Lecture 1 - Introduction to Quantum Computation

The Eigensolvers

July 5, 2021

1 What is Quantum?

The idea of quantum started with some experiments that could not be explained using the tools of classical physics. Some of these experiments were the black body radiation, photoelectric effect and double slit experiment. These problems were solved by assuming the existence of discrete amounts of matter, the so called **quanta**, as if matter came in small packets instead of being a continuum.

1.1 Ultraviolet catastrophe

In the late 19th century, physicists were studying how bodies emitted light when they were heated up. When they applied classical laws, they obtained that the body must emit an infinite amount of ultraviolet light. This is known as the ultraviolet catastrophe or Rayleigh-Jeans catastrophe. (See Fig. 1) This problem was solved by Planck introducing a small energy, so that any possible energy of the system was an integer multiple of that energy.

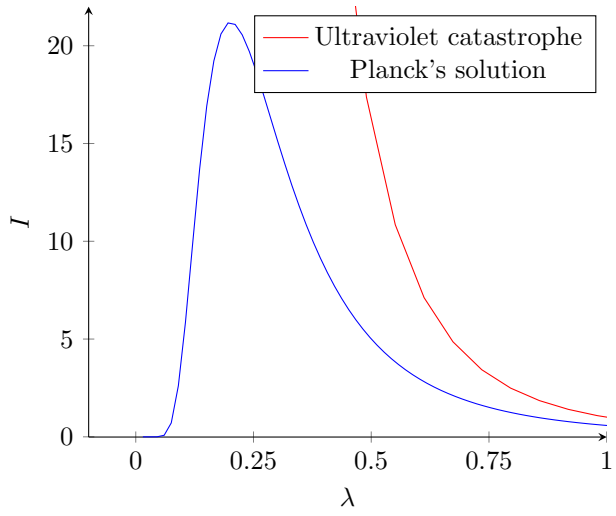


Figure 1: Classical and modern models for radiation due to heat.

1.2 Photoelectric effect

In 1887, Heinrich Hertz discovered that sometimes light can draw electrons from some materials. This is called the photoelectric effect and it could not be explained using the classical theory of electromagnetism, which treat light as a wave. This problem was solved in 1905 by Albert Einstein, he postulated that light consisted of small packets, which are called photons.

1.3 Double-slit experiment

This experiment was carried out in 1801 by Young and shows how light could also behave as a wave. Just like water waves, light exhibited interference patterns (See Fig. 2)

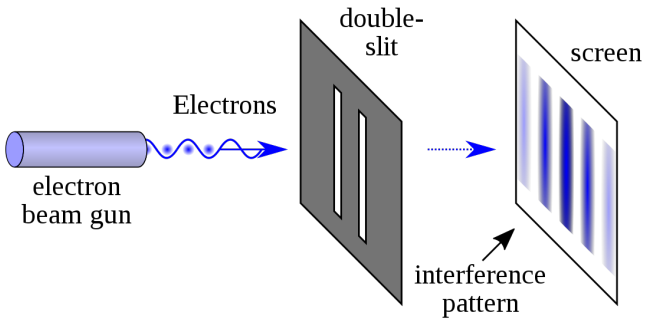


Figure 2: A light source is placed against a double-slit plate, followed by a screen where the photons are measured. The brighter the color on the screen, the higher the intensity.

This experiment was then generalized to electrons and some other types of matter. So, matter displays a wave-particle duality.

The double-slit experiment might seem contrary to the theory of quantum objects we have been mentioning. However, if the light source is used at low intensities (single photons) the interference pattern remains there. This is due to superposition, which will be mentioned later.

After looking at these "failures" of classical mechanics, you may wonder what came to replace it. Well, you probably already know the answer: quantum mechanics. We would love to introduce quantum mechanics now, but we need to first dive into some mathematics. Don't worry, we won't go into *everything* used in quantum mechanics. We will only look at the fundamental topics essential to understand the basic postulates of quantum mechanics.

2 Mathematical interlude

2.1 Complex numbers

A complex number can be expressed as $z = a + bi$, where a and b are real numbers and i satisfies the equation $i^2 = -1$. Since no real number satisfies this equation, i is called an imaginary number. As you can see, complex numbers have two parts: a called the real part and bi called the imaginary part.

You can do the normal operations you do with real numbers in the complex realm quite easy. Suppose you have two complex numbers: $z = a+bi$ and $w = c+di$. You can easily add them, subtract them, and multiply them as in the equations below.

$$z + w = (a + bi) + (c + di) = (a + c) + (b + d)i \tag{1}$$
$$z - w = (a + bi) - (c + di) = (a - c) + (b - d)i \tag{2}$$
$$z \cdot w = (a + bi) \cdot (c + di) = ac + (ad + bc)i - bd \tag{3}$$

Division is a bit more tricky, but we won't need it here.

2.2 Dirac notation

Also referred to as bra-ket notation, you will find Dirac notation all over quantum mechanics, not only quantum computing. It consists of bras and kets, which are represented by $\langle\alpha|$ and $|\alpha\rangle$, respectively. Kets are used to represent column vectors as

$$|\alpha\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \tag{4}$$

where n is the dimension of the vector. And bras are used to represent the conjugate transpose of these column vectors, as

$$\langle\alpha| = \begin{pmatrix}\alpha_1^* & \alpha_2^* & \cdots & \alpha_n^*\end{pmatrix} \tag{5}$$

where * denotes complex conjugation.

In quantum computing, you are going to be using $|0\rangle$ and $|1\rangle$ a lot, let's see what they mean. These two kets represent the two orthonormal vectors that make up the computational basis. In vector notation, they are represented as

$$|0\rangle = \begin{pmatrix}1\\0\end{pmatrix} \text{ and } |1\rangle = \begin{pmatrix}0\\1\end{pmatrix}. \tag{6}$$

You can see these two are orthogonal by taking their dot product, topic of next section.

2.3 Dot product

Taking the dot product of two vectors is a very common operation in quantum computing. You will do this when you want to find out the probability of finding a state $|\psi\rangle$ in one of the vectors making up the computational basis. But let's not go too far yet. First, let's see how to compute the dot product. We will do this first using vector notation and then we will translate it to Dirac notation.

Suppose you have two vectors

$$\vec{a} = \begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix} \text{ and } \vec{b} = \begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}, \tag{7}$$

where n is once again the dimension of these vectors. Then, the dot product (denoted by $\cdot$) is defined as follows

$$\vec{a}\cdot\vec{b} = \begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}^\dagger \begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix} = \begin{pmatrix}a_1^* & a_2^* & \cdots a_n^*\end{pmatrix} \begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix} = a_1^*b_1 + a_2^*b_2 + \cdots a_n^*b_n = \sum_{i=1}^{n} a_i^*b_i. \tag{8}$$

As you can see, we take the conjugate transpose (denoted by †) of the first vector and multiply it by the second vector. And we end up with a single scalar number.

At this point, you may have already noticed how to express the dot product in Dirac notation. Anyways, let's see how to do it. Suppose you have two kets $|\alpha\rangle$ and $|\beta\rangle$. Since these two correspond to column vectors, we need to make one a row vector, plus complex conjugate, to take the dot product. We can do this by converting one of them into a bra. Let's see this in action.

$$|\alpha\rangle\cdot|\beta\rangle = \langle\alpha|\cdot|\beta\rangle = \begin{pmatrix}\alpha_1^* & \alpha_2^* & \cdots \alpha_n^*\end{pmatrix} \begin{pmatrix}\beta_1\\\beta_2\\\vdots\\\beta_n\end{pmatrix} = \sum_{i=1}^{n} \alpha_i^*\beta_i \tag{9}$$

As you can see, we are basically performing the same operation we would do with normal vectors, but the notation is much more convenient. Instead of writing $\langle\alpha|\cdot|\beta\rangle$, we write a new symbol called bra-ket $\langle\alpha|\beta\rangle$. And we usually don't write the matrix representation when working in Dirac notation, I just showed it so the operation we clear. Therefore, a normal dot product in Dirac notation would look like

$$\langle\alpha|\beta\rangle = \sum_{i=1}^{n} \alpha_i^*\beta_i, \tag{10}$$

which is much more pleasing!

2.4 Orthonormal basis

An orthonormal basis is a set of linearly independent vectors $\{\vec{v_1}, \cdots, \vec{v_n}\}$ whose norm equals 1. These vectors satisfy $\vec{v_i} \cdot \vec{v_j} = \delta_{ij}$. In Dirac's notation, $\langle v_i | v_j \rangle = \delta_{ij}$. Where $\delta_{ij} = 0$ if $i = j$ and $\delta_{ij} = 0$ otherwise.

2.5 Matrices

A matrix is an array of numbers aligned in rows and columns. You can think of the vectors we introduced earlier on as a special case of matrices, in which the number of rows or columns is just 1.

We can define another product using two vectors, in which we end up with a matrix instead of a complex number. How is that possible? Consider

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} b_1^* & b_2^* & \cdots b_n^* \end{pmatrix} = \begin{pmatrix} a_1 b_1^* & a_1 b_2^* & \cdots & a_1 b_n^* \\ a_2 b_1^* & a_2 b_2^* & \cdots & a_2 b_n^* \\ \vdots & \vdots & \ddots & \vdots \\ a_n b_1^* & a_n b_2^* & \cdots & a_n b_n^* \end{pmatrix} \tag{11}$$

which is called the outer product, and is represented in Dirac's notation as $|\alpha\rangle\langle\beta|$. As you can see, the meaning in Dirac notation is the same as the one showed in the equation above, a column vector is multiplied by a row vector.

2.6 Eigenvalues and eigenvectors

Let M be a matrix, $\vec{v}$ a vector and λ a complex number, Then, if the following equation is satisfied

$$M\vec{v} = \lambda\vec{v} \tag{12}$$

we say $\vec{v}$ is an eigenvector of M with eigenvalue λ. These concepts are very important in quantum mechanics, since they're directly related to measurement. The eigenvectors of a matrix are those vectors that don't change their direction after they're multiplied by M.

To get the eigenvalues of a matrix, you can solve the characteristic equation. This is

$$|A - \lambda I| = 0, \tag{13}$$

where A is the matrix you want the eigenvalues and eigenvectors of, λ is a variable, I is the identity matrix (1s on the main diagonal and 0s elsewhere), and $|\cdot|$ denotes the determinant of a matrix. For the 2 by 2 case, the determinant of a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is given by $ad - bc$.

Once you have the eigenvalues $\{\lambda\}$ of a matrix, you can solve a simple equation to get the eigenvectors. Consider $\vec{v} = \begin{pmatrix} x \\ y \end{pmatrix}$. The equation you need to solve is $A\vec{v} = \lambda_i\vec{v}$ for each eigenvalue λ_i in $\{\lambda\}$. This equation expanded looks as follows.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda_i \begin{pmatrix} x \\ y \end{pmatrix} \tag{14}$$

Note that the values a, b, c, d and λ_i will be known. Therefore, you only need to solve for x and y, and $\vec{v}$ will be an eigenvector corresponding to the eigenvalue λ_i. And $a \cdot \vec{v}$ for any scalar a will also be a valid eigenvector.

2.7 Matrix vector multiplication

Let M be a matrix,

$$\begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn} \end{pmatrix} \tag{15}$$

These terms M_{ij} are called *matrix elements*. Besides, notice that

$$M_{ij} = \begin{matrix} \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 \end{pmatrix} \\ \begin{pmatrix} 1 & \cdots & i & \cdots & n \end{pmatrix} \end{matrix} \begin{pmatrix} M_{11} & \cdots & M_{1j} & \cdots & M_{1n} \\ \vdots & \ddots & & & \vdots \\ M_{i1} & \cdots & M_{ij} & \cdots & M_{in} \\ \vdots & & & \ddots & \vdots \\ M_{n1} & \cdots & M_{nj} & \cdots & M_{nn} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ \vdots & \vdots \\ 1 & j \\ \vdots & \vdots \\ 0 & n \end{pmatrix} \tag{16}$$

This can be written much simpler using Dirac's notation, $M_{ij} = \langle v_i | M | v_j \rangle$.

Notice that this has the same meaning as the equation above. First, we have a bra which corresponds to a row vector, then we have the matrix and finally we have a ket that corresponds to a column vector. Since $\langle v_i|$ and $|v_j\rangle$ are both part of an orthonormal basis (the most simple one), they only have one non-zero element and that element has value 1.

3 Postulates of Quantum Mechanics

The aim of the three subsections of Section 1 was to make one point: the classical laws of physics do not appropriately describe certain systems[1], namely, microscopic systems[2]. A new theoretical framework had to be developed to fill these gaps, which meant that the Royal Swedish Academy of Sciences had to get their Nobel Prizes ready because physicists all around the globe were prepared to begin investigating this. The field of quantum mechanics is hard to introduce and would require a full course[3]. However, it is ubiquitously summarized by a non-deterministic model described by the following postulates [4]:

3.1 State of a system

Postulate 1: The state of a system is described by its statevector, $|\psi\rangle \in \mathcal{H}$.

In classical physics, the state of a system is defined by a set of physical quantities such as position and momentum which could be used to derive the equations of motion of that system. In quantum mechanics, a system's *state* is treated differently and is rather described by a vector with components called amplitudes, that, when squared, yield the probability of obtaining some state. For instance, let's say a particle can be in one of three states: blue, green or red. Its state is then represented by the following vector:

$$|\psi\rangle = \begin{pmatrix} \psi_{blue} \\ \psi_{green} \\ \psi_{red} \end{pmatrix} \tag{17}$$

where $P_{blue} = |\psi_{blue}|^2 = \psi^*_{blue}\psi_{blue}$, $P_{green} = |\psi_{green}|^2 = \psi^*_{green}\psi_{green}$, and $P_{red} = |\psi_{red}|^2 = \psi^*_{red}\psi_{red}$. One thing to note is that the sum of all probabilities (i.e the *norm*[5] of the statevector) must be equal to 1, or

$$\sum_{i=0}^{n} P_i = \sum_{i=0}^{n} \psi_i^* \psi_i = \sum_{i=0}^{n} |\psi_i|^2 = 1 \tag{18}$$

Similarly, the position of a particle (a continuous property that could have an infinite number of values) could be described in terms of a statevector $|\psi\rangle \in C^\infty$, where $\langle x|\psi\rangle$ yields a complex amplitude function called the wavefunction, $\psi(x)$ ($|x\rangle$ is the basis vector of the position basis). The wavefunction, when squared, could be used to find the probability of a particle being in any x value[6].

[1] Although act as a great approximations for others. See, Apollo 13.

[2] As well as systems with low temperatures, high velocities, etc. But microscopic should be fine for the sake of this lecture.

[3] Let us know if you want that.

[4] Note that these are not the original postulates proposed, but rather the reduced postulates found in 2.2 of *Nielsen and Chuang*

[5] You should know this from Section 3 to 5 in pre-requisites chapter 1

[6] This will not be used for this level of QM/QC

3.2 State Evolution

Postulate 2: The evolution of a state, $|\psi\rangle \rightarrow |\psi'\rangle$, could be described by a unitary operator.

The word mechanics implies that change is involved, and given that quantum mechanics treats a state as a complex vector (See Postulate 1), it must also describe how that vector evolves. Postulate Two states that the evolution of a state $|\psi\rangle$ can be described by a unitary operator acting on it, or

$$|\psi'\rangle = U\,|\psi\rangle \tag{19}$$

where $|\psi'\rangle \in C^n$ is the evolved state, $|\psi\rangle \in C^n$ is the initial state and $U \in C^{n\times n}$ is a unitary matrix, where $UU^\dagger = I$.

An important property of unitary matrices is norm-preservation, which means that the total probability of the system will remain 1 regardless of how the state evolves

3.3 Measurement

Postulate 3: The probability of measuring the system to be in some basis state $|\,|a_n\rangle$ is given by $\langle a|\psi\rangle\,|^2$ Quantum mechanics states that until a system is measured, one can't know anything about the physical state of that system. However, once measurement is performed, one can obtain any basis state $\{|a_1\rangle, |a_2\rangle, |a_3\rangle, \ldots, |a_n\rangle\}$ with probability of each being $|\,\langle a_i|\psi\rangle\,|^2$, which equates to $|a_i\rangle$'s amplitude squared, or $\psi_i^*\psi_i$

3.4 Composite Quantum Systems

Postulate 4: The state of composite quantum systems can be described by the tensor product of its elements, or $|\psi_1\psi_2\psi_3\ldots\psi_n\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes |\psi_3\rangle \otimes \ldots |\psi_n\rangle$

For instance, if there was a system with two parts, $|\psi_1\rangle = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ and $|\psi_2\rangle = \begin{pmatrix}\gamma\\ \delta\end{pmatrix}$, then the state of the total system could be described by

$$|\psi_1\psi_2\rangle = |\psi_1\rangle \otimes |\psi_2\rangle = \begin{pmatrix}\alpha\\ \beta\end{pmatrix} \otimes \begin{pmatrix}\gamma\\ \delta\end{pmatrix} \tag{20}$$

which equates to

$$|\psi_1\psi_1\rangle = \begin{pmatrix}\alpha\gamma\\ \alpha\delta\\ \beta\gamma\\ \beta\delta\end{pmatrix} \tag{21}$$

The normality of the elements (i.e $||\psi_1\rangle| = ||\psi_2\rangle| = 1$) is preserved (i.e $||\psi_1\psi_2\rangle| = 1$)

4 What is quantum computing?

4.1 What is a qubit?

The qubit, or quantum bit, is the fundamental unit of quantum computing. As the name shows, it is the quantum analog of a classical bit. Physically, it is a two-state quantum device, e.g. spin. What this means is that we can implement a qubit using something that has two mutually exclusive states. The spin, for example, can be up, down (assuming we look at it from the z axis) or anything in between. But when we measure it in some basis, we will only get up or down (z axis), left or right (x axis), or inside or outside (y axis).

Then, why is quantum computing a thing if at the end we only measure two states as in classical computing? Well, while we are on a superposition of these two states, we will exploit quantum properties like entanglement and interference.

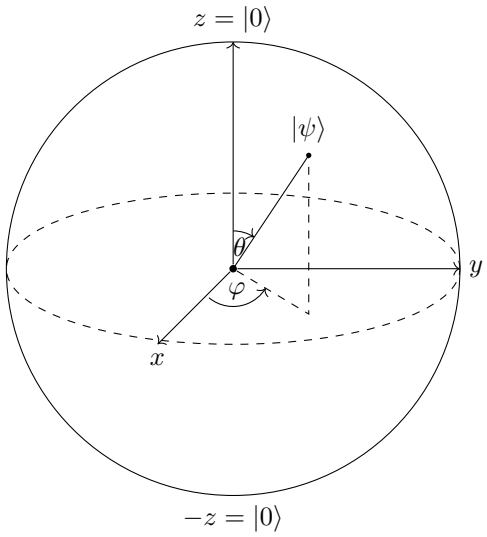
4.2 Bloch sphere

There is a way to represent a qubit geometrically, which turns out to be really helpful when doing operations with a single qubit. As mentioned before, an arbitrary state $|\psi\rangle$ is a superposition of the two mutually exclusive states, that is

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \tag{22}$$

In order to represent this in a sphere, we rewrite it as

$$|\psi\rangle = \cos\frac{\theta}{2}\,|0\rangle + \sin\frac{\theta}{2}e^{i\varphi}\,|1\rangle \tag{23}$$



In simple terms, single qubit operations (which are analogous to the application of a logic gate on a classical bit) could be described as rotations of the vector $|\psi\rangle$ around the Bloch sphere.

As mentioned above $|\psi\rangle$ could be described as a 2D linear combination of basis states $|0\rangle$ and $|1\rangle$ with amplitudes $\cos(\theta/2)$ and $\sin(\theta/2)e^{i\phi}$, respectively. It could also be described by a 3D Cartesian position vector, which, by translation of coordinate system would be equal to

$$\vec{r} = \begin{pmatrix} \cos\phi\sin\theta \\ \sin\phi\sin\theta \\ \cos\theta \end{pmatrix} = \cos\phi\sin\theta\hat{x} + \sin\phi\sin\theta\hat{y} + \cos\theta\hat{z} \tag{24}$$

This vector is known as the *Bloch Vector* and can be used to find the density matrix[7] of a qubit.

4.3 Its role

"A few weeks into researching about quantum computing, I had the privilege of meeting Prof. Scott Aaronson with a few friends of mine over a web conference. My memory is a bit faint but I remember one of the first questions asked went along the lines of: 'can you solve [some problem] with a quantum computer?' Dr. Aaronson responded the following in a manner so instinctive, it would perhaps be safe for one to assume that he gets asked similar questions on a daily basis: 'Well, yeah you can do that. Except there wouldn't be any noticable speedup. Quantum Computing doesn't change the theory of computability, meaning that everything that can be solved on a quantum computer, can also be solved on a classical computer, it's just that for some problems, a classical computer may take exponentially more time to compute it'. In my opinion, this is a great 'one-liner' which explains a lot about why quantum computing is being so heavily investigated, because by using the logic that will later on be described, one can potentially design an algorithm which highly outperforms the current classical versions in terms of efficiency.

[7] You'll learn what that means soon.

Richard Feynman noticed this in a talk he gave in 1982, where he said that the current model of computation (i.e classical computers) cannot efficiently model quantum systems in a scalable manner. He gave the following example: Suppose you have an electron which could either be spin up or spin down. To represent every possible states, one would need two bits (1 for spin up and 1 for spin down), now if there are two electrons, one would need 4 bits, and for n electrons, one would need 2^n bits... this gets exponentially less practical and is not an issue of scalability, but rather in the model itself. With a series of quantum bits (i.e bits that could be in a superposition of 1 and 0 at the same time), the number of qubits necessary to model n-electron spins would be reduced to n.

However, problems outside of modeling quantum systems are also being investigated at the moment and there is a whole field dedicated towards what can and can't be efficiently solved by a quantum computer, and whilst this field is rapidly evolving, the main applications seem to lie within three main 'pillars' of application: search[8], factoring[9] and simulation[10]. Contrary to popular interpretations of quantum computing, it will not replace current classical computers." – Adam.

4.4 Measurement

Although qubits can be in a superposition of the $|0\rangle$ and $|1\rangle$ states, we can only see (measure) them in one of these two states. This means that from a state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, we will never be able to get the exact values of α and β. More formally, this can be referred to as Holevo's theorem, which establishes an upper bound on the amount of information we can get from a quantum state. But since we are just getting started, let's just say that the act of measuring (we will explain what actually is measurement later on) a qubit takes it to the $|0\rangle$ or $|1\rangle$ state with some *probability*, which is known as the computational basis. There are other bases, but this is the most common one in quantum computation and the one we are going to focus on.

But how do we get that probability? This is determined by the amplitudes of the states, i.e., α and β. More specifically, the probability of measuring the qubit in the $|0\rangle$ state is $|\alpha|^2$ and the probability of measuring it in the $|1\rangle$ state is $|\beta|^2$. This fact allows amplitudes to interfere with one another while making sure that the probability of an individual state is always ≥ 0.

[8] See Grover's Algorithm.
[9] See Shor's Algorithm.
[10] See the Variational Quantum Eigensolver and more.