QUANTUM INTERFERENCE

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.

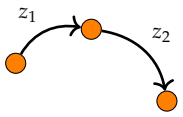
The classical theory of computation usually does not refer to physics. Pioneers such as Alan Turing, Alonzo Church, Emil Post and Kurt Gödel managed to capture the correct classical theory by intuition alone and, as a result, it is often falsely assumed that its foundations are self-evident and purely abstract. They are not! The concepts of information and computation can be properly formulated only in the context of a physical theory – information is stored, transmitted and processed always by *physical* means. Computers are physical objects and computation is a physical process. Indeed, any computation, classical or quantum, can be viewed in terms of physical experiments, which produce *outputs* that depend on initial preparations called *inputs*. Once we abandon the classical view of computation as a purely logical notion independent of the laws of physics it becomes clear that whenever we improve our knowledge about physical reality, we may also gain new means of computation. Thus, from this perspective, it is not very surprising that the discovery of quantum mechanics in particular has changed our understanding of the nature of computation. In order to explain what makes quantum computers so different from their classical counterparts, we begin with the rudiments of quantum theory.

Computation is a physical process!
Computation is a physical process!
Computation is a physical process!
Computation is...
Computation...

1.1. **Two basic rules.** Quantum theory, at least at some instrumental level, can be viewed as a modification of 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:

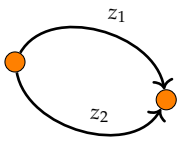
Born's Rule

The correspondence between probability amplitude z and probability $p = |z|^2$ is known as **Born's Rule**.



$z = z_1 z_2$

Whenever something can happen in a sequence of independent steps, we multiply the amplitudes of each step.



$z = z_1 + z_2$

Whenever something can happen in several alternative ways, we add the amplitudes for each separate way.

That's it! These two rules are basically all you need to manipulate amplitudes in any physical process, no matter how complicated. (we will amend the two rules later on when we touch upon the particle statistics). They are universal and apply to any physical system, from elementary particles through atoms and molecules to white dwarfs stars. They also apply to information for, as we have already emphasised, information is physical. The two rules look deceptively simple but, as you will see in a moment, their consequences are anything but trivial.

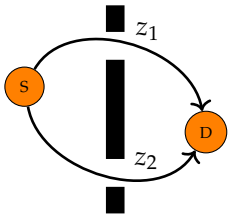
1.2. **Quantum interference and the failure of probability theory.** Modern mathematical probability theory is based on three axioms, proposed by Andrey Nikolaevich Kolmogorov (1903–1987) in his monograph with the impressive German title *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Foundations of Probability Theory). The Kolmogorov axioms are simple and intuitive. Once you identify all elementary outcomes, or events, you may then assign probabilities to them. Probability is a number

between 0 and 1, and an event which is certain has probability 1. These are the first two axioms. There is one more. The probability of any event can be calculated using a deceptively simple rule - the additivity axiom:

> Whenever an event can occur in several mutually exclusive ways, the probability for the event is the sum of the probabilities for each way considered separately.

Obvious, isn't it? So obvious, in fact, that probability theory was accepted as a mathematical framework theory, a language that can be used to describe actual physical phenomena. Physics should be able to identify elementary events and assign numerical probabilities to them. Once this is done we may revert to mathematical formalism of probability theory. The Kolmogorov axioms will take care of the mathematical consistency and will guide us whenever there is a need to calculate probabilities of more complex events. This is a very sensible approach apart from the fact that it does not always work! Today, we know that probability theory, as ubiquitous as it is, fails to describe many common quantum phenomena. 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 S can reach detector D by taking two different paths, e.g. through an upper or a lower slit. After sufficiently many repetitions of this experiment we can evaluate the frequency of clicks in the detector D and show that it is inconsistent with the predictions based on the probability theory. Let us use the quantum approach to show how the discrepancy arises.

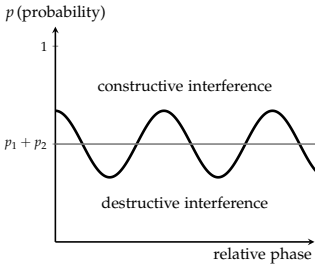
The particle emitted from a source S can reach detector D by taking two different paths, e.g. through an upper or a lower slit, with amplitudes z_1 and z_2 respectively. We may say that the upper slit is taken with probability $p_1 = |z_1|^2$ and the lower slit with probability $p_2 = |z_2|^2$. These are two mutually exclusive events. With the two slits open, probability theory declares (the additivity axiom) that the particle should reach the detector with probability $p_1 + p_2 = |z_1|^2 + |z_2|^2$. Wrong! Following the "quantum rules", first we add the amplitudes and then we square the absolute value of the sum to get the probability. Thus, the particle will reach the detector with probability



$z = z_1 + z_2$

$$\begin{aligned} p = |z|^2 = |z_1 + z_2|^2 &= |z_1|^2 + |z_2|^2 + z_1^\star z_2 + z_1 z_2^\star, \\ &= p_1 + p_2 + |z_1||z_2|(e^{i(\varphi_2-\varphi_1)} + e^{-i(\varphi_2-\varphi_1)}), \\ &= p_1 + p_2 + 2\sqrt{p_1 p_2}\,\cos(\varphi_2 - \varphi_1), \\ &= p_1 + p_2 + \text{interference terms}, \end{aligned} \tag{1}$$

where we have expressed the amplitudes in their polar forms $z_1 = |z_1|e^{i\varphi_1}$ and $z_2 = |z_2|e^{i\varphi_2}$. 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, $2\sqrt{p_1 p_2}\,\cos(\varphi_2 - \varphi_1)$. Depending on the relative phase $\varphi_2 - \varphi_1$, the interference term can be either negative (destructive interference) or positive (constructive interference), leading to either suppression or enhancement of the total probability p.

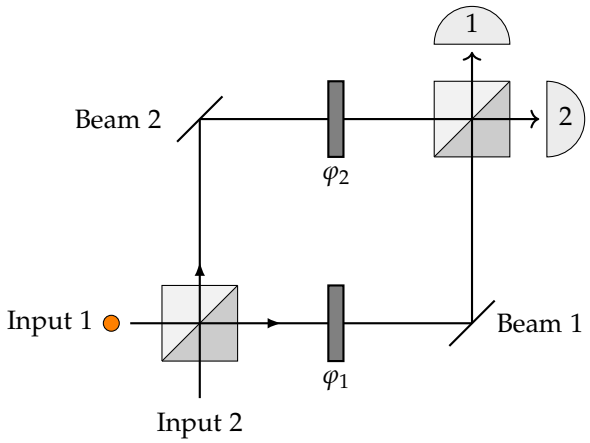


The algebra is simple, our focus is on the physical interpretation. Firstly, note that the important quantity here is the relative phase $\varphi_2 - \varphi_1$ rather than the absolute values φ_1 and φ_2. This observation is not trivial at all. If a particle reacts only to the difference of the two phases, each pertaining to a separate path, then it must have, somehow, experienced the two paths, right? 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! Secondly, what has happened to the axiom of additivity in probability theory, 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, as we have just mentioned, the two transitions *both occur*, simultaneously. However, we cannot learn this from probability theory, or any other a priori mathematical construct. There is no fundamental reason why Nature should conform to the additivity axion. We find out how nature works by making intelligent guesses, running experiments, checking what happens and formulating physical theories. If our guess disagrees with experiments it is wrong, so we try another intelligent guess, and another, etc. Right now quantum theory is the best guess we have; it offers good explanations and predictions that have not been falsified by any of the existing experiments. This said, be assured that one day quantum theory will be falsified and we will have to start guessing again.

According to the philosopher Karl Popper (1902–1994) a theory is genuinely scientific only if it is possible, in principle, to establish that it is false. Genuinely scientific theories are never finally confirmed because no matter how many confirming observations have been made observations that are inconsistent with the empirical predictions of the theory are always possible.

1.2.1. *Example:* One of the simplest quantum devices in which quantum interference can be controlled is a Mach-Zehnder interferometer.

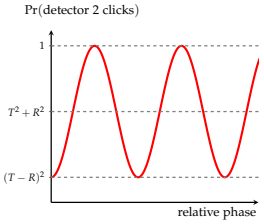


It consists of two beam-splitters (the square boxes, bottom left and top right) and two slivers of glass of different thickness which are inserted into each of the optical paths connecting the two beam-splitters. The slivers are usually referred to as "phase shifters" and their thicknesses, φ_1 and φ_2, are measured in units of the photon's wavelength multiplied by 2π. The two inputs ports of the interferometer are labelled as 1 and 2, and each of the two output ports, also labelled as 1 and 2, terminates in a photodetector. A photon (the orange dot) impinges on the first beam-splitter from one of the two input ports, here input 1, and begins its journey towards one of the two photodetectors. Let U_{ij} denotes the probability amplitude that the photon initially in input port $j = 1, 2$ ends up in detector $i = 1, 2$ (here, and in the following, index i should not be confused with the imaginary unit). At each of the two beam-splitters the photon is transmitted with the probability amplitude $\sqrt{T}$ and reflected with the probability amplitude $i\sqrt{R}$, $(R + T = 1)$, and the two phase shifters modify the amplitudes by phase factors, $e^{i\varphi_1}$ and $e^{i\varphi_2}$, respectively. In quantum theory we almost always start with the amplitudes and once we have a full expression for the amplitude of a given outcome we square its absolute value to get the corresponding probability. For example, let us calculate U_{11}. We notice that there are two alternative ways for the photon in the input port 1 to end up in the output port 1. It can take the lower path, through the phase shifter φ_1, or the upper path, through the phase shifter φ_2. The lower path implies two consecutive transmissions at the beamsplitters and the phase factor $e^{i\varphi_1}$, whereas the upper path implies two consecutive reflections and the phase factor $e^{i\varphi_2}$. Once the photon ends in the output port 1 there is no way of knowing which path was taken, thus we add the amplitudes pertaining to each path. The resulting amplitude is

$$U_{11} = \sqrt{T}e^{i\varphi_1}\sqrt{T} + i\sqrt{R}e^{i\varphi_2}i\sqrt{R},$$

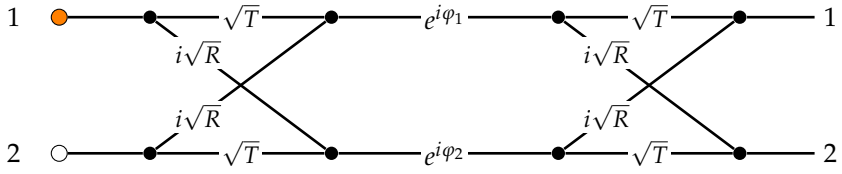
and the corresponding probability $P_{11} = |U_{11}|^2$ reads

$$\begin{aligned} P_{11} &= |\sqrt{T}e^{i\varphi_1}\sqrt{T} + i\sqrt{R}e^{i\varphi_2}i\sqrt{R}|^2 = |Te^{i\varphi_1} - Re^{i\varphi_2}|^2 \\ &= T^2 + R^2 - 2TR\cos(\varphi_2 - \varphi_1). \end{aligned}$$



The "classical" part of this expression, $T^2 + R^2$, basically says that the photon undergoes either two consecutive transmissions with probability T^2, or two consecutive reflections with probability R^2. The probability of being transmitted through any phase shifter is always 1, hence the phase shifters play no role in the classical description of this process. But the classical description is not correct, as the experiments show, and hence the interference term $2TR\cos(\varphi_2 - \varphi_1)$, in which the phase shifters play the essential role. Depending on the *relative* phase $\varphi = \varphi_2 - \varphi_1$ the probability that the detector 1 "clicks" can vary from $(T-R)^2$, for $\varphi = 0$, to 1 for $\varphi = \pi$.

If we do not care about the experimental details, we can represent the action of the Mach-Zehnder interferometer in terms of a diagram:



first beamsplitter phase shifts second beamsplitter

Here, we can follow, from left to right, the multiple different paths that a photon can take in between specific input and output ports. The amplitude for any given path is just the product of the segments, while the overall amplitude is the sum of the amplitudes for the many different paths. You can, for example, see that the probability amplitude U_{21} is given by

$$U_{21} = \sqrt{T}e^{i\varphi_1}i\sqrt{R} + i\sqrt{R}e^{i\varphi_2}\sqrt{T},$$

and the corresponding probability

$$\begin{aligned} P_{21} &= |\sqrt{T}e^{i\varphi_1}i\sqrt{R} + i\sqrt{R}e^{i\varphi_2}\sqrt{T}|^2 \\ &= 2RT + 2RT\cos(\varphi_2 - \varphi_1). \end{aligned}$$

Again, the first term is of "classical" origin and represents probabilities corresponding to each path, one reflection followed by one transmission plus one transmission followed by one reflection, that is, $RT + TR = 2RT$. The second term is the interference term. Clearly, the photon entering port 1 will end up in one of the two detectors, hence,

$$P_{11} + P_{21} = R^2 + 2RT + T^2 = (T+R)^2 = 1.$$

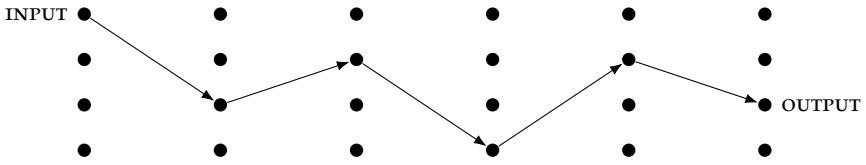
The action of the interferometer is fully described by the four probability amplitudes U_{ij} $(i, j = 1, 2)$. The most popular instance of a Mach-Zehnder interferometer involves only symmetric beamsplitters $(R = T = \frac{1}{2})$ and is fully described by the matrix

In general, any isolated quantum device, including a quantum computer, can be described by a matrix of probability amplitudes U_{ij} that input j generates output i. Watch the order of indices.

$$U = \begin{bmatrix} -\sin\varphi/2 & \cos\varphi/2 \\ \cos\varphi/2 & \sin\varphi/2 \end{bmatrix},$$

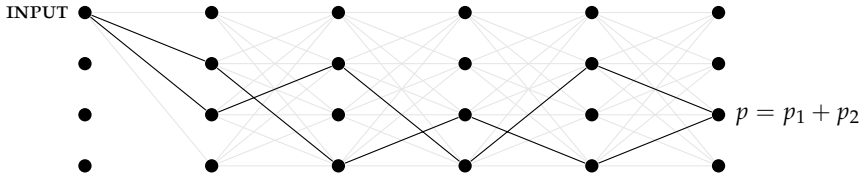
where $\varphi = \varphi_2 - \varphi_1$. In fact, when you do all the calculations you obtain $ie^{i\frac{\varphi_1+\varphi_2}{2}}U$ rather than U, but the global phase factor $ie^{i\frac{\varphi_1+\varphi_2}{2}}$ is common to all the amplitudes in the matrix and as such it does not contribute to the resulting probabilities (why?).

1.3. **Computation.** Think about computation as a physical process that evolves a prescribed initial configuration of a computing machine, called INPUT, into some final configuration, called OUTPUT. We shall refer to the configurations as *states*. The diagram below shows five consecutive computational steps performed on four distinct states.



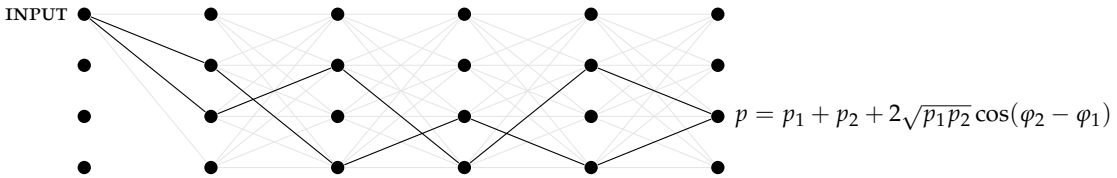
DETERMINISTIC

That computation was *deterministic* – every time you run it with the same input, you get the same output. Such a computation does not have to be deterministic – we can augment a computing machine by allowing it "to toss an unbiased coin" and to choose its steps randomly. It can then be viewed as a directed, tree-like graph where each node corresponds to a state of the machine, and each edge represents one step of the computation.



PROBABILISTIC

The computation starts from some initial state (INPUT) and it subsequently branches into other nodes representing states reachable with non-zero probability from the initial state. The probability of a particular final state (OUTPUT) being reached is equal to the sum of the probabilities along all mutually exclusive paths which connect the initial state with that particular state. The diagram above shows only two computational paths, but, in general, there could be many more of them (here, up to 256) paths contributing to the final probability. Quantum computation can be represented by a similar graph:



QUANTUM

We associate with each edge in the graph the probability *amplitude* that the computation follows that edge. The probability amplitude of a particular path to be followed is the product of amplitudes pertaining to transitions in each step. The probability amplitude of a particular final state being reached is equal to the sum of the amplitudes along all mutually exclusive paths which connect the initial state with that particular state,

$$z = \sum_{\text{all paths k}} z_k.$$

The resulting probability, as we have just seen, is the sum of the probabilities pertaining to each computational path p_k modified by the interference terms,

$$p = |z|^2 = \sum_{k,j} z_j^\star z_k = \sum_k p_k + \sum_{k \neq j} \sqrt{p_k p_j}\cos(\varphi_k - \varphi_j).$$

Quantum computation can be viewed as a complex multiparticle quantum interference involving many computational paths through a computing device. The art of quantum computation is to shape quantum interference, through a sequence of computational steps, enhancing probabilities of correct outputs and suppressing probabilities of the wrong ones.

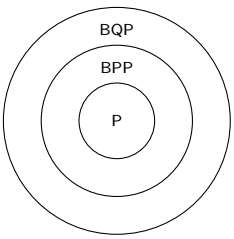
1.4. **Computational Complexity.** Is there a compelling reason why we should care about quantum computation? It may sound like an extravagant way to compute something that can be computed anyway. Indeed, your standard laptop, given enough time and memory, can simulate pretty much any physical process. In principle, it can also simulate any quantum interference and compute everything that quantum computers can compute. The snag is, this simulation, in general, is very inefficient. And efficiency does matter, especially if you have to wait more than the age of the Universe for your laptop to stop and deliver an answer!

The age of the Universe is currently estimated at 13.772 billion years

In order to solve a particular problem, computers (classical or quantum) follow a precise set of instructions — an algorithm. Computer scientists quantify the efficiency of an algorithm according to how rapidly its running time, or the use of memory, increases when it is given ever larger inputs to work on. An algorithm is said to be *efficient* if the number of elementary operations taken to execute it increases no faster than a polynomial function of the size of the input. We take the input size to be the total number of binary digits (bits) needed to specify the input. For example, using the algorithm taught in elementary school, one can multiply two n digit numbers in a time that grows like the number of digits squared, n^2. In contrast, the fastest-known method for the reverse operation—factoring an n-digit integer into prime numbers—takes a time that grows exponentially, roughly as 2^n. That is considered inefficient.

Notice that the technological progress alone, such as increasing the speed of classical computers, will never turn an inefficient algorithm (exponential scaling) into an efficient one (polynomial scaling). Why?

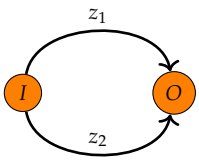
The class of problems that can be solved by a deterministic computer in polynomial time is represented by the capital letter P, for *polynomial* time. The class of problems that can be solved in polynomial time by a probabilistic computer is called BPP, for *bounded-error probabilistic polynomial* time. It is clear that BPP contains P, since a deterministic computation is a special case of a probabilistic computation in which we never consult the source of randomness. When we run a probabilistic, aka randomised, computation many times on the same input, we will not get the same answer every time, but the computation is useful if the probability of getting the right answer is high enough. Finally, the complexity class BQP, for *bounded-error quantum polynomial*, is the class of problems that can be solved in polynomial time by a quantum computer. Since a quantum computer can easily generate random bits and simulate a probabilistic classical computer, BQP certainly contains the class BPP. Here we are interested in problems that are in BQP but not known to be in BPP. The most popular example of such a problem is factoring. A quantum algorithm, discovered by Peter Shor in 1994, can factor n-digit numbers in a number of steps that grows only as n^2. Since the intractability of factorisation underpins the security of many methods of encryption Shor's algorithm was soon hailed as the first 'killer application' for quantum computation, something very useful that only a quantum computer could do. Since then, the hunt has been on for interesting things for quantum computers to do, and at the same time, for the scientific and technological advances that could allow us to build quantum computers.



It must be stressed that not all quantum algorithms are so efficient, in fact many are no faster than their classical counterparts. Which particular problems will lend themselves to quantum speed-ups is an open question.

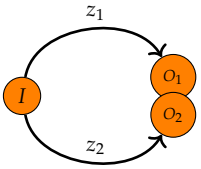
1.5. **Quantum decoherence.** In principle we know how to build quantum computers out of simple components, such as qubits (quantum bits) and quantum logic gates. We will describe these components in detail in the subsequent lectures. However, as the number of quantum components increases, we quickly run into some serious practical problems. The more interacting components are involved, the harder it tends to be to engineer the interactions that would cause the necessary gate operations and induce quantum interference without introducing errors. The more components there are,

the more likely it is that quantum interference will spread outside the quantum computer, to the surrounding environment, thus spoiling the computation. This process is called decoherence. In order to understand the essence of decoherence consider the following two different scenarios in which a quantum computer is prepared in some input state I and generates output O



$p = |z_1 + z_2|^2$

The computer is isolated and quantum computation does not affect the environment. The computer and the environment evolve independently from each other and, as a result, the environment does not hold any physical record of how the computer reached output O. In this case we add the amplitudes for each of the two alternative computational paths.



$p = |z_1|^2 + |z_2|^2$

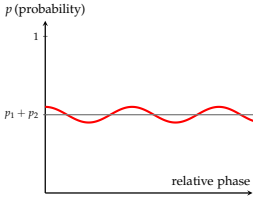
Quantum computation affects the environment. The environment now holds a physical record of how the computer reached output O, which results in two final states of the composed system (computer + environment) which we denote O_1 and O_2. We add the probabilities for each of the two alternative computational paths.

The addition of probability amplitudes, rather than probabilities, applies to physical system which are completely isolated. When quantum computation affects the environment we have to include the environment in our analysis for it now takes part in the computation, i.e. our isolated system is now composed of a quantum computer and its environment. Depending on which computational path was taken the environment may end up in two distinct states. The computer itself may show output O but when we include the environment we have not one but two outputs, O_1 and O_2, denoting, respectively, "computer shows output O and the environment knows that path 1 was taken" and "computer shows output O and the environment knows that path 2 was taken". There are no alternative ways of reaching O_1 or O_2 hence there is no interference and the corresponding probabilities read $p_1 = |z_1|^2$ for O_1, and $p_2 = |z_2|^2$ for O_2. The probability that the computer shows output O, regardless the state of the environment, is the sum of of the two probabilities $p = p_1 + p_2$. We have lost the interference term and with it any advantages of quantum computation. In the presence of decoherence the interference formula Eq.(1) is modified and reads,

$$p = p_1 + p_2 + 2v\sqrt{p_1 p_2}\,\cos(\varphi_2 - \varphi_1),$$

where the parameter v, called the "visibility" of the interference pattern, ranges from 0 (the environment can perfectly distinguish between the two paths, total decoherence, no interference) to 1 (the environment cannot distinguish between the two paths, no decoherence, full interference), with the values in between corresponding to partial decoherence. We shall derive this formula later on and you will see that v quantifies the degree of distinguishability between O_1 and O_2. The more environment knows about which path was taken the less interference we see.

Decoherence suppresses quantum interference.



Decoherence is chiefly responsible for our classical description of the world – without interference terms we may as well add probabilities instead of amplitudes. While decoherence is a serious impediment to building quantum computers, depriving us of the power of quantum interference, it is not all doom and gloom; there are clever ways around decoherence such as the quantum error correction and fault-tolerant methods we will meet later.

1.6. **Outlook.** When the physics of computation was first investigated, starting in the 1960s, one of the main motivations was a fear that quantum-mechanical effects might place fundamental bounds on the accuracy with which physical objects could render the properties of the abstract entities, such as logical variables and operations, that

appear in the theory of computation. But it turned out that quantum mechanics imposes no significant limits but does break through some of those that classical physics imposed. The quantum world has a richness and intricacy that allows new practical technologies, and new kinds of knowledge. In this course we will merely scratch the surface of the rapidly developing field of quantum computation. We will concentrate mostly on the fundamental issues and skip many experimental details. However, it should be mentioned that quantum computing is a serious possibility for future generations of computing devices. At present it is not clear how and when fully-fledged quantum computers will eventually be built; but notwithstanding this, the quantum theory of computation already plays a much more fundamental role in the scheme of things than its classical predecessor did. I believe that anyone who seeks a fundamental understanding of either physics, computation or logic must incorporate its new insights into his world view.

Notes & Exercises

(1) I always found it an interesting coincidence that the two basic ingredients of modern quantum theory, namely probability and complex numbers, were discovered by the same person, an extraordinary man of many talents, a gambling scholar by the name of Girolamo Cardano (1501–1576).

(2) Complex numbers have many applications in physics, however, not until the advent of quantum theory was their ubiquitous and fundamental role in the description of the actual physical world so evident. Even today, their profound link with probabilities appears to be a rather mysterious connection. Mathematically speaking, the set of complex numbers is a field. This is an important algebraic structure used in almost all branches of mathematics. You do not have to know much about algebraic fields to follow these lectures, but still, you should know the basics. Look them up.

(3) (a) The sets of rational and real numbers are all fields, but the set of integers is not. Why?
(b) What does it mean that the field of complex numbers is algebraically closed?
(c) Evaluate each of the following quantities $1+e^{-i\pi}$, $|1+i|$, $(1+i)^{42}$, $\sqrt{i}$, 2^i and i^i.
(d) Here is a simple proof that $+1=-1$,

$$1=\sqrt{1}=\sqrt{(-1)(-1)}=\sqrt{-1}\sqrt{-1}=i^2=-1$$

What is wrong with it?

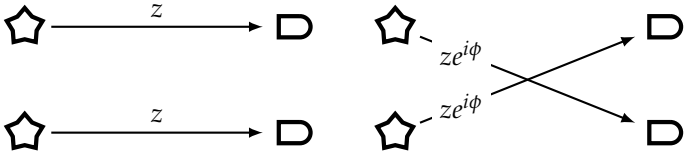
(4) A quantum computer starts calculations in some initial state, then follows n different computational paths which lead to the final output. The computational paths are followed with probability amplitudes $\frac{1}{\sqrt{n}}e^{ik\varphi}$, where φ is a fixed angle $0<\varphi<2\pi$ and $k=0,1,...n-1$. Show that the probability of generating the output is

$$\frac{1}{n}\left|\frac{1-e^{in\varphi}}{1-e^{i\varphi}}\right|^2=\frac{1}{n}\frac{\sin^2(n\frac{\varphi}{2})}{\sin^2(\frac{\varphi}{2})}.$$

for $0<\varphi<2\pi$ and 1 for $\varphi=0$. Plot the probability as a function of φ.

$1+z+z^2+\ldots+z^n=\frac{1-z^{n+1}}{1-z}$

(5) Imagine two distant stars that emit *identical* photons. If you point a single detector towards them you will register a click every now and then, but you never know which star the photon came from. Now prepare two detectors and point them towards the stars. Assume the photons arrive with the probability amplitudes specified below.

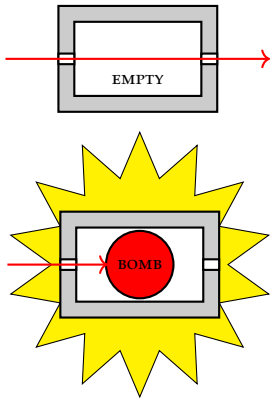


Every now and then you will register a coincidence – the two detectors will fire.
(a) Calculate the probability of a coincidence.
(b) Now, assume that $z\approx\frac{1}{r}e^{i\frac{2r\pi}{\lambda}}$, where r is the distance between detectors and the stars. How can we use this to measure r?

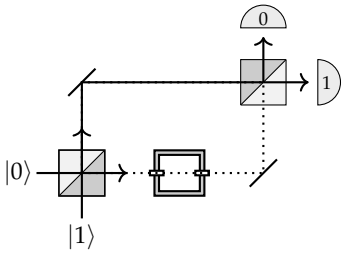
(6) **Quantum Bomb Tester** You have been drafted by the government to help in the demining effort in a former war-zone. In particular, retreating forces have left very sensitive bombs in some of the sealed rooms. The bombs are configured such that if even one photon of light is absorbed by the fuse (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

This 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).

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. You would like to know with certainty whether a particular room had a bomb in it.



Hint: Consider the setup where the input and output ports are hooked up in one of the arms of a Mach-Zehnder interferometer.



(a) To start with, consider the setup (see the margin) where the input and output ports are hooked up in the lower arm of a Mach-Zehnder interfereometer.
   (i) Assume an empty room. Send a photon to input port $|0\rangle$. Which detector, at the output port, will register the photon?
   (ii) Now assume that the room does contain a bomb. Again, send a photon to input port $|0\rangle$. Which detector will register the photon and with which probability?
   (iii) 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?

(b) Assume that the two beam splitters in the interferometer are different. Say the first beamsplitter reflects incoming light with probability r and transmits with probability $t = 1 - r$ and the second one transmits with probability r and reflects with probability t. Would the new setup improve the overall success rate of the detection of a bomb without blowing it up?

(c) There exists a scheme, involving many beamsplitters and something called "quantum Zeno effect", such that the success rate for detecting a bomb without blowing it up approaches 100%. Try to work it out or find a solution on internet.

(7) A quantum machine has N perfectly distinguishable configurations. What is the maximum number of computational paths connecting a specific input with a specific output after k steps of the machine? Suppose you are using your laptop to add together amplitudes pertaining to each of the paths. As k and N increase you may need more time and more memory to complete the task. How does the execution time and the memory requirements grow with k and N? Will you need more time or more memory or both?

(8) The classical theory of computation is essentially the theory of the universal Turing machine - the most popular mathematical model of classical computation. Its significance relies on the fact that given a large but finite amount of time the universal Turing machine is capable of any computation that can be done by any modern classical digital computer, no matter how powerful. The concept of Turing machines may be modified to incorporate quantum computation, but we will not follow this path. It is much easier to explain the essence of quantum computation talking about quantum logic gates and quantum Boolean networks or circuits. The two approached are computationally equivalent, even though certain theoretical concepts, e.g. in computational complexity, are easier to formulate precisely using the Turing machine model. The main advantage of quantum circuits is that they relate far more directly to proposed experimental realisations of quantum computation.

(9) In computational complexity the basic distinction is between polynomial versus exponential algorithms. Polynomial growth is good and exponential growth is bad, especially if you have to pay for it. There is an old story about the legendary inventor of chess who asked the Persian king to be paid only by a grain of cereal, doubled on each of the 64 squares of a chess board. The king placed one grain of rice on the first square, two on the second, four on the third, and he was supposed to keep on doubling until the board was full. The

One light year (the distance that light travels through a vacuum in one year) is 9.4607×10^{15}m.

last square would then have $2^{63} = 9,223,372,036,854,775,808$ grains of rice, more than has been ever harvested on planet Earth, to which we must add the grains of all previous squares, making the total number about twice as large. If we placed that many grains in an unbroken line we would reach the nearest star Alpha Centauri, our closest celestial neighbour beyond the solar system, about 4.4 light-years away. The moral of the story: if whatever you do requires an exponential use of resources you are in trouble.

(10) In order to make qualitative distinctions between how different functions grow we will often use the asymptotic big-O notation. For example, suppose an algorithm running on input of size n takes $an^2 + bn + c$ elementary steps, for some positive constants a, b and c. These constants depend mainly on the details of the implementation and the choice of elementary steps. What we really care about is that for large n the whole expression is dominated by its quadratic term. We then say that the running time of this algorithm grows as n^2, and we write it as $O(n^2)$, ignoring the less significant terms and the constant coefficients. More precisely, let $f(n)$ and $g(n)$ be functions from positive integers to positive reals. You may think of $f(n)$ and $g(n)$ as the running times of two algorithms on inputs of size n. We say $f = O(g)$, which means that f grows no faster than g, if there is a constant $c > 0$ such that $f(n) \leq cg(n)$ for all sufficiently large values of n. Essentially, $f = O(g)$ is a very loose analog of $f \leq g$. In addition to the big-O notation, computer scientists often use Ω for lower bounds: $f = \Omega(g)$ means $g = O(f)$. Again, this is a very loose analog of $f \geq g$.

$f = O(g)$ is pronounced as "f is big-oh of g".

(11) (a) When we say that $f(n) = O(\log n)$, why don't we have to specify the base of the logarithm?
(b) Let $f(n) = 5n^3 + 1000n + 50$, is $f(n) = O(n^3)$ or $O(n^4)$ or both?
(c) Which of the following statements are true?
(i) $n^k = O(2^n)$ for any constant k
(ii) $n! = O(n^n)$
(iii) if $f_1 = O(g)$ and $f_2 = O(g)$ then $f_1 + f_2 = O(g)$

(12) There exists a randomised algorithm which tests whether a given number N is prime. The algorithm always returns YES when N is prime and the probability it returns YES when N is not prime is ϵ, which not greater than half (independently, each time you run the algorithm). You run this algorithm (for the same N) r times and each time the algorithm returns YES. What is the probability that N is not prime?

Primality used to be given as the classic example of a problem in BPP but not P. However, in 2002 a deterministic polynomial time test for primality was proposed by Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. Thus, since 2002, primality has been in P.

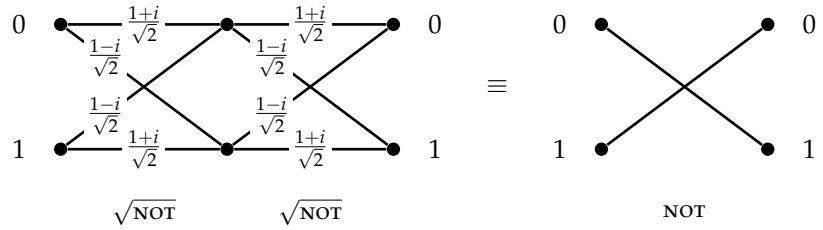
(13) Suppose a randomised algorithm solves a decision problem, returning YES or NO answers. It gets the answer wrong with a probability not greater than $\frac{1}{2} - \delta$, where $\delta > 0$ is a constant.
(a) If we perform this computation r times, how many possible sequences of outcomes are there?
(b) Give a bound on the probability of any particular sequence with w wrong answers.
(c) If we look at the set of r outcomes, we will determine the final outcome by performing a majority vote. This can only go wrong if $w > r/2$. Give an upper bound on the probability of any single sequence that would lead us to the wrong conclusion.
(d) Using the bound $1 - x \leq e^{-x}$, conclude that the probability of our coming to the wrong conclusion is upper bounded by $e^{-2r\delta^2}$.

Chernoff Bound

This result is known as the Chernoff bound. If we are willing to accept a probability of error no larger than ϵ, then it suffices to run the computation a number of times $r = O(\log 1/\epsilon)$.

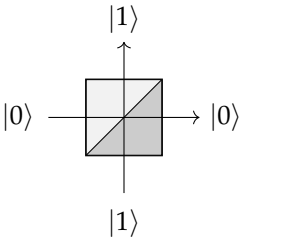
Appendix: Physics against logic
Explained with a beamsplitter

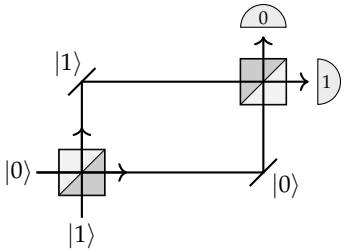
Consider the following task: design a logic gate that operates on a single bit such that when it is followed by another, identical, logic gate the output is always the negation of the input. Let us call this logic gate the square root of NOT ($\sqrt{\text{NOT}}$). A simple check, such as an attempt to 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.

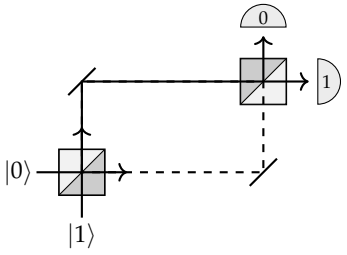


Here is a simple computation, two identical computational steps performed on two states labelled as 0 and 1, i.e. on one bit. An interplay of constructive and destructive interference makes some transitions impossible and the result is the logical NOT. Thus, quantum theory declares, the square root of NOT is possible. And it does exist! Experimental physicists routinely construct this and many other "impossible" gates in their laboratories. In fact, the square root of NOT can be as simple as a symmetric beam-splitter.

A symmetric beam-splitter is a cube of glass which reflects half the light that impinges upon it, while allowing the remaining half to pass through unaffected. For our purposes it can be viewed as a device which has two input and two output ports which we label as $|0\rangle$ and $|1\rangle$. When we aim a single photon at such a beam-splitter using one of the input ports, we notice that 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 possible 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 $|1\rangle$. It may seem obvious that at the very least, the photon is *either* in the transmitted beam $|0\rangle$ *or* in the reflected beam $|1\rangle$ during any one run of this experiment. Thus we may be tempted to think of the beam-splitter as a random binary switch which, with equal probability, transforms any binary input into one of the two possible outputs. However, that is not necessarily the case. Let us introduce a second beam-splitter and place two normal mirrors so that both paths intersect at the second beam-splitter (see diagrams in the margin).

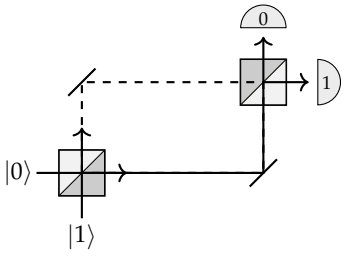






Two consecutive reflections give amplitude $\frac{i}{\sqrt{2}}\frac{i}{\sqrt{2}} = -\frac{1}{2}$

Now, the axiom of additivity in probability theory, says that whenever something can happen in several alternative ways we add probabilities for each way considered separately. We might argue that a photon fired into the input port $|0\rangle$ can reach the detector 0 in two *mutually exclusive* ways: either by two consecutive reflections or by two consecutive transmissions. Each reflection happens with probability $1/2$ and each transmission happens with probability $1/2$ thus the total probability of reaching detector 0 is a sum of the probability of the two consecutive reflections ($1/2 \times 1/2 = 1/4$) and the probability of the two consecutive transmissions ($1/2 \times 1/2 = 1/4$) which gives probability $1/2$. This makes perfect sense – a random switch followed by a random switch should give nothing else but a random switch. However, if we set up such an experiment, that is not what happens! When the optical paths between the two beam-splitters are the same, the photon fired from input port $|0\rangle$ *always* strikes detector 1 and *never* detector 0 (and the photon fired from input port $|1\rangle$ *always* strikes detector 0 and *never* detector 1). Thus a beam-splitter acts as the square root of NOT gate.

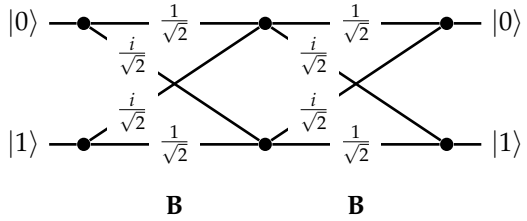


The action of the beamsplitter – in fact, the action of any quantum device – can be described by tabulating the amplitudes of transitions between its input and output ports.

$$B = \begin{bmatrix} B_{00} & B_{01} \\ B_{10} & B_{11} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

There is no reason why probability theory or any other a priori mathematical construct should make any meaningful statements about outcomes of physical experiments.

The matrix element B_{lk}, where $k, l = 0, 1$, represents the amplitude of transition from input $|k\rangle$ to output $|l\rangle$ (watch the order of indices). Each reflection (entries B_{01} and B_{10}) happens with amplitude $i/\sqrt{2}$ and each transmission (entries B_{00} and B_{11}) happens with amplitude $1/\sqrt{2}$. Thus the total amplitude that a photon fired from input port $|0\rangle$ will reach detector 0 is the sum of the amplitude of the two consecutive reflections ($i/\sqrt{2} \times i/\sqrt{2} = -1/2$) and the amplitude of the two consecutive transmissions ($1/\sqrt{2} \times 1/\sqrt{2} = 1/2$) which gives the total amplitude 0. The resulting probability is then zero. Unlike probabilities, amplitudes can cancel out each other out. We can now go on and calculate the amplitude that the photon will reach detector 1. In this case we will get i, which gives probability 1. We can then switch to input $|1\rangle$ and repeat our calculations. All possible paths and associated amplitudes are shown in the diagram below.



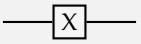
However, instead of going through all the paths in this diagram and linking specific inputs to specific outputs, we can simply multiply the transition matrices,

$$BB = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = iX.$$

As you can see, the matrix multiplication in one swoop takes care of multiplication and addition of amplitudes corresponding to different alternatives. You can now inform you colleagues logicians that they are now entitled to propose a new logical operation $\sqrt{\text{NOT}}$ for a faithful physical model for it exists in nature!

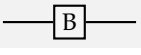
LOGICAL NOT

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$



BEAM SPLITTER

$$B = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$$



Note that gate B is not the same square root of NOT as the one described in the first diagram in this section. There are infinitely many ways of implementing this "impossible" logical operation.