INTRODUCTION TO QUANTUM INFORMATION SCIENCE

ARTUR EKERT

About complex numbers, called probability amplitudes, that, unlike probabilities, can cancel each other out leading to quantum interference and qualitatively new ways of processing information.

Quantum mechanics, at least at some instrumental level, can be viewed as a modification of the probability theory. We replace positive numbers (probabilities) with complex numbers \( z \) (probability amplitudes) such that the squares of their absolute values, \( |z|^2 \), are interpreted as probabilities. The rules for combining amplitudes are very reminiscent of the rules for combining probabilities.

\[
\begin{align*}
    z_1 z_2 &= z_1 + z_2 \\
    z &= z_1 z_2
\end{align*}
\]

Whenever something can happen in a sequence of independent steps we multiply the amplitudes of each step. Whenever something can happen in several alternative ways we add amplitudes for each way considered separately.

That’s it! The two rules are basically all you need to manipulate amplitudes in any, no matter how complicated, physical process (we will amend the two rules later on when we touch upon the particle statistics). The two simple rules are universal and apply to any physical system, to the big and to the small, from elementary particles through atoms and molecules to white dwarfs stars. They also apply to information for information is physical; it is both represented and processed by physical means. Here we will use the two rules to explain a distinctive power of quantum information processing.

1.1. Quantum Interference. In order to see the need for quantum theory let us consider a simple experiment in which probability theory fails to give the right predictions. In a double slit experiment a particle emitted from a source \( A \) can reach detector \( B \) by taking either an upper or a lower slit, with amplitudes \( z_1 \) and \( z_2 \) respectively. We may say that the upper slit is taken with probability \( p_1 = |z_1|^2 \) and the lower slit with probability \( p_2 = |z_2|^2 \). These are two mutually exclusive events. With the two slits open, probability theory declares, the particle will reach the detector with probability

\[
\begin{align*}
    p &= |z|^2 = |z_1 + z_2|^2 \\
    &= |z_1|^2 + |z_2|^2 + z_1^\dagger z_2 + z_1 z_2^\dagger,
\end{align*}
\]

where we have expressed the amplitudes in their polar form \( z_1 = |z_1| e^{i\phi_1} \) and \( z_2 = |z_2| e^{i\phi_2} \). The appearance of the interference terms marks the departure from the classical theory of probability. The probability of any two seemingly mutually exclusive events is the sum of the probabilities of the individual events, \( p_1 + p_2 \), modified by the interference term, \( 2\sqrt{p_1 p_2} \cos(\phi_1 - \phi_2) \). Depending on the relative
phase \( \phi_1 - \phi_2 \), the interference term can be either negative (destructive interference) or positive (constructive interference), leading to either suppression or enhancement of the total probability \( p \). If we can control relative phases we can take advantage of this effect. For example, quantum computation can be viewed as a complex multiparticle quantum interference, involving many computational paths, which is designed to enhance probabilities of correct outputs and to suppress probabilities of the wrong ones.

1.2. What is wrong with the additivity axiom? You may be wondering what has happened to the axiom of additivity in probability theory, which says that if \( E_1 \) and \( E_2 \) are mutually exclusive events then the probability of the event (\( E_1 \) or \( E_2 \)) is the sum of the probabilities of the constituent events, \( E_1, E_2 \). So what is wrong with the additivity axiom? One thing that is wrong is the assumption that the processes of taking the upper or the lower slit are mutually exclusive. In reality, the two transitions both occur, simultaneously. We cannot learn about this fact from probability theory or any other a priori mathematical construct. We learn it from the best physical theory available at present, namely quantum theory. This knowledge was created as the result of conjectures, experimentation, and refutations.

1.3. Relative Phase. Note that the important quantity here is the relative phase \( \phi_1 - \phi_2 \) rather than the absolute values \( \phi_1 \) and \( \phi_2 \). This observation is not trivial at all. In simplistic terms - if a particle reacts only to the difference of the two phases, each pertaining to a separate path, then it must have, somehow, experienced the two paths. Thus we cannot say that the particle has travelled either through the upper or the lower slit, it has travelled through both. In the same way quantum computers follow, in some tangible way, all computational paths simultaneously, producing answers that depend on all these alternative calculations. Weird but this is how it is!

1.4. Decoherence. If this is how it is, so why we do not see quantum interference on a daily basis? This is because phases of probability amplitudes tend to be very fragile and may fluctuate rapidly due to spurious interactions with the environment. This has influence on the interference term; it may average to zero and we recover the classical addition of probabilities. This phenomenon is known as decoherence. We will discuss the origin of decoherence later on, for now let me only mention that it is very conspicuous in physical systems made out of many interacting components and it is chiefly responsible for our classical description of the world – without interference terms we may as well add probabilities instead of amplitudes. Decoherence, as you can imagine, is a serious impediment to building quantum computers; it deprives us of the power of quantum interference. This said, it is not all doom and gloom, there are clever ways around decoherence and later on we will touch upon quantum error corrections and quantum fault-tolerant computations.

1.5. Quantum bits called qubits. In order to see how quantum interference leads to new ways of processing information, let us first take a closer look at a basic chunk of information as traditionally conceived: one bit. From a physicist's point of view a bit is a physical system that can be prepared in one of two different states, representing two logical values: no or yes, false or true, or simply 0 or 1. For example, in digital computers, the voltage between the plates in a capacitor represents a bit of information: a charge on the capacitor denotes 1 and an absence of charge denotes 0. A bit can also be encoded using two different polarisations of light or two different states of an electron in an atom, e.g. state \( |0 \rangle \) may be chosen to be the lowest energy state, the ground state, and state \( |1 \rangle \) a higher energy state, the excited state. The bracket symbol \( | ... \rangle \), about which later, is commonly used to denote states of a quantum system. Pulses of light of appropriate frequency, duration and intensity can take the atom back and forth between the states \( |0 \rangle \) and \( |1 \rangle \) (implementing logical not). Some other pulses, say, half the duration or
intensity, will take the atom into states which have no classical analogue. Quantum theory tells us that if a bit can exist in either of two distinguishable states, it can also exist in coherent superpositions of them. These are further states in which the bit represents both values, 0 and 1, simultaneously, just as the particle that may take both the upper and the lower path in the double slit experiment. Any object in which such states can be reliably prepared, manipulated and measured is called a quantum bit or a qubit.

1.6. State vectors. Technically, the best way to describe a quantum state of a qubit is to say that the qubit is in state \( |0\rangle \) with some amplitude \( \alpha_0 \) and in state \( |1\rangle \) with some other amplitude \( \alpha_1 \). This is conveniently represented by a vector

\[
|\alpha_0 |0\rangle + \alpha_1 |1\rangle \leftrightarrow \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix}.
\]

(2)

We square the absolute values of amplitudes to get probabilities and so it is that if we choose to measure the bit value (say, energy of the atom) in such a superposition we will find either 0 (the energy of the ground state) or 1 (the energy of the excited state), with probabilities \( |\alpha_0|^2 \) and \( |\alpha_1|^2 \), respectively. As no other states are involved these are the only two possible and mutually exclusive outcomes of the measurement hence we require that \( |\alpha_0|^2 + |\alpha_1|^2 = 1 \). By analogy with the elementary geometry we can view this expression as the square of the length of the state vector and declare that all state vectors are of unit length.

1.7. Unitary evolutions. Coherent superpositions are created and modified by quantum evolutions, e.g. by atomic transitions induced by pulses of light, which redistribute amplitudes between the basis states. Mathematically, any quantum evolution \( U \) on a closed system is represented by a linear operator which sends state \( |k\rangle \) to state \( |l\rangle \) \((k,l = 0,1)\) but only with some probability amplitude \( U_{lk} \) (watch the order of the indices). We write this as

\[
|k\rangle \rightarrow \sum_l U_{lk} |l\rangle.
\]

(3)

The operator is succinctly specified by the matrix \( U_{lk} \), which tabulates all possible transition amplitudes between states \( |0\rangle \) and \( |1\rangle \). For example, pulses of light affecting the two states of our atom, may be described by a class of matrices with a real parameter \( \theta \) (related to the shape of the pulse)

\[
U(\theta) = \begin{bmatrix} U_{00} & U_{01} \\ U_{10} & U_{11} \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.
\]

(4)

The action of the operator on the basis states,

\[
|0\rangle \rightarrow \cos (\theta) |0\rangle + \sin (\theta) |1\rangle, \tag{5}
\]

\[
|1\rangle \rightarrow -\sin (\theta) |0\rangle + \cos (\theta) |1\rangle, \tag{6}
\]

can be extended, by linearity, to any other quantum state, \( |\Psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle \).

You recognise \( U(\theta) \) as a rotation matrix, thus, whenever \( \alpha_0 \) and \( \alpha_1 \) are real, the result of \( U(\theta) \) acting on \( |\Psi\rangle \) is state \( |\Psi'\rangle \), which is \( |\Psi\rangle \) rotated by the angle \( \theta \).

As state vectors are unit vectors, all admissible quantum evolutions must be represented by isometries, that is, operators that preserve the length. This essentially restricts us to unitary operators, i.e. operators that are described by unitary matrices. Matrix \( U \) is called unitary if

\[
U^\dagger U = UU^\dagger = I,
\]

(7)

where the adjoint or Hermitian conjugate \( U^\dagger \) of any matrix \( U \) with complex entries \( U_{ij} \) is obtained by taking the complex conjugate of every element in the matrix and then interchanging rows and columns (\( U^\dagger_{kl} = U^{*}_{lk} \)).

If you prefer to write the matrices explicitly, in terms of their components, then the unitarity condition above can be expressed as \( \sum_l U^*_{im} U_{lk} = \sum_l U_{mj} U^*_{kl} = \delta_{mk} \).
where δ_{mk}, known as the “Kronecker delta”, is a symbol that is defined to be zero for k ≠ m and to be one for k = m. Operator U sends state |Ψ⟩, with components α_k, into state |Ψ’⟩ = U |Ψ⟩, with components α’_l = \sum_k U_{lk} α_k,

|Ψ⟩ = \sum_k α_k |k⟩ → \sum_k α_k \left( \sum_l U_{lk} |l⟩ \right) = \sum_l \left( \sum_k U_{lk} α_k \right) |l⟩ = \sum_l α’_l |l⟩ = |Ψ’⟩. \tag{8}

Unitary operators redistribute amplitudes between the basis states in such a way that state vectors evolve into state vectors. They can take basis states into coherent superpositions and manipulate them in all kind of interesting ways. Here we will be interested in their computational abilities. As you will see in a moment, unitary operations will extend our repertoire of elementary logic gates.

1.8. Superpositions are not statistical mixtures. On the first encounter coherent superpositions are somewhat mysterious. How do we know that, say, our atom, after being treated with a carefully shaped pulse, ends up in a qualitatively new state representing both 0 and 1? Take, for example, pulse R, defined as

|0⟩ → \frac{1}{\sqrt{2}} |0⟩ + \frac{1}{\sqrt{2}} |1⟩, \tag{9}

|1⟩ → -\frac{1}{\sqrt{2}} |0⟩ + \frac{1}{\sqrt{2}} |1⟩. \tag{10}

You recognise R as U(\theta) with \theta = \pi/4, but it is a rather special pulse so I have decided to give it a separate label R. Quantum mechanics tells us that R takes the ground state of the atoms into an equally weighted superposition \((|0⟩ + |1⟩)/\sqrt{2}\). If we choose to measure the energy of the atom in such a superposition, we will find either the energy of the ground state or the energy of the excited state, with equal probability. Perhaps the the atom is in a statistical mixture, that is, either the ground or the excited state with probability 1/2 each? It is not!

We can demonstrate this by sending a second R–pulse to the atom. If it were merely the case that there was a 50% chance that the atom was in the ground state and a 50% chance that it was in the excited state, then the second pulse would make no difference; we should still find the atom either in the ground or the excited state with equal probability. The R–pulse in this case would act as a random switch between the two states, and the application of two successive random switches is equivalent to a random switch. Any explanation, which assumes that, while in between the two pulses, the atom is exactly one of the two states, leads to the conclusion that the atom should end up either in the ground or the excited state, with equal probability. But experiments show otherwise. The atom, initially in state |0⟩, after being exposed to two identical R–pulses, is always found in state |1⟩! The inescapable conclusion is that the atom, while in a superposition, must, in some sense, be both in state |0⟩ and |1⟩ at the same time.

Quantum calculations explain the process.

The diagram above shows all the relevant transitions, and the corresponding amplitudes, induced by the two R–pulses. You can, for example, calculate amplitudes alongs the two paths connecting the input state |0⟩ with the output state |0⟩ and see how they cancel each other out, preventing the atom from ending in state |0⟩.
You can also step through the diagram and follow the evolution of the state vector (shown at the bottom of the diagram). Finally, if you prefer to work with column vectors and matrices, you can write the two consecutive application of $R$ to state $|0\rangle$ as

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (11)$$

Following a well established convention the formula above should be read from right to left. Whichever way you do the calculations the conclusion is the same: coherent superpositions are very different from statistical mixtures.

1.9. **Impossible logic.** From the computational point of view $R$ implements a classically impossible logical operation. Consider a following task: design a logic gate that operates on a single bit and such that when it is followed by another, identical, logic gate the output is always the negation of the input. Let us call this logic gate the square root of $\text{NOT} \ (\sqrt{\text{NOT}})$. A simple check - such as an attempt to construct a truth table - should persuade you that there is no such operation in logic. It may seem reasonable to argue that since there is no such operation in logic, $\sqrt{\text{NOT}}$ is impossible. But think again. A $R$–pulse implements this “impossible” logic gate. The result of sending a $R$–pulse on an atom in state $|0\rangle$ (or $|1\rangle$) is an equally weighted coherent superposition of $|0\rangle$ and $|1\rangle$, but two $R$–pulses one after another, acting independently, implement the logical operation $\text{NOT}$. Quantum theory explains the behaviour of the square root of $\text{NOT}$, hence, reassured by the physical experiments that corroborate this theory, logicians are now entitled to propose a new logical operation $\sqrt{\text{NOT}}$. Why? Because a faithful physical model for it exists in nature. This purely quantum operation has no counterpart in classical logic and is one of the elementary operations of a quantum computer.

1.10. **Quantum circuits.** Needless to say, a carefully shaped pulse of light is just one way of implementing the square root of $\text{NOT}$ – there are many others. Atoms, trapped ions, molecules, nuclear spins and many other quantum objects can be prepared in two distinct states, labelled as $|0\rangle$ and $|1\rangle$, and manipulated so that transition amplitudes between these states are the same as in matrix $R$. There is no need to learn about physics behind these diverse technologies if all you want is to understand the square root of $\text{NOT}$. You may conveniently forget about any specific experimental realisation (hardware) and draw a diagram which represents them all,

$$|0\rangle \quad R \quad R \quad |1\rangle$$

This quantum network, or quantum circuit, diagram should be read from left to right. The horizontal line represents a qubit that is inertly carried from one operation to another. Various icons on its path represent elementary logical operations, that is, logic gates. Any sequence of unitary operators, say $U$ followed by $V$,

$$|\Psi\rangle \quad U \quad V \quad |\Psi'\rangle$$

is a unitary operator which can be represented, as you can easily guess, by the matrix product $VU$ (note the order in which we multiply the matrices),

$$|\Psi\rangle \quad VU \quad |\Psi'\rangle$$

Indeed, if $U$ and $V$ are defined by

$$|k\rangle \rightarrow \sum_l U_{lk} |l\rangle, \quad |l\rangle \rightarrow \sum_m V_{ml} |m\rangle,$$  

$$R = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$
then
\[ |k\rangle \rightarrow \sum_l U_{lk} \left( \sum_m V_{ml} |m\rangle \right) = \sum_l \left( \sum_m V_{ml} U_{lk} \right) |m\rangle. \] (13)

If you want to hone your quantum intuition think about it this way. The amplitude that input \( |k\rangle \) evolves to \( |m\rangle \) via a specific intermediate state \( |l\rangle \) is given by \( V_{ml} U_{lk} \) (evolutions are independent so the amplitudes are multiplied). This done we have to sum over all possible values of \( l \) (the transition can occur in several mutually exclusive ways so the amplitudes are added) to obtain \( \sum_l V_{ml} U_{lk} \). As you can see, the bookkeeping tools are useful, the matrix multiplication in one swoop takes care of multiplication and addition of amplitudes corresponding to different alternatives.

**1.11.** Addition of probability amplitudes, squaring the absolute value of the resulting sum and discussing the role of the interference terms is the essence of most quantum calculations. The rest is just a set of convenient mathematical tools developed for the purpose of good bookkeeping of contributing amplitudes. However, truth be told, these bookkeeping tools are often indispensable. It is much easier to manipulate amplitudes once they are tabulated into matrices and vectors. Consider, for example, the following circuit

\[ |\Psi\rangle \xrightarrow{U(\theta_1)} \xrightarrow{U(\theta_2)} |\Psi'\rangle \]

Given input \( |\Psi\rangle \), you want to know the probabilities of 0 and 1 at the output. You can, of course, figure this out in a pedestrian way, by multiplying and adding amplitudes along “paths” connecting the basis states \( |0\rangle \) and \( |1\rangle \) at the input with those at the output, but it is easier to solve the problem using vectors and matrices. A bit of linear algebra combined with some geometric intuition (you should recognize \( U(\theta) \) as a rotation matrix) tells you that the state vector \( |\Psi\rangle \) undergoes two consecutive rotations by angles \( \theta_1 \) and \( \theta_2 \), respectively. The net effect of the circuit is rotation by \( \theta_1 + \theta_2 \). The two basis states, \( |0\rangle \) and \( |1\rangle \), evolve as
\[ |0\rangle \rightarrow \cos(\theta_1 + \theta_2) |0\rangle + \sin(\theta_1 + \theta_2) |1\rangle, \]
\[ |1\rangle \rightarrow -\sin(\theta_1 + \theta_2) |0\rangle + \cos(\theta_1 + \theta_2) |1\rangle, \]
which, by linearity, gives you evolution of any input state \( a_0 |0\rangle + a_1 |1\rangle \). Not all physically admissible operations on quantum states have such a simple geometric interpretation but many of them do, in particular operations on a qubit. In any case, from now on we will switch to vectors and operators but do remember that behind this linear algebra there are still the two simple rules for the multiplication and addition of the amplitudes.

**1.12.** There are essentially four components of the mathematical structure of quantum theory. We need to know how to represent (1) states, (2) physically admissible operations or quantum evolutions, (3) measurements, and (4) composite systems. I will describe them in stages, starting with a very simple case of a single qubit, and working towards the most general case. So far we know that

- **Quantum states are described by unit vectors with complex components.** Any two state vectors which differ by a multiplicative phase factor \( e^{i\phi} \), e.g. \( |\Psi\rangle \) and \( e^{i\phi} |\Psi\rangle \), represent the same quantum state. Global phase factors can be ignored. In contrast, all relative phase factors are very important; the two vectors \( |\Psi\rangle = a_0 |0\rangle + a_1 |1\rangle \) and \( |\Psi\rangle = a_0 |0\rangle + e^{i\phi}a_1 |1\rangle \) represent two different states.
- **States evolve in time.** Any admissible physical evolution \( U \) of a closed system is represented by a unitary operator.
- **We have tacitly assumed that measurements are associated with the choice of the basis vectors.** For example, if we choose the computational basis, \( |0\rangle \) and \( |1\rangle \), the measurement of the bit value correspond to checking if the
qubit is in state $|0\rangle$ or in state $|1\rangle$. The two states are perfectly distinguishable so the outcomes of the measurement are mutually exclusive. If the state of the qubit just prior to the bit value measurement is $a_0 |0\rangle + a_1 |1\rangle$, then the measurement will find the qubit in state $|0\rangle$ with probability $|a_0|^2$ or in state $|1\rangle$ with probability $|a_1|^2$.

Quantum measurement is a bit of a philosophical issue, which we shall revisit few times later on. Right now let me only comment on the infamous “collapse” of the state vector. As you can see from the diagram above, after the measurement $M$, if the result is 0 the post-measurement state of the qubit is no longer $a_0 |0\rangle + a_1 |1\rangle$, but $|0\rangle$, and if the result is 1 the post-measurement state is $|1\rangle$. If the measurement is immediately repeated, then according to this rule the same outcome is obtained again, with probability one. The original state $a_0 |0\rangle + a_1 |1\rangle$ is irretrievably lost. This sudden change of state, to $|0\rangle$ with probability $|a_0|^2$ and to $|1\rangle$ with probability $|a_1|^2$, due to a measurement is often called a “collapse” or a “reduction” of the state vector. It looks like there are two distinct ways for a quantum state to change. On the one hand we have unitary transition matrices and on the other hand we have an abrupt change during the measurement process. Let me stress – the measurement process is not governed by any different laws of physics. A measurement is a physical process and can be explained in terms of unitary evolutions, without any “collapse”. It is usually a complicated process in which one complex system (a measuring apparatus or an observer) interacts and gets correlated with a physical system being measured. However, in order not to get lost in discussing this issue at this stage let us accept a “collapse” as a convenient mathematical shortcut. Just remember that the status of this shortcut in the formulation of quantum mechanics is still very much debated.

1.13. Quantum interference revisited. Let me conclude this lecture with a typical quantum interference experiment on a single qubit. Any such experiment can be represented by a sequence of three elementary operations (quantum logic gates). The most common sequence is the Hadamard gate, followed by a phase shift gate, and followed by the Hadamard gate. We represent it graphically as

Remember, quantum network diagrams are read from left to right. The horizontal line represents a quantum wire, which inertly carries a qubit from one quantum operation to another. The wire may describe translation in space, e.g. atoms travelling through cavities, or translation in time, e.g. between operations performed on a trapped ion. If we want to signify that a particular unitary evolution is to be enacted on our qubit, then we put a box with a symbol describing this unitary operation along the quantum wire. In particular, the Hadamard gate and the phase shift gate are defined as

$$
|0\rangle \rightarrow \frac{1}{\sqrt{2}} |0\rangle + \frac{1}{\sqrt{2}} |1\rangle \\
|1\rangle \rightarrow \frac{1}{\sqrt{2}} |0\rangle - \frac{1}{\sqrt{2}} |1\rangle \\
|0\rangle \rightarrow |0\rangle \\
|1\rangle \rightarrow e^{i\theta} |1\rangle
$$
In terms of matrices (in the computational basis)

\[
H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad P_\phi = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix}
\]

There are different approaches to analysing this quantum interference circuit (or any other circuit). For example, you can multiply the three matrices \( H P_\phi H \); the product is a unitary matrix that describes the action of the whole circuit; it gives transition amplitudes between states \(|0\rangle\) and \(|1\rangle\) at the input and the output.

\[
\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = e^{i\phi/2} \begin{bmatrix} \cos\phi/2 & -i\sin\phi/2 \\ -i\sin\phi/2 & \cos\phi/2 \end{bmatrix}.
\]

Thus the action of the circuit is described by the operator

\[
|0\rangle \rightarrow \cos \frac{\phi}{2} |0\rangle - i \sin \frac{\phi}{2} |1\rangle, \quad (16)
\]

\[
|1\rangle \rightarrow -i \sin \frac{\phi}{2} |0\rangle + \cos \frac{\phi}{2} |1\rangle. \quad (17)
\]

Given that our input state is almost always \(|0\rangle\) it is sometimes much easier to step through the execution of this circuit and follow the evolving state. The interference circuit effects the following sequence of transformations

\[
|0\rangle \xrightarrow{H} \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \xrightarrow{P_\phi} \frac{1}{\sqrt{2}} (|0\rangle + e^{i\phi} |1\rangle) \xrightarrow{H} \cos \frac{\phi}{2} |0\rangle - i \sin \frac{\phi}{2} |1\rangle.
\]

The Hadamard gate prepares an equally weighted superposition of \(|0\rangle\) and \(|1\rangle\) and it closes the interference by bringing the interfering paths together. The phase shift \(\phi\) effectively controls the evolution and determines the output. The probabilities of finding the qubit in state \(|0\rangle\) or \(|1\rangle\) at the output are, respectively,

\[
Pr(0) = \cos^2 \frac{\phi}{2}, \quad Pr(1) = \sin^2 \frac{\phi}{2}. \quad (19)
\]

This simple quantum process contains, in a nutshell, the essential ingredients of quantum computation. The sequence, Hadamard - phase shift - Hadamard, will appear over and over again. It reflects a natural progression of quantum computation: first we prepare different computational paths, then we evaluate a function which effectively introduces phase shifts into different computational paths, then we bring the computational paths together at the output.

**Notes**

(1) Back in 1926 Max Born simply postulated the connection between amplitudes and probabilities. However, it is worth pointing out, that he did not get it quite right on his first approach. In the original paper proposing the probability interpretation of the state vector (wavefunction) \(^1\) he wrote:

...If one translates this result into terms of particles only one interpretation is possible. \(\Theta_{\eta,\tau,m}(\alpha, \beta, \gamma)\) [the wavefunction for the particular problem he is considering] gives the probability\(^*\) for the electron arriving from the \(z\) direction to be thrown out into the direction designated by the angles \(\alpha, \beta, \gamma\)...

\(^*\) Addition in proof: More careful considerations show that the probability is proportional to the square of the quantity \(\Theta_{\eta,\tau,m}(\alpha, \beta, \gamma)\).

\(^1\)Max Born, Zur Quantenmechanik der Stoßvorgänge, Zeitschrift für Physik, 37, 863–867 (1926).
(2) All physical evolutions \( \mathcal{U} \) are represented by operators that map unit state vectors \( \sum_k \alpha_k |k\rangle \) into unit state vectors \( \sum_l \alpha'_l |l\rangle \), hence
\[
1 = \sum_l |\alpha'_l|^2 = \sum_l \left( \sum_m \mathcal{U}_{lm} \alpha_m \right) \left( \sum_n \mathcal{U}_{ln} \alpha'_n \right)^* = \sum_{m,n} \left( \sum_l \mathcal{U}_{lm}^* \mathcal{U}_{ln}^\dagger \right) \alpha_m \alpha'_n. \tag{20}
\]
This equality must hold for all admissible values of \( \alpha_m \) and \( \alpha'_n \). This is possible only if
\[
\sum_l \mathcal{U}_{lm} \mathcal{U}_{ln}^\dagger = \sum_l \mathcal{U}_{lm}^\dagger \mathcal{U}_{ln} = \delta_{mn},
\]
i.e. the operators must be unitary.

Exercises

(1) A quantum computer starts calculations in some initial state then follows \( n \) different computational paths and finally reaches the final output. Computational paths are followed with probability amplitudes \( \frac{1}{\sqrt{n}} e^{i k \phi} \), where \( \phi \) is a fixed angle \( 0 < \phi < 2\pi \) and \( k = 0, 1, \ldots, n - 1 \). Show that the probability of reaching the final state is
\[
\frac{1}{n} \left| \frac{1 - e^{i n \phi}}{1 - e^{i \phi}} \right|^2 = \frac{\sin^2(\frac{n \phi}{2})}{n \sin^2(\frac{\phi}{2})},
\]
for \( 0 < \phi < 2\pi \) and 1 for \( \phi = 0 \).

(2) The four matrices below describe all possible computations on a single classical bit.
\[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix} \quad \begin{bmatrix}
0 & 1 \\
1 & 0
\end{bmatrix} \quad \begin{bmatrix}
1 & 1 \\
0 & 0
\end{bmatrix} \quad \begin{bmatrix}
0 & 0 \\
1 & 1
\end{bmatrix}
\]

Which of the four operations can be also viewed as quantum operations?

(3) A beamsplitter is a cube of glass, made out of two triangular glass prisms which are glued together at their base. This simple device reflects half the light that impinges upon it, while allowing the remaining half to pass through unaffected. We shall view it as a quantum device which operates on the two beams of light, here labelled as \( |0\rangle \) and \( |1\rangle \). The two beams intersect at the beamsplitter and a photon from either beam can be transmitted (and remain in the same beam) with amplitude \( \frac{1}{\sqrt{2}} \), or reflected to the other beam with amplitude \( i/\sqrt{2} \). We shall assume that there is only one photon involved and summarise the action of the beamsplitter by tabulating the amplitudes of transitions between the two beams.
\[
B = \begin{bmatrix}
B_{00} & B_{01} \\
B_{10} & B_{11}
\end{bmatrix} = \begin{bmatrix}
\frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\
\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}}
\end{bmatrix}.
\]
Let us aim a single photon at the beam-splitter using beam \( |0\rangle \). The photon doesn’t split in two: we can place photo-detectors wherever we like in the apparatus, fire in a photon, and verify that if any of the photo-detectors registers a hit, none of the others do. In particular, if we place a photo-detector behind the beam-splitter in each of the two exit beams, the photon is detected with equal probability at either detector, no matter whether the photon was initially fired from input port \( |0\rangle \) or \( |1\rangle \). You may conclude that the beam-splitter is just a random switch, a device that directs incoming photons to randomly chosen output ports. How would you convince someone that this is not the case? Show, by performing appropriate calculations, that the beamsplitter \( B \) implements the \( \sqrt{\text{NOT}} \) gate. How would you demonstrate this in a lab?
(4) A Mach-Zehnder interferometer is composed of two beamsplitters and two mirrors that redirect the two output beams from the first beamsplitter to the input of the second beamsplitter. Two photodetectors, labeled 0 and 1, are placed at the output of the second beamsplitter. The two optical paths in between the two beamsplitter are of the same length. A single photon enters the first beamsplitter in beam $|1\rangle$. Which detector will register the photon? Suppose one of the two beams in between the beamsplitters is blocked by an absorbing screen. Which detector will register the photon now?

(5) The Quantum Bomb Tester. You have been drafted by the government to help in the demining effort in a former war-zone. In particular, retreating forces have left very sensitive bombs in some of the sealed rooms. The bombs are configured such that if even one photon of light is absorbed by the fuse (i.e. if someone looks into the room), the bomb will go off. Each room has an input and output port which can be hooked up to external devices. An empty room will let light go from the input to the output ports unaffected, whilst a room with a bomb will explode if light is shone into the input port and the bomb absorbs even just one photon. Your task is to find a way of determining whether a room has a bomb in it without blowing it up, so that specialised (limited and expensive) equipment can be devoted to defusing that particular room. Design a scheme that allows you – at least part of the time – to decide whether a room has a bomb in it without blowing it up. If you iterate the procedure, what is its overall success rate for the detection of a bomb without blowing it up?

(6) Assume that the two beam splitters in the interferometer are different. A beamsplitter which reflects incoming light with probability $r$ and transmits with probability $1 - r$ is described by the matrix

$$B = \begin{bmatrix} \cos \theta & i \sin \theta \\ i \sin \theta & \cos \theta \end{bmatrix},$$

where $r$ is parametrised as $\cos^2 \theta$. Let the probability of reflection in the first beamsplitter be $r$ and in the second one $1 - r$. Would the new setup improve the overall success rate of the detection of a bomb without blowing it up?

(7) Consider a setup of $N$ identical beam splitters (probability of reflection $r$) where $N$ is an even number. Can you design a scheme (i.e. choose a suitable $r$) such that the success rate for detecting a bomb without blowing it up approaches 100%?

(8) A quantum interference experiments is described by the following circuit

$$|0\rangle \xrightarrow{B} |\Psi\rangle$$

where the gate $B$ is a symmetric beamsplitter defined in the previous exercises and the central gate is a phase shift gate which introduces a relative phase $\phi$. Step through the execution of this circuit and follow the evolving input state. What is the state of the qubit at the output, here represented as $|\Psi\rangle$? Suppose you can control the input of the circuit and measure the output, but you do not know the phase shift $\phi$ introduced by the phase gate. You are promised, however, that $\phi$ is either 0 or $\pi$. You can run the circuit only once to find out which of the two phases was chosen. How will you do it?

(9) Show, by expressing matrices in terms of their components or otherwise, that $$(UV)^\dagger = V^\dagger U^\dagger$$ and that the product of two unitary matrices is another unitary matrix. The set of all unitary $N \times N$ matrices with the matrix multiplication forms a non-Abelian group called $U(N)$.

---

2This is a slightly modified version of a bomb testing problem described by Avshalom Elitzur and Lev Vaidman in Quantum-mechanical interaction-free measurement, Found. Phys. 47, 987-997 (1993).
(10) Any unitary $N \times N$ matrix has $N^2$ complex entries, or, if we were to view each complex number as a pair of real numbers, $2N^2$ real entries. They are not independent from each other because of the unitarity condition. Show that any unitary $N \times N$ matrix can be parametrised by $N^2$ independent real parameters.

(11) Check that any matrix of the form,

$$e^{i\phi} \begin{bmatrix} \cos \theta e^{i\alpha} & -\sin \theta e^{i\beta} \\ \sin \theta e^{-i\beta} & \cos \theta e^{-i\alpha} \end{bmatrix}$$

where $\alpha, \beta, \theta$ and $\phi$ are independent real parameters, is unitary. In fact, any $2 \times 2$ unitary matrix can be written in this form.

(12) The exponent of matrix $A$ is defined as

$$e^A = \mathbb{1} + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots = \sum_{n=0}^{\infty} \frac{(A)^n}{n!}$$

Show that if $H$ is self-adjoint (Hermitian), that is $H = H^\dagger$, then $U = e^{iHt}$ is unitary for any real $t$. Many quantum evolutions are expressed in this way. This is because matrix $H$, known as the Hamiltonian, is related to energies, which are measurable physical quantities, and $t$ stands for time.

(13) Show that for any real $\alpha$ and for any $A$ such that $A^2 = \mathbb{1}$

$$e^{i\alpha A} = \cos \alpha \mathbb{1} + i \sin \alpha A.$$  

(14) A qubit (spin one-half particle) initially in state $|0\rangle$ (spin up) is placed in a uniform magnetic field. The interaction between the field and the qubit is described by the Hamiltonian

$$H = \omega \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},$$

where $\omega$ is proportional to the strength of the field. What is the state of the qubit after time $t = \pi/4\omega$?