a natural consequence that emerges when one discusses more than one particles or one partice but with multiple degrees of freedom. Entanglement is an key concept in modern quantum information theory, but it was Erwin Schrödinger who coined the term in 1935 in a letter to Einstein [28]. You may have already learned of these principles in your quantum physics or quantum mechanics class. But we want to stress here that later when you analyze quantum algorithms, these three ingredients are used over and over again.

(I) Superposition. Quantum states can have superposition (and later entanglement when there are more degrees of freedom). Superposition is not a new concept; it is already there in classical wave phenomena. Interference is a result of superpositon, either constructive or destructive. But everyday object like a coin cannot be in a superposition of head and tail. A quantum coin, if it exists, would be able to do so.

We have actually seen that a qubit can be a superposition of up and down, with respective weights or more precisely, amplitudes, written as,

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle,$$

e.g. a Q coin: α $+\ \beta$.

We have seen the Bloch sphere earlier and some other superpositions.

But how do you put it in such a superposition (e.g. if we begin with up)? Answer: by using quantum gates (e.g. the Hadamard gate H).

$$H|\uparrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) =: |+\rangle, \quad H|\downarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle) =: |-\rangle.$$

But how are quantum gates implemented? One key approach is to let quantum states evolve (under the so-called Hamiltonian), and the evolution gives rise to the action of a quantum gate. This naturally leads to the next principle—evolution.

(II) Unitary evolution. The evolution in time under the Hamiltonian is linear and unitary. The 'driver' of the evolution is the Hamiltonian (unfortunately has same symbol H as the Hadamard, so we will put a hat ˆ over it if needed for distinction). How it drives the evolution is via the Schrödinger's equation (see the interesting story about how Schrödinger discovered this equation in the article by Felix Bloch in Ref. [29]),

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

Let us not worry about this, as we will not dwell on how to solve the equation (this is what do we in PHY251 Modern Physics or PHY308 Quantum Physics at Stony Brook University). But there is a formal solution (for time independent Hamiltonian):

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}\hat{H}t}|\psi(0)\rangle.$$

This in effect gives us a gate (if we time the evolution well),

$$U_H(t) \equiv e^{-\frac{i}{\hbar}\hat{H}t}.$$

(Note that if the Hamiltonian is time-dependent, then evolution operator is formally written as a time-ordered integration: $U_H(t) \equiv \mathcal{T}e^{-\frac{i}{\hbar}\int dt\hat{H}(t)}$, as $\hat{H}(t)$ and $\hat{H}(t')$ may not commute.)

There are two important properties of this operator U_H: unitary and linear. First, it is easy to see that $U_H U_H^\dagger = U_H^\dagger U_H = 1$, showing that U_H is unitary [30]. Second, the linearity implies that

$$U_H(a|\psi\rangle + b|\phi\rangle) = a\left(U_H|\psi\rangle\right) + b\left(U_H|\phi\rangle\right).$$

Exercise. How do you use a magnetic field to implement the effect of the Hadamard gate on the spin of a spin-1/2 particle?

(III) Measurement. Strong measurement projects wavefunction; the outcome is often probabilistic. This is one mystical part of quantum mechanics, but is easy to illustrate with a quantum coin. Suppose we measure in the 'classical' or 'computational' basis to reveal up or down on a Q coin:

α $+\ \beta$ or $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$

You will obtain an outcome randomly. Sometimes it's up (we will give a score of +1) and sometimes it's down (we give a score of -1). What we know is that it occurs according to some distribution, which is known as the Born rule,

$$P_\uparrow = |\alpha|^2 = |\langle\uparrow|\psi\rangle|^2, \ \sigma = +1;$$

$$P_\downarrow = |\beta|^2 = |\langle\downarrow|\psi\rangle|^2, \ \sigma = -1.$$

Now we frame the understanding into the standard QM language and this usually requires the introduction of an 'observable', which is an Hermitian matrix. The notion of 'observables' is tightly related to the 'basis' of measurement, in this case is the $Z = \sigma_z$ operator (as the observable), written in a way to express the essence of the measurement postulate:

$$Z = (+1)|\uparrow\rangle\langle\uparrow| + (-1)|\downarrow\rangle\langle\downarrow|.$$

The basis is defined by the eigenstates of the observable. The 'eigenvalues' are what we 'read out' and the 'eigenstates' define the measurement basis. The act of measurement will project the system randomly into one of the eigenstates of the observable. The average 'score' represents the expected value of the observable over many repeated measurements,

$$\langle\psi|Z|\psi\rangle = P_\uparrow \cdot (+1) + P_\downarrow \cdot (-1).$$

Measurement is important, as the result of a quantum computation needs to be read out—by measurement. The actual process of quantum measurement in real life, however, is not as clear cut as the postulate. There are other kinds of measurement, such as the weak measurement, as opposed the strong measurement here, and the so-called positive-operator-value-measure measurement (POVM). As we shall see in a later unit, measurement along can indeed allows universal quantum computation. But it is worth the wait and sometimes need some kind of resource, i.e., the next topic—entanglement.

Entanglement. The true quantum-ness comes at two qubits or more, where you can have 'entanglement'. Superposition also occurs at classical waves, but entanglement is "the characteristic feature of quantum mechanics" according to Schrödinger. It is a natural consequence of superposition in the context of two or more particles or degrees of freedom. We will also see the advantage of Dirac's 'bra-ket' notation.

For two qubits, there are four basis states (we may omit 'tensor product' ⊗ notation)

$$|\uparrow\uparrow\rangle \equiv |\uparrow\rangle \otimes |\uparrow\rangle, \ |\uparrow\downarrow\rangle, \ |\downarrow\uparrow\rangle, \ |\downarrow\downarrow\rangle.$$

There are entangled states (which cannot be written as a product form), four such examples that form the basis of two qubits are,

$$|\Phi^\pm\rangle \equiv \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle \pm |\downarrow\downarrow\rangle), \ \ |\Psi^\pm\rangle \equiv \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle \pm |\downarrow\uparrow\rangle).$$

These are call Bell states. We will see later that they are useful resources for many quantum tasks (e.g., teleportation, superdense coding, and key distribution, just to name a few). Notation wise, it is cumbersome to write N-qubit states using vectors, as it requires 2^N components; so we see the advantage of the Dirac's notation.

$$|\uparrow\rangle = |0\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \ |\downarrow\rangle = |1\rangle = \begin{pmatrix}0\\1\end{pmatrix}, \ |\uparrow\rangle \otimes |\downarrow\rangle = \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}.$$

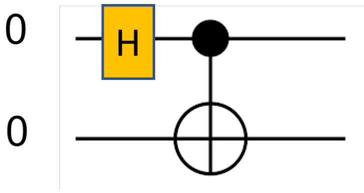
We now illustrate how to obtain one such entangled state from applying gates to the product state $|\uparrow\uparrow\rangle = |00\rangle$; we introduce the CNOT (Controlled-NOT or Controlled-X) gate,

$$\mathrm{CNOT}_{12} = |\uparrow\rangle\langle\uparrow| \otimes I + |\downarrow\rangle\langle\downarrow| \otimes X = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right).$$

Then

$$|\uparrow\uparrow\rangle \xrightarrow{H_1} \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \otimes |\uparrow\rangle \xrightarrow{\mathrm{CNOT}_{12}} \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle),$$

where the subscript $_1$ in H indicates that H acts on the first qubit. In terms of a quantum circuit (which we introduce now), it can be represented as:



where we use 0 and 1 instead of up and down arrows, and we have introduced the diagram for the CNOT gate.

Challenge exercise. How do you realize a CNOT gate? What form of a Hamiltonian for two qubits is needed and what is the evolution time?

By extending the above circuit to three qubits, we can arrive a three-particle entangled state like $|\mathrm{GHZ}\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$, which is the so-called famous Greenberger-Horne-Zeilinger (GHZ) state. There are other three-qubit entangled states, such as the W state, $|\mathrm{W}\rangle = (|001\rangle + |010\rangle + |100\rangle)/\sqrt{3}$. The more qubits we have, the more types of entangled states we can have [31].

IV. A TASTE OF QUANTUM POWER: SIMPLE QUANTUM ALGORITHMS

Even if you never learn quantum mechanics before, you can still learn quantum information and computation provided you know matrices and vectors (linear algebra). You need just to remember the three basic rules of QM and you can understand them in terms of linear algebra. We are ready for the first quantum algorithm.

A. Deutsch algorithm—constant or balanced?

Consider a function f mapping from one bit to one bit, $f : 0, 1 \to 0, 1$. There are four possibilities, but they can be classified into two categories, (1) constant: $f(0) = f(1)$ and (2) balanced: $f(1) = f(0) + 1$ (modulo 2); see Fig. 2. This problem and its quantum solution was proposed by David Deutsch in 1985 [32].

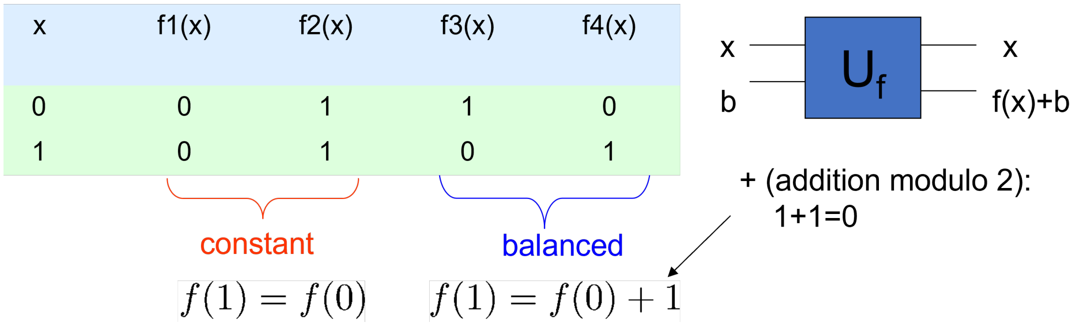
The question that David Deutsch asked is whether the function (unknown to us) is constant or balanced. This is mathematically equivalent to whether $f(0) \oplus f(1)$ is 0 or 1. It is obvious that classical computers need two function evaluations to determine which case f is. However, for quantum computers, it only takes one evaluation.

Useful observation/trick: 'phase kickback'. In the circuit shown in Fig. 2, we suppose its effect is to compute $f(x)$ of the first register and add it (modulo 2) to the second register,

$$|x\rangle \otimes |b\rangle \to |x\rangle \otimes |f(x) + b\rangle.$$

There is a special input of the second register: $|-\rangle \equiv (|0\rangle - |1\rangle)/\sqrt{2}$. If we send in $|x\rangle \otimes (|0\rangle - |1\rangle)$, ignoring the normalization, we have

$$\begin{aligned} |x\rangle \otimes (|0\rangle - |1\rangle) &\to |x\rangle \otimes (|f(x)\rangle - |f(x) + 1\rangle) \text{ by superposition and linearity} \\ &= |x\rangle \otimes (-1)^{f(x)}(|0\rangle - |1\rangle) \text{ by using } f(x) = 0 \text{ or } 1. \end{aligned}$$



x	f1(x)	f2(x)	f3(x)	f4(x)
0	0	1	1	0
1	0	1	0	1

FIG. 2. Illustration of function f and a circuit to perform $f(x) + b$ (modulo 2).

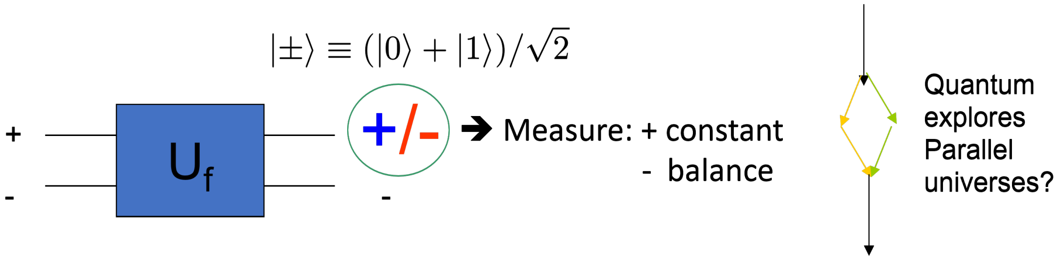


FIG. 3. Illustration of the Deutsch algorithm. It is as if the quantum computer is exploring two parallel universes by superposition.

Since the second register remains the same, for the purpose of analysis, we do not need to display it and we write (for $x = 0$ or 1, i.e. in the computational basis),

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

The effect is that a phase $(-1)^{f(x)}$ multiplies $|x\rangle$, which is called the phase kickback. This arises by using a superposition in the second register.

The solution. Now for the first register, instead of either $|0\rangle$ or $|1\rangle$, we can send in a superposition as well: $|+\rangle \equiv (|0\rangle + |1\rangle)/\sqrt{2}$, which is a superposition of all possible arguments of the function f. Using the phase kickback, we find that after the circuit the first register becomes,

$$|0\rangle + |1\rangle \to (-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle = \begin{cases} (-1)^{f(0)}(|0\rangle + |1\rangle) & \text{constant case} \\ (-1)^{f(0)}(|0\rangle - |1\rangle) & \text{balanced case} \end{cases}$$

To tell which case the function $f(x)$ is, we only need to measure the first register in a different basis $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$, as illustrated in Fig. 3. It is as if the quantum computer is exploring two parallel universes by superposition; Deutsch originally proposed this algorithm to test the multiverse or many-world theory.

Comment on input and readout. We have seen that we need one evaluation only but need to send in states with '+/-' and at the end of the circuit measure in '+/-' basis in order to determine whether the function is constant or balanced. Usually, qubits are assumed to be initialized to 0 and assumed to be measured in 0/1 basis. For the initialization of $|+\rangle$, we can apply the Hadamard gate to $|0\rangle$, yielding $|+\rangle = H|0\rangle$. For the initialization of $|-\rangle$, we first use X gate to flip $|0\rangle$ to $|1\rangle$, followed by a Hadamard gate: $|-\rangle = HX|0\rangle$.

For readout in $|\pm\rangle$ basis, we can apply a Hadamard gate right before the readout in $|0/1\rangle$ basis, which can be understood by the following identity,

$$\langle \pm|\psi\rangle = \langle 0/1|H|\psi\rangle.$$

We can include all these gates into the circuit and arrive at one shown in Fig. 4.

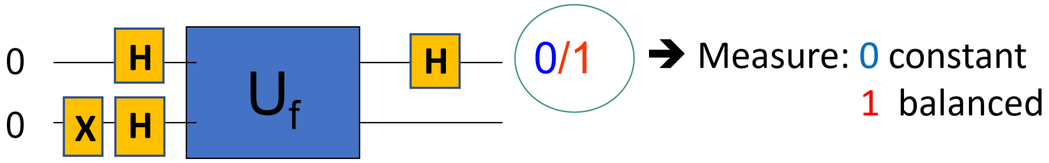


FIG. 4. Illustration of the circuit for the Deutsch algorithm with the assumption that qubits are initialized in $|0\rangle$ and measured in $|0/1\rangle$ basis.

B. Quantum Parallelism—the source of power?

It seems that quantum computers can solve problems more efficiently than classical computers (we will study other more powerful algorithms later, as such Shor's factoring and Grover's search). One perspective is the massive parallelism of quantum mechanics: one can input a superposition of 'questions' and the quantum computer output the superposition of pairs of 'questions' and 'answers'. For example, we consider the same form of circuit previously in the Deutsch algorithm, but allow the input x to be N qubits and the other input b to be M qubits. Then the function f maps from $\{0,1\}^{\otimes N}$ to $\{0,1\}^{\otimes M}$. Assume we have a unitary circuit to implement that:

$$|x\rangle \otimes |b\rangle \rightarrow |x\rangle \otimes |f(x)+b\rangle.$$

With this, we can create a superposition of all possible classical inputs (ignoring normalization) and a blank second register (in M-qubit $|0\rangle$ state),

$$(|0\rangle + |1\rangle + |2\rangle + \dots) \otimes |0\rangle.$$

After the circuit, the total state becomes,

$$\rightarrow \big(|0\rangle \otimes |f(0)\rangle + |1\rangle \otimes |f(1)\rangle + |2\rangle \otimes |f(2)\rangle + \dots\big).$$

The quantum computer thus contains a massively entangled state that contains all possible information about the function.

The dilemma. However, to read out some answer, we need to probe or measure the qubits of the quantum computer, and this creates some problems. For example, we can measure first register and we may get a number k (which is random, recall the measurement postulate). Then the second register will collapse to $f(k)$. So we read out randomly an argument of the function and its corresponding function value, which is not so useful.

However, we can try to be smart and measure in different basis or/and the second register. Suppose we measure the second register and obtain a value f_0, then according to the measurement postulate the first register will collapse to a superposition of those arguments evaluating to f_0, i.e.,

$$|\mathrm{Reg}_1\rangle = \sum_{x: f(x)=f_0} |x\rangle,$$

which is useful for probing the symmetry properties of the function f. In essence, Shor's algorithm exploits this and extracts the periodicity of the modular exponentiation $x^{k+r} \equiv x^k \bmod N$, which we study later in this course.

C. Other simple algorithms

(1) **Deutsch-Josza Algorithm** [33]: which is an extension of Deutsch algorithm to multiple qubits; see Fig. 5. Unlike the one-qubit case, a function of n-bit input and n-bit output can be both non-constant and non-balanced. So here, we are promised that the function in question is either (i) constant (evaluted to 0 on all inputs or 1 on all inputs) or (ii) balanced (evaluated to 1 for exactly half of the input and 0 for the other half). Non-constant and non-balanced cases will appear.

The solution is similar to the single-qubit case. We will send in a state in the first register,

$$|\mathrm{in}\rangle = [\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)]^{\otimes n} = \underbrace{H \otimes H \otimes \dots H}_{n} \underbrace{|0\rangle \otimes |0\rangle \otimes \dots |0\rangle}_{n},$$

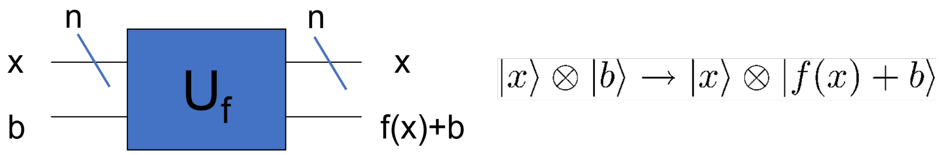


FIG. 5. Illustration of the circuit for the Deutsch-Josza algorithm. The 'slashes' on the first register indicates that it is a multiple-qubit register.

and the second register will be supplied with a $|b\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ state, as we have seen in the phase kickback. One can convince him/herself that in the case the same phase kickback works, i.e. for x in the 'computational' basis (e.g. $|00\ldots0\rangle$, ..., $|11\ldots1\rangle$), we have

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

Exercise. As an exercise, you will work out the following:
a) Show the quantum state of the first register after the circuit.
b) Show that if f is constant, the first register is always $|+\ldots+\rangle$.
c) Show that if f is balanced, the first register is always orthogonal to $|+\ldots+\rangle$.

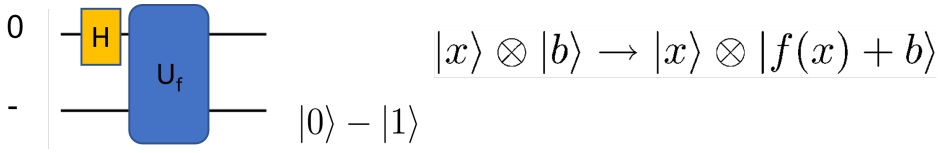


FIG. 6. Illustration of the circuit for the Bernstein-Vazirani algorithm.

(2) **Bernstein-Vazirani algorithm** [34]: to determine $a = 0, 1$ in the promised function $f(x) = a \cdot x$. This also has an n-qubit version (original problem). For the single-qubit case, we have the action on the first regsiter,

$$\text{first register}: |0\rangle \xrightarrow{H} |0\rangle + |1\rangle \xrightarrow{U_f} (-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle \begin{cases} (|0\rangle + |1\rangle), & a = 0 \\ (|0\rangle - |1\rangle), & a = 1 \end{cases}$$

The diagram for Bernstein-Varzirani algorithm is similar to the Deutsch algorithm; see Fig. 6. Thus, in a similar way to the Deutsch algorithm, by measuring the first register, we can distinguish whether $a = 0$ or $a = 1$. Note that there a related generalization—the so-called the hidden linear function problem, which has a shallow-depth quantum solution [35, 36].

Exercise. As an exercise, if $f(x) = a \cdot x + c_0$, how do you find the value of c_0?

Recent generalization. As mentioned above, the Bernstein-Varzirani problem was recently generalized (a non-oracular version) by Bravyi, Gosset, and König to a 2D hidden linear function problem (HLF) [35], which is specified by a quadratic form **q** that maps n-bit strings to integers modulo four , with the goal being to identify a linear boolean function which describes the action of **q** on a certain subset of n-bit strings. They show that any classical probabilistic circuit composed of bounded fan-in gates that solves the 2D HLF problem with high probability must have depth logarithmic in n, while a constant-depth quantum circuit using local gates on a 2D grid can solve this, i.e., with constant time. This inspired later development that showed that there is an exponential separation between shallow quantum circuits and unbounded fan-in shallow classical circuits [36]. Finding a provable separation between classical computation and quantum computation is still an active research.

(3) **Simon's algorithm** [37]: Consider a function $f : \{0,1\}^n \to$ finite set X. We are promised that there is some"hidden" string $s = s_1 s_2 .. s_n$ such that $f(x) = f(y)$ if and only if $x = y$ or $x = y \oplus s$ (bitwise XOR). The goal is to find the string s. We begin with an observation of Hadamard gates acting on all 0's,

$$|\mathbf{0} \equiv 0^{\otimes n}\rangle \xrightarrow{H^{\otimes n}} \frac{1}{2^{n/2}} \sum_{z_i's} |z_1\rangle \otimes |z_2\rangle \otimes \cdots |z_n\rangle = \frac{1}{2^{n/2}} \sum_{\mathbf{z}} |\mathbf{z}\rangle.$$

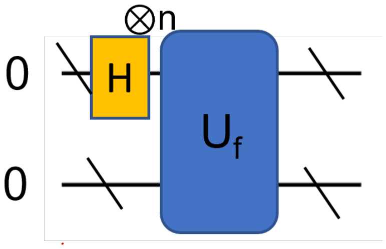


FIG. 7. Illustration of the circuit for the Simon's algorithm.

The case on an arbitrary bit string s is left as an exercise,

$$|\mathbf{s} \equiv s_1 \dots s_n\rangle \xrightarrow{H^{\otimes n}} \frac{1}{2^{n/2}} \sum_{\mathbf{z}} (-1)^{\mathbf{s}\cdot\mathbf{z}} |\mathbf{z}\rangle.$$

With these two identities, we can consider how the Hadamard gates transform their superposition,

$$\frac{1}{\sqrt{2}}(|\mathbf{0}\rangle + |\mathbf{s}\rangle) \xrightarrow{H^{\otimes n}} \frac{1}{2^{(n+1)/2}} \sum_{\mathbf{z}} (1 + (-1)^{\mathbf{s}\cdot\mathbf{z}})|\mathbf{z}\rangle = \frac{1}{2^{(n-1)/2}} \sum_{\mathbf{z}\in\{\mathbf{s}\}^{\perp}} |\mathbf{z}\rangle.$$

The resultant state has no amplitude with any bit-string state $|z\rangle$ that $s \cdot z = 1 \pmod 2$. It only contains components z's orthogonal to s.

We can generalize the above two superposition to a more general case, which is left for you to prove,

$$\frac{1}{\sqrt{2}}(|\mathbf{x}\rangle + |\mathbf{x}\oplus\mathbf{s}\rangle) \xrightarrow{H^{\otimes n}} \frac{1}{2^{(n-1)/2}} \sum_{\mathbf{z}\in\{\mathbf{s}\}^{\perp}} (-1)^{\mathbf{x}\cdot\mathbf{z}} |\mathbf{z}\rangle.$$

Similar to the previous case, it also only contains components that are orthogonal to s. The strategy to solve for s is to obtain enough bit strings z that are orthogonal to s so that s can be uniquely determined by linear algebra.

A dilemma. The stratgey of this algorithm nicely illustrates what we have said earlier about the dilemma between the quantum parallelism and measurement. We prepare the first register in a superposition of all bit strings $\sum_x |x\rangle$ and the second register in $|0\rangle$. After the circuit, the state of the whole system is in a superposition of all pairs of bit strings and values: $|\psi\rangle = \sum_x |x\rangle \otimes |f(x)\rangle$ (un-normalized). As we mention earlier, we could be smart and measure the second register, which randomly collapses it to $|f_0\rangle$. Then the first register is collapsed to $\sum_{x,f(x)=f_0} |x\rangle$, which can only have two components, as promised. Then we apply Hadamard gates and perform additional measurement to obtain a bit string z that is orthogonal to s. By repeating this whole procedure a few times, we will obtain enough number of z's that are orthogonal to s. Then the remaining task is solved by a classical computer.

A quick peek into Shor's factoring algorithm. [Can skip this comment for your first reading.] Note that Simon's algorithm can be regarded as a precursor of Shor's factoring algorithm, which will be covered in a later unit, but the essence is to find the integer period r, such that an integer x coprime with N (the number we want to factorize) is taken to 1 modulo N, i.e., $x^r \equiv 1 \pmod N$. We have quantum computer to take the input of an integer $k = 0, \dots, N-1$ in an equal-weight supersposition (in the first register) and compute x^k and store the result in a second register. We can measure the second register and see the first register being collapsed to a superposition of a subset of k's, i.e., $\sim |k_0\rangle + |k_0 + r\rangle + |k_0 + 2r\rangle + \cdots$, which has a spacing of r. Can a quantum computer extract this spacing or 'period'? The answer is affirmative, and it is the task of the quantum Fourier transform. But based on your prior knowledge on the discrete Fourier transform, in the transformed space, there is a discrete 'momentum' which is a multiple of $2\pi/r$, showing up in the 'phase' encoded in a quantum state. Fortunately, this information can be extracted. A secodn quantum: why does r allow us to factorize N? This is because $x^r - 1 = mN$, which is equivalent to $(x^{r/2}+1)(x^{r/2}-1) = mN$. If r is even, then the two factors $(x^{r/2} \pm 1)$ may each have a common factor with N.

Quantum Algorithm Zoo. There are many quantum algorithms that have been developed and more are being developed. You can find the list at the website of the "Quantum Algorithm Zoo" at `https://quantumalgorithmzoo.org/`. Two most well-known and by now classic quantum algorithms are Shor's factoring and Grover's search, which will be discussed later in this course. We will also discuss other algorithms, such as those for solving linear equations (e.g. the HHL algorithm to solve for $\vec{x}$ in the equation $A\vec{x} = \vec{b}$) and those of the type using variational ansatzes. The point is that by completing this unit, you are now ready to explore quantum computing and quantum algorithms.

Recent Development. It turns out that there is a shift in thinking and developing new quantum algorithms from the point of view of 'quantum signal processing' [38], something the author did not know previously. There was also the progress in quantum singular value transformation [39] (QSVT). This has evolved to a *grand unifying* approach that can incorporate many known quantum algorithms [40] and has the potential to improve existing and construct new algorithms.

V. REFLECTION

After going through this unit and studying the lecture slides, the recorded videos, this set of notes, and perhaps homework exercise, have you achieved the learning outcomes?
(1) You'll be ready to embark on a QIS journey. (2) You'll understand why quantum computing has a lot of potential and why it may also not be all-powerful.

Suggested reading: Nielsen and Chuang (N&C) 1.2-1.4, 2.2, 2.4; Kaye, Laflamme and Mosca (KLM) 1.4, 1.6, chapter 3, 6.2-6.4; Qiskit book (Qb) chapter 1, 2.1-2.3.

[1] P. Di Francesco, P. Mathieu, and D. Sénéchal, Conformal invariance in two dimensions, in *Conformal Field Theory* (Springer, 2011) pp. 111–149.

[2] J. L. Miller, Does quantum mechanics need imaginary numbers?, Physics Today **75**, 14 (2022).

[3] M.-O. Renou, D. Trillo, M. Weilenmann, T. P. Le, A. Tavakoli, N. Gisin, A. Acín, and M. Navascués, Quantum theory based on real numbers can be experimentally falsified, Nature **600**, 625 (2021).

[4] T. Rudolph and L. Grover, A 2 rebit gate universal for quantum computing, arXiv preprint quant-ph/0210187 (2002).

[5] D. Aharonov, A simple proof that toffoli and hadamard are quantum universal, arXiv preprint quant-ph/0301040 (2003).

[6] D. J. Griffiths and D. F. Schroeter, *Introduction to quantum mechanics* (Cambridge university press, 2018).

[7] L. Susskind and A. Friedman, *Quantum mechanics: the theoretical minimum* (Basic Books, 2014).

[8] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, Physical review **47**, 777 (1935).

[9] J. S. Bell, On the Einstein Podolsky Rosen paradox, Physics Physique Fizika **1**, 195 (1964).

[10] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed experiment to test local hidden-variable theories, Physical Review Letters **23**, 880 (1969).

[11] R. P. Feynman, Simulating physics with computers 1981, International Journal of Theoretical Physics **21**, 467 (1982).

[12] R. P. Feynman, Quantum mechanical computers, Optics news **11**, 11 (1985).

[13] A. Aspect, P. Grangier, and G. Roger, Experimental realization of einstein-podolsky-rosen-bohm gedankenexperiment: a new violation of bell's inequalities, Physical review letters **49**, 91 (1982).

[14] C. H. Bennett and B. G., Quantum cryptography: Public key distribution and coin tossing, Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing **175**, 8 (1984).

[15] P. Benioff, Quantum mechanical hamiltonian models of turing machines, Journal of Statistical Physics **29**, 515 (1982).

[16] Y. I. Manin, Vychislimoe i nevychislimoe [computable and noncomputable], Sov.Radio , 13 (1980).

[17] A. K. Ekert, Quantum cryptography based on bell's theorem, Phys. Rev. Lett. **67**, 661 (1991).

[18] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels, Phys. Rev. Lett. **70**, 1895 (1993).

[19] P. W. Shor, Algorithms for quantum computation: discrete logarithms and factoring, in *Proceedings 35th annual symposium on foundations of computer science* (Ieee, 1994) pp. 124–134.

[20] P. W. Shor, Scheme for reducing decoherence in quantum computer memory, Phys. Rev. A **52**, R2493 (1995).

[21] L. K. Grover, A fast quantum mechanical algorithm for database search, in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing* (1996) pp. 212–219.

[22] L. K. Grover, Quantum mechanics helps in searching for a needle in a haystack, Phys. Rev. Lett. **79**, 325 (1997).

[23] L. K. Grover, Quantum computers can search rapidly by using almost any transformation, Phys. Rev. Lett. **80**, 4329 (1998).

[24] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[25] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. Van Den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, *et al.*, Evidence for the utility of quantum computing before fault tolerance, Nature **618**, 500 (2023).

[26] M. S. Byrd and N. Khaneja, Characterization of the positivity of the density matrix in terms of the coherence vector representation, Phys. Rev. A **68**, 062322 (2003).

[27] S. K. Goyal, B. N. Simon, R. Singh, and S. Simon, Geometry of the generalized bloch sphere for qutrits, Journal of Physics A: Mathematical and Theoretical **49**, 165203 (2016).

[28] See `https://en.wikipedia.org/wiki/Quantum_entanglement`.

[29] F. Bloch, Heisenberg and the early days of quantum mechanics, Physics Today **29**, 23 (1976), https://pubs.aip.org/physicstoday/article-pdf/29/12/23/8280756/23_1_online.pdf.

[30] In quantum mechanics, anti-unitary action is also possible, such as the time reversal and charge conjugation.

[31] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde, Four qubits can be entangled in nine different ways, Phys. Rev. A **65**, 052112 (2002).

[32] D. Deutsch, Quantum theory, the church–turing principle and the universal quantum computer, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **400**, 97 (1985).

[33] D. Deutsch and R. Jozsa, Rapid solution of problems by quantum computation, Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences **439**, 553 (1992).

[34] E. Bernstein and U. Vazirani, Quantum complexity theory, SIAM Journal on computing **26**, 1411 (1997).

[35] S. Bravyi, D. Gosset, and R. König, Quantum advantage with shallow circuits, Science **362**, 308 (2018).

[36] A. B. Watts, R. Kothari, L. Schaeffer, and A. Tal, Exponential separation between shallow quantum circuits and unbounded fan-in shallow classical circuits, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (2019) pp. 515–526.

[37] D. R. Simon, On the power of quantum computation, SIAM journal on computing **26**, 1474 (1997).

[38] G. H. Low and I. L. Chuang, Optimal hamiltonian simulation by quantum signal processing, Physical review letters **118**, 010501 (2017).

[39] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe, Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (2019) pp. 193–204.

[40] J. M. Martyn, Z. M. Rossi, A. K. Tan, and I. L. Chuang, Grand unification of quantum algorithms, PRX Quantum **2**, 040203 (2021).