

Lectures on Quantum Computation

David Deutsch

Last modified 2020/12/03 at 23:38:22 GMT

Contents

1	The Qubit	3
2	Interference	27
3	Measurement	47
4	The Schrödinger Picture	63
5	A Quantum Algorithm	77
6	Grover's Search Algorithm	91

© DAVID DEUTSCH. Permission to make this transcript available was graciously granted by the author who remains the exclusive holder/owner of all relevant copyrights to everything presented herein. Material was prepared by Ben Webb and Adrian German between 12/25/19 and 11/29/20.



Table 0.1: Our intention was/is to promote the videos (sponsored by Quiproc Network of Excellence and Hewlett-Packard) by capturing their essence (content plus style) in a format that is friendlier to browsing and as a reference. (Lectures available at http://www.quiprocone.org/Protected/DD_lectures.htm)

Chapter 1

The Qubit

Quantum theory and Einstein's general theory of relativity are the two great fundamental theories of contemporary physics. Between them they provide the conceptual framework and the mathematical language in which we express all other theories in physics, and they provide the basic principles to which all known laws of nature conform. The deeper and more general a theory is, the further away it tends to be from every day experience, so it's not surprising that our deepest theories involve some very unfamiliar counter-intuitive phenomena not least of which are the phenomena of quantum computation, the subject of these lectures. But in this first lecture I won't describe any phenomena, I'll give an overview of how quantum theory describes the world and physical processes. And then, I will introduce you to the simplest of all quantum systems, which is also the centerpiece of quantum computation, the qubit (or quantum bit¹).

01:59

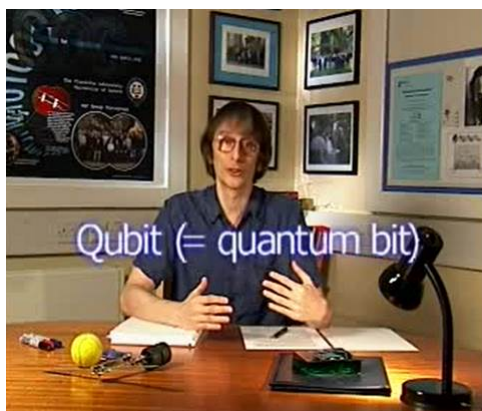


Figure 1.1: [...] and then I will introduce you to the simplest of all quantum systems which is also the centerpiece of quantum computation: the qubit [...]

¹Qubit (= quantum bit).

Quantum computation isn't something that existing microchips do even if they rely on quantum mechanical phenomena. Today's computers² don't count as quantum computers because their repertoire of computations is still the same as that of the abstract Universal Turing Machine which was the prototype of all classical computers devised by the mathematician Alan Turing in 1936. Quantum computers will be capable of new modes of computation—which classical computers are incapable of even in principle.

02:38



Figure 1.2: Quantum computation isn't something that existing microchips do.

The equations that predict the outcomes of quantum mechanical experiments are all uncontroversial, but what the underlying explanation is, what's happening physically to bring about those outcomes, is still very controversial.

02:58



Figure 1.3: The various rival explanations are called interpretations [...]

The various rival explanations are called *interpretations* of quantum theory.

²The year is 2006, same year that the book by Dasgupta, Papadimitriou and Vazirani is finished and published by McGraw Hill Education.

The one I'll be using sounds like science-fiction—at first—it was proposed by Hugh Everett (in 1957) and it's called the Many-Universes Interpretation. It says that the Universe, the space we see around us with all the galaxies and stars and matter doesn't constitute the whole [of] reality. In fact, it's just a small slice of physical reality as a whole, where there are among other things vast numbers of coexisting universes similar to ours. If you are new to this idea

03:41



Figure 1.4: In fact, it's just a small slice of physical reality as a whole [...].

and skeptical—that's good—but I ask you to go along with it for the purpose of learning the theory. In the course of that I expect to persuade you that this very fruitful way of understanding quantum theory makes sense, in fact that it's the only way that makes sense. Anyway, if there are many universes we need a new word to denote physical reality as a whole. And that word is, instead of Universe: Multiverse.

04:18



Figure 1.5: [...] we need a new word to denote physical reality as a whole.

Our Universe then is to some approximation a self-contained entity within the Multiverse. This approximation is called Classical Physics (pre-Quantum

Physics), and in computational theory it's called Classical Computation (that is to say, Turing-type computation). As we'll see Turing's theory is a complete model for computations that happen within individual universes. The quantum theory of computation is the full theory which has the Multiverse as its arena.

04:54

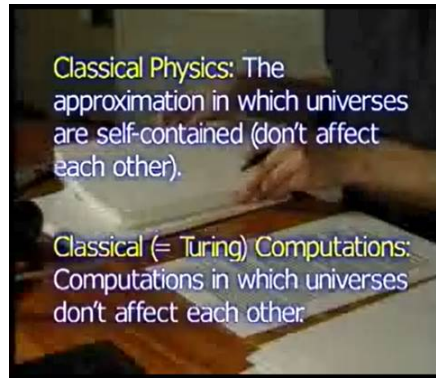


Figure 1.6: Turing's theory is a complete model for computations that happen within individual universes. The quantum theory [...] is the full theory.

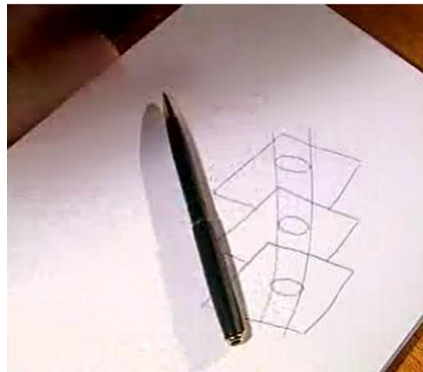


Figure 1.7: [I]n reality physical objects aren't confined to just one universe [...] every object in one universe has counterparts in a range of other universes [that] can behave differently from each other and can affect each other.

In many physical phenomena, especially on the microscopic scales, the classical approximation just breaks down because in reality physical objects aren't confined to just one universe, they have a certain extension across the multiverse or, to put that in another way, every object in one universe has counterparts in a range of other universes and these counterparts can behave differently from each other and they can affect each other. Such effects are called *quan-*

tum interference. They constitute our evidence of the existence of a reality beyond our universe. Under certain circumstances they permit fundamentally new modes of information processing which we call: quantum computation and quantum communication.

05:57

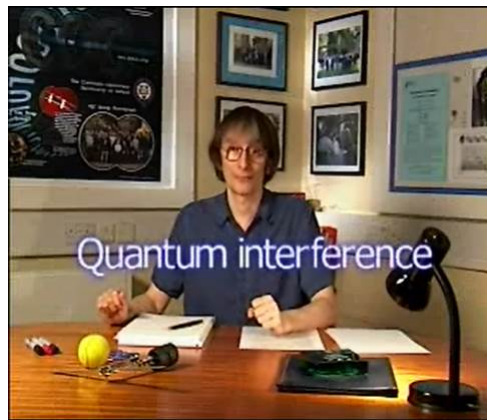


Figure 1.8: [O]ur evidence of the existence of a reality beyond our universe.

The theory of computation was originally conceived of as a branch of pure mathematics. It has been incorporated into physics via the quantum theory of computation which is now *the* theory of computation. The previous abstract theory developed by Turing and others lives on only as the classical approximation though as I said that's good enough to describe what all computers currently on the market do.

06:39

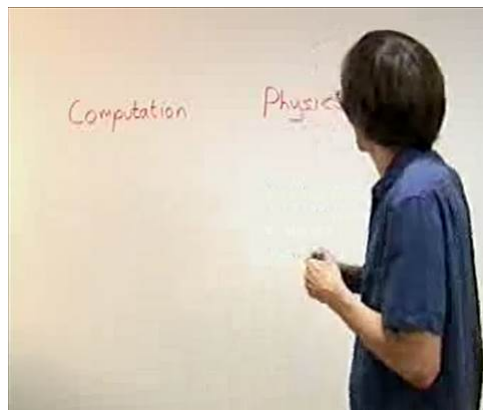


Figure 1.9: With the benefit of hindsight we can see that [...]

With the benefit of hindsight we can see that the theory of computation always did have a lot in common conceptually with Physics. When a **computer**

performs a **computation** it starts with some **input** information which it modifies according to definite **rules** which are characteristic of the hardware of that computer. So the **output** depends on the input and on the rules by which the computer operates. A **physical system** is roughly speaking some part of Nature that could in principle be experimented on, such as this.

07:50



Figure 1.10: A physical system is roughly speaking [...] such as this.

Physical systems undergo **motion** or change in other words we can pick any two times and say that the system has changed from an **initial state** to a **final state** between those two times according to **laws of motion**—which are the laws of physics as specialized to that system.

08:27

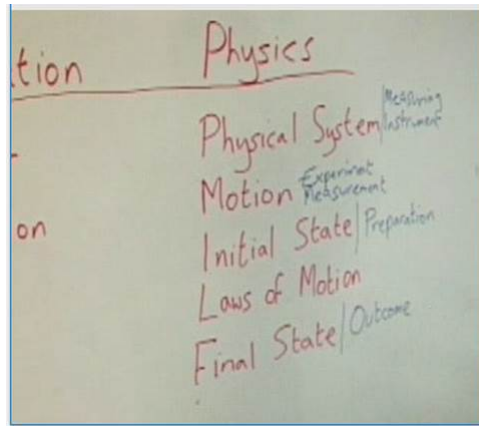


Figure 1.11: The system and the measuring instrument interact [via] the laws of Physics which makes the measuring instrument display the outcome [...].

Experiment and measurement are just forms of motion. They involve both a system we're experimenting on and some measuring instrument or observer.

We find the system in an initial state or we prepare it in some way and we prepare a measuring instrument. The system and the measuring instrument then interact according to the laws of Physics which makes the measuring instrument display the outcome of the experiment. You can see that everything in the left-hand column here is a special case of the corresponding thing on the right. But you can also think of that the other way around: any final state contains information about the system's initial state and about what it happened to it since. So the motion of any physical system, because it obeys definite laws, can be regarded as information processing.

09:38

Computation	Physics
Computer	Physical System ^{Measuring Instrument}
Computation	Motion ^{Experiment Measurement}
Input	Initial State ^{Preparation}
Rules	Laws of Motion
Output	Final State ^{Outcome}

Figure 1.12: [T]he motion of any physical system [is] information processing.

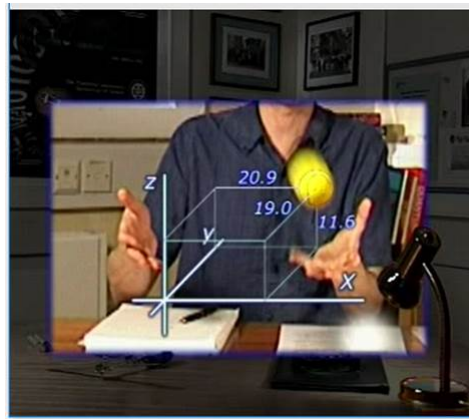


Figure 1.13: The central idea of computer science is [...] a degree of freedom.

In this first lecture I'm going to describe the simplest of all quantum physical systems: the qubit (short for quantum bit). To do that I first have to explain how physical systems are described in quantum theory. The central idea of computer science is that of a computational variable or memory location, a place where information can be stored at one time and perhaps processed and later retrieved. Classical physics has a very similar concept to that, a degree of

freedom. The degrees of freedom of a classical system are the real numbers that specify its configuration, for instance, a point particle would have three degrees of freedom, because three real numbers are necessary to specify its position in space at a given instant. In quantum physics the closest thing to a degree of freedom is an observable.

10:55

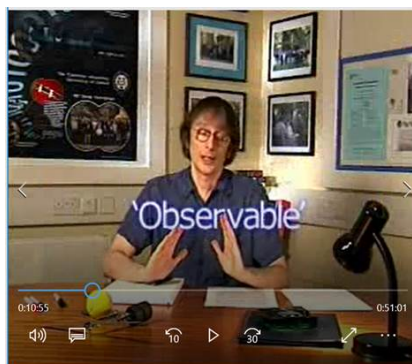


Figure 1.14: In quantum physics the closest thing to [that] is an observable.

Just as in classical physics, any attribute of a physical system that could in principle be prepared with a value that could in principle be measured is a quantum observable—but a quantum observable can't be summed up as a mere number like the value of a degree of freedom at a given time there's a lot more to it than that and it takes a while to get to grips with this concept. The word observable might even be misleading since what we see of a physical object is part of a larger object extending across many universes. A quantum observable refers to what we see and its counterparts in other universes and it contains information about the structure of the multiversal object.

10:55



Figure 1.15: So [this] angle is in fact a quantum observable.

So, the angle between these two rods, which would be deemed a degree of freedom if this system were described in classical physics is in fact a quantum

observable. This protractor is a measuring instrument that can be used either to prepare that observable with a given value or to measure its value to some degree of accuracy. Let me call that observable $\hat{\Theta}(t)$. I will always use this hat symbol (or caret) for observables to stress that they are not numbers³.

12:27

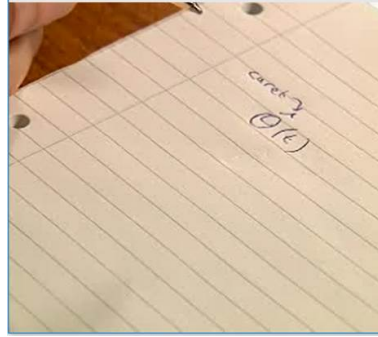


Figure 1.16: Let me call that observable: $\hat{\Theta}(t)$

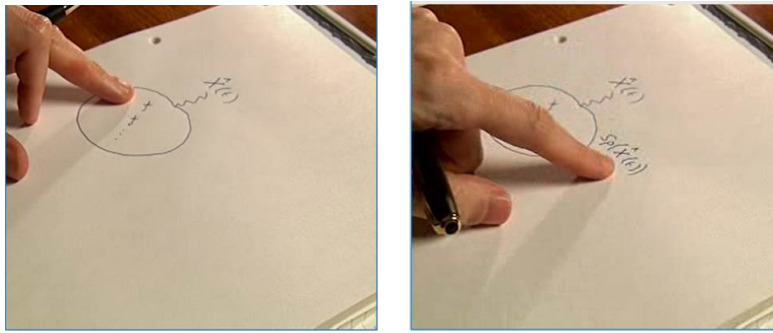


Table 1.1: In quantum theory each observable \hat{X} is associated with a set of [...] distinct labels called the spectrum of \hat{X} and [...] written like this: $\text{Sp}(\hat{X}(t))$.

For each memory location in a computer there's a finite set of possible values that can be stored in it. For instance, one bit can hold two possible values, a byte (consisting of eight bits) can hold any one of 2^8 (or 256) different values. In quantum theory similarly each observable \hat{X} is associated with a set of possible ways (x_1, x_2, \dots) in which it could be prepared and which could

13:24

³If I were to measure the angle at time t and the outcome was, say, 37 degrees it would still be quite false to write $\hat{\Theta}(t) = 37$ because the left hand side of the "equality" is an observable (it refers to the whole multiversal object in many universes) while the right hand side is a number (representing a certain dimension, in this case degrees) and it refers to just the universe in which the outcome was 37. Strictly speaking the measuring instrument also includes the light which I used to align the protractor with the rods. That's because for any measurement to work something has to be affected by the physical system in question and in this case it's light that's affected by the rods and then by the protractor and carries information about them to my eye[s].

later be distinguished from each other by measuring X . Each of these ways of preparing X or possible outcomes of measuring X is given a distinct label, which is a real number. This set of labels is called the spectrum of X and it's written like this: $\text{Sp}(\hat{X}(t))$. For example the spectrum of the angle $\hat{\Theta}$ measured in degrees might be the set of all real numbers between say 10 degrees and 170 degrees, so:

14:41

$$\text{Sp}(\hat{\Theta}) = \{x \in \mathbb{R} | 10 \leq x \leq 170\}$$

Now that's quite a big set in terms of number of members it has uncountably many members. No real experiment could prepare Θ with an arbitrary value in that set or measure precisely what it was. What we could really measure is something like the angle between the rods rounded down to the nearest degree at time t . Let's call that observable $\hat{\Phi}$ ($\hat{\Phi}$). Its spectrum is a discrete set, not a continuum—it's the integers from 10 to 170:

15:36

$$\text{Sp}(\hat{\Phi}) = \{x \in \mathbb{Z} | 10 \leq x \leq 170\}$$

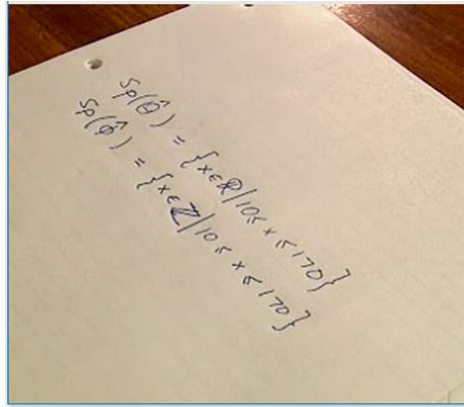


Figure 1.17: Many quantum observables have a discrete spectrum [...].

Many quantum observables have a discrete spectrum from the outset, they are not approximations to anything continuous at all. That's what gave quantum theory its name⁴. Notice that a physical system has to be envisaged as having two contrasting properties. On the one hand it can undergo motion, including being experimented on and measured—it can undergo changes over time—but on the other hand for the very idea of a physical system that changes to make sense it must retain its identity over time which means that it has some characteristics that remain invariant and identifiable throughout any possible change. This angle (see Figure 1.18 on the next page) can change with time but the fact that the system has an angle observable, that it's defined in a certain way in terms of these rods, is an invariant feature of the system. If

⁴Quantum means discrete chunk.

this were disassembled into components it might still be possible to define Θ and to measure it but ultimately if the object were melted down, say, it could become physically impossible to identify which atom was which and then Θ would definitely no longer exist, and nor would this physical system.

17:05



Figure 1.18: This angle can change with time [...].

To some extent it's arbitrary where we draw the line between a physical system having merely changed its configuration and having ceased to exist. But what's important is that in quantum physics it's always possible to analyze phenomena in terms of physical systems that undergo changes and interactions with other systems but nevertheless retain their identity over time because some things about them remain invariant. In particular: what observables they have, how they can be measured, and what their spectra are.

17:44

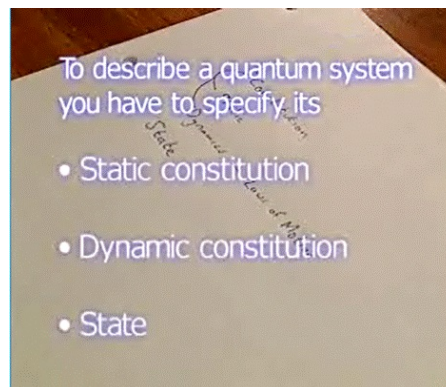


Figure 1.19: What else is invariant?

What else is invariant? Well—the system's laws of motion. It's actual motion can be different on different occasions but the laws that determine how the

system behaves in isolation or how it would behave under any given external circumstances these two are invariant features of a physical system. Let me call that which is invariant about the physical system its **constitution**. Its **static** constitution is all its invariant characteristics apart from laws of motion and its **dynamics** are its laws of motion. To complete the description of a quantum physical system we need to specify a third thing, namely its **state**—what it's actually doing during a particular experiment in all the universes in which it exists. In classical physics the analog of specifying the state would be specifying which trajectory the system is on. So, in the quantum case the state specifies which of its trajectories in the multiverse the whole multiversal object is on.

19:18

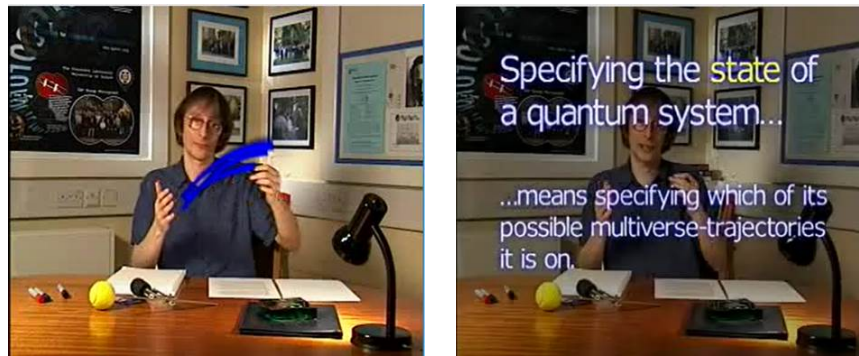


Table 1.2: In the quantum case the state specifies which of its trajectories in the multiverse the whole multiversal object is on. [... T]he state is timeless.



Table 1.3: In this sense of the word “state” the state of a system doesn’t change.

In this sense of the word (state) the state of a system doesn’t change. Quantum observables change under the laws of motion but the state is timeless. This sort of state is known as the Heisenberg state of the system. I am just telling you that in case you come across an alternative way of specifying what a quantum

system is doing, called the Schrödinger state, which does change with time. I'll be using the Schrödinger state a lot in later lectures but for the moment by state I mean the constant Heisenberg state. In a moment I'll tell you how to specify all three things: the static constitution of a system, its laws of motion and its Heisenberg state, but first some more about observables. We can express the relationship that I defined between the observables Θ and Φ like this:

20:29

$$\hat{\Phi}(t) = \text{floor}(\hat{\Theta}(t))$$

Now $\text{floor} : \mathbb{R} \rightarrow \mathbb{Z}$ is the function that takes any real number to the nearest integer less than or equal to it. Since this function is initially defined for real numbers we have to be careful about what we mean by applying it to an observable. This brings us to our first encounter with the algebra of observables.

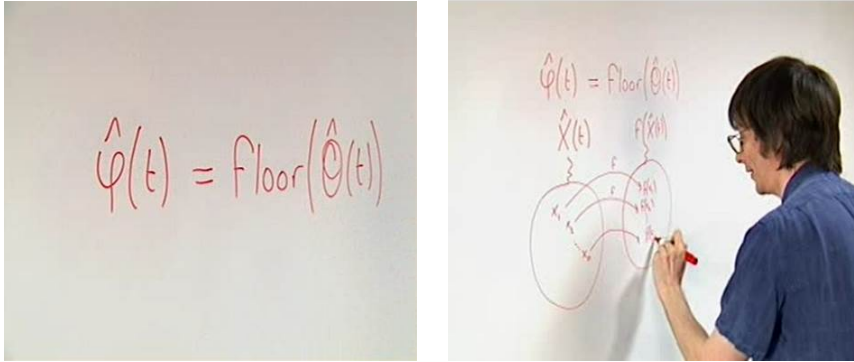


Table 1.4: This brings us to our first encounter with the algebra of observables.

Given any observable, say $\hat{X}(t)$ and any function f that's defined for every element in the spectrum of X (let's say $f : \{x_1, x_2, \dots, x_n\} \rightarrow \dots$) quantum theory says that there exists another⁵ observable $f(\hat{X}(t))$ whose spectrum consists of elements which are f of elements in the spectrum of $\hat{X}(t)$:

20:57

$$\{f(x_1), f(x_2), \dots, f(x_n)\}$$

Okay, this is the spectrum of $f(\hat{X}(t))$ but what is $f(\hat{X}(t))$ physically? Well, given a way of measuring $\hat{X}(t)$ there are guaranteed to be at least two ways of measuring $f(\hat{X}(t))$:

22:16

- One way is to measure $\hat{X}(t)$ and then to compute the function f of the outcome⁶ of that measurement.
- Another way of measuring the observable $f(\hat{X}(t))$ would be to relabel the measuring instrument that's used to measure $\hat{X}(t)$.

⁵For example we can multiply $\hat{X}(t)$ by a real number, say λ , and get another observable $\lambda\hat{X}(t)$ whose spectrum consists of λ times elements in the spectrum of $\hat{X}(t)$.

⁶That whole operation of measuring $\hat{X}(t)$ and then performing that computation constitutes a measurement of the observable $f(\hat{X}(t))$.

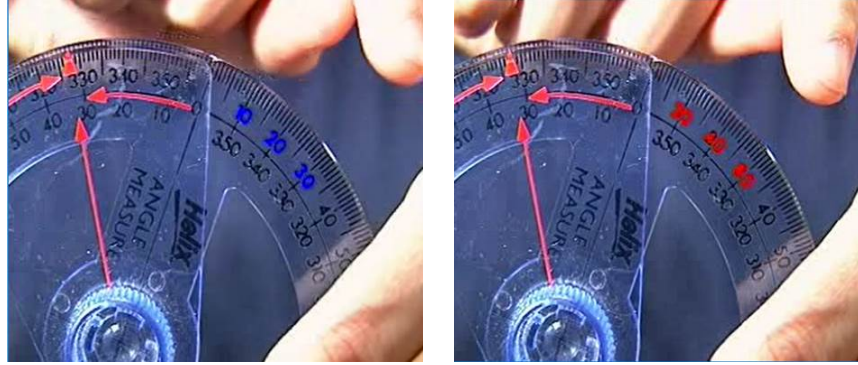
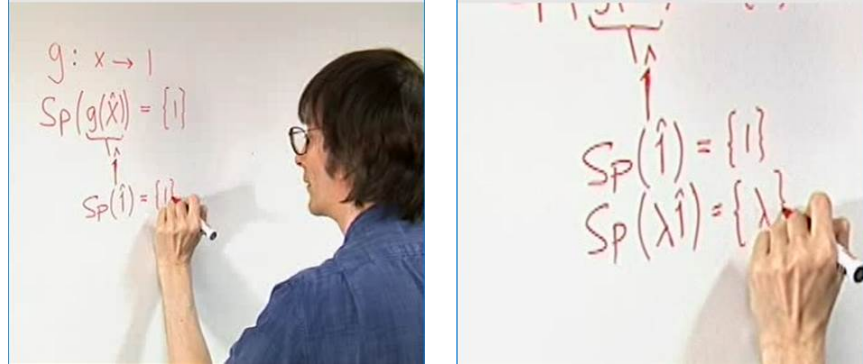


Table 1.5: For example, if we took this protractor and relabeled it [...]

23:02 For example, if we took this protractor and relabeled it so that where it now says 10, 20, 30, and so on, we would write 20, 40, 60 and so on, then the act of lining up that protractor with these rods, and reading off the number on that scale would constitute measuring the observable $2\hat{\Phi}(t)$ instead of $\hat{\Phi}(t)$. In general, there are also other ways of measuring the observable $f(\hat{X}(t))$ that don't involve $\hat{X}(t)$ at all, for instance, my measurement of $\hat{\Phi}(t)$ here certainly does not involve first measuring $\hat{\Theta}(t)$ and then rounding the outcome to the nearest integer, nor does it involve recalibrating an instrument that measures $\hat{\Theta}(t)$ [in fact] my whole reasoning for introducing $\hat{\Phi}(t)$ was that there just isn't an instrument that's capable of literally measuring $\hat{\Theta}(t)$.

Table 1.6: Now let $g : \mathbb{R} \rightarrow \{1\}$ be a function that maps [...]

24:01 Now let $g : \mathbb{R} \rightarrow \{1\}$ be the function that maps any real number $x \in \mathbb{R}$ to a constant, say, 1. Consider the observable $g(\hat{X})$. Its spectrum just contains a single element $\text{Sp}(g(\hat{X})) = \{1\}$ and therefore to measure $g(\hat{X})$ you don't even have to do an experiment all you have to do is write down "the outcome was 1". We call this trivial observable "the unit observable" and write $\hat{1}$ instead

of $g(\hat{X})$ thus $\text{Sp}(\hat{1}) = \{1\}$. Similarly λ times the unit observable where λ is any real number must be another trivial observable the one where the only possible outcome of measuring it is λ : $\text{Sp}(\lambda\hat{1}) = \{\lambda\}$. I said that the spectrum of an observable is part of the system's invariant identity. We can express this invariance in an algebraic way. Let the values in the spectrum of some observable $\hat{X}(t)$ be x_1, x_2 , and so on, up to x_n . 25:28

$$\text{Sp}(\hat{X}(t)) = \{x_1, x_2, \dots, x_n\}$$

These values don't change with time. Consider this function

$$P: x \rightarrow (x - x_1)(x - x_2) \cdots (x - x_n)$$

This is a polynomial. It will have an expansion, like this (the degree n of the polynomial is the same as the number of elements in the spectrum of \hat{X}):

$$P(x) = a_0 + a_1x + \dots + a_nx^n$$

Now since P is a function that's defined for all elements of the spectrum of \hat{X} , [as] in fact it's defined for all real numbers, $P(\hat{X}(t))$ must be an observable and it's easy to find out what observable it is if we first work out its spectrum. 26:50

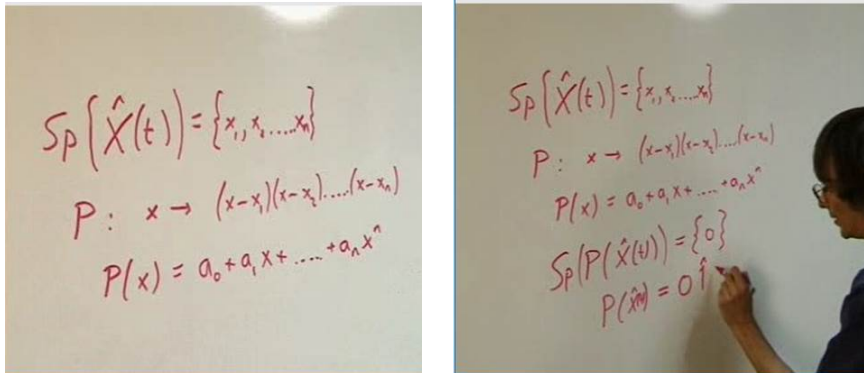


Table 1.7: This is a polynomial. It will have an expansion, like this [...]

The elements of the spectrum are $P(x_1)$ and $P(x_2)$ and so on up to $P(x_n)$ and those values are all zero, because whenever (the little) x is in the spectrum there is a term in this product that's zero so the spectrum of $P(\hat{X}(t))$ only contains one element, namely 0 (zero):

$$\text{Sp}(P(\hat{X}(t))) = \{0\}$$

Measure it and you will always get the outcome 0 (zero) and therefore $P(\hat{X}(t))$ itself must be 0 times the unit observable which we also write as just 0: 27:52

$$P(\hat{X}(t)) = 0 \cdot \hat{1} = 0$$

So, although the observable \hat{X} itself changes with time we found an algebraic equation that it satisfies at all times. Let me write it down explicitly:

$$a_0 \hat{1} + a_1 \hat{X}(t) + \dots + a_n (\hat{X}(t))^n = 0$$

28:48 Therefore this algebraic statement about $\hat{X}(t)$ is a statement about the static constitution of the quantum system that $\hat{X}(t)$ belongs to. And here's a general truth about quantum systems: every algebraic relation among observables at one time is true of the same observables at every other time too and so it's part of the static constitution of the system. In fact the set of all true algebraic relations among observables at any one time defines the static constitution of the system so suppose you found an algebraic equation that related some of the observables of the system at time t say $f(\hat{A}(t), \hat{B}(t), \dots) = 0$ then no matter how much the system may change or interact with other systems—so long as it remains in existence at all—quantum theory says that its observables will only change in ways that keep that equation, and all such equations, true. The set of all true algebraic equations among the observables of a system, is called the algebra of the observables.

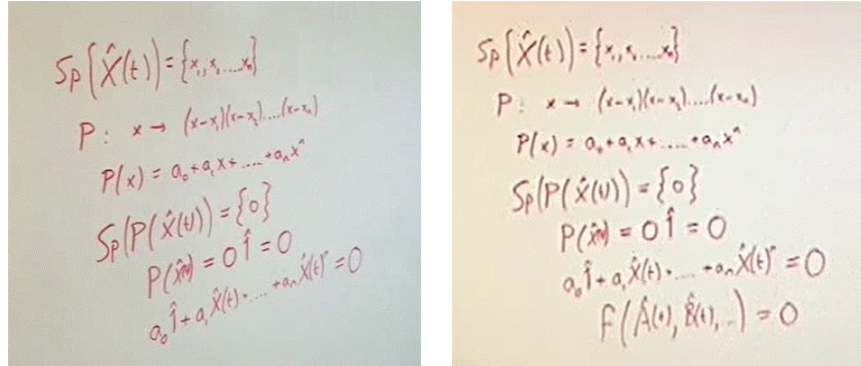


Table 1.8: Static constitution: the algebra of the observables at any one time.

30:31 The algebra of the observables at any one time, like this: $f(\hat{A}(t), \hat{B}(t), \dots) = 0$ is what specifies the system's static constitution. The algebraic relationships between observables at different times specify the system's dynamics, or laws of motion. These can usually be summarized as differential equations in other words algebraic relationships between observables at infinitesimally different times. So, in summary, the whole constitution of a quantum system is defined by the whole algebra of its observables including observables at different times.

31:14 What about the state? To specify the state of a quantum system you have to specify a function called the Expectation-Value Function that maps each observable $\hat{X}(t)$ to a real number called its expectation value and written⁷:

$$\langle \hat{X}(t) \rangle$$

⁷“the expectation value of the observable $\hat{X}(t)$, or the expectation value of \hat{X} at time t ”

Quantum theory places certain constraints on this function namely:

- (a) that the expectation value of an observable is never lower than the lowest element in its spectrum and never higher than the highest element, and
- (b) that the expectation-value function has to be a linear function.

We write the first condition as follows:

$$\min(\text{Sp}(\hat{X}(t))) \leq \langle \hat{X}(t) \rangle \leq \max(\text{Sp}(\hat{X}(t)))$$

The other condition is that it has to be a linear function, so:

32:48

$$\langle \lambda \hat{X} + \mu \hat{Y} \rangle = \lambda \langle \hat{X} \rangle + \mu \langle \hat{Y} \rangle$$

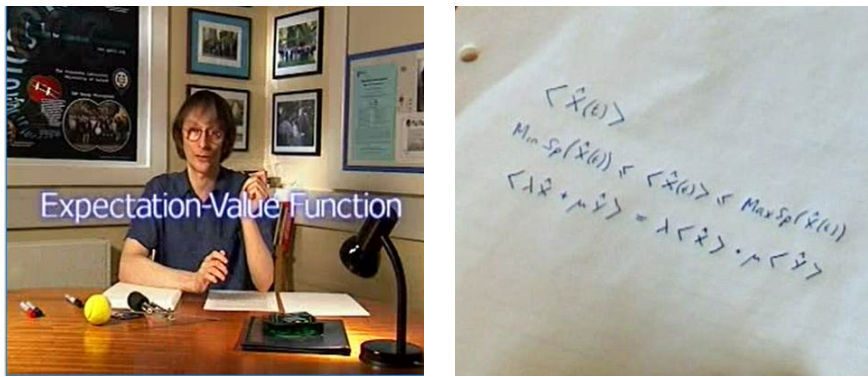


Table 1.9: Quantum theory places certain constraints on this function [...]

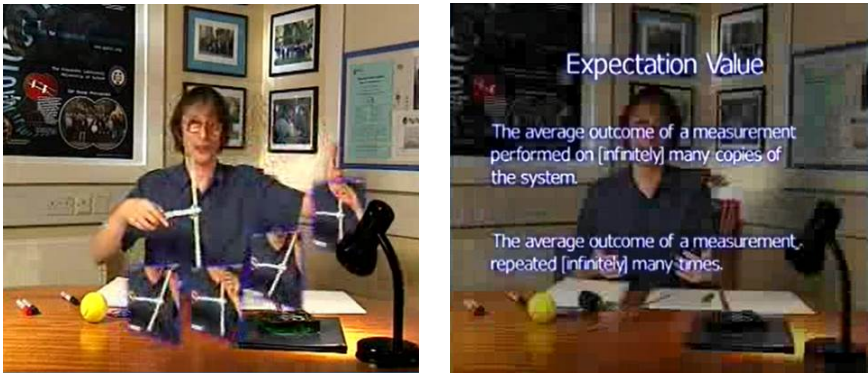


Table 1.10: [...] let me give an indication of what expectation values mean.

I'll give an example of an expectation-value function in a moment, but first let me give an indication of what expectation values mean. If you do an experiment you're doing it in a range of universes. You, and the system, and the

33:13

measuring instrument are all multiversal objects. In general, all the outcomes in the spectrum of \hat{X} actually occur in different universes. Therefore, it is in general impossible to predict a specific outcome for a measurement, that's where expectation values come in. They are what quantum theory makes predictions about. For instance, suppose that you perform the same experiment repeatedly. What's meant by the "same experiment"? Well, you keep preparing the system in the same way again and again and on each occasion you measure \hat{X} at a time t later. Then the average of all the outcomes of all those measurements tends towards the expectation value of $\hat{X}(t)$. In the limit of an infinite sequence of preparations and measurements the average outcome would be exactly the expectation value of $\hat{X}(t)$. That's one rough and ready operational meaning of the expectation value. Another rough and ready operational meaning is this: suppose you make many copies of the system. And prepare them all in identical ways. And then measure \hat{X} and its counterparts in the copies then the average outcome tends to the expectation value of \hat{X} as the number of copies tends to infinity. These meanings cease to be operational meanings as soon as you insert the qualification "infinitely". But they cease to be strictly true if you leave it out. Well, the real meaning of the expectation value is that it's the average value of \hat{X} over a region of the multiverse, the region where the system is prepared and measured in the given way. But I'm getting ahead of myself. In due course I'll explain what a region of the multiverse is and how one averages over it and how that relates to these more operational meanings of the expectation value of an observable.

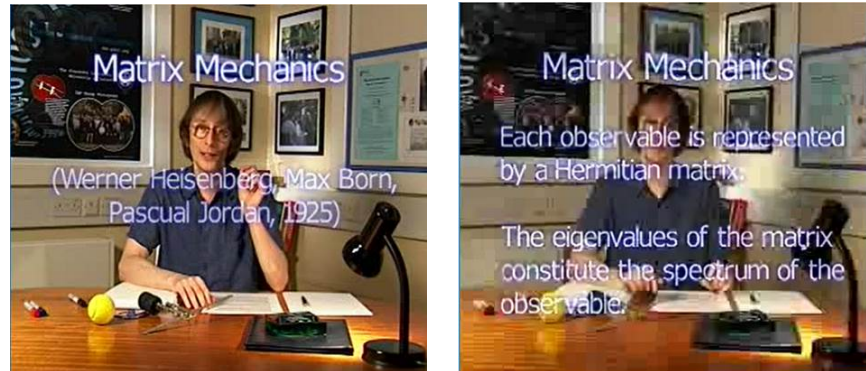


Table 1.11: Historically, the first formulation of [...] Quantum Theory [...]

Historically, the first formulation of what we would today call Quantum Theory was proposed by Werner Heisenberg with the help of Max Born and Pascual Jordan in 1925. It was called Matrix Mechanics because in it each observable is represented by a Hermitian matrix (check the accompanying notes if you want a summary of the basic properties of Hermitian matrices. The eigenvalues of the matrix (which are real numbers) constitute the spectrum of the observable represented by the matrix. We're going to be using a lot of matrices but it's

important to keep at the back of your mind that it's not the matrices but the algebra of the matrices (the algebra of the observables) that defines a physical system. If we found any set of matrices that satisfy the same algebraic relationships as the observables of a given physical system then those matrices would do as a description of the constitution of the system. So there's a lot of freedom to choose quite different sets of matrices to describe the same physical system. One constraint is that all the matrices representing observables of a given system must have the same dimension otherwise we couldn't add and subtract them and do arithmetic with them as I described. For each physical system there's a minimum dimension of matrix required and that's determined by the largest spectrum that any [... observable]⁸ whose spectrum has n elements then the matrix representing that observable has n different eigenvalues so it must be at least an $n \times n$ matrix. Furthermore not only is every observable of the system represented by an $n \times n$ Hermitian matrix, every such matrix represents an observable of the system.

So, that's how Quantum Theory describes the world.

Now, what is the simplest possible quantum physical system?

In classical computation the simplest possible memory location is one that can hold either one of exactly two values: that's called a bit. Considered as a variable a bit is co[...]thing⁹ like a degree of freedom that has only two possible values though such things don't fit comfortably into the scheme of classical dynamics, which is based on continuously varying degrees of freedom. The simplest possible quantum observable is a boolean observable, defined as an observable with exactly two eigenvalues.

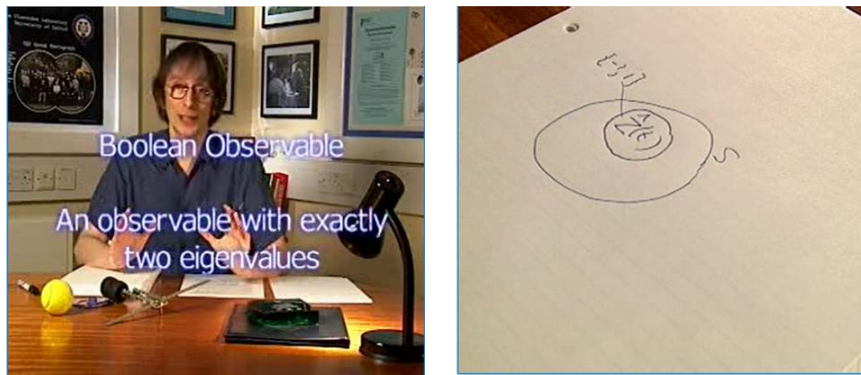


Table 1.12: The simplest quantum observable is the boolean observable.

Any observable simpler than that having only one eigenvalue would be trivial; it would be a multiple of the unit observable. Every physical system has boolean observables for instance whether Θ is less than 90 degrees or greater than or equal to 90 degrees is a boolean observable, it's the observable $\text{floor}(\frac{\Theta(t)}{90})$

⁸Five second gap in video at 0:37:20.

⁹Eight second gap in video at 0:38:20.

whose eigenvalue 0 (zero) stands for less than 90 and 1 (one) stands for greater than or equal to 90. The fact that an observable has exactly two eigenvalues is physically much more significant than what those eigenvalues actually are. Remember that the eigenvalues are just labels for possible outcomes of measurements (and we can always relabel those) so for any observable say \hat{X} with eigenvalues a and b (so, $\{a, b\}$) there exists another observable of the form $\alpha\hat{X} + \beta\hat{1}$ that has any other two eigenvalues we like and measuring that observable isn't going to be much different from measuring \hat{X} .

40:16

So, consider any physical system, let's say S and let's choose a boolean observable of S , call it $\hat{Z}(t)$ that has eigenvalues ± 1 (that turns out to be slightly more elegant for our purposes than choosing 1 and 0). Now, measuring $\hat{Z}(t)$ has only those two possible outcomes. We could measure it by first measuring some more complicated observable of S and then evaluating some function of the outcome, that ranges over plus and minus 1, but usually there are easier and more direct ways of measuring $\hat{Z}(t)$ which involve ignoring most of S in other words we need only interact with a subsystem of S in which case $\hat{Z}(t)$ is an observable of that sub-system.

How simple could that sub-system be? In other words, what's the simplest kind of physical system that could hold one bit of information?

41:26

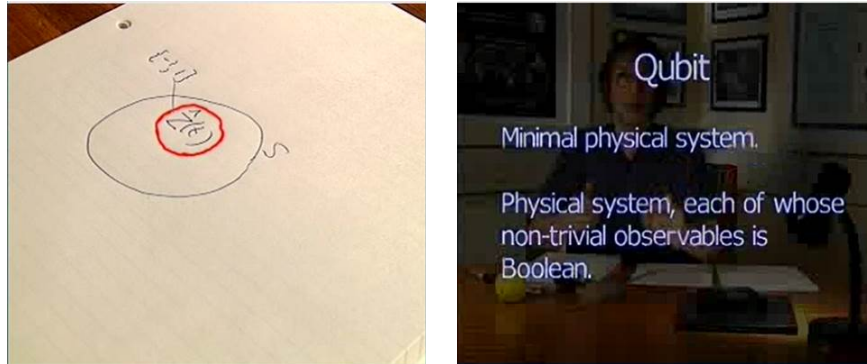


Table 1.13: [A]ny physical system that has $\hat{Z}(t)$ as an observable also has [...] and that's the minimum [...]. [A] physical system with that property, that every one of its non-trivial observables is a boolean observable is called a Qubit.

Well, if $\hat{Z}(t)$ is represented by a certain matrix it has to be at least a 2×2 matrix because it must have exactly two eigenvalues. Then, as I've said, every other Hermitian matrix of the same dimension also represents an observable of the same system. So we know that any physical system that has $\hat{Z}(t)$ as an observable also has a whole continuum of other observables, one for every 2×2 Hermitian matrix. And that's the minimum set of observables that a physical system can have. Every 2×2 matrix has either one or two distinct eigenvalues and therefore every observable of this minimal type of physical system is either a boolean observable or a multiple of the unit observable. A physical system with that property, that every one of its non-trivial observables is a boolean

42:27

observable is called a Qubit. Bear in mind the difference between a qubit, a boolean observable and a bit. A bit (classical) is a degree of freedom, that can take one of two possible values. A boolean observable is the quantum generalization of that—in any one Universe it resembles a bit, but it can take two different values simultaneously in different universes. A qubit is a physical system—a minimal physical system that contains boolean observables and I'll describe several such systems in future lectures.

43:16

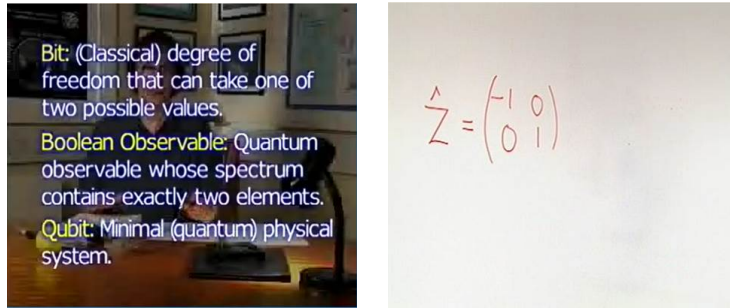


Table 1.14: Now consider a qubit at a given time, say, $t = 0$.

Now consider a qubit at a given time¹⁰, say, $t = 0$. This qubit will have many boolean observables. Let's pick one that has eigenvalues ± 1 and is represented by a diagonal matrix. Let's call that observable $\hat{Z} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. Now I'll define a state that this qubit could be in, during some experiment. To do that, remember, I have to specify a function on the set of all its observables which means on all 2×2 Hermitian matrices.

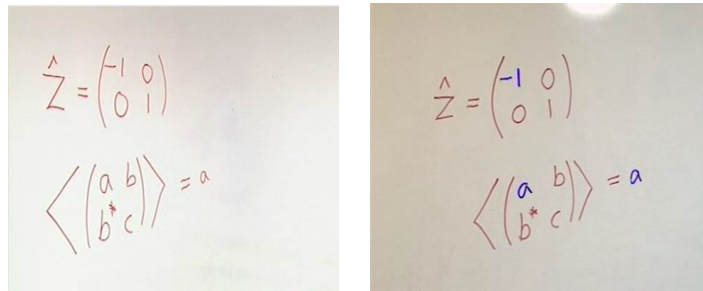


Table 1.15: And this is the function I'll specify [...]

And this is the function I'll specify: the expectation value of a , b , b^* , c [the 2×2 matrix with elements read left to right and from top down] equals just a .

44:31

$$\left\langle \begin{pmatrix} a & b \\ b^* & c \end{pmatrix} \right\rangle = a$$

¹⁰We won't consider other times so I can drop the t .

In other words, the expectation value of any observable is the top left element of the matrix representing that observable in this state that I defined.

If you look at the worked examples for this lecture you can verify that this function has all the properties that I specified for expectation value functions and therefore that it specifies a possible state. So, what does this tell us about what's happening to our qubit? Well, what's happening to its observable \hat{Z} ?

Well, the expectation value of \hat{Z} is -1 . So, for instance, if we repeated the whole experiment many times, repeated it including the initial preparation and everything, then the average value of all the outcomes would be -1 . OK, but look at the spectrum of \hat{Z} . It only contains the values -1 and 1 . So each individual outcome must be one of those two values. And so, if the average over many outcomes is -1 it follows that every single outcome must in fact be -1 therefore in this case we don't have to bother with repeating the experiment many times we don't have to set up an ensemble of identically prepared copies of the system we don't have to worry about what's happening in other universes quantum theory predicts that the outcome of a measurement of \hat{Z} will be -1 .

46:15

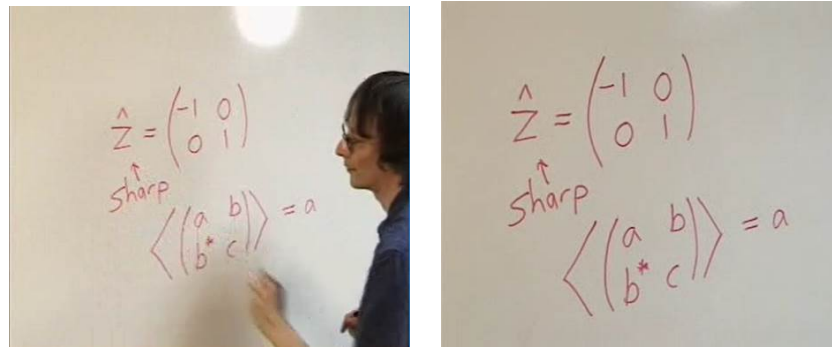


Table 1.16: [...] then the observable is said to be *sharp* in that state.

In a given state when an observable has this property—that if it were measured the same outcome would occur in all universes—[then] the observable is said to be *sharp* in that state. Now let's think about measuring a different observable of the same qubit at time 0 (zero). Let's say this one:

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

So we can read off the expectation value (that's the top left corner: the expectation value of \hat{X} is zero) so if we make many copies of this experiment measuring \hat{X} the average value of the outcome will be 0 (zero). Same with repeating it many times, same is true with the average value over all the universes in which \hat{X} is measured. No particular outcome will be 0 (zero) though. That's because, as you can verify in the worked examples the eigenvalues of \hat{X} are also ± 1 so each outcome will be either $+1$ or -1 . And if their average is going to tend to 0 (zero) well you can see that half of them will have to be

47:51

+1 and the other one -1 in the long run and this translates to a probabilistic prediction: the probability of each of these two outcomes is $\frac{1}{2}$.

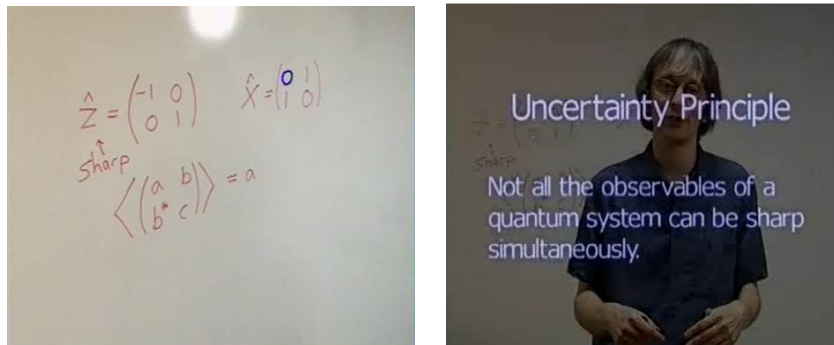


Table 1.17: \hat{Z} was sharp, \hat{X} isn't, and that's no accident.

So, the observable \hat{X} is not sharp in fact it's as far from being sharp in this state as a boolean observable can get. \hat{Z} was sharp, \hat{X} isn't, and that's no accident. There's a principle of quantum mechanics called rather misleadingly, the Uncertainty Principle, which says that not all the observables of a quantum system (at a given instant) can be sharp simultaneously. If some of them are, others will not¹¹ be. Heisenberg introduced a quantitative version of this principle which is named after him: the Heisenberg Uncertainty Principle.

48:51



Figure 1.20: I hope you're beginning to see a qubit is a seriously weird thing.

In this lecture you've seen how quantum systems are described in terms of observables, which are matrices, and states, which are real-valued functions of matrices. And in particular, you've seen the simplest type of quantum system, the qubit, described in that way. I hope you are beginning to see that a qubit is a seriously weird thing. I haven't even discussed any actual experiments yet or told you about any phenomena by which quantum theory is tested and which

¹¹You can prove this for yourself from the properties of expectation values (again, see the worked examples).

provides quantum computation with its power. But suppose the theory is true. Here we have an entity, the qubit, that's literally not of this universe. If we try to pin it down, if we prepare it carefully so that a particular boolean observable is sharp and has the same value in all the universes in which we measure it then other observables of the same qubit cease to be sharp, there's no way we can make the qubit as a whole homogeneous across universes. It's an unequivocally multiversal object [...]¹² Every boolean observable is part of a qubit and every question of whether something measurable is so or not, is in reality a boolean observable, and therefore the complete answer to such a question is not in reality just one of those yes/no values, not even both of them in parallel, but a quantum observable—something that can be represented as a matrix. Even as you measure one of those observables and perceive one of those eigenvalues as “the outcome” the other outcomes are generically also present in the wider reality and are affecting each other. What we perceive to some degree of approximation as a world of single-valued variables is actually something much larger and richer corresponding to a great algebra in which there is a matrix whenever we perceive a variable. Yet despite this rich structure there's an overall unity and simplicity to the quantum world. The complete description of the essence of a physical system reduces to its characteristic algebra. The complete specification of what the system is doing across the multiverse and over time is a certain function defined on the elements of that algebra.

50:00

51:14



52:16

It's a beautiful theory, but more to the point it's how the world is. You and I are collections of not just particles at particular positions at each instant but of matrices—walking around, performing measurements, perceiving the world and ourselves. You may not want to be a bunch of matrices—I quite like the idea. But either way we have no choice. If we want to understand the physical world at the deepest level currently known to human beings it has to be via quantum theory. I hope you want to do

that—and I hope you'll join me in the subsequent lectures of these series¹³.

¹²Five second gap at 50:09 in video.

¹³Video ends at 52:28

Chapter 2

Interference



Table 2.1: In this lecture I will use that theory of observables and expectation values to analyze the archetypal quantum phenomenon [...] of [...] Interference.

In the last lecture I gave an overview of how physical systems in general are described in quantum theory via entities called observables which can be represented by matrices and via expectation-value functions that map observables to real numbers. I also told you about the simplest kind of observables, boolean observables, which are those whose spectrum contains exactly two elements, and I told you about the simplest kind of quantum system, the qubit, a physical system all [of] whose non-trivial observables are boolean. But I didn't describe any phenomena. In this lecture I will use that theory of observables and expectation values to analyze the archetypal quantum phenomenon, that of Quantum Interference. It can be defined as a phenomenon in which an easily measurable observable is first sharp, and then becomes unsharp, and then becomes sharp again. To put that in parallel universes terms an interference phenomenon is one in which the observed outcome depends on what has been happening in more than one Universe.

01:08

Quantum interference can be demonstrated in a classic set of experiments involving a single qubit. In the case I'll describe, the qubit is a subsystem of a single photon, or particle of light. What I mean by a subsystem of a single

02:33

photon is that only some of the observables of the photon are involved. And these observables evolve independently of all the others during these experiments. So it's permissible to regard them as constituting a physical system in their own right.

03:13



Table 2.2: The single qubit experiment I'm going to describe [...]

Like all experiments, this one can be regarded as a computation on a single qubit. Now, a classical computer with only a single bit of memory couldn't do very much. In fact, there are exactly four possible computations that one can do with a single bit: set it to one value, set it to its other possible value, flip the value (that's the NOT computation), or leave it alone (that's the "null" computation). Or does any sequence of those, though the net effect of any sequence would still be the same as one of those four.

03:57

That's classical computations, but a single qubit already allows for quite a variety of quantum computations. Some of them are useful as computations, some of them are interesting in the theory of quantum computation, and some of them are interesting as physical processes. The single qubit experiment I'm going to describe has a bit of all three.

Now I'll specify a qubit within an individual photon using the method I introduced in the previous lecture—namely, I'll pick a boolean observable of the photon, and then I'll define our qubit as the minimal physical system containing that boolean observable.

The observable in question is: which of two particular directions the photon is traveling in. We'll make sure that in this experiment, it never does travel in any direction other than very close to one or other of these two directions. So, measuring which one it is, we're measuring a boolean observable. And I'll call that observable "Z-hat" with eigenvalues plus and minus one standing for the two possible directions of travel.

05:32

Here's a source¹ of photons; a laser. Here are some of them streaming out of it. About 3 times ten to the 15 of them per second, in fact. If we make this beam a meter long, let's say, well, the speed of light is about $3 \times 10^8 \frac{m}{s}$, so

¹See picture on the next page.

at any one instant there would be about 10 to the 7, or 10 million, photons present in the beam.

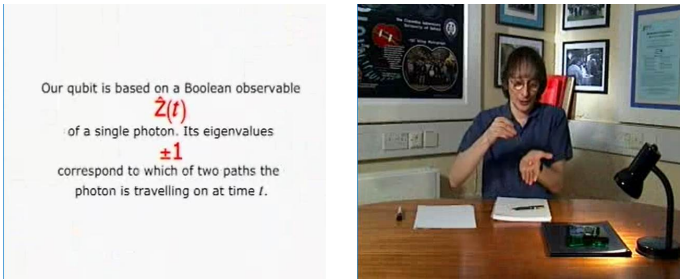


Table 2.3: Here’s a source of photons: a laser.

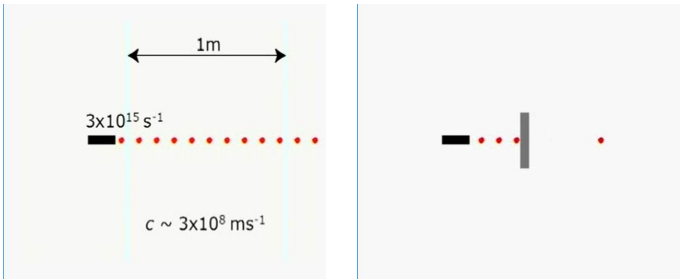


Table 2.4: Here are some of the[photons], streaming out.

If we want to experiment on one photon at a time, we put a dark filter in front of the laser which absorbs most of the photons. If the filter reduced the intensity of the light by a factor of 10 million, then there’d be only one photon at a time on average on this meter long path. In the actual experiment I’m going to show you in a moment, we’re going to make the intensity even lower than that. So the effective distance between successive photons entering the apparatus will be not one meter but about 35 meters.

06:04

06:42



Table 2.5: This is the quantum optics lab [...]

This is the quantum optics lab where the experiment’s going to be done.

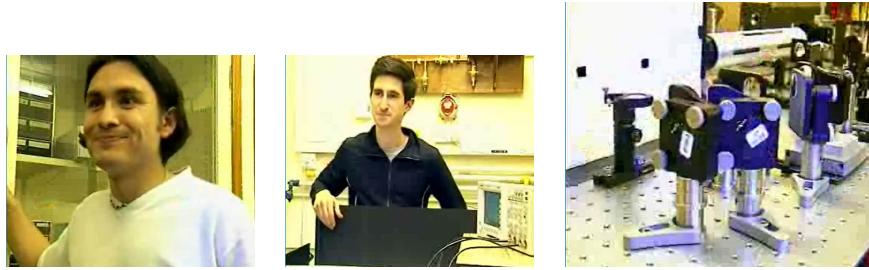


Table 2.6: This is Manuel and this is Ehran.

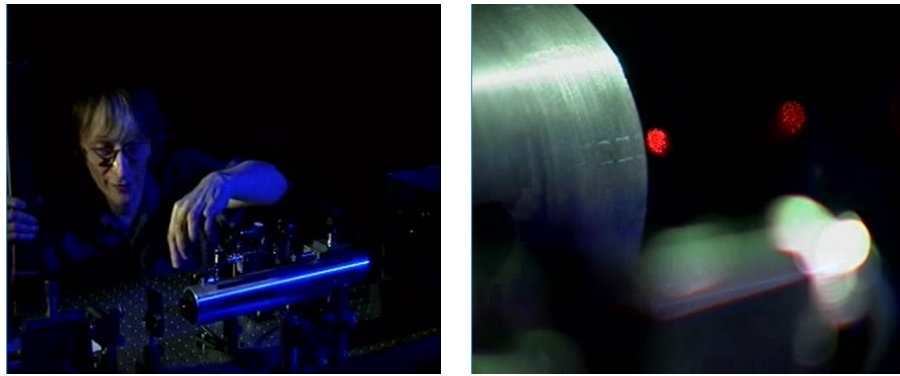


Table 2.7: This is a source of photons, it's a laser [...].

This is Manuel, and this is Ehran. This is a source of photons, it's a laser, and it's emitting about 5 mW in the form of photons here. Now we want to do this experiment on one photon at a time. So we pass the beam through dark filters which consist of glass with embedded metal in it. There's one of them.



Table 2.8: There's one of them [...] then the light passes [...] into this [...]

And then the light passes into this fiber optic cable, around here, and into this, past this dark filter whose effect you can see here, making the beam far too weak for the naked eye to see. The net effect of these filters is to attenuate

the beam by a factor of 10 to the 10, so that, here, there are only 10 to the 6 photons entering the apparatus per second: one per microsecond.

08:05

I said that a photon is a particle. What I mean is that in a single universe it has many of the properties that particles would have in classical physics. For instance, the photons here are moving through space at a constant speed in a straight line, just as Newton’s laws would require, or actually a very slightly curved line because of gravity. And here they are bouncing off a barrier with a neat specular reflection that the conservation laws of classical mechanics would predict. Crossing two beams shows that to the extent that they resemble classical particles, photons must be really tiny—point particles to a good approximation. Because the beams just don’t notice each other. Even with 10 to the 15th odd photons passing by each other every second here, we don’t notice any absorption or deflection out of the beam through collisions.

09:05

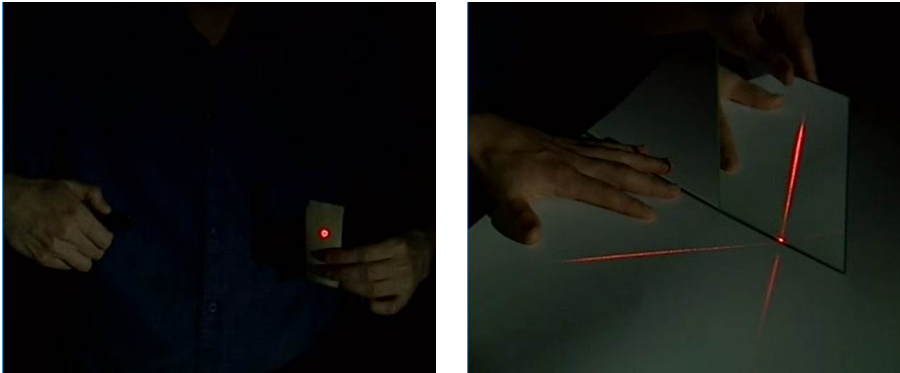


Table 2.9: I said that the photon is a particle [...] What I mean is that [...]

But we’ll see in a moment that in other ways they don’t resemble classical particles and they’re not located at just one point. We could verify that none of them is deflected by putting a sensitive photon detector like this one somewhere outside the two beams.



Table 2.10: [...] a sensitive photon detector, like this one [...]

This is a solid state photon detector. If a photon goes in here, it creates a small electric current which is then amplified and can be displayed on an oscilloscope. So in an experiment involving only one photon, this device measures a boolean observable; does the photon strike a given location where we put

09:55 the detector or not? Incidentally, this is a destructive sort of measurement. The photon, including the qubit we are interested in, is always destroyed in the course of detecting its presence with this device. That's a limitation that won't matter in these experiments, but in future lectures when I discuss more sophisticated quantum computations, we'll need qubits that remain in existence after each computational step so that they can participate in subsequent steps. So we can only use this device as the last step of a computation: reading the output. It wouldn't do as an internal component of a quantum computer.

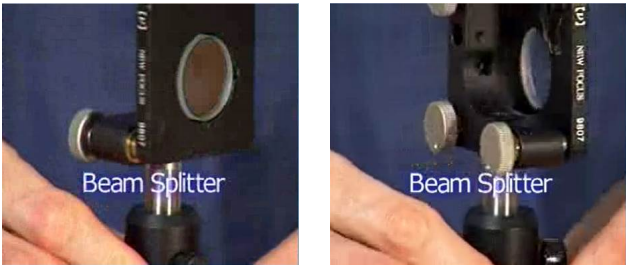


Table 2.11: Here's the key component [... i]t's called a beam splitter [...]

10:44 But this device would. Here's the key component in the little one-qubit quantum computer we are going to build. It's called a beam splitter.

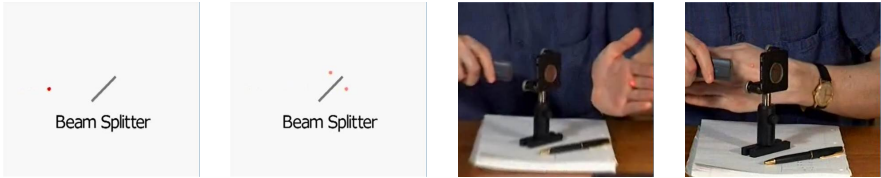


Table 2.12: What happens is, that in half [of] the universes it [...]

When a photon with a sharp direction of motion strikes it, its direction of motion becomes unsharp. What happens is that in half the universes it just carries on straight through, and in the other half it bounces off as if from a mirror. It turns out that it can do the reverse of that too.



Table 2.13: Then [...] its direction of motion is sharp again.

If we were to put ordinary mirrors here and here to reflect the photon back the way it came, whichever way that was in each universe then, after it strikes the

beam splitter again, its direction of motion is sharp again. Let me just remind you that this isn't a case of two photons merging or becoming alike. This is a case of a single photon whose direction of motion is not sharp, striking the beam splitter and its direction of motion becoming sharp again. So the beam splitter is also a beam joiner. That phenomenon of an unsharp direction of motion becoming sharp is our quantum interference phenomenon.

Both parts of the photon that have traveled on different paths in different universes participate in the final interference process. How can we tell? Well if we were to remove one of the mirrors so that whenever the photon takes this path, it never returns, then the photon never would become sharp. In fact, it would end up with a three fold unsharpness going in that direction in some universes and then the beam splitter would have nothing to join so in other universes it would end up going in this direction and in others in that direction.

12:09

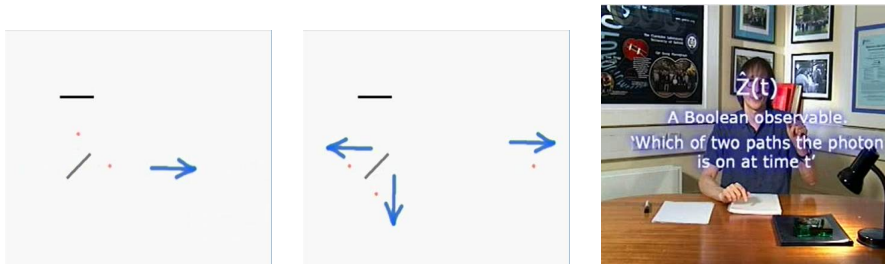


Table 2.14: We start with a photon, we pass it through a beam splitter [...]

If we want to experiment on a single qubit we have to keep things simpler than that. A qubit has only boolean observables. So we have to make sure that none of the observables of the system or subsystem that we are investigating, ever take on more than two values at a time. So the basic form of all the single qubit experiments that I'm going to describe is this. We start with a photon, we pass it through a beam splitter after which it's travelling on two paths. We define a boolean observable of the photon as being which of the two paths it's traveling on. Then, using the method I introduced in the previous lecture, we'll define our qubit as the minimal physical system containing that boolean observable. I'll call that observable "Z-hat" with spectrum plus or minus one, the eigenvalue minus one standing for one path and plus one for the other. Since $\hat{Z}(t)$ has spectrum ± 1 , it will be represented by a Hermitian matrix which may change with time but always in such a way that its eigenvalues are plus or minus one. To give a full description of the qubit and what happens to it in these experiments, we need to specify as always three things. Its static constitution, which is what its observables are and what the algebra is at any one time; the dynamics is how the observables change with time—remember, they change in such a way as to keep their algebra at any given time constant; and the state, which means how the system is programmed—how it is prepared on a particular occasion. And that is all summed up in the expectation value

12:53

14:46

function.



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

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

The Pauli matrices

Table 2.15: There's a nice way of working with [...]

So, static constitution first. As I indicated last time, the set of all observables of a qubit at any one instant has the same algebra as the set of all 2×2 Hermitian matrices. Now, that algebra is most straightforwardly represented by 2×2 Hermitian matrices. But there are higher dimensional matrix representations of the same algebra—and it turns out that in most experiments we'd have to use one of those². Anyway, for our present purposes it's sufficient to represent the observables of our qubit as 2×2 matrices. There's a nice way of working with 2×2 Hermitian matrices, and more generally $2^n \times 2^n$ ones as well where we don't have to bother getting our hands dirty with the actual components, the complex numbers. Instead we use four fixed standard Hermitian matrices, namely the unit matrix I , and three other matrices known as the Pauli matrices: σ_x , σ_y and σ_z . Every 2×2 Hermitian matrix can be expressed as a linear combination with real coefficients of these four matrices.

16:22

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$$

$$\begin{array}{l|l} \sigma_x \sigma_y = i \sigma_z & \sigma_y \sigma_x = -i \sigma_z \\ \sigma_y \sigma_z = i \sigma_x & \sigma_z \sigma_y = -i \sigma_x \\ \sigma_z \sigma_x = i \sigma_y & \sigma_x \sigma_z = -i \sigma_y \end{array}$$

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$$

$$\sigma_x \sigma_y = i \sigma_z \quad \sigma_y \sigma_x = -i \sigma_z$$

$$\sigma_y \sigma_z = i \sigma_x \quad \sigma_z \sigma_y = -i \sigma_x$$

$$\sigma_z \sigma_x = i \sigma_y \quad \sigma_x \sigma_z = -i \sigma_y$$

Table 2.16: [...] That's all encoded in these formulae.

Here are some more properties of Pauli matrices. First, you can easily verify that the eigenvalues of a Pauli matrix are ± 1 . And that the square of each

²There's a simple reason for that which I'll get to in the next lecture.

Pauli matrix, therefore, is the unit matrix.

Another property: $\sigma_x \sigma_y = i\sigma_z$ and the other two cyclic permutations of that. And also the Hermitian conjugates of those three: $\sigma_y \sigma_x = -i\sigma_z$ and so on. These formulae completely define the algebra of 2×2 Hermitian matrices, in the sense that if we express matrices as linear combinations of these, we don't need to know the rules for adding or multiplying matrices or finding their inverses. That's all encoded in these formulae. So, I'll express the observables of our qubit in terms of Pauli matrices. By the way, Pauli matrices are often called the Pauli spin matrices because they are useful in the quantum mechanics of spin and rotation. But nothing in today's experiment has anything to do with spin or rotation. Nowadays when I think of Pauli matrices I don't think of spin, I think of qubits.

17:34

$\hat{1} = I$ $\hat{Z}(0) = \sigma_z$	$\hat{1} = I$ $\hat{Z}(0) = \sigma_z$ $\hat{X}(0) = \sigma_x$ $\hat{Y}(0) = \sigma_y$	$\hat{X}^2(t) = \hat{Y}^2(t) = \hat{Z}^2(t) = \hat{1}$ $\hat{X}(t) \hat{Y}(t) = i\hat{Z}(t)$ $\hat{Y}(t) \hat{Z}(t) = i\hat{X}(t)$ $\hat{Z}(t) \hat{X}(t) = i\hat{Y}(t)$ $\hat{Y}(t) \hat{X}(t) = -i\hat{Z}(t)$ $\hat{Z}(t) \hat{Y}(t) = -i\hat{X}(t)$ $\hat{X}(t) \hat{Z}(t) = -i\hat{Y}(t)$	<p>Let $\hat{A}(t)$ be an arbitrary observable of a qubit. Then</p> $\hat{A}(t) = c_0 \hat{1} + c_1 \hat{X}(t) + c_2 \hat{Y}(t) + c_3 \hat{Z}(t),$ <p>where c_0, c_1, c_2, c_3 are constants.</p>
--	--	---	---

Table 2.17: [...] From this [...] we [...] know what their algebra is at any time.

Well, now I can tell you the matrix representation of the observable “Z-hat” that I’ve defined. “Z-hat” is represented at time 0 by the matrix σ_z . And the unit observable, “one-hat”, is as always represented by the 2×2 unit matrix I . These equal signs should really be “is represented by” not “equals”, but we save ourselves a lot of useless mathematical baggage if we gloss over that difference when we are doing calculations. Now, all the other 2×2 Hermitian matrices also represent observables of the qubit at time zero, or at any other time, but in particular at time zero. So, there is an observable whose representation at time zero is σ_x . I’ll call that observable “X-hat”. And I’ll call the observable represented by σ_y at time zero: “Y-hat”. What are “X-hat” and “Y-hat” physically? It will become clear in the course of this analysis what they are, how one would measure them. But from this definition we already know what their algebra is at any time. Since X, Y and Z have the same algebra as the Pauli matrices at time zero, and since the algebra of observables is an invariant feature of any quantum system, they must have the same algebra at any other time t as well. So, for instance, $\hat{X}(t)$ squared must just equal $\hat{1}$ and the same for Y and Z. And $\hat{X}(t)$ [times] $\hat{Y}(t)$ equals I [times] $\hat{Z}(t)$ and so on. All other observables must be linear combinations of $\hat{X}(t)$, $\hat{Y}(t)$ and the unit observable, with real coefficients. And the coefficients must not change with time. You can check the worked examples to verify that.

20:29

Well, next, I'll specify the state of our qubit. As I said, we are going to start each of these experiments by shining the laser at the beam splitter in the Z equals +1 direction. So at time zero, Z is sharp with a value plus 1. And so it's expectation value must be 1. In terms of matrices, that just says the expectation value of σ_z is 1. And in that way we define the expectation value function for matrices as well as observables. You'll see in the worked examples that it follows from this that the expectation value of σ_x and the expectation value of σ_y are both zero. And as always, we have that the expectation value of the unit observable is one. So the expectation value of the unit matrix I is one as well. Since we can express any observable of the qubit as a linear combination of Pauli matrices, and since the expectation value function is a linear function, we can use these formulae to find the expectation value of any observable at any time. So I've completely specified the state of the qubit. Note that these formulae refer to one particular state, the one with Z initially sharp with the value plus one. They're not inherent properties of Pauli matrices; in a different state these expectation values would be different.

22:37

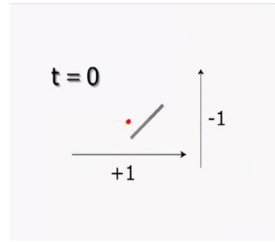
	$\Leftrightarrow \begin{aligned} \langle \hat{Z}(0) \rangle &= 1 \\ \langle \sigma_z \rangle &= 1 \end{aligned}$	$\begin{aligned} \langle \sigma_z \rangle &= 1 \\ \langle \sigma_x \rangle &= \langle \sigma_y \rangle = 0 \\ \langle I \rangle &= 1 \end{aligned}$
--	--	---

Table 2.18: So I've completely specified the state of the qubit.

Okay, now the dynamics. The dynamics are the laws of motion of all the qubit's observables. If we think of the apparatus as a quantum computer, with our qubit as its sole working register, the laws of motion are the rules defining what the computer does step by step to an arbitrary single qubit input.

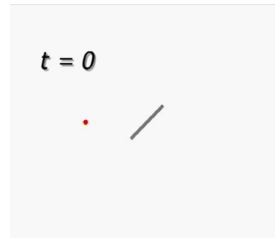
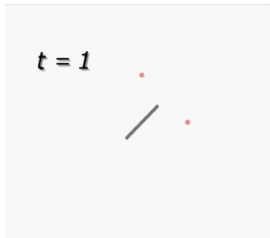
		<p>Effect of beam splitter:</p> $\begin{aligned} \hat{Z}(t+1) &= \hat{X}(t) \\ \hat{Y}(t+1) &= \hat{Y}(t) \\ \hat{X}(t+1) &= -\hat{Z}(t) \\ \hat{I}(t+1) &= \hat{I}(t) \end{aligned}$
---	--	---

Table 2.19: As in classical computation [...]

Well the first thing that happens is that the photon passes through the beam splitter. As in classical computation, it's convenient to analyze computations as proceeding in steps, each of which is a relatively simple computation. For

present purposes, we are not interested in what happens during the operation; only in its net effect on observables. So, here's the law of motion in that sense for the photon, or for the qubit, passing through a beam splitter^{3,4}.

Effect of beam splitter:	Effect of beam splitter:	Effect of beam splitter:
$\hat{Z}(t+1) = \hat{X}(t)$ $\hat{Y}(t+1) = \hat{Y}(t)$ $\hat{X}(t+1) = -\hat{Z}(t)$ $\hat{I}(t+1) = \hat{I}(t)$	$\hat{Z}(t+1) = \hat{X}(t)$ $\hat{Y}(t+1) = \hat{Y}(t)$ $\hat{X}(t+1) = -\hat{Z}(t)$ $\hat{I}(t+1) = \hat{I}(t)$	$\hat{Z}(t+1) = \hat{X}(t)$ $\hat{Y}(t+1) = \hat{Y}(t)$ $\hat{X}(t+1) = -\hat{Z}(t)$ $\hat{I}(t+1) = \hat{I}(t)$

Table 2.20: So here's the law of motion in that sense [...]

So, “Z-hat” of 1 we said was equal to “X-hat” of zero, which equals σ_x . And so the expectation value of “Z-hat” of 1 equals the expectation value of σ_x which I said earlier is zero.

$\hat{Z}(t+1) = \hat{X}(t)$ $\Rightarrow \hat{Z}(1)$	$\hat{Z}(t+1) = \hat{X}(t)$ $\Rightarrow \hat{Z}(1) = \hat{X}(0) = \sigma_x$	$\hat{Z}(t+1) = \hat{X}(t)$ $\Rightarrow \hat{Z}(1) = \hat{X}(0) = \sigma_x$ $\Rightarrow \langle \hat{Z}(1) \rangle = \langle \hat{X}(0) \rangle = \langle \sigma_x \rangle$	$\hat{Z}(t+1) = \hat{X}(t)$ $\Rightarrow \hat{Z}(1) = \hat{X}(0) = \sigma_x$ $\Rightarrow \langle \hat{Z}(1) \rangle = \langle \hat{X}(0) \rangle = \langle \sigma_x \rangle$ $= 0$
---	---	---	--

Table 2.21: So $\hat{Z}(t)$ at time $t = 1$ is no longer sharp.

So, Z at time 1 is no longer sharp. Its expectation value goes to zero which means physically that the photon goes straight through on path plus one in half the universes, and bounces back on path minus one in the other half. 25:04

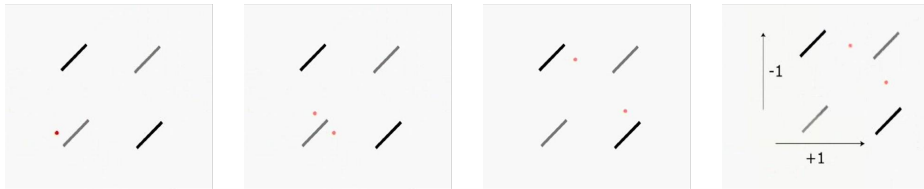


Table 2.22: [N]ext the photon bounces off an ordinary mirror, on either path.

Okay, next the photon bounces off an ordinary mirror, on either path. And you can see what that does, the operation of bouncing off either of these mirrors is

³“Z-hat” at $t + 1$ equals “X-hat” at t ; “Y-hat” at $t + 1$ equals “Y-hat” at t ; and “X-hat” at $t + 1$ equals minus “Z-hat” at t . The unit observable of course never changes.

⁴By taking linear combinations of these, we can determine the law of motion for any observable of the qubit.

just the quantum generalization of the simplest single bit classical operation, namely the NOT operation—converting minus one to plus one, and vice versa.

Effect of the mirrors:	Effect of the mirrors:	Effect of the mirrors:	Effect of the mirrors:
$\hat{Z}(t+1) = -\hat{Z}(t)$ $\hat{Y}(t+1) = \hat{Y}(t)$ $\hat{X}(t+1) = -\hat{X}(t)$ $\hat{I}(t+1) = \hat{I}(t)$	$\hat{Z}(t+1) = -\hat{Z}(t)$ $\hat{Y}(t+1) = \hat{Y}(t)$ $\hat{X}(t+1) = -\hat{X}(t)$ $\hat{I}(t+1) = \hat{I}(t)$	$\hat{Z}(t+1) = -\hat{Z}(t)$ $\hat{Y}(t+1) = \hat{Y}(t)$ $\hat{X}(t+1) = -\hat{X}(t)$ $\hat{I}(t+1) = \hat{I}(t)$	$\hat{Z}(t+1) = -\hat{Z}(t)$ $\hat{Y}(t+1) = \hat{Y}(t)$ $\hat{X}(t+1) = -\hat{X}(t)$ $\hat{I}(t+1) = \hat{I}(t)$

Table 2.23: The law of motion for the NOT operation.

26:07 The law of motion for the NOT operation⁵: “Z-hat” of $t+1$ equals, as you’d expect, minus “Z-hat” of t , and “Y-hat” of $t+1$ equals “Y-hat” of t , and “X-hat” of $t+1$ equals minus “X-hat” of t . In general, I’d have to describe the dynamics of one further elementary operation before I could describe the experiment as a whole. It’s simply the operation of the photon traveling unimpeded for a certain distance in a straight line, in between bits of apparatus.

Travelling a distance d :	Travelling a distance d :	Travelling a distance d :
$\hat{Z}(t+d/c) = \hat{Z}(t)$ $\hat{Y}(t+d/c) = \cos(2\pi d/\lambda) \hat{Y}(t) - \sin(2\pi d/\lambda) \hat{X}(t)$ $\hat{X}(t+d/c) = \cos(2\pi d/\lambda) \hat{X}(t) + \sin(2\pi d/\lambda) \hat{Y}(t)$ $(\lambda \text{ is a constant wavelength})$	$\hat{Z}(t+d/c) = \hat{Z}(t)$ $\hat{Y}(t+d/c) = \cos(2\pi d/\lambda) \hat{Y}(t) - \sin(2\pi d/\lambda) \hat{X}(t)$ $\hat{X}(t+d/c) = \cos(2\pi d/\lambda) \hat{X}(t) + \sin(2\pi d/\lambda) \hat{Y}(t)$ $(\lambda \text{ is a constant wavelength})$	$\hat{Z}(t+d/c) = \hat{Z}(t)$ $\hat{Y}(t+d/c) = \hat{Y}(t)$ $\hat{X}(t+d/c) = \hat{X}(t)$ $(d \text{ is a multiple of } \lambda)$

Table 2.24: To make the calculation easy I’m going to pretend [...]

Well, the observables change periodically with distance. So, to make the calculation easy, I’m going to pretend that all the distances travelled are a whole number of periods - a whole number of wavelengths, so that our whole qubit will remain unchanged in between mirrors.

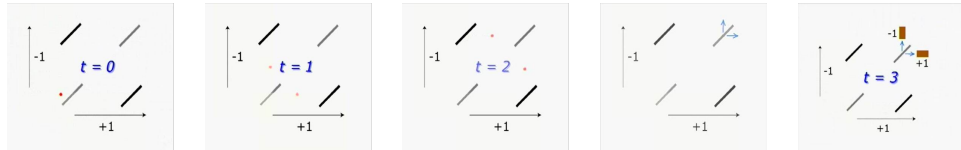
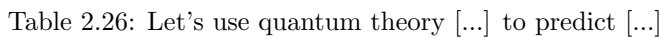


Table 2.25: Now, here’s the experiment.

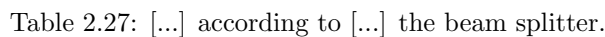
Now, here’s the experiment. It starts at time zero, with the photon entering the apparatus traveling in the plus one direction. It strikes a beam splitter, and

⁵I’m just telling you this by fiat, just as I did for the beam splitter.

27:59



Now, the first thing that happens is that our photon strikes the beam splitter. And so, what happens to X, Y, and Z? Well by time 1 it will have hit the beam splitter, and so the observables of the qubit will have changed according to the formulae that I have given for the effect of the beam splitter. So, we can read off “X-hat” of 1 equals minus “Z-hat” of 0, which equals $-\sigma_z$ and, similarly, for “Y-hat” of 1 and “Z-hat” of 1. Well, then the photon strikes the mirrors. And the effect is the NOT operation according to the[se] formulae I gave. From which, we read off the observables at time 2. Finally, the photon hits the second beam splitter and again we can read off “X-hat” of 3 equals minus “Z-hat” of 2, which makes it σ_x , and so on. It turns out that all the observables of the qubit at time 3 are the same as they were initially at time



Now, the first thing that happens is that our photon strikes the beam splitter. And so, what happens to X, Y, and Z? Well by time 1 it will have hit the beam splitter, and so the observables of the qubit will have changed according to the formulae that I have given for the effect of the beam splitter. So, we can read off “X-hat” of 1 equals minus “Z-hat” of 0, which equals $-\sigma_z$ and, similarly, for “Y-hat” of 1 and “Z-hat” of 1. Well, then the photon strikes the mirrors. And the effect is the NOT operation according to the[se] formulae I gave. From which, we read off the observables at time 2. Finally, the photon hits the second beam splitter and again we can read off “X-hat” of 3 equals minus “Z-hat” of 2, which makes it σ_x , and so on. It turns out that all the observables of the qubit at time 3 are the same as they were initially at time

zero. In particular, $\hat{Z}(3)$ equals $\hat{Z}(0)$. And so $\hat{Z}(3)$ is also sharp. So finally, the photon is travelling only in the plus one direction and there's our prediction.

t	\hat{X}	\hat{Y}	\hat{Z}
0	σ_x	σ_y	σ_z
beam splitter			
1	$-\sigma_z$	σ_y	σ_x
mirrors			
2	σ_z	σ_y	$-\sigma_x$
$\hat{Z}(t+1) = -\hat{Z}(t)$			
$\hat{Y}(t+1) = \hat{Y}(t)$			
$\hat{X}(t+1) = -\hat{X}(t)$			

t	\hat{X}	\hat{Y}	\hat{Z}
0	σ_x	σ_y	σ_z
beam splitter			
1	$-\sigma_z$	σ_y	σ_x
mirrors			
2	σ_z	σ_y	$-\sigma_x$
beam splitter			
3	σ_x	σ_y	σ_z

t	\hat{X}	\hat{Y}	\hat{Z}
0	σ_x	σ_y	σ_z
beam splitter			
1	$-\sigma_z$	σ_y	σ_x
mirrors			
2	σ_z	σ_y	$-\sigma_x$
beam splitter			
3	σ_x	σ_y	σ_z

Table 2.28: It turns out that all the observables [...] at time $t = 3$ [...]

30:48 Remember, there's only one photon participating in this experiment. Consider the moment just before it strikes the second beam splitter. In different universes it's coming from different directions. But just consider the universes it's coming along here, let's say in the North direction. It's approaching the beam splitter.

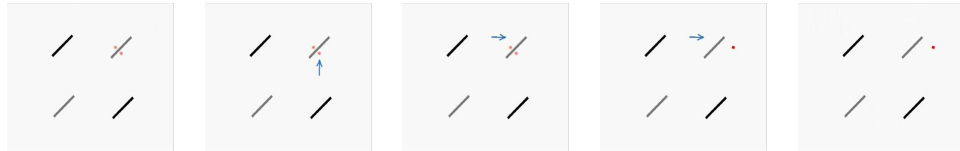


Table 2.29: In this experiment, the photon necessarily [...]

And usually when a photon is approaching a beam splitter it goes through in half of the universes and bounces off in the other half. So usually, if we observe what happens, half the time we detect that it's passed straight through.



Table 2.30: [...] the net effect is that in all the universes [...]

Not in this experiment.

In this experiment, the photon necessarily takes a sharp right turn and is never observed to pass straight through. And similarly, in universes where the photon is approaching from the eastward direction it ignores the possibility of being reflected and just passes straight through. And the net effect is that in all the universes in which the experiment is done, the photon is observed at the East detector and never at the North detector.

32:00

Now I'll show you this experiment in action.

Before I do, I better explain a detail that might be confusing otherwise. The experimentalists like to use the same physical beam splitter for both ends of the experiment. That way they don't have to bother with finding two of them with precisely matched properties and making them exactly parallel to each other and so on. It's hard enough aligning everything as it is. They eliminate the need for one of the beam splitters by slightly changing the geometry like this. Now, this is still the plus 1 direction and this is the minus 1 direction for Z and there are the two corresponding detectors. And our prediction again is that all the photons will come out here in this detector and none in[to] this one.

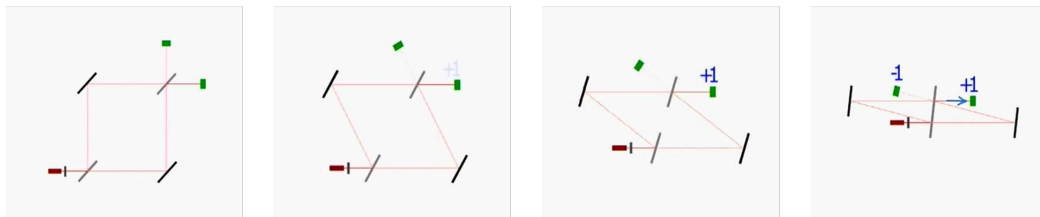


Table 2.31: They eliminate the need for one of the beam splitters [...]

And when we do the experiment that's exactly what we see.

33:03



Table 2.32: And [...] that's exactly what we see.

Each spike here on the oscilloscope represents one photon. The upper trace, the yellow one, is from the detector at the plus 1 position, and the lower trace (the green one) from the minus one position. The few stray spikes here on the minus one trace are due to various imperfections in the apparatus, especially the detector.

What is it that has made the photon in the universes where it approaches the beam splitter northwards turn right? It is the presence of its counterparts in the other universes striking the same point at the same time going East. And we can verify that because if we put an opaque barrier on the eastward path, and let some photons come through, certainly both detectors will be

firing; open up this path again, [and] the North detector goes quiet. Close it off again, and the photons again start reaching the North detector.

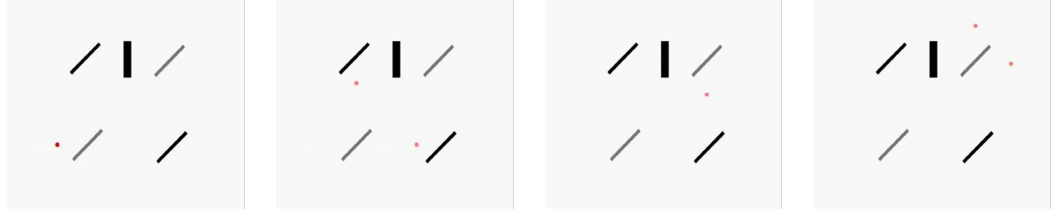


Table 2.33: [...] if we put an opaque barrier [...]



Table 2.34: [...] barrier [...] open up path [...] close it off [...]

Now, let's consider this whole process as a computation.

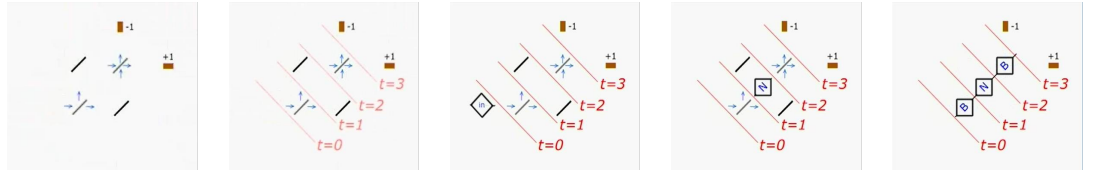


Table 2.35: In this experiment, the photon necessarily [...]

34:45 The input was 1 (one). The computation as a whole consisted of three elementary computations:

- the middle one was a NOT operation.
- the outer two were the same⁶ as each other

And the output was again 1 (one). And you can see from here that the whole computation is an elaborate way of performing the unit operation, otherwise known as doing nothing at all—is that an interesting computation?

⁶Let me call it B for “beam splitter”.

Yes, because just look at what happened here: we performed B followed by NOT followed by B, and the result is the unit operation. So, B, [dot], NOT, [dot], B, equals unit operation. “[dot]” here just means performed after. It’s easy to prove that there is no classical computation using just one bit with this property—no computation such that if you perform it then flip the value of the bit then perform it again, the value of the bit will be unchanged. So B is an example of an elementary computation—an elementary quantum computation—that has no classical analogue. B also has an inverse which formally you can see from here⁷, is the square root of NOT. I’ll show you in later lectures that this isn’t just a manner of speaking. In quantum computation, NOT really does have a square root. Many of the new modes of computation that quantum computers are capable of make use of the root of NOT.

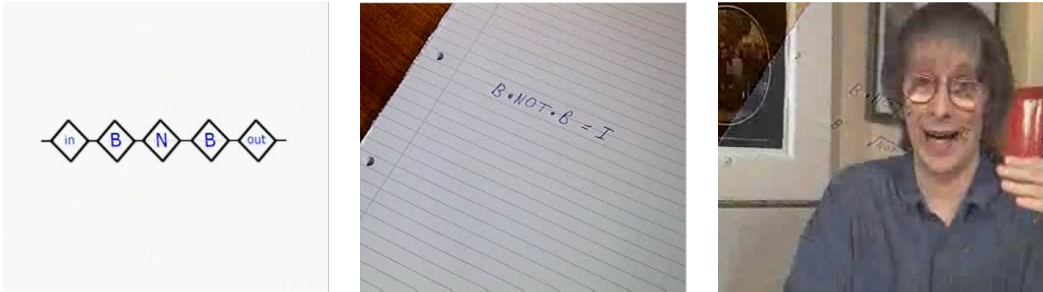


Table 2.36: B also has an inverse which formally [...] is $\sqrt{\text{NOT}}$

In this experiment, we only ever measured \hat{Z} . How would we measure \hat{X} or \hat{Y} ?
Or the qubit’s other observables? We’ve already seen the answer—there. 36:37

<p>$\hat{Z}(t)$</p> <p>A Boolean observable.</p> <p>‘Which of two paths the photon is on at time t’</p>	<p>Effect of beam splitter:</p> <p>$\hat{Z}(t+1) = \hat{X}(t)$</p> <p>$\hat{Y}(t+1) = \hat{Y}(t)$</p> <p>$\hat{X}(t+1) = -\hat{Z}(t)$</p> <p>$\hat{I}(t+1) = \hat{I}(t)$</p>	<p>Effect of beam splitter:</p> <p>$\hat{Z}(t+1) = \hat{X}(t)$</p> <p>$\hat{Y}(t+1) = \hat{Y}(t)$</p> <p>$\hat{X}(t+1) = -\hat{Z}(t)$</p> <p>$\hat{I}(t+1) = \hat{I}(t)$</p>
---	---	---

Table 2.37: B also has an inverse which formally [...] is $\sqrt{\text{NOT}}$

Look at the effect of beam splitter: $\hat{Z}(1) = \hat{X}(0)$ So if we measure \hat{Z} at time 1 (one), that’s the same as measuring \hat{X} at time 0 (zero). So, the act of passing a photon through a beam splitter and then measuring \hat{Z} amounts

⁷ $B^{-1} = \sqrt{\text{NOT}}$

to a measurement of \hat{X} . So, the beam splitter followed by the two detectors constitute a measuring instrument for measuring \hat{X} .

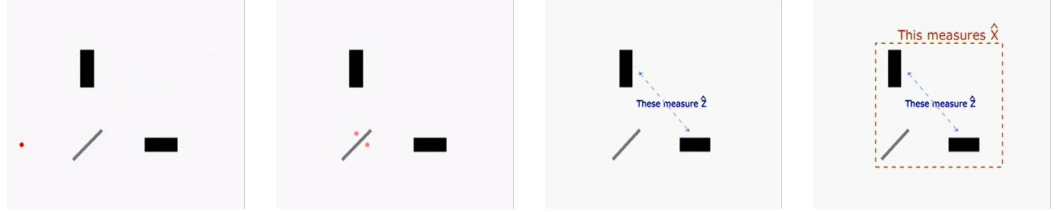


Table 2.38: Look at the effect of [the] beam splitter.

Here again I've been making the simplifying assumption that our qubit observables don't change spontaneously with time.

38:03

But they do, according to this:

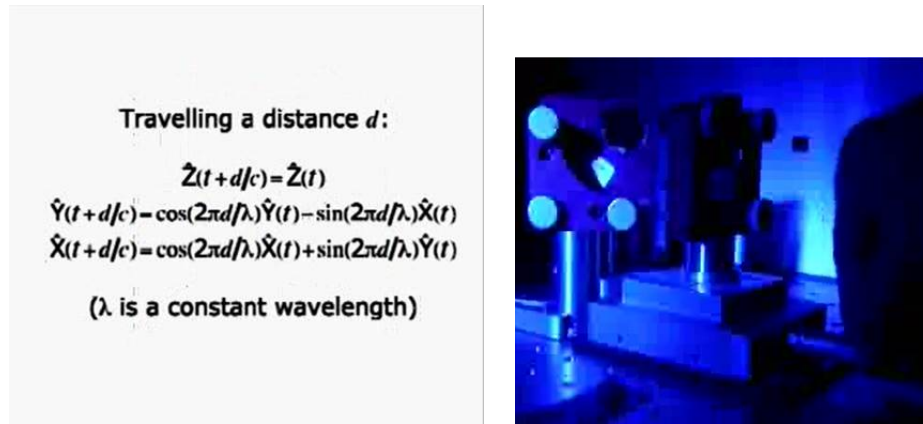


Table 2.39: [...] and in the lab the fine adjustments always [...]

Don't worry about the details, but it follows from this that if we slightly change the path lengths in one of these experiments, and in the lab the fine adjustments always involve doing that anyway, we can easily convert an instrument from measuring \hat{X} into one for measuring \hat{Y} .

That's how to measure \hat{Y} . But you may still be wondering "what is \hat{Y} ?". In the sense that \hat{Z} is the boolean observable for whether the particular photon is traveling in a particular direction. What is the boolean observable \hat{Y} ? Well, I could answer " \hat{Y} is the boolean observable for which of the two directions the photon would travel in if it first passed through a certain beam splitter."

Here we are coming up against the fact that since quantum systems are far richer than everyday language and intuition are adapted to describing, there isn't always a satisfactory terminology for describing quantum systems in ev-

39:14

everyday language. There's a lot more to be said on this issue, but it would take us far beyond the scope of this lecture.

You'll see in the worked examples that similar methods using phase shifts as well as mirrors and beam splitters allow us to measure any of the other observables of the qubit in a similar way. Note that in the general case where the observable being measured is not sharp, we have to make sure that all the instances of the photon in all universes strike a beam splitter, say, at the same moment at the same spot from their two possible directions. Otherwise the photon will be on four possible paths. And the observable for which path it is on is no longer boolean, and so we are no longer talking about a qubit.

40:06

I described this as a single qubit experiment. But to perform it we did need other physical systems aside from the qubit itself. For start, the photon has other subsystems, with observables such as its polarization, or components of its velocity in directions that we made sure it never traveled in. Another physical system we used was the laser to generate the photon. And then the filter to make sure that only one photon at a time was in the apparatus. And then there were the mirrors and the beam splitters. These are all physical systems involving vast numbers of atoms. And even larger numbers of observables, all engaged in intricate interactions. This experiment was carefully arranged so that the net effect of all those interactions on the qubit during the computation could be expressed as a law of motion for the qubit alone. For reasons that I'll explain next time, this would not have been possible, for instance, with the detectors. As I've said, that didn't matter in this experiment because we are not interested in what happens to the photon after its \hat{Z} observable has been measured at the end of the experiment. But to understand experiments involving repeated measurements of the same quantum systems within a quantum computation, we need a quantum analysis of the measurement process.

And that's what the next lecture will be about.

41:48



Table 2.40: On to the next chapter: “The Quantum Theory of Measurement”.

Chapter 3

Measurement

Today, I'm going to talk about the Quantum Theory of Measurement.

Measurement is an important type of quantum computation. It's also, of course, what links quantum theory with experiment. But, for historical reasons, unfortunately, the quantum theory of measurement also plays a prominent role in what we could call the quantum mechanics folklore: the informal misunderstandings of quantum physics that range from careless remarks in textbooks through bad philosophy, and all the way down to complete nonsense.

01:44

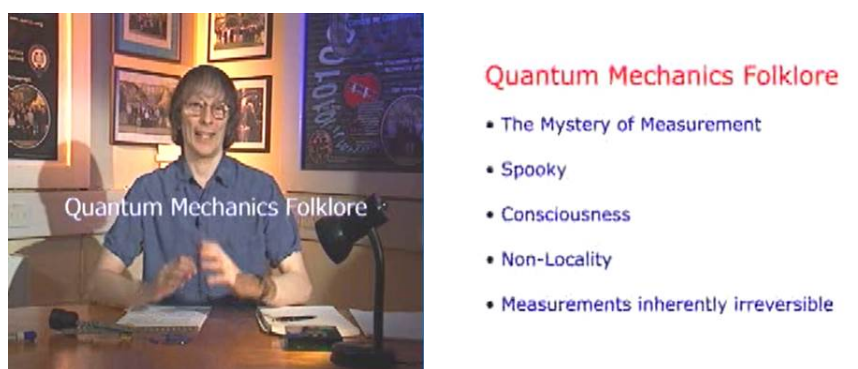


Table 3.1: [...] for historical reasons, unfortunately [...]

You may have heard that the measurement process in quantum theory is deeply mysterious and spooky. That it has a special role for the conscious observer. You may have heard that when you make a measurement in one place, there's an instantaneous effect on quantum systems in a different place. You may have heard that measurements are inherently irreversible processes, even though the laws of motion in fundamental physics are all reversible.

As you'll see in this lecture and later lectures, none of that is true. In fact, quantum measurement theory is pretty well understood and has been a powerful tool for understanding quantum physics. Nowadays, it's part of the

quantum theory of computation. The most elementary, but important, thing to understand about measurements, is that they are physical processes that we understand using the same theory and laws of physics as we do for all other physical processes.

03:22

A measurement is a process in which one physical system, a measuring instrument, interacts with another physical system that is being measured and some observable of the measuring instrument is affected by some observable of the system being measured, and ends up ideally having the same value. Let’s confine our attention for the moment to the useful idealization of a perfect measurement. We can define a perfect measurement process by the effect it has when the observable being measured is sharp—has a single value.

<p>Measurement</p> <p>A physical process in which</p> <p>one physical system (a measuring instrument) interacts</p> <p>with another physical system in such a way that</p> <p>an observable of the first system comes to depend on an observable of the second. (Ideally they become equal.)</p>	<p>Perfect Measurement</p> <ul style="list-style-type: none">• Defined by its effect when the observable being measured is sharp• Its outcome is then the value of that observable• It leaves that value unchanged
---	---

Table 3.2: Let’s confine our attention for the moment to [...]

First, the outcome of a perfect measurement of a sharp observable is the value of that observable. And second, at the end of the measurement, the observable being measured is still sharp with the same value it had before. So, a perfect measurement records sharp values accurately and leaves them unchanged. The simplest possible measurement is the measurement of a boolean observable, where the outcome is stored in the second boolean observable.

Let’s think about that first in the case of a classical computation. Let’s make it a reversible classical computation, partly to substantiate what I said just now about measurement not being inherently irreversible. But mainly because in fact, classical reversible computations are a special case of quantum computation. Imagine that we have a single bit that we are going to measure. And the bit has an unknown value that is either plus one or minus one. Call that unknown value a . And we have a second bit somewhere in an apparatus, in which we want to record a copy of the value in the first bit. The second bit is called the target bit. So the target bit has to end up with the value a . And the bit being measured has to keep the value a .

Considered as a computation, perfect measurement is a process of faithfully copying information, so that where there was one copy of the sharp value of the observable, there are now two: the second one being in the measuring

apparatus. As far as the definition of a perfect measurement is concerned, we are not interested in what the initial value of the target bit was, so long as the final value is a . But in a reversible computation, different inputs must always produce different outputs. So it follows that if the computation only involves those two bits, it can only be a perfect measurement for one initial value of the target bit—let's say for the value $+1$.

06:31

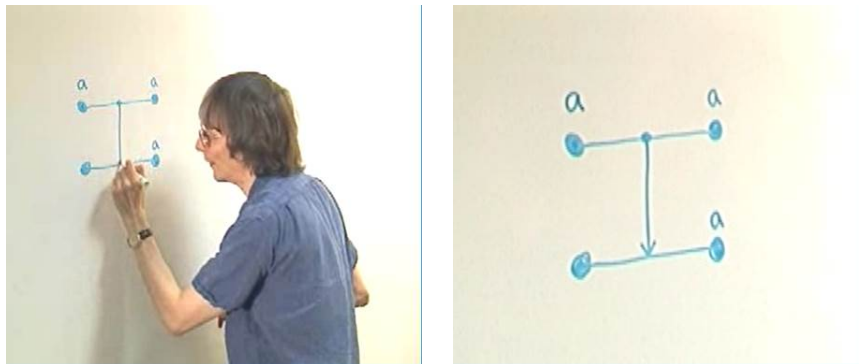


Table 3.3: As far as the definition of a perfect measurement is concerned [...]

So it's not a perfect measurement when the target bit starts at -1 . What does happen if it's -1 depends on the law of motion. And one interesting law of motion is that of a very useful operation called "Controlled NOT". The ordinary NOT operation is the single bit, or single qubit operation, that figured in the interference experiment I discussed last time. It flips $+1$ to -1 and vice versa.

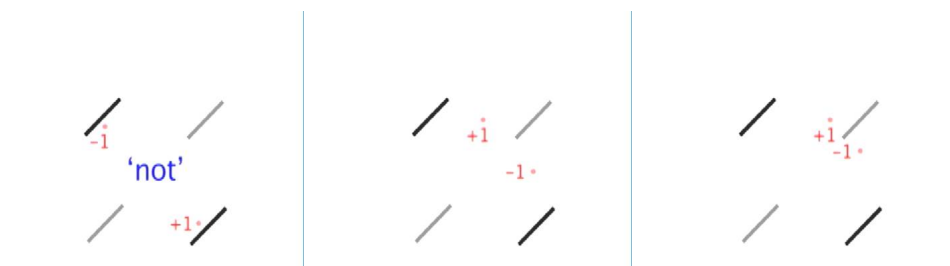


Table 3.4: The ordinary NOT operation is the [...] single qubit operation [...]

Controlled NOT means flip the value of the target bit only if the first bit is -1 . The first bit is now called the control bit. If the control bit is $+1$, the target bit remains unaffected. So, we can summarize this controlled NOT operation like this: if the input values of the control and target bits are a and b , then the outputs are a and ab . A computer like this that performs a simple computation on a fixed number of bits, and completes it in a fixed time, is called a computational gate, also known as a logic gate. And in this case it's

a reversible logic gate because it's performing a reversible computation—the controlled NOT operation.

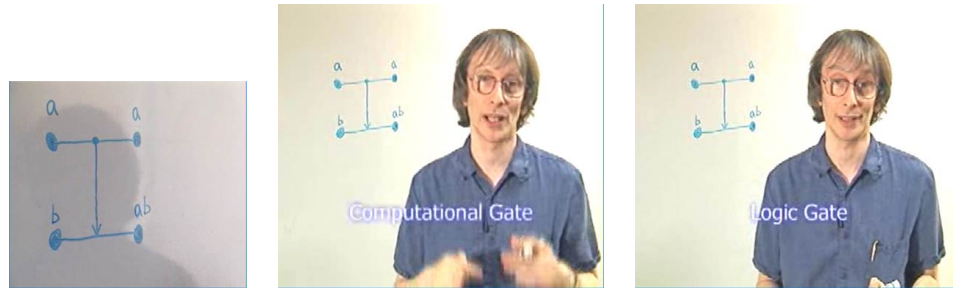


Table 3.5: A computer like this [...] is called [...]

08:20 Let me just summarize some of the ways in which we can think about what the controlled NOT gate does. First of all, if $b = +1$, then considering it as a physical process it's a perfect measurement. The first bit starts with an unknown value, a , and the second bit ends up holding the measured value of the first bit, which would also be a . Second, considered as a computation, if $b = +1$, then the process copies information. The information a , which starts out in only one of the bits, ends up in both of them.

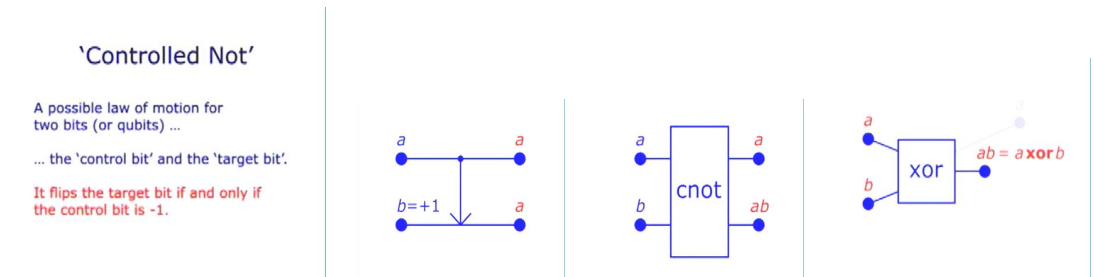


Table 3.6: And, as I said, like all classical reversible gates[...]

Third, for an arbitrary b , well, that's the controlled NOT operation: the target bit is flipped if and only if the control bit is -1 . And a fourth interpretation, this gate is a reversible version of the classical exclusive OR gate. If we consider $+1$ as standing for false and -1 as for true, then a times b is the same thing as the exclusive OR of a and b . And the target bit ends up as the exclusive OR of the two inputs. And as I said, like all classical reversible gates, the controlled NOT gate has a quantum implementation that works on two qubits instead of two bits—that's the quantum controlled NOT gate.

In some implementations, there's a physical object you can identify as the gate. As with the mirror in the interference experiment last time, which was the NOT gate, and the beam splitter which was a fractional power of NOT. But

speaking more precisely, the computational gate is not the object that makes the process happen, but the dynamical process itself that is undergone by the qubit or qubits that physically realizes a rule that defines each output in terms of the inputs, and nothing else.

10:49

A quantum gate...

Is a physical process whose law of motion implements a rule that defines each output qubit in terms of the input qubits.

The diagram shows a CNOT gate. On the left, two input qubits are labeled \hat{Z}_1 and \hat{Z}_2 . They enter a rectangular box labeled 'cnot'. Above the box, the time is marked as $t=0$ on the left and $t=1$ on the right. Two output qubits exit the box to the right. The top output line is connected to the top input line, and the bottom output line is connected to the bottom input line, representing a controlled operation.

$$\hat{X}_1(t), \hat{Y}_1(t), \hat{Z}_1(t), \hat{1}$$

$$[\hat{X}_1(t), \hat{Y}_1(t)] = i\hat{Z}_1(t)$$

$$[\hat{Y}_1(t), \hat{Z}_1(t)] = i\hat{X}_1(t)$$

$$[\hat{Z}_1(t), \hat{X}_1(t)] = i\hat{Y}_1(t)$$

Qubit 1

$$\hat{X}_1(t), \hat{Y}_1(t), \hat{Z}_1(t), \hat{1}$$

$$[\hat{X}_1(t), \hat{Y}_1(t)] = i\hat{Z}_1(t)$$

$$[\hat{Y}_1(t), \hat{Z}_1(t)] = i\hat{X}_1(t)$$

$$[\hat{Z}_1(t), \hat{X}_1(t)] = i\hat{Y}_1(t)$$

Qubit 2

$$\hat{X}_2(t), \hat{Y}_2(t), \hat{Z}_2(t), \hat{1}$$

$$[\hat{X}_2(t), \hat{Y}_2(t)] = i\hat{Z}_2(t)$$

$$[\hat{Y}_2(t), \hat{Z}_2(t)] = i\hat{X}_2(t)$$

$$[\hat{Z}_2(t), \hat{X}_2(t)] = i\hat{Y}_2(t)$$

$$[\hat{Z}_1(t), \hat{X}_2(t)] = ?$$

Table 3.7: So, as in the one-qubit interference experiment [...]

Now, to describe the quantum physical system consisting of two qubits passing through a quantum controlled NOT gate, we are going to make the \hat{Z} observable of the control qubit control the \hat{Z} observable of the target qubit. To describe that quantum system, I have to tell you as always the static constitution, the dynamics, and on each particular occasion when the experiment is done, the state which summarizes how the qubits were prepared as the input. Now I'll start as usual, with the static constitution. For this, I have to tell you the algebra of all the observables at any one time, which will then be the same as the algebra at any other time because the algebra summarizes the time invariant features of a quantum system. So, here we have a two-qubit quantum system. We already know what the algebra of the observables of either one of the two qubits is. Because that algebra is itself an invariant—it's the same for all qubits and regardless of what interactions the qubit happens to be undergoing. So, as in the one-qubit interference experiment, the first qubit will have observables \hat{X} , \hat{Y} and \hat{Z} which have the same algebra as the Pauli matrices, together with the unit observable, and linear combinations of those with constant coefficients. Let me call these observables \hat{X}_1 , \hat{Y}_1 and \hat{Z}_1 now, where the suffix-es indicate that they are observables of qubit number 1 (one), the controlled qubit of the quantum controlled NOT gate. We don't have to put a suffix on the unit observable, because as you will remember the unit observable can be measured without even referring to the system it is an observable of. The second qubit, the target qubit, will have different observables, \hat{X}_2 , \hat{Y}_2 , \hat{Z}_2 and so on, but they will also have the Pauli algebra.

12:39

Now, I have to specify all the additional algebraic relations that may hold involving observables from both qubits. And here's a universal rule that defines the algebra of any composite quantum system given the algebra of its

constituent systems. It just says that the observables of different quantum systems commute with each other. That is to say that if \hat{A}_1 is an observable of one system and \hat{B}_2 is an observable of another, then $\hat{A}_1 \cdot \hat{B}_2$ equals $\hat{B}_2 \cdot \hat{A}_1$.

Qubit 1	Qubit 2	Qubit 1	Qubit 2	Qubit 1	Qubit 2
$\hat{X}_1(t), \hat{Y}_1(t), \hat{Z}_1(t), \hat{1}$	$\hat{X}_2(t), \hat{Y}_2(t), \hat{Z}_2(t), \hat{1}$	$\hat{X}_1(t), \hat{Y}_1(t), \hat{Z}_1(t), \hat{1}$	$\hat{X}_2(t), \hat{Y}_2(t), \hat{Z}_2(t), \hat{1}$	$\hat{X}_1(t), \hat{Y}_1(t), \hat{Z}_1(t), \hat{1}$	$\hat{X}_2(t), \hat{Y}_2(t), \hat{Z}_2(t), \hat{1}$
$[\hat{X}_1(t), \hat{Y}_1(t)] = i\hat{Z}_1(t)$	$[\hat{X}_2(t), \hat{Y}_2(t)] = i\hat{Z}_2(t)$	$[\hat{X}_1(t), \hat{Y}_1(t)] = i\hat{Z}_1(t)$	$[\hat{X}_2(t), \hat{Y}_2(t)] = i\hat{Z}_2(t)$	$[\hat{X}_1(t), \hat{Y}_1(t)] = i\hat{Z}_1(t)$	$[\hat{X}_2(t), \hat{Y}_2(t)] = i\hat{Z}_2(t)$
$[\hat{Y}_1(t), \hat{Z}_1(t)] = i\hat{X}_1(t)$	$[\hat{Y}_2(t), \hat{Z}_2(t)] = i\hat{X}_2(t)$	$[\hat{Y}_1(t), \hat{Z}_1(t)] = i\hat{X}_1(t)$	$[\hat{Y}_2(t), \hat{Z}_2(t)] = i\hat{X}_2(t)$	$[\hat{Y}_1(t), \hat{Z}_1(t)] = i\hat{X}_1(t)$	$[\hat{Y}_2(t), \hat{Z}_2(t)] = i\hat{X}_2(t)$
$[\hat{Z}_1(t), \hat{X}_1(t)] = i\hat{Y}_1(t)$	$[\hat{Z}_2(t), \hat{X}_2(t)] = i\hat{Y}_2(t)$	$[\hat{Z}_1(t), \hat{X}_1(t)] = i\hat{Y}_1(t)$	$[\hat{Z}_2(t), \hat{X}_2(t)] = i\hat{Y}_2(t)$	$[\hat{Z}_1(t), \hat{X}_1(t)] = i\hat{Y}_1(t)$	$[\hat{Z}_2(t), \hat{X}_2(t)] = i\hat{Y}_2(t)$
$[\hat{Z}_1(t), \hat{X}_2(t)] = ?$		$[\hat{Z}_1(t), \hat{X}_2(t)] = 0$		$[\hat{A}_1(t), \hat{B}_2(t)] = 0$	

Table 3.8: And here's a universal rule that defines the algebra of [...]

14:12

I told you last time that 2×2 matrix representations of a qubit algebra wouldn't always be enough. Here's why: This [entire] set of algebraic relations (for both qubit 1 and qubit 2) can't be faithfully represented by 2×2 matrices. This set (for just qubit 1) by itself can, and so can this set (for qubit 2). But because this computation relation (between \hat{A}_1 and \hat{B}_2) for the combined system has to be represented as well, qubits 1 and 2 can't use the same set of matrices. Take \hat{Z}_2 for instance: it has to commute for every observable of qubit 1. But the only 2×2 matrices that commute with every 2×2 matrix are multiples of the unit matrix. And \hat{Z}_2 can't be represented by a multiple of the unit matrix because it's a boolean observable—it has to have 2 distinct eigenvalues and multiples of the unit matrix only have one eigenvalue.

Qubit 1	Qubit 2	Qubit 1	Qubit 2
$\hat{X}_1(t), \hat{Y}_1(t), \hat{Z}_1(t), \hat{1}$	$\hat{X}_2(t), \hat{Y}_2(t), \hat{Z}_2(t), \hat{1}$	$\hat{X}_1(t), \hat{Y}_1(t), \hat{Z}_1(t), \hat{1}$	$\hat{X}_2(t), \hat{Y}_2(t), \hat{Z}_2(t), \hat{1}$
$[\hat{X}_1(t), \hat{Y}_1(t)] = i\hat{Z}_1(t)$	$[\hat{X}_2(t), \hat{Y}_2(t)] = i\hat{Z}_2(t)$	$[\hat{X}_1(t), \hat{Y}_1(t)] = i\hat{Z}_1(t)$	$[\hat{X}_2(t), \hat{Y}_2(t)] = i\hat{Z}_2(t)$
$[\hat{Y}_1(t), \hat{Z}_1(t)] = i\hat{X}_1(t)$	$[\hat{Y}_2(t), \hat{Z}_2(t)] = i\hat{X}_2(t)$	$[\hat{Y}_1(t), \hat{Z}_1(t)] = i\hat{X}_1(t)$	$[\hat{Y}_2(t), \hat{Z}_2(t)] = i\hat{X}_2(t)$
$[\hat{Z}_1(t), \hat{X}_1(t)] = i\hat{Y}_1(t)$	$[\hat{Z}_2(t), \hat{X}_2(t)] = i\hat{Y}_2(t)$	$[\hat{Z}_1(t), \hat{X}_1(t)] = i\hat{Y}_1(t)$	$[\hat{Z}_2(t), \hat{X}_2(t)] = i\hat{Y}_2(t)$
$[\hat{A}_1(t), \hat{B}_2(t)] = 0$		$[\hat{A}_1(t), \hat{B}_2(t)] = 0$	



Table 3.9: And here's a universal rule that defines the algebra of [...]

So, there is no 2×2 matrix representation of this algebra as a whole. The simplest representation turns out to be a 4×4 representation. And the way it is constructed applies to any two quantum systems, not just a pair of qubits, that you want to consider as a single system. It uses the tensor product of matrices. The tensor product of two matrices of dimensions m and n is an

n -dimensional matrix consisting of all possible products of pairs of elements, one from the first matrix, and one from the second, like this:

$$\begin{array}{|c|c|c|c|} \hline \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix} \\ \hline \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix} \\ \hline \end{array}$$

Table 3.10: The tensor product of two matrices [...]

That \otimes symbol denotes the tensor product. So the tensor product of a pair of 2×2 matrices is a matrix of all 16 possible products consisting of an element of the first matrix multiplied by an element of the second. You can verify in the worked examples that the tensor product of two Hermitian matrices is a Hermitian matrix. So the tensor product of two observables, one from each system, is an observable of the combined system.

16:42

What observable is it? Well, if you have two systems, and \hat{A} is a matrix representing an observable of System 1 by itself, and \hat{B} is a matrix representing System 2 by itself, then the tensor product of \hat{A} cross \hat{B} is an observable you measure by measuring \hat{A} on the first system and \hat{B} on the second, and then multiplying the result together.

$$\begin{array}{|c|c|c|c|} \hline \begin{array}{cc} \boxed{\hat{A}} & \boxed{\hat{B}} \\ \text{System 1} & \text{System 2} \end{array} & \begin{array}{cc} \hat{A} \otimes \hat{B} & \\ \boxed{\phantom{\hat{A}}} & \boxed{\phantom{\hat{B}}} \\ \text{System 1} & \text{System 2} \end{array} & \begin{array}{cc} \hat{A} \otimes \hat{B} \neq \hat{B} \otimes \hat{A} & \\ \boxed{\phantom{\hat{A}}} & \boxed{\phantom{\hat{B}}} \\ \text{System 1} & \text{System 2} \end{array} & (\hat{A} \otimes \hat{B})(\hat{C} \otimes \hat{D}) = (\hat{A}\hat{C}) \otimes (\hat{B}\hat{D}) \\ \hline \end{array}$$

Table 3.11: Well, if you have two systems [...]

You can verify that the tensor product operation is associative, but it's not commutative. In other words, \hat{A} cross \hat{B} is not the same as \hat{B} cross \hat{A} . You can also verify that the tensor product has the following property with regard

18:04 to ordinary matrix multiplication: \hat{A} cross \hat{B} times \hat{C} cross \hat{D} equals \hat{A} [times] \hat{C} cross \hat{B} [times] \hat{D} . Now, it follows from all that, that using the tensor product we can make a 4×4 representation of the algebra of observables of qubit 1 (one) as follows: just multiply each matrix of any 2×2 representation by the 2×2 unit matrix on the right. These are now 4×4 matrices.

$\begin{array}{c} \text{2x2 representation} \\ \hat{X}_1(0) = \sigma_x \\ \hat{Y}_1(0) = \sigma_y \\ \hat{Z}_1(0) = \sigma_z \end{array}$	$\begin{array}{c} \text{2x2 representation} \\ \hat{X}_1(0) = \sigma_x \otimes \mathbf{I} \\ \hat{Y}_1(0) = \sigma_y \otimes \mathbf{I} \\ \hat{Z}_1(0) = \sigma_z \otimes \mathbf{I} \end{array}$	$\begin{array}{c} \text{4x4 representation} \\ \text{2x2 representation} \\ \hat{X}_1(0) = \sigma_x \otimes \mathbf{I} \\ \hat{Y}_1(0) = \sigma_y \otimes \mathbf{I} \\ \hat{Z}_1(0) = \sigma_z \otimes \mathbf{I} \end{array}$
---	--	---

Table 3.12: Now it follows from all that [...]

And because of this property¹, it's easy to show that they do indeed represent the standard algebra for one qubit. For qubit 2 (two), we can do a similar thing, this time multiplying by the unit matrix on the left.

$\begin{array}{c} \text{4x4 representation} \\ \hat{X}_2(0) = \mathbf{I} \otimes \sigma_x \\ \hat{Y}_2(0) = \mathbf{I} \otimes \sigma_y \\ \hat{Z}_2(0) = \mathbf{I} \otimes \sigma_z \end{array}$	$\begin{array}{c} \text{4x4 representation} \\ \hat{X}_2(0) = \mathbf{I} \otimes \sigma_x \\ \hat{Y}_2(0) = \mathbf{I} \otimes \sigma_y \\ \hat{Z}_2(0) = \mathbf{I} \otimes \sigma_z \end{array}$	$\begin{array}{c} \text{4x4 representation} \\ \hat{X}_1(0) = \sigma_x \otimes \mathbf{I} \\ \hat{Y}_1(0) = \sigma_y \otimes \mathbf{I} \\ \hat{Z}_1(0) = \sigma_z \otimes \mathbf{I} \end{array}$	$\begin{array}{c} \text{4x4 representation} \\ \hat{X}_2(0) = \mathbf{I} \otimes \sigma_x \\ \hat{Y}_2(0) = \mathbf{I} \otimes \sigma_y \\ \hat{Z}_2(0) = \mathbf{I} \otimes \sigma_z \end{array}$
		$\hat{X}_1(0)\hat{Y}_2(0) = (\sigma_x \otimes \mathbf{I})(\mathbf{I} \otimes \sigma_y) = \sigma_x \otimes \sigma_y$	$\hat{Y}_2(0)\hat{X}_1(0) = (\mathbf{I} \otimes \sigma_y)(\sigma_x \otimes \mathbf{I}) = \sigma_x \otimes \sigma_y$

Table 3.13: For qubit 2 (two) we can do a similar thing [...]

They too form a 4×4 representation of the standard algebra of a single qubit. And, the object of the exercise: any observable of the first qubit commutes with any observable of the second. For instance, $\hat{X}_1\hat{Y}_2$ at time 0 (zero) is σ_x cross 1 (one) times one cross σ_y which equals σ_x cross σ_y and, taking the other way around, $\hat{Y}_2\hat{X}_1$ equals $\mathbf{I} \otimes \sigma_y$ times $\sigma_x \otimes \mathbf{I}$ which also equals $\sigma_x \otimes \sigma_y$.

19:24 By the way, like in the single qubit 2×2 case, we need never work with the actual components of these matrices. Every 4×4 Hermitian matrix can be

¹ $(\hat{A} \otimes \hat{B})(\hat{C} \otimes \hat{D}) = (\hat{A}\hat{C}) \otimes (\hat{B}\hat{D})$

expressed as a linear combination of tensor products of the form sigma cross sigma, and sigma \otimes I (one) and $I \otimes \sigma$, and $I \otimes I$, which is the 4×4 unit matrix.

Every 4x4 Hermitian matrix is a linear combination, with real coefficients, of these 16 matrices:			
$\sigma_x \otimes \sigma_x$	$\sigma_x \otimes \sigma_y$	$\sigma_x \otimes \sigma_z$	$\sigma_x \otimes I$
$\sigma_y \otimes \sigma_x$	$\sigma_y \otimes \sigma_y$	$\sigma_y \otimes \sigma_z$	$\sigma_y \otimes I$
$\sigma_z \otimes \sigma_x$	$\sigma_z \otimes \sigma_y$	$\sigma_z \otimes \sigma_z$	$\sigma_z \otimes I$
$I \otimes \sigma_x$	$I \otimes \sigma_y$	$I \otimes \sigma_z$	$I \otimes I$

Table 3.14: By the way [...] we need never work with [...]

So again, we can just express all 4×4 matrices, and so all observables in terms of Pauli matrices and do all matrix arithmetic in terms of the algebra of Pauli matrices. So, I've told you the static constitution of a pair of qubits, which would be the same for any two-qubit gate. Now, the dynamics of the controlled NOT gate in particular. As usual, we will imagine that it performs its operation in one unit of time—one computational step. For the moment, we are considering the gate as an elementary quantum computation so we are not interested in its internal workings, we are only interested in how the output depends on the input. So, for present purposes, the dynamics of the gate just means how the observables at time 1 after the gate has acted depend on the observables at time 0. Well, the dynamics of a controlled NOT gate are defined by a set of equations, most of whose details are not relevant for present purposes, but I'll put them on the screen anyway just so I can point out the highlights that are relevant.

21:37

<p>Dynamics of a Controlled-Not Gate</p> $\begin{aligned} \hat{X}_1(1) &= \hat{X}_1(0)\hat{X}_2(0) \\ \hat{Y}_1(1) &= \hat{Y}_1(0)\hat{X}_2(0) \\ \hat{Z}_1(1) &= \hat{Z}_1(0) \\ \hat{X}_2(1) &= \hat{X}_2(0) \\ \hat{Y}_2(1) &= \hat{Z}_1(0)\hat{Y}_2(0) \\ \hat{Z}_2(1) &= \hat{Z}_1(0)\hat{Z}_2(0) \end{aligned}$	

Table 3.15: [...] the dynamics of a controlled NOT gate [...]

These six equations tell us how six representative observables change—three of them from one qubit and three from the other. We don't have to list the

unit observable exclusively because it never changes. Using these equations, we can find out how any other observable behaves as well by expressing it at time 0 in terms of these representative observables, and then using the fact that algebraic relations between observables at a given time don't change with time. These equations are actually redundant. We could make due with just four of them, because for instance, the static constitution for any qubit already defines its observable $\hat{Z}(t)$ at any given time in terms of its observables $\hat{X}(t)$ and $\hat{Y}(t)$ at that time. Now, take a look at the equations for the time evolution of just the \hat{Z} observables. You can see that \hat{Z}_1 and \hat{Z}_2 at time 1 depend only on \hat{Z}_1 and \hat{Z}_2 at time 0. The \hat{X} 's and \hat{Y} 's don't affect the evolution of the \hat{Z} 's, so the \hat{Z} observables almost form a physical system in their own right—a little subsystem of the qubit system that evolves independently of the rest. It doesn't quite count as a separate physical system because you can never change \hat{Z} without changing \hat{X} or \hat{Y} .

<div>Dynamics of a Controlled-Not Gate</div> <div>$\hat{X}_1(1) = \hat{X}_1(0)\hat{X}_2(0)$$\hat{Y}_1(1) = \hat{Y}_1(0)\hat{X}_2(0)$$-i[\hat{X}_1(1), \hat{Y}_1(1)] = \hat{Z}_1(1) = \hat{Z}_1(0)$</div> <div>$\hat{X}_2(1) = \hat{X}_2(0)$$\hat{Y}_2(1) = \hat{Z}_1(0)\hat{Y}_2(0)$$-i[\hat{X}_2(1), \hat{Y}_2(1)] = \hat{Z}_2(1) = \hat{Z}_1(0)\hat{Z}_2(0)$</div>	<div>Dynamics of a Controlled-Not Gate</div> <div>$\hat{Z}_1(1) = \hat{Z}_1(0)$</div> <div>$\hat{Z}_2(1) = \hat{Z}_1(0)\hat{Z}_2(0)$</div>	<div>Dynamics of a Controlled-Not Gate</div> <div>$\hat{Y}_2(1) = \hat{Z}_1(0)\hat{Y}_2(0)$</div>
--	--	--

Table 3.16: These equations are actually redundant.

Nevertheless, in a controlled NOT computation, the \hat{Z} observables are autonomous: they are evolving independently of all other observables. I've drawn your attention to that because it will come up again in later lectures, but for the moment its only significance is that it makes it easy for us to check that this gate really does perform a perfect measurement.

23:50

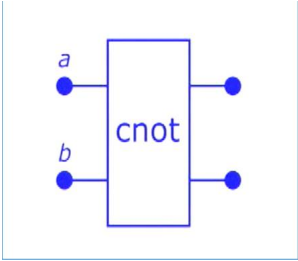
<div>Perfect Measurement</div> <div><ul style="list-style-type: none">Defined by its effect when the observable being measured is sharpIts outcome is then the value of that observableIt leaves that value unchanged</div>	<div></div>
---	---

Table 3.17: [...] but, for the moment, the only significance [...]

Remember, a perfect measurement interaction is defined by what it does when the observable being measured is sharp. So, now it's time for me to specify a

convenient state for our two qubits in which the \hat{Z} observables of both qubits are sharp. There are only four possible cases of this, so we may as well deal with all four of them wholesale. Let the observable \hat{Z}_1 be sharp with the value a , and let \hat{Z}_2 be sharp with the value b ; a and b are each both plus and/or minus 1 (one). So, to specify the state, or actually four possible states, we have to specify the expectation values of all the observables. And I've just said that the expectation value of $\hat{Z}_1(0)$ is equal to a , and the expectation value of $\hat{Z}_2(0)$ equals b . As I mentioned in the last lecture, and you will have proved in the work examples, \hat{Z} being maximally sharp implies that \hat{X} and \hat{Y} are both minimally sharp. In other words, that their expectation values are both zero.

This in turn tells us how to find expectation values of Pauli matrices in this state, in the representation we've chosen for the observables. We have:

<p>The general state in which the Z-observables of two qubits are sharp</p> $\begin{aligned}\langle \hat{X}_1(0) \rangle &= 0 \\ \langle \hat{Y}_1(0) \rangle &= 0 \\ \langle \hat{Z}_1(0) \rangle &= a \\ \langle \hat{X}_2(0) \rangle &= 0 \\ \langle \hat{Y}_2(0) \rangle &= 0 \\ \langle \hat{Z}_2(0) \rangle &= b\end{aligned}$	<p>The general state in which the Z-observables of two qubits are sharp</p> $\begin{aligned}\langle \sigma_x \otimes \mathbf{I} \rangle &= \langle \hat{X}_1(0) \rangle = 0 \\ \langle \sigma_y \otimes \mathbf{I} \rangle &= \langle \hat{Y}_1(0) \rangle = 0 \\ \langle \sigma_z \otimes \mathbf{I} \rangle &= \langle \hat{Z}_1(0) \rangle = a \\ \langle \mathbf{I} \otimes \sigma_x \rangle &= \langle \hat{X}_2(0) \rangle = 0 \\ \langle \mathbf{I} \otimes \sigma_y \rangle &= \langle \hat{Y}_2(0) \rangle = 0 \\ \langle \mathbf{I} \otimes \sigma_z \rangle &= \langle \hat{Z}_2(0) \rangle = b\end{aligned}$	<p>The general state in which the Z-observables of two qubits are sharp</p> $\begin{aligned}\langle \sigma_x \otimes \mathbf{I} \rangle &= \langle \hat{X}_1(0) \rangle = 0 \\ \langle \sigma_y \otimes \mathbf{I} \rangle &= \langle \hat{Y}_1(0) \rangle = 0 \\ \langle \sigma_z \otimes \mathbf{I} \rangle &= \langle \hat{Z}_1(0) \rangle = a \\ \langle \mathbf{I} \otimes \sigma_x \rangle &= \langle \hat{X}_2(0) \rangle = 0 \\ \langle \mathbf{I} \otimes \sigma_y \rangle &= \langle \hat{Y}_2(0) \rangle = 0 \\ \langle \mathbf{I} \otimes \sigma_z \rangle &= \langle \hat{Z}_2(0) \rangle = b\end{aligned}$
--	--	--

Table 3.18: [...] but, for the moment, the only significance [...]

By linearity, this tells us all the expectation values of matrices of the form ‘something’ \otimes (cross) I (one) and I (one) \otimes ‘something’. You’ll see in the work examples that these relations also tell us that in this state, the expectation values of any tensor product of the form $A \otimes B$ where A and B are 2×2 matrices, is the product of the expectation values.

<p>In our state, and in the representation we have chosen...</p> $\langle A_1 \otimes B_2 \rangle = \langle A_1 \otimes \mathbf{I} \rangle \langle \mathbf{I} \otimes B_2 \rangle$
--

Table 3.19: [...] these relations also tell us that in, this state [...]

With these rules, we can find expectation values of general matrices and hence of general observables in this state. So, what do we predict for the output of the controlled NOT gate where the inputs were sharp? Well, the controlled observable \hat{Z}_1 has the same expectation value at time 1 as it did at time 0, namely a . And since a is plus or minus one, that means that \hat{Z}_1 is still sharp with the value a , which is correct. \hat{Z}_2 is sharp as well, but its value has changed—it's now ab —just like the target bit of a classical controlled NOT gate. OK, the next obvious thing to work out is: ‘what happens if we perform a perfect measurement of an observable that’s not sharp?’ Well, you can prove in the work examples that if \hat{Z}_1 isn’t sharp at time 0, it will be just as unsharp at time 1. And also \hat{Z}_2 will have become just as unsharp as well.

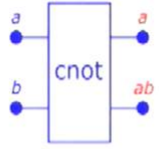
 $\begin{aligned}\langle \hat{Z}_1(\mathbf{1}) \rangle &= \langle \hat{Z}_1(\mathbf{0}) \rangle = \langle \sigma_z \otimes \mathbf{I} \rangle = a \\ \langle \hat{Z}_2(\mathbf{1}) \rangle &= \langle \hat{Z}_1(\mathbf{0}) \hat{Z}_2(\mathbf{0}) \rangle \\ &= \langle \sigma_z \otimes \sigma_z \rangle = \langle \sigma_z \otimes \mathbf{I} \rangle \langle \mathbf{I} \otimes \sigma_z \rangle = ab\end{aligned}$	<p>What happens if the observable $\langle \hat{Z}_1(\mathbf{0}) \rangle$ of the control qubit of a controlled-not gate is not sharp?</p> <p>(I.e. if $\langle \hat{Z}_1(\mathbf{0}) \rangle \neq \pm 1$)</p> $\begin{aligned}\langle \hat{Z}_1(\mathbf{1}) \rangle &= \langle \hat{Z}_1(\mathbf{0}) \rangle \\ \langle \hat{Z}_2(\mathbf{1}) \rangle &= \langle \hat{Z}_1(\mathbf{0}) \rangle\end{aligned}$
---	---

Table 3.20: [...] we predict for the output of the [...]

27:30 So, the measurement interaction propagates unsharpness from one system to the other. But there’s more: consider the observable $\hat{Z}_1\hat{Z}_2$.

$$\hat{Z}_1(\mathbf{1})\hat{Z}_2(\mathbf{1})$$

Table 3.21: In other words, it’s a boolean observable [...]

That is an observable because \hat{Z}_1 and \hat{Z}_2 commute at any one time, so $\hat{Z}_1\hat{Z}_2$ is Hermitian; and from what I’ve said, its operational meaning is that it’s the observable for what you’d get if you measured \hat{Z}_1 and \hat{Z}_2 and multiplied the outcomes together. In other words, it’s a boolean observable whose eigenvalue

-1 means that the outcomes of measuring \hat{Z}_1 and \hat{Z}_2 would be different, and $+1$ means that those outcomes would be the same. Now, here's a remarkable thing. Starting in a state where \hat{Z}_1 is not sharp at time 0, and hence both \hat{Z}_1 and \hat{Z}_2 are unsharp at time 1, [let's] calculate the expectation value of $\hat{Z}_1\hat{Z}_2$. You'll find it's $+1$, which means that the observable $\hat{Z}_1\hat{Z}_2$ is sharp. In other words, \hat{Z}_1 and \hat{Z}_2 are equal at the end of the measurement—sharply equal. Even though neither of them have a sharp value. How can two things be perfectly equal without either of them being sharp? Well, like this of course:


$\hat{Z}_1(1)\hat{Z}_2(1)$	$ \begin{aligned} \langle \hat{Z}_1(1)\hat{Z}_2(1) \rangle &= \langle \hat{Z}_1(1)(\hat{Z}_1(0)\hat{Z}_2(0)) \rangle \\ &= \langle \hat{Z}_1(0)^2\hat{Z}_2(0) \rangle \\ &= \langle \hat{Z}_2(0) \rangle \\ &= \langle \mathbf{I} \otimes \sigma_z \rangle \\ &= 1 \end{aligned} $	
----------------------------	--	---

Table 3.22: How can two things be perfectly equal without either of them [...]

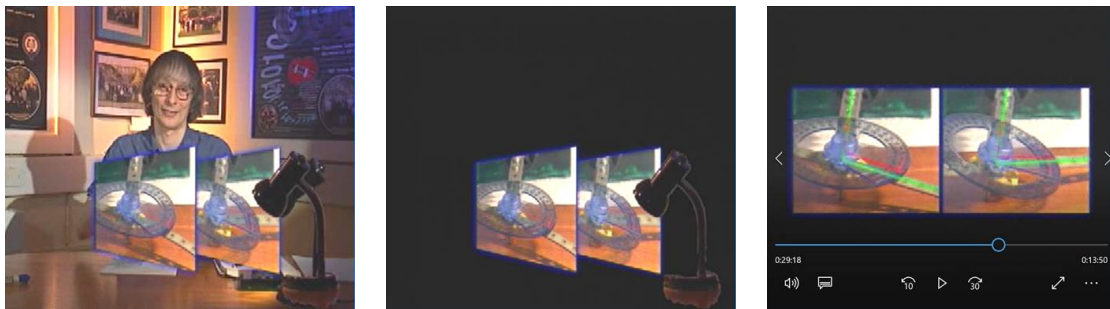


Table 3.23: [...] Well, like this, of course!

Finally, let's look at the case where this measurement of an unsharp observable takes place in the middle of an attempted interference experiment.

29:28

Let's work out what happens when we combine the analysis of the previous lecture of the interference experiment with this time's analysis of the dynamics of measurement. Let's take a qubit, based on the photons direction of motion, and pass it through a beam splitter to make $\hat{Z}(1)$ non-sharp.

So, from the equations of motion, $\hat{Z}(1)$ takes this form in the 2×2 representation of last time. But now, we are going to measure the direction of motion

so we need a second qubit and therefore a 4×4 representation, like this:

Effect of beam splitter:	Effect of beam splitter:	Effect of beam splitter:	Effect of beam splitter:
$\hat{Z}(t+1) = \hat{X}(t)$ $\hat{Y}(t+1) = \hat{Y}(t)$ $\hat{X}(t+1) = -\hat{Z}(t)$	$\hat{Z}(1) = \hat{X}(0)$ $\hat{Y}(1) = \hat{Y}(0)$ $\hat{X}(1) = -\hat{Z}(0)$	$\hat{Z}(1) = \hat{X}(0) = \sigma_x$ $\hat{Y}(1) = \hat{Y}(0)$ $\hat{X}(1) = -\hat{Z}(0)$	$\hat{Z}(1) = \hat{X}(0) = \sigma_x \otimes \mathbf{I}$ $\hat{Y}(1) = \hat{Y}(0)$ $\hat{X}(1) = -\hat{Z}(0)$

Table 3.24: But now we're going to measure [...]

30:31 Second qubit will be a subsystem of some instrument that detects the direction of motion. In principle that could be another subatomic particle, but equally it could be a pair of photon detectors like the ones we actually used in the interference experiment.

		<p>Photon direction at time 2</p> $\langle \hat{Z}_1(2) \rangle = 0$ <p>Detector reading at time 2</p> $\langle \hat{Z}_2(2) \rangle = 0$
--	--	--

Table 3.25: And so, the expectation value of $\hat{Z}_2(2)$ is [...]

The essence of any such instrument is that somewhere in there is an observable that will be the target of the controlled NOT or perfect measurement operation—a boolean observable measuring the \hat{Z} observable as we defined it for the photon's direction of motion. So, we can just analyze those two qubits and forget all the other degrees of freedom just as we did before. The details, again, are in the worked examples.

We find that the expectation value of \hat{Z}_1 , the photon direction of motion, at time 2 (just after the measurement) is zero just as it was in the original experiment. Because in half the universes the photon travels on one path and in the other half it travels on the other. And so the expectation value of $\hat{Z}_2(2)$ is also zero. Just as we would expect from something that has measured the correct value of \hat{Z}_1 in each universe.

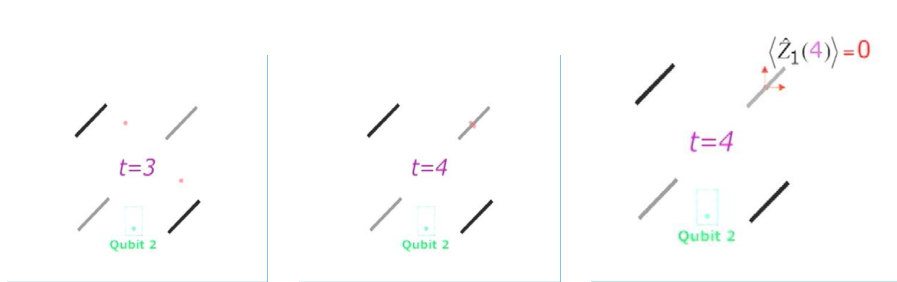


Table 3.26: The essence of any such instrument [...]

Then we let the photon go on and do exactly what it did before, namely, let it bounce off the mirrors and strike the second beam splitter. Which previously had the effect of making the direction of motion sharp again. But look, not in this experiment, the expectation value of \hat{Z}_1 is still zero at the end of the experiment. The fact of having made a measurement, even a perfect measurement, of the value of \hat{Z}_1 at an intermediate stage of the experiment, has spoiled the interference phenomenon!

31:54

This is why we don't see very clear examples of quantum interference in everyday life. It's because undergoing an interaction in which even one qubit from the outside is affected by a physical system is enough to suppress interference—and it doesn't have to be a measurement interaction, you will see that almost any interaction affecting an outside qubit will do. Specifically, what prevents interference is when something carries off information about the system. And the reason why that makes a difference is that any process that transfers information out of a system always changes the system itself.

If the process is a perfect measurement, it doesn't change the observable being measured, say \hat{Z} , but it changes the other observables like \hat{X} or \hat{Y} that are inextricably linked to it by the uncertainty principle and by the dynamics of quantum physics. And that can make the subsequently behave differently.

34:10

A process in which information is carried off in this way is known as a decoherence process. In practical implementations, decoherence is the great enemy of quantum computation—and I'll be saying more about that.

So far, when I've described the dynamics of quantum systems to you, they've all been quantum gates. I've always just told you the laws of motion of the gate by fiat—just a set of equations for how the representative observables of the qubits are changed by passage through the gate.

In the next lecture, I'll show you a general framework for quantum dynamics, and I'll tell you in principle which sets of such equations describe processes that can occur in nature and which can't.

35:09







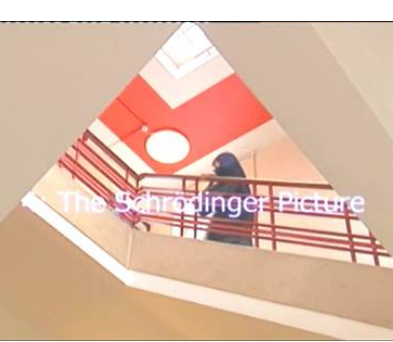

	
	
	
	

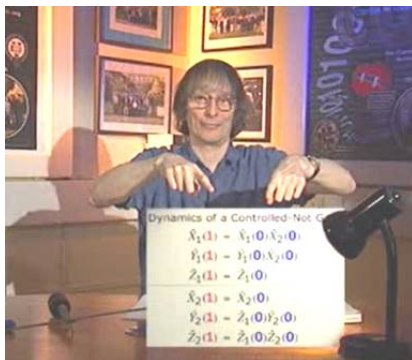
Table 3.27: [N]ext: [...] a general framework for quantum dynamics [...]

Chapter 4

The Schrödinger Picture

So far whenever I described the dynamics of quantum systems to you, and they've all been quantum gates, I've always just told you the laws of motion of each gate by fiat—just a set of equations for how the representative observables of the qubit are changed by passage through that particular gate. Now, I'll go to the other extreme and show you the most general framework for quantum dynamics, and tell you in principle which sets of such equations describe processes that occur in nature and which don't.

01:16



$$\begin{aligned}\hat{X}^2(t) &= \hat{Y}^2(t) = \hat{Z}^2(t) = \hat{1} \\ \hat{X}(t)\hat{Y}(t) &= -\hat{Y}(t)\hat{X}(t) = i\hat{Z}(t) \\ \hat{Y}(t)\hat{Z}(t) &= -\hat{Z}(t)\hat{Y}(t) = i\hat{X}(t) \\ \hat{Z}(t)\hat{X}(t) &= -\hat{X}(t)\hat{Z}(t) = i\hat{Y}(t)\end{aligned}$$

Table 4.1: [...] by fiat; [...] this Pauli algebra is [...]

Laws of motion are about how things change with time. But as I said in the first lecture, the notion of a system changing itself implies that some things about it are invariant. That's what I called the constitution of the system. The laws of motion, the dynamics, are part of the constitution. The other part is the algebra of the observables at any one time, which in a system of n qubits is always the same as the algebra of all Hermitian matrices of dimension 2^n . So for instance, you will recall that this Pauli algebra is one way of summarizing the algebra of 2×2 Hermitian matrices, representing the observables of a single qubit at any given time.

02:45

$$\begin{array}{ll}
\begin{array}{l}
[\hat{X}(t), \hat{Y}(t)] = i\hat{Z}(t) \\
\textcolor{red}{[\hat{X}(t), \hat{Y}(t)] - i\hat{Z}(t) = 0}
\end{array}
&
\begin{array}{l}
[\hat{X}(t), \hat{Y}(t)] = i\hat{Z}(t) \\
[\hat{X}(t), \hat{Y}(t)] - i\hat{Z}(t) = 0 \\
\textcolor{red}{\frac{\partial \hat{X}}{\partial t} \hat{Y} + \hat{X} \frac{\partial \hat{Y}}{\partial t} - \frac{\partial \hat{Y}}{\partial t} \hat{X} - \hat{Y} \frac{\partial \hat{X}}{\partial t} - i \frac{\partial \hat{Z}}{\partial t} = 0}
\end{array}
\end{array}$$

Table 4.2: [...] take this equation for the [...]

Now, since the algebra is invariant, the laws of motion have to be such that even though the observables change, their algebra does not. And that's obviously a constraint on what the laws of motion can be. For instance, take this equation for the commutator of the observables \hat{X} and \hat{Y} of a qubit. Rearrange it and you get a quantity which the laws of motion have to prohibit from ever changing at all. It has to stay fixed at zero for all time. Differentiate it with respect to time, that too has to be zero. So, there we have a relationship between the time derivatives of observables and we know that this relationship must be not just compatible but deducible from the law of motion of any qubit that exists anywhere in nature. And we can get similar constraints from any other equation holding among the observables at any given time. It may sound as though its going to be tricky to find any law of motion that enforces all of these intricate constraints on how observables can change with time. But no, it's easy to find one. Here's one:

$$\hat{H}_t = a(t)\hat{X}(t) + b(t)\hat{Y}(t) + c(t)\hat{Z}(t) + d(t)\hat{\mathbf{1}}$$

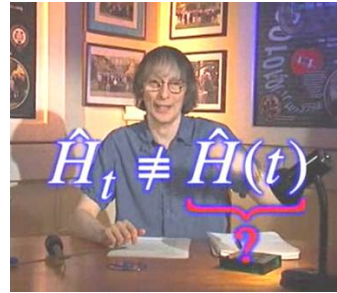


Table 4.3: [...] a whole one-parameter family [...]

Pick any observable \hat{H} of a quantum system at a given time. If the system is a qubit, then like all observables \hat{H} can be expressed as a linear combination

with real coefficients of the representative observables \hat{X} , \hat{Y} and \hat{Z} at that time, and the unit observable. Do the same for every other time. In other words, pick a whole one parameter family of observables, \hat{H} subscript t , one for each time (t). The reason I'm not calling that \hat{H} of t , with the t in parentheses, is that we're already using that notation to refer to the time evolution of a single observable, whereas in this construction we are allowed to choose a different observable at each time. Now, given that one-parameter family of observables, consider the following law of motion: for each observable \hat{A} of t of the system $d\hat{A}$ by dt equals i times the commutator of \hat{H} at t with $\hat{A}(t)$: 05:30

$$\frac{d\hat{A}(t)}{dt} = i[\hat{H}_t, \hat{A}(t)]$$

Table 4.4: now [...] consider this law of motion [...]

We can express that law in a form that's more like the specific laws of motion I've told you in previous lectures by just integrating it over an infinitesimal period, dt , so we get $\hat{A}(t + dt)$ equals $\hat{A}(t) + idt$ times the commutator of \hat{H} with \hat{A} and in that form, it expresses each observable at one time, $t + dt$, in terms of observables at another time, t .

$$\frac{d\hat{A}(t)}{dt} = i[\hat{H}_t, \hat{A}(t)]$$
$$\Rightarrow \hat{A}(t + dt) = \hat{A}(t) + idt[\hat{H}_t, \hat{A}(t)]$$

$$\frac{d\hat{A}(t)}{dt} = i[\hat{H}_t, \hat{A}(t)]$$
$$\Rightarrow \hat{A}(t + dt) = \hat{A}(t) + idt[\hat{H}_t, \hat{A}(t)]$$

Table 4.5: [...] integrating [...] over an infinitesimal period [...]

\hat{H} at t is by construction some observable at time t , and so is \hat{A} of t . For gates, we were only interested in the relationship between observables before and after the gate acted, so we used laws of motion referring to unit time instead of an infinitesimal time. Such laws could be obtained by integrating this differential equation (Table 4.4) of motion between t and $t + 1$.

OK, now does our new law of motion preserve all the algebraic relationships that may exist between observables? Well, consider any such relationship at time t . We can always write it in the form of \hat{F} of \hat{A} of t , \hat{B} of t ... and so on, equals 0 (zero) where \hat{A} and \hat{B} and so on are observables at time t and the function \hat{F} uses any of the operations of matrix algebra, namely: matrix addition, matrix multiplication, and multiplication by a constant.

$$\hat{F}(\hat{A}(t), \hat{B}(t), \dots) = 0$$

Where the function \hat{F} is composed of:

- matrix addition
- matrix multiplication
- multiplication by a (scalar) constant

To Prove: $\frac{d\hat{F}(\hat{A}(t), \hat{B}(t), \dots)}{dt} = i[H_t, \hat{F}]$
 (Provided $\frac{d\hat{A}(t)}{dt} = i[H_t, \hat{A}]$, $\frac{d\hat{B}(t)}{dt} = i[H_t, \hat{B}]$, ...) $\hat{F}(t) = \lambda \hat{P}(t)$

$$\hat{F}(t) = \lambda \hat{P}(t) \Rightarrow \frac{d\hat{F}(t)}{dt} = \lambda \frac{d\hat{P}}{dt} = i\lambda[H_t, \hat{P}] = i[H_t, \lambda\hat{P}] = i[H_t, \hat{F}]$$

Table 4.6: [...] consider any such relationship at time t [...]

08:39 Its sufficient to prove that \hat{F} itself obeys the law of motion. Well, what can \hat{F} be? Suppose \hat{F} is a real multiple of something that does obey the law of motion. Then, \hat{F} obeys it too. Suppose \hat{F} is the sum of two things, that each obey the law of motion. Then again, \hat{F} also obeys it. What if \hat{F} is the matrix product of two things that obey the law of motion? Well, then [...] since \hat{P} and \hat{Q} obey the laws of motion, multiply that out and you'll see that it's i times the commutator of \hat{H} with the product $\hat{P}\hat{Q}$ which is i times the commutator of \hat{H} with \hat{F} again.

$$\hat{F}(t) = \hat{P}(t) + \hat{Q}(t) \Rightarrow \frac{d\hat{F}(t)}{dt} = \frac{d\hat{P}}{dt} + \frac{d\hat{Q}}{dt} = i[H_t, \hat{P}] + i[H_t, \hat{Q}] = i[H_t, \hat{P} + \hat{Q}] = i[H_t, \hat{F}]$$

$$\hat{F}(t) = \hat{P}(t)\hat{Q}(t) \Rightarrow \frac{d\hat{F}(t)}{dt} = \frac{d\hat{P}}{dt}\hat{Q} + \hat{P}\frac{d\hat{Q}}{dt} = i[H_t, \hat{P}]\hat{Q} + i\hat{P}[H_t, \hat{Q}] = i[H_t, \hat{P}\hat{Q}] = i[H_t, \hat{F}]$$

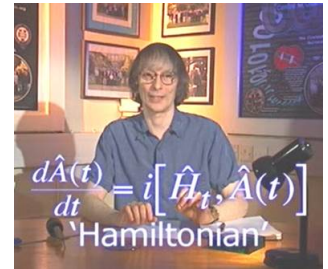


Table 4.7: [...] multiply that out, and you'll see that [...]

So the upshot is, however \hat{F} is composed out of observables obeying our equation of motion, \hat{F} obeys it too. And hence, it remains zero over time. And so, any equation of motion of this form leaves any algebra of observables invariant. And actually, we're done. Because, as you can prove in the worked

examples, the converse is also true: any law of motion that leaves the algebra of observables of some quantum system invariant takes this form. \hat{H} is called the Hamiltonian of the system. This (Table 4.4) is called the Heisenberg equation of motion. And the way of representing observables which I've been using is called the Heisenberg picture of Quantum Theory

Consider the special case where the Hamiltonian is actually the same observable at all times. So, then \hat{H}_t suffix t does equal some \hat{H} of t where \hat{H} of t is a bona fide observable. In such cases, we say that the Hamiltonian has no explicit time dependence. How does this observable change with time?

$$\begin{array}{ccc}
 \text{If } \hat{H}_t = \hat{H}(t) \text{ for all } t, & \hat{H}_t = \hat{H}(t) \Rightarrow & \frac{d\hat{H}_t}{dt} = 0 \quad \& \quad [\hat{H}_t, \hat{A}(t)] = 0 \\
 \text{'}\hat{H}_t \text{ has no explicit time-dependence'} & & \underbrace{\hspace{10em}} \\
 \frac{d\hat{H}_t}{dt} = \frac{d\hat{H}(t)}{dt} = i[\hat{H}(t), \hat{H}(t)] = 0 & & \downarrow \\
 & & \frac{d\hat{A}(t)}{dt} = 0
 \end{array}
 \qquad
 \begin{array}{c}
 \frac{d\hat{H}_t}{dt} = 0 \\
 \& \\
 \frac{d\hat{A}(t)}{dt} = 0
 \end{array}
 \quad
 \begin{array}{c}
 \text{[}\hat{H}_t, \hat{A}(t)\text{]} = 0 \\
 \text{[}\hat{H}_t, \hat{A}(t)\text{]} = 0
 \end{array}$$

Table 4.8: [...a]nd, again the converse is also true [...]

Well, as you can see, it doesn't. When the Hamiltonian has no explicit time dependence it has no time dependence at all. So under those circumstances, its not only an observable, it's a conserved quantity. This is a conservation law. And in fact, the Hamiltonian is up to a constant of proportionality, the Energy observable. You can also see that any other observable that commutes with the Hamiltonian is a conserved quantity too. And if an observable commutes with a conserved Hamiltonian, at any one instant, even one instant, then because the algebra is invariant it will commute at all instants. And will be conserved too. And again, the converse is also true, in a system with a conserved Hamiltonian, every conserved observable commutes with that Hamiltonian.

Now, in the case where the Hamiltonian is not only the same observable at all times, but one of the systems own observables, the system is said to be isolated. The idea of an isolated system is an idealization. There are no physical systems that remain unaffected by the outside world for all possible states of the outside world. Remember that the idea of a physical system itself is not perfectly well defined because there is no physical system that even remains in existence in all states. And for both of these reasons, the idea of the Hamiltonian for a given physical system is always an idealization or an approximation. Except perhaps for the multiverse as a whole; that's not interacting with anything outside itself. So the multiverse is the only truly isolated system. Nevertheless, in real physical phenomena, its often a very good approximation to say that a system is isolated for a period. For instance, I said that the qubit in our interference experiment in lecture two was isolated during periods when the photon was traveling between gates. In fact it was well

12:02

described as having a Hamiltonian of zero during those periods, even though its real Hamiltonian—ultimately the Hamiltonian of the multiverse—would have contained terms describing how the photon would interact with a passing bumblebee that just happened to fly through the apparatus at the same time, and indeed how it did interact with one in some universes.

Now, even when a system isn't isolated, it is sometimes possible to encode the whole influence of the outside world on that system into a law of motion that involves only that system's observables. But with a time dependent Hamiltonian. Such systems are said to be 'coherent'. Another way of stating the definition of 'coherent' is that, a coherent quantum system is one in which the observables at one time are functions only of the observables of the same system at another time. And that would be in a given state, always, in a given class of states. So in summary, in states where the dynamics of a system can be well described by a Hamiltonian that's simply one of its own observables, it's said to be isolated. When it can be well described by a Hamiltonian built out of its own observables but possibly with explicit time dependence, it's said to be coherent. So an isolated system is also coherent.

 <p>Isolated System One whose Hamiltonian is an observable of that system alone.</p>	 <p>Coherent System One whose Hamiltonian is a (time-dependent) function of observables of that system alone.</p>	<p>Isolated System: One whose Hamiltonian is an observable of that system alone.</p> <p>Coherent System: One whose Hamiltonian is a (time-dependent) function of the observables of that system alone.</p>
---	---	--

Table 4.9: [...] So an isolated system is also coherent.

When it can only be described by a Hamiltonian involving other systems, it is said to be 'decoherent.' But ultimately, all these terms are approximative. Short of the multiverse as a whole, there are no isolated systems and every system is decoherent to some extent. And there are no time dependent Hamiltonians. Every time dependent Hamiltonian is just an approximate way of taking into account the effects of other systems in states where the decoherence that they cause in the system of interest is negligible. In our single photon interference experiment, for example, the qubit was effectively isolated during its motion between the beam splitters and the mirrors. But as it interacted with those, its Hamiltonian changed briefly as I described, but only to some function of the qubit's own observables. And then it changed back to zero. So the qubit remained coherent throughout the whole experiment until the moment when it interacted with the detector at the end—at that point the qubit's Hamiltonian would have depended on observables of the detector and

vice versa, thus causing decoherence. And also as we saw, if the qubit interacted with another qubit during the experiment, then its Hamiltonian would depend on observables of both qubits, and we saw that the interference was thus reduced or eliminated. And it's true in general, that for an interference phenomenon to occur, or for a quantum computation to be performed, the system doesn't have to be isolated but it does have to be coherent.

17:24

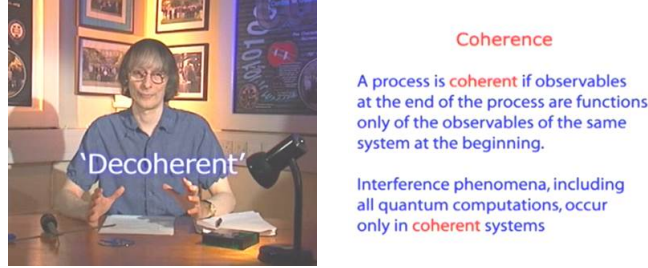


Table 4.10: [...] ultimately, all these terms are approximative.

Okay, now, for an isolated system, we can solve the equation of motion in closed form. It's just $\hat{A}(t)$ equals e to the $i\hat{H}t$ [times] $\hat{A}(0)$ [times] e to the $-i\hat{H}t$. So that gives \hat{A} at any time in terms of \hat{A} at time zero, and \hat{H} . We can also express the constant Hamiltonian \hat{H} in terms of observables at time zero, and so, this equation would express \hat{A} , at an arbitrary time t , in terms of observables of the system at time zero.

<p>Isolated: $\hat{H}_t = \hat{H}$</p> <p>Equation of Motion: $\frac{d\hat{A}(t)}{dt} = i[\hat{H}, \hat{A}(t)]$</p> <p>$\Rightarrow \hat{A}(t) = e^{i\hat{H}t} \hat{A}(0) e^{-i\hat{H}t}$</p>	<p>Isolated: $\hat{H}_t = \hat{H}$</p> <p>Equation of Motion: $\frac{d\hat{A}(t)}{dt} = i[\hat{H}_t, \hat{A}(t)]$</p> <p>$\Rightarrow \hat{A}(t) = e^{i\hat{H}t} \hat{A}(0) e^{-i\hat{H}t}$</p>	<p>Isolated: $\hat{A}(t) = e^{i\hat{H}t} \hat{A}(0) e^{-i\hat{H}t}$</p> <p>Coherent: $\hat{A}(t) = U_t^\dagger \hat{A}(0) U_t$</p> <p>$U_t^\dagger U_t = \hat{1}$</p> <p>$\frac{dU_t}{dt} = -iU_t \hat{H}_t$</p>
--	--	--

Table 4.11: [...] for an isolated system [...] solve the equation of motion [...]

For a general coherent system, the solution will be a slight generalization of that. It will be [...] where U_t is a one-parameter family of unitary matrices¹ whose own equation of motion is [...]. U is called the evolution matrix between time zero and time t . The state of a system is defined by giving its expectation value function. Which is, a linear function mapping its observables to real numbers. I've used this notation to denote the expectation value of an

¹Unitary means $U^\dagger U = \hat{1}$, where $\hat{1}$ is the unit matrix.

observable \hat{A} of t : $\langle \hat{A}(t) \rangle$. But because its a linear function, it must also be possible to write it in any given matrix representation like this, where alpha and beta are matrix indices, and rho is some matrix that doesn't change with time. In other words, the expectation values is Trace \hat{A} of t [times] ρ .



$$\begin{aligned} \langle \hat{A}(t) \rangle &= \sum_{\alpha, \beta} \hat{A}_{\alpha\beta}(t) \rho_{\beta\alpha} & \langle \hat{A}(t) \rangle &= \sum_{\alpha, \beta} \hat{A}_{\alpha\beta}(t) \rho_{\beta\alpha} & \langle \hat{\mathbf{1}} \rangle &= \sum_{\alpha, \beta} \hat{A}_{\alpha\beta}(t) \rho_{\beta\alpha} \\ & & &= \text{Tr } \hat{A}(t) \rho & &= \text{Tr } \hat{\mathbf{1}} \rho = 1 \end{aligned}$$

Table 4.12: [...] but because it is a linear function [...]

Since the expectation value of the unit observable is always 1, Trace of ρ itself has got to be 1. ρ is called the density matrix of the system. It has to satisfy certain conditions to make sure that the expectation value can never be higher than the highest eigenvalue of \hat{A} nor lower than the lowest. You can find out what those conditions are in the worked examples. Now, look at the expression for the expectation values of an arbitrary observable. Its often more convenient to write this in the form of [...] see below]:

20:11



$$\begin{aligned} \langle \hat{A}(t) \rangle &= \text{Tr } \hat{A}(t) \rho & \langle \hat{A}(t) \rangle &= \text{Tr } \hat{A}(t) \rho \\ &= \text{Tr } U_t^\dagger \hat{A}(0) U_t \rho & &= \text{Tr } U_t^\dagger \hat{A}(0) U_t \rho \end{aligned}$$

Isolated: $\hat{A}(t) = e^{i\hat{H}t} \hat{A}(0) e^{-i\hat{H}t}$


Coherent: $\hat{A}(t) = U_t^\dagger \hat{A}(0) U_t$

$$\begin{aligned} U_t^\dagger U_t &= \hat{\mathbf{1}} \\ \frac{dU_t}{dt} &= -iU_t \hat{H}_t \end{aligned}$$

Table 4.13: [...] look at the expression for the expectation value of [...]

Why? Because in this expression, the only quantity that changes with time is the unitary evolution matrix U , which is the same in the corresponding expression for any observable. Therefore, to track everything that the system is doing over time, we don't have to solve the equations of motion for all the observables in terms of their values at previous times, we need only solve this one equation for the evolution matrix U suffix t , given the Hamiltonian, and with that we can read off the time evolution of any observable of the system evolving under that Hamiltonian. Actually we can do better than that. The way the dynamics of quantum systems are encoded in these unitary matrices allows for a whole alternative way for describing quantum systems that's often

much more efficient. What I've described so far is the Heisenberg picture, and the alternative way I'll show you now is called the Schrödinger picture.



$$\begin{aligned}
 \langle \hat{A}(t) \rangle &= \text{Tr} \hat{A}(t) \rho \\
 &= \text{Tr} \underbrace{U_t \rho U_t^\dagger}_{\rho_t} \underbrace{\hat{A}(0)}_{\hat{A}} \\
 &= \text{Tr} \underbrace{U_t \rho U_t^\dagger}_{\rho_t} \hat{A} \\
 &= \text{Tr} \underbrace{U_t \rho U_t^\dagger}_{\rho_t} \hat{A} \\
 \frac{d\rho_t}{dt} &= -i[\hat{H}_t, \rho_t]
 \end{aligned}$$

Table 4.14: [... let's] look again at the expression [...]

First, look again at this expression for the expectation value of a general observable at an arbitrary time. We can rewrite it using the cyclic invariance of the Trace like this: the expectation value of [... see picture above]. Let's call this quantity ρ subscript t . ρ subscript t is called the density matrix in the Schrödinger picture. So to evaluate an arbitrary expectation value, we need only know this density matrix in the Schrödinger picture, and all the observables at any one time, say, $t = 0$. So for each observable \hat{A} , we only need to know one matrix, $\hat{A}(0)$. Which in the Schrödinger picture we just call \hat{A} . To summarize the Schrödinger picture then, each observable is represented by a constant matrix, while the density matrix changes with time. This is to be compared with the Heisenberg picture, where the observables are functions of time and the density matrix is a constant.

From this expression, we can read off the law of motion for the Schrodinger picture density matrix: it's [... see pics above]. This has almost the same form as the law of motion for an observable in the Heisenberg picture. But with the opposite sign. I'll come back to that sign in a moment. Now consider the eigenvectors of the Schrödinger density matrix at time t . The standard notation for vectors in quantum theory is called the 'Dirac notation'. It uses a symbol called the 'ket' which looks like this, to denote vectors.

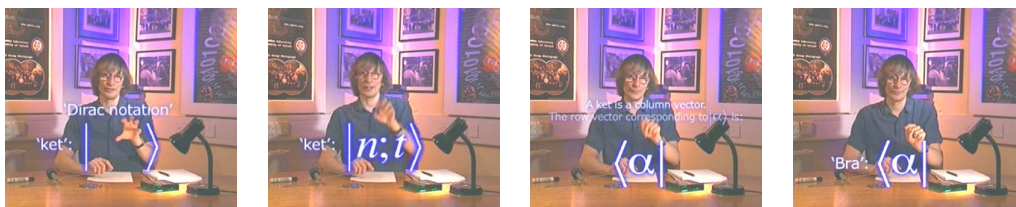


Table 4.15: The standard notation for vectors in [...]

We typically write inside the ket symbol the information specifying which vector it is. So for instance, we can call the n -th eigenvector of ρ of t the ket with label n and t . The ρ vector, that is the Hermitian transpose or dual of a given

ket, say with label alpha, is written like this: $\langle\alpha|$ this symbol is called a ‘bra’. The origin of this terminology is that the scalar up product of a bra with a ket is written like this: $\langle\alpha|\beta\rangle$ the whole thing being a bra-cket. So the left half of it is a ‘bra’ and the right half a ‘ket’. The scalar product of a ket with the corresponding bra, say: $\langle\alpha|\alpha\rangle$ is by definition the squared magnitude or norm of the ket alpha.



Table 4.16: The origins of this terminology is that [...]

The vector space of kets is therefore a ‘Hilbert space’—basically a vector space with [a] norm—and it’s usually called the Hilbert space of the given system. Observables represented by matrices are linear operators on the Hilbert space. The set of all eigen-kets of the density matrix at any given time, or of any observable at any time, constitutes a basis for the system’s Hilbert space. So, we can expand the density matrix in terms of its eigen-values and eigen-vectors like this [below]. And you can prove in the worked examples that the eigenvalues p_n remain constant, and that the equation of motion for the eigenvector s is this [below]. Or more generally, if Ψ is any eigenvector of the density matrix [below]. This is called the Schrödinger equation, and again, over a period t its general solution will have this form, where the U_t are the same unitary matrices as in the solution of the Heisenberg equation of motion.

25:58

$\rho_t = \sum_n p_n n;t\rangle\langle n;t $			The Schrödinger Equation
$\frac{d}{dt} n;t\rangle = -i\hat{H}_t n;t\rangle$	$\frac{d}{dt} n;t\rangle = -i\hat{H}_t n;t\rangle$	The Schrödinger Equation	$\frac{d}{dt} \psi(t)\rangle = -i\hat{H}_t \psi(t)\rangle$
$\frac{d}{dt} \psi(t)\rangle = -i\hat{H}_t \psi(t)\rangle$	$\frac{d}{dt} \psi(t)\rangle = -i\hat{H}_t \psi(t)\rangle$	$\frac{d}{dt} \psi(t)\rangle = -i\hat{H}_t \psi(t)\rangle$	$ \psi(t)\rangle = U_t \psi(0)\rangle$

Table 4.17: The origins of this terminology is that [...]

Now, the significance of that sign in the equation of motion of the density matrix. The density matrix is not an observable. Both in the Schrödinger and Heisenberg pictures, it’s equal to a different observable at each instant. So why did it have a similar equation of motion to an observable, but with the opposite sign? Well, in the Schrödinger picture, the state changes and the observables are constant matrices. In the Heisenberg picture, the state is invariant and the observables change. But in both of them, the expectation

value of an observable \hat{A} at a given time is Trace of $\rho\hat{A}$ and the Heisenberg \hat{A} of t , we know, is U dagger \hat{A} of zero [times] U . Remember, Schrödinger \hat{A} equals Heisenberg \hat{A} of 0. And also, Heisenberg ρ equals Schrödinger ρ of 0. What's happening, is that the unitary transformation, U_t , that defines the motion, amounts to a rigid rotation in Hilbert space. It's rigid in that it preserves the scalar product between any two kets that it acts on.

Heisenberg:				
$\langle \hat{A}(t) \rangle = \text{Tr } \rho \hat{A}(t)$	$\langle \hat{A}(t) \rangle = \text{Tr } \rho \hat{A}(t)$	$\langle \hat{A}(t) \rangle = \text{Tr } \rho U_t^\dagger \hat{A}(0) U_t$	$\langle \hat{A}(t) \rangle = \text{Tr } \rho U_t^\dagger \hat{A}(0) U_t$	
Schrödinger:				
$\langle \hat{A}(t) \rangle = \text{Tr } \rho_t \hat{A}$	$\langle \hat{A}(t) \rangle = \text{Tr } \rho_t \hat{A}$	$\langle \hat{A}(t) \rangle = \text{Tr } \rho_t \hat{A}$	$= \text{Tr } U_t \rho U_t^\dagger \hat{A}(0)$	

Table 4.18: But in both of them [...]

When ket alpha goes to U alpha, bra alpha goes to bra alpha U dagger. And so, any scalar product alpha beta is unchanged. And that's rigid rotation.

$ \alpha\rangle \rightarrow U_t \alpha\rangle$	$ \alpha\rangle \rightarrow U_t \alpha\rangle$	$ \alpha\rangle \rightarrow U_t \alpha\rangle$
$ \beta\rangle \rightarrow U_t \beta\rangle$	$ \beta\rangle \rightarrow U_t \beta\rangle$	$ \beta\rangle \rightarrow U_t \beta\rangle$
$\langle\alpha \rightarrow \langle\alpha U_t^\dagger$	$\langle\alpha \rightarrow \langle\alpha U_t^\dagger$	$\langle\alpha \beta\rangle \rightarrow \langle\alpha U_t^\dagger U_t \beta\rangle$
		$= \langle\alpha \beta\rangle$

Table 4.19: [...] that's a rigid rotation.

The density matrix represents the state of the world. In the Heisenberg picture, the state is constant and the observables rotate in the sense that their eigenvectors rotate in Hilbert space. In the Schrödinger picture, the observables are fixed. And so to maintain the same relative orientation, the state rotates rigidly in the opposite sense. Thus ensuring that both pictures make the same predictions for the physical quantities—the expectation values of observables.



Table 4.20: In the Heisenberg picture [...]. In the Schrödinger picture [...]

The Schrödinger equation is the Schrödinger picture way of defining the dynamical evolution of a quantum system. In the general case, we have to solve

the equation for each eigen-vector of the density matrix. From which we can reconstitute the density matrix itself.

$$\begin{array}{llll}
 \rho_t = \sum_n p_n |n;t\rangle\langle n;t| & & & \text{Pure state:} \\
 \frac{d}{dt}|n;t\rangle = -i\hat{H}_t|n;t\rangle & & \frac{d}{dt}|n;t\rangle = -i\hat{H}_t|n;t\rangle & \text{One in which the} \\
 \frac{d}{dt}|\psi(t)\rangle = -i\hat{H}_t|\psi(t)\rangle & \frac{d}{dt}|\psi(t)\rangle = -i\hat{H}_t|\psi(t)\rangle & \frac{d}{dt}|\psi(t)\rangle = -i\hat{H}_t|\psi(t)\rangle & \text{density matrix is sharp} \\
 & & & \text{(i.e. equal to a sharp observable).}
 \end{array}$$

Table 4.21: In the general case we have to solve the equation [...]

But, if ρ is ever sharp—well ρ is not an observable but what I mean is—if it's ever equal at some instant to a sharp observable, then the system is said to be in a pure state. And we're often particularly interested in systems in pure states. For instance, if the \hat{Z} observables of a set of qubits are all sharp—so the qubits together are like a register of a classical computer holding a single integer—then the density matrix is sharp and the system is in a pure state. You can prove that in the worked examples. And you can also prove that in a pure state, in the eigen-vector representation of the density matrix, all the coefficients vanish except for one, which takes the value of 1. And since the eigen-values of the density matrix don't change with time, if a system is in a pure state at any instant, then it remains in a pure state so long as the evolution remains coherent. And the density matrix in such a case just takes this form, where Ψ of t is a single one-parameter family of kets obeying the Schrödinger equation with respect to t . Ψ at t is then called the Schrödinger state vector. For a quantum system in a pure state, all the motion of the system is summed up in the motion of this state vector.

31:54

$$\begin{array}{llll}
 \rho_t = \sum_n p_n |n;t\rangle\langle n;t| & \rho_t = & |n;t\rangle\langle n;t| & \rho_t = |\psi(t)\rangle\langle\psi(t)| \\
 & & & \text{Pure state} \\
 & & \text{Schrödinger equation} & \text{Schrödinger state vector} \\
 \frac{d|\psi(t)\rangle}{dt} = -i\hat{H}_t|\psi(t)\rangle & & \frac{d|\psi(t)\rangle}{dt} = -i\hat{H}_t|\psi(t)\rangle & \text{Schrödinger equation}
 \end{array}$$

Table 4.22: For a quantum system in a pure state [...]

Okay, this equation for pure states and this one for non-pure states are in the majority of cases the most convenient ways of analyzing the motion of quantum systems, and in particular the computations of quantum computers. And I'll mostly use it from now on. But beware. As we've just seen, this economy of calculation comes at the expense of additional layers of abstraction. Heisenberg observables, which are the dynamical quantities in the Heisenberg picture of quantum physics, have a lot in common with the variables of clas-

sical physics. Yes, they're matrices rather than real numbers, but at least there's one such matrix corresponding to each quantity we can observe. And the particular number corresponding to what we do directly observe is in there somewhere—its one of the eigen-values. And in the Heisenberg picture, the changing quantities—the observable—are located in particular systems at particular locations. They obey an equation of motion that expresses how they affect other observables and are affected by them, and thereby carry information from one place to another. The Schrödinger state doesn't have any of that. There's only one of them. It's not located anywhere, but refers to the whole system or ultimately the whole multiverse.

$$\frac{d\rho_t}{dt} = -i[\hat{H}_t, \rho_t]$$

$$\frac{d|\psi(t)\rangle}{dt} = -i\hat{H}_t|\psi(t)\rangle \qquad \frac{d|\psi(t)\rangle}{dt} = -i\hat{H}_t|\psi(t)\rangle$$

Table 4.23: [An] equation for pure states and [...] one for non-pure states.

What is the Hamiltonian of the whole multiverse?

Traditionally, fundamental physics has been about what types of systems exist in nature, what their observables are, and what their Hamiltonians are.

That's what elementary particle physicists call "the theory of everything". But there is another way of looking at what is elementary or fundamental. Instead of asking "what types of Hamiltonian are found in natural systems?" which means "what sorts of changes can occur in elementary systems over an infinitesimal time?", one could go all the way to the bottom line and ask "which transformations can be realized in nature by some quantum system evolving for some time, and which cannot?".

And the short answer is, all changes in which observables of the system undergo unitary evolution with every observable going under the same transformation so as to preserve their algebra. Every one of those can occur in nature. And nothing else can. Every unitary matrix is the evolution matrix for some quantum system evolving over some time. Or at least we think it is. Because it turns out that the vast majority of these possible evolutions can only be realized in a very special type of physical system: a universal quantum computer. They don't occur naturally because they require a complex computer program to bring them about.

34:30

So here's a fascinating situation: in terms of Hamiltonians, the laws of physics are very finely tuned. The multiverse has its Hamiltonian, and subsystems of the multiverse only have very special Hamiltonians. Most Hermitian

matrices cannot be realized in nature as Hamiltonians. But when we ask a slightly different question, “which unitary evolution operators can be realized in nature?”, the answer is “all of them, IF a universal quantum computer can exist”. This special type of object, the universal quantum computer, in a sense contains within itself all the diversity in nature. No other system does. Except perhaps systems that are capable of constructing a universal quantum computer. Suddenly we find ourselves unavoidably playing a role at the deepest level of the structure of physical reality.

36:38

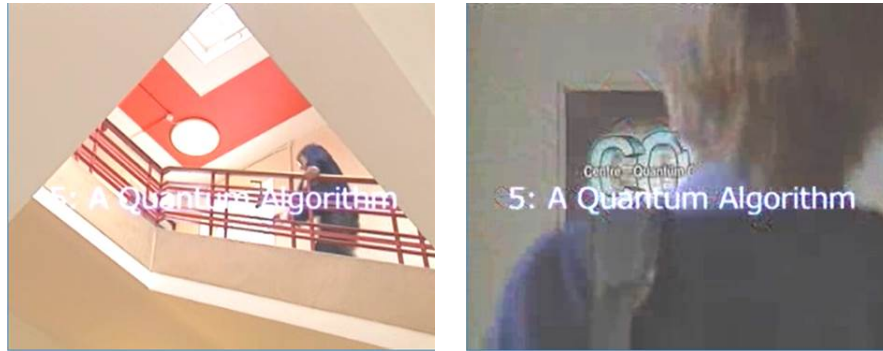


Table 4.24: Suddenly we find ourselves unavoidably playing a role [...]

Chapter 5

A Quantum Algorithm

An algorithm is a hardware independent specification of a computation. By hardware independent I mean that it doesn't specify what the computer should be made of, only the effect that it should have on information that's provided to it as input. So that if you had access to any technology that allowed you to build a computer, a universal computer, you could translate the algorithm into a program for that computer which would perform the information processing that the algorithm specifies. An algorithm is a way of performing a computational task like sorting a list or factorizing a number. In general, to specify a computational task, you specify properties that the output information should have as a function of the input information. So, a computational task is a problem to which an algorithm is the solution.

Computational task:	Problem to which an algorithm is the solution
Algorithm:	Hardware-independent specification of a computation
Program:	Way of preparing a computer to perform a computation




Table 5.1: The algorithm that I'm going to describe today [...]

The algorithm that I'm going to describe today is probably the simplest of all quantum algorithms. So, I'll first describe the problem that it solves. You're given a black box containing a computer that's dedicated to computing a particular function f using reversible classical computation. It operates coherently though, because part of our task is going to be to use this black box as a component in a quantum computational network. You're not allowed to look inside

the box. All you're allowed to do is feed qubits into it at one end and then retrieve them at the other end after a fixed time that's independent of the input. Such a box is known as an 'oracle'.



Table 5.2: All you are allowed to do is feed qubits into it at [...]

The idea of having an oracle in the specification of a computational task is a trick much beloved by complexity theorists, both classical and quantum, because it can greatly simplify the analysis of algorithms. That's because often, looking inside the oracle doesn't help you to perform the task, yet it's quite hard to prove that it doesn't. Why that's so often the case is quite a deep question. It's beyond the scope of these lectures. But I'll explain why it's plausible in this case in a moment.

Okay, the oracle in this problem computes a boolean function f with one boolean parameter. That is to say, f maps the set containing minus 1 and 1 to itself. Since the oracle performs a reversible computation, it can't just do this: x to $f(x)$. That wouldn't be reversible. The oracle must have an auxiliary input and output, and when it computes f what it must really be doing is something like this: x and y goes to x and $yf(x)$. That's a reversible computation. And if you want to use the oracle to compute $f(x)$, then you just have to make sure that y is initialized to the value $+1$.



Table 5.3: [...] the Oracle in this problem computes a [...]

Now, there are only four functions that map a single bit to a single bit. The identity, the NOT operation, or delivering a constant output -1 or 1 . Internally, the oracle may be computing something arbitrarily complex, or rather one of two arbitrarily complex things, depending on the input. For instance, it might be doing the traveling salesman problem or some other hard problem for one of

two graphs depending on the input, and then reporting some boolean property of the answer such as whether the shortest path has an even or odd number of steps. But, however complex the oracle is internally, it will be computing one of these four functions. That’s why its plausible that looking inside the oracle won’t in general help. The oracle might contain a complicated network with a large number of qubits doing things in parallel. So it might still be a lot faster to run it twice than to work out what it does once. Anyway, a simple computational task, given the oracle, would be “find out what f is”. Well, the only way to do that without looking inside is to run the oracle once for each possible input. The resulting four pairs of outputs tabulate f .

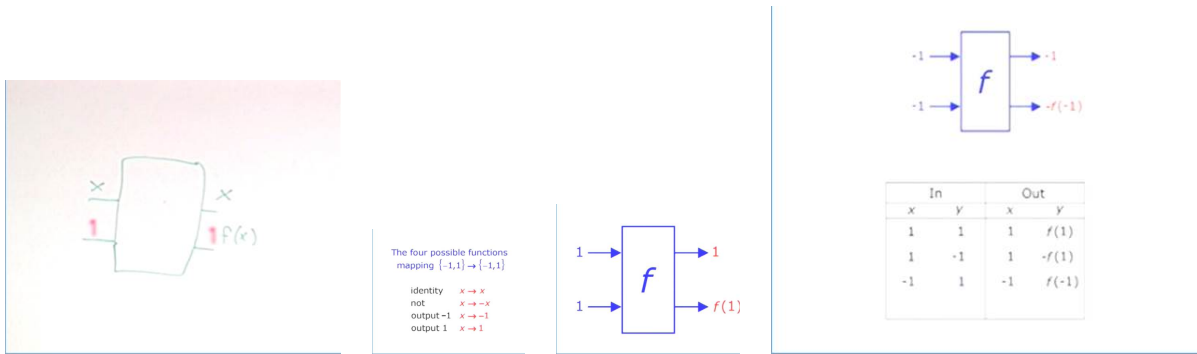


Table 5.4: [...] the only way to do that without looking inside [...]

Is it possible to perform that task using only one run of the oracle? Well we can prove that it isn’t. If you’re only allowed to run the oracle once, then whatever else you do you’ll have to pick a particular pair of values for its input x and y . And you’ll get output values x , which you already knew, and $yf(x)$. Only the second of those two outputs even depends on f . So that can’t fully specify what f is because f can be one of four different functions.

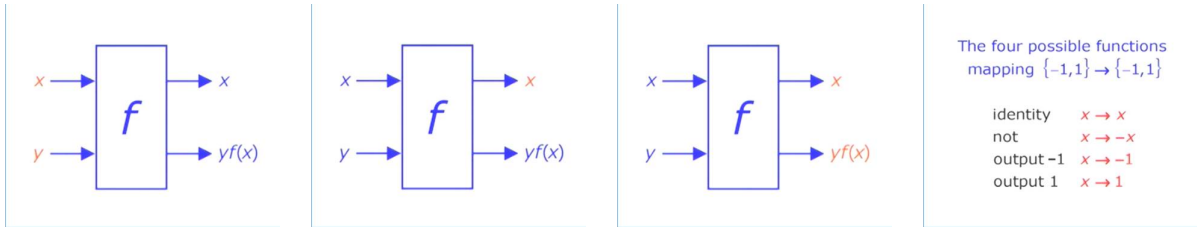


Table 5.5: If you’re only allowed to run the oracle once [...]

And it turns out that that is true in the quantum case too. You can’t work out what f is without invoking the oracle twice. So quantum computation doesn’t help with that task, which illustrates the general fact that quantum computation doesn’t speed up all computational tasks, only some of them. And

that's one reason why quantum complexity theory is fundamentally different from classical complexity theory. Anyway, the problem that we are going to solve with a quantum algorithm is not to find out what f is, but to determine a property of f . Specifically, whether $f(1)$ equals $f(-1)$ or not. Or more concisely, the task is to compute the product $f(1)f(-1)$ without looking inside the oracle.

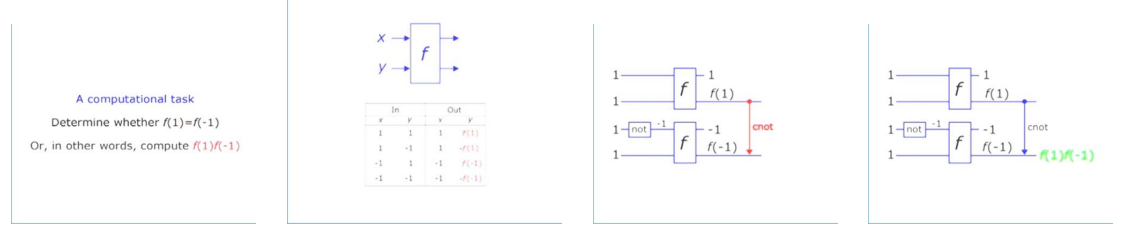


Table 5.6: So is there any way [...]

Knowing the quantity $f(1)f(-1)$ certainly doesn't tell you everything about f . It just tells you one bit of information about it, which is half the information in the table of f . So, the proof I gave that the previous task requires two calls of the oracle doesn't hold for this task. So is there any way of computing the single boolean value $f(1)f(-1)$ using only one invocation of the oracle?

For the case of classical computation, it's again easy to prove that there is no way. Even though it's only one bit of information this time. Because again, there are only four possible ways of invoking the oracle in a classical computation, and in all four the first output is again independent of f . And the second depends on only one of the two values $f(1)$ or $f(-1)$, never both. So $f(1)f(-1)$ can never be deduced from it. Here's a classical network that finds the answer using two instances of the oracle. This is a controlled NOT gate, which we are using in its capacity as an exclusive OR gate. Alternatively, we could find the answer by invoking one instance of the oracle twice, doubling the running time, like this:

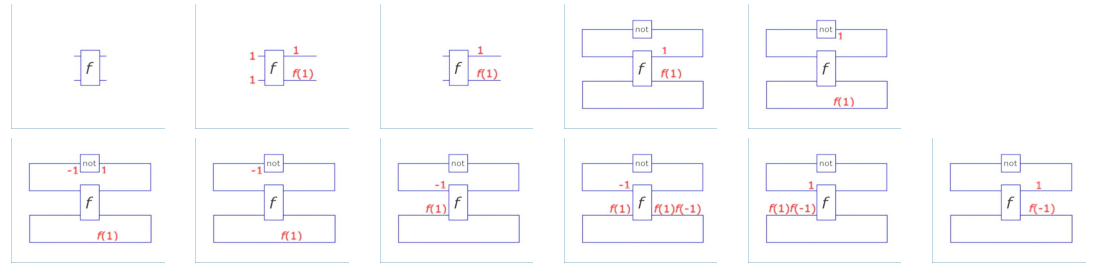


Table 5.7: Here's a classical network [...]

The input bits both start at 1, they pass through the oracle making them +1 and $f(+1)$. Then, they each go around a loop, the first one passes through

a NOT gate that flips it to -1 , and the second goes in again as the auxiliary bit. So the final output is $f(1)f(-1)$. Strictly speaking, this network runs forever so I'll leave it as an exercise to modify it so that it delivers an output after exactly two oracle invocations. What the quantum algorithm does, in short, is that it uses only one instance of the oracle and it invokes it only once, but it invokes it with a different input in different universes. Which means, that the oracle performs different computations in different universes yielding potentially different outputs. And the single qubit that holds those two outputs in different universes combines them in an interference phenomenon—it exclusive ORs them. A bit like this, except that the invocations of the oracle occur in different universes.

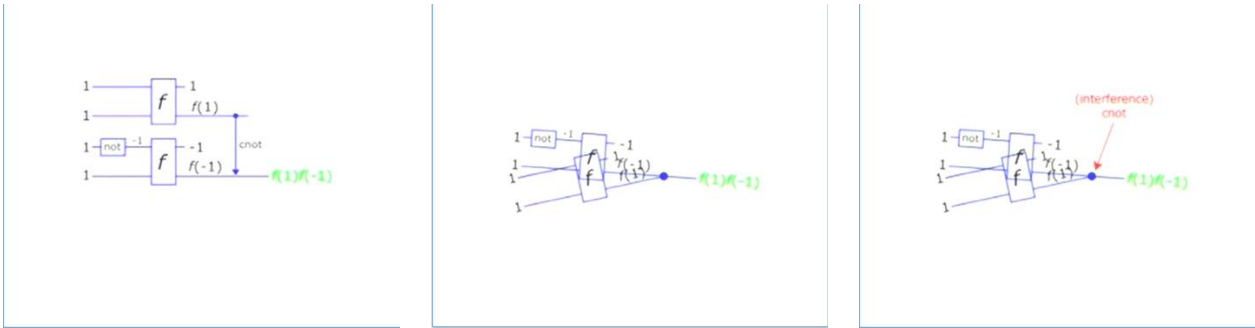


Table 5.8: What the quantum algorithm does, in short [...]

To describe the quantum algorithm that does that, as with most quantum algorithms, it's simplest to work in the Schrödinger picture, where as you will recall, the observables are constant matrices and the state changes with time. In this problem we can make the further simplification of considering only pure states. Remember that a quantum system is said to be in a pure state when its density matrix takes the form [reads equations].

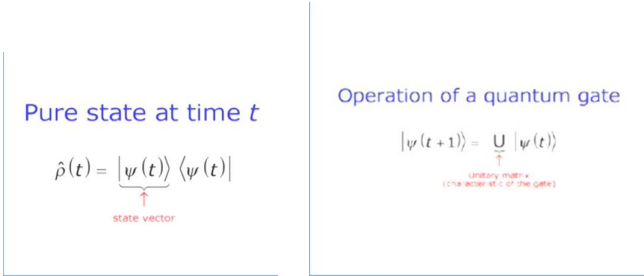


Table 5.9: What the quantum algorithm does, in short [...]

ψ is then said to be the state vector of the system and we often refer to the state vector as just the ‘state’ of such a system. The effect of a quantum gate

operating between times t and $t+1$ on an arbitrary pure state is [reads equation] where U is a constant unitary matrix characteristic of the gate. Like the reversible classical algorithms for performing this task, this algorithm involves two qubits, and that's not counting whatever qubits might be inside the oracle. Combining two quantum systems in the Schrödinger picture is much the same as doing it in the Heisenberg picture. Any tensor product of observables, one from each system, is an observable of the combined system, and if both systems are in pure systems—say ket ψ and ket ϕ —then the combined system is in a pure state, and its state vector is the tensor product of ψ and ϕ . We write the tensor product of kets without any explicit multiplication symbol like this.

And that's a ket of the combined system.

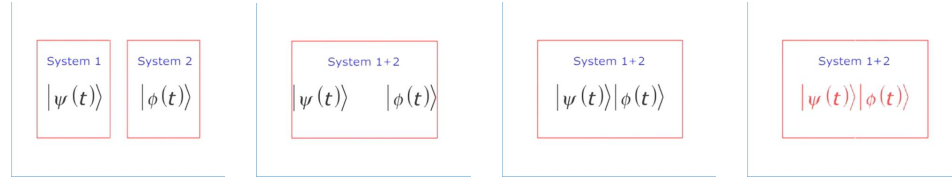


Table 5.10: Combining two quantum systems in the Schrödinger picture [...]

Consider a system of two qubits then. In an eigen-state of both their \hat{Z} observables, \hat{Z}_1 and \hat{Z}_2 , the four such eigen-states ket a ket b , where a and b range over the eigenvalues plus and minus 1, form an orthonormal basis in the four dimensional vector space of all pure states of the combined system. The space of all pure states that a system could be in is technically a Hilbert space—basically that's just a vector space with a norm—and we usually refer to it as "the Hilbert space" of the given system. Strictly speaking, pure states are always unit vectors. That follows from the definition of a pure state that I've given, because the trace of the density matrix has to be 1. But it's no big deal, some people like to work with un-normalized vectors, and they insert a normalization factor in this formula.

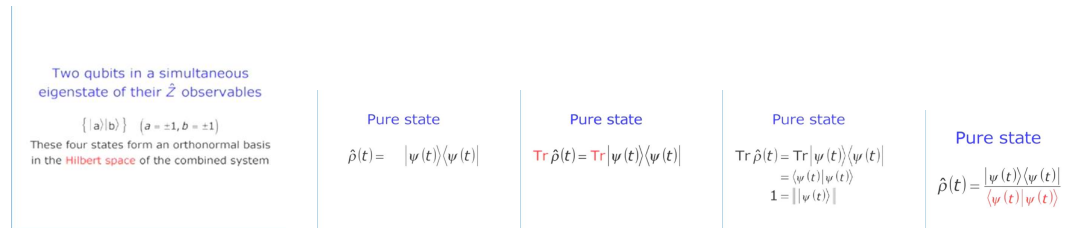


Table 5.11: Strictly speaking, pure states are always [...]

So un-normalized kets are usually called states too, and eigen-vectors of observables are called 'eigen-states'. The normalized pure states lie on a unit

sphere in the Hilbert space. And all the non-zero states along any straight line through the origin represent the same physical state.

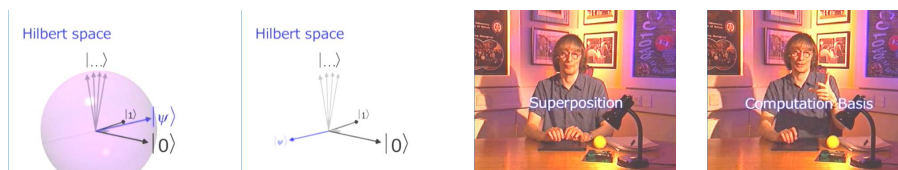


Table 5.12: A linear combination of states is called [...]

When working with pure states, we usually pick the fixed basis in the system's Hilbert space and express the evolving state of the system as a linear combination with time varying coefficients. Time varying complex coefficients of those basis states. A linear combination of states is called a 'superposition'. In quantum computation, we call whatever basis we choose for doing that the 'computation basis'. We can choose any basis as the computation basis but usually some choices are more convenient than others. Usually the most convenient one is determined by the classical operations that form part of the computation. I'll explain.

<p>In practice, the most convenient choice for the computation basis is usually determined by the classical operations that occur in the computation.</p> <p>Why?</p>	<p>When a quantum computer is executing classical operations, the set of all the \hat{Z}-observables of the qubits evolves autonomously.</p> <p>This makes it convenient to use the simultaneous eigenstates of those observables as the computation basis.</p> <p>Why...?</p>	<p>Elementary classical operations take computation basis states to computation basis states...</p> <p>To specify a gate, it is enough to specify its action on an arbitrary pure state...</p> <p>...or on an arbitrary computation basis state.</p>
--	--	---

Table 5.13: And that's why the basis of eigenstates [...]

In a classical computer, the state of the computation at each step is specified by the sharp values of a set of observables—the computational variables. So when a quantum computer simulates a classical computer, there's a set of observables which are sharp at the beginning and end of every step. And that sharpness is maintained because, as I mentioned last time, the elementary classical operations in a quantum computer have the property that a particular set of observables evolve autonomously. We usually use a notation in which these are the \hat{Z} -observables of the qubits. So that means that during periods when our quantum computer is executing only classical operations, the \hat{Z} -observables of the qubits are distinguished by being the observables on which those gates perform classical operations. Note that in general, the \hat{Z} -observables will not be sharp during those classical operations because they need not have been sharp when the classical part of the computation began. Some earlier quantum operations will in general have made them unsharp. Nevertheless, the \hat{Z} -observables form an autonomous evolving system under classical operations.

And that's why the basis of eigen-states of those said observables is often the most convenient one to use as the computation basis. I'll explain.

Classical reversible computations, starting in a computation basis state, proceed from computation basis state to computation basis state. For this and other reasons, the dynamics of simple quantum gates are particularly easy to represent in the Schrödinger picture. It's enough to specify the behavior of each member of a set of basis states such as the computation basis. For instance, the NOT gate is a single qubit gate. A single qubit has a 2-dimensional Hilbert space. So we can completely specify the behavior of the NOT gate by giving its effect on just two states, like this: ket a goes to ket $-a$, where a is plus or minus one and these are eigenstates of the cubit's \hat{Z} -observable.


<p>Dynamics of the not gate in the Schrödinger Picture</p> $ a\rangle \rightarrow -a\rangle \quad (a = \pm 1)$ $\hat{Z} a\rangle = a a\rangle$	<p>Dynamics of the not gate in the Schrödinger Picture</p> $ \psi(0)\rangle \rightarrow \psi(1)\rangle$ $ a\rangle \rightarrow -a\rangle \quad (a = \pm 1)$	 <p>Heisenberg</p> $\hat{Z}(t+1) = -\hat{Z}(t)$ $\hat{Y}(t+1) = \hat{Y}(t)$ $\hat{X}(t+1) = -\hat{X}(t)$ $\hat{I}(t+1) = \hat{I}(t)$ <p>Schrödinger</p> $ a\rangle \rightarrow -a\rangle \quad (a = \pm 1)$
--	--	--

Table 5.14: So we can completely specify the behavior of the [...]

This statement means that if the state at time 0 is the eigenvalue a eigenstate of \hat{Z} , then the state at time 1 after the gate has acted is the eigenvalue $-a$ eigenstate. Compare that with the description of the NOT gate in the Heisenberg picture and you begin to see why the Schrödinger picture is preferred for most calculations in the quantum theory of computation.


 <p>Hadamard gate</p>	<p>Hadamard Gate, H</p> $ 1\rangle \rightarrow \frac{1}{\sqrt{2}}(1\rangle + -1\rangle)$ $ -1\rangle \rightarrow \frac{1}{\sqrt{2}}(1\rangle - -1\rangle)$ <p>hence $H^2 = I$</p>	<p>Hadamard Gate, H</p> $ 1\rangle \rightarrow \frac{1}{\sqrt{2}}(1\rangle + -1\rangle)$ $ -1\rangle \rightarrow \frac{1}{\sqrt{2}}(1\rangle - -1\rangle)$ <p>hence $H^2 = I$</p>
--	--	--

Table 5.15: So we can completely specify the behavior of the [...]

Another important single qubit gate that I haven't mentioned before, and which is used in the algorithm I'm going to describe, is the so-called Hadamard gate,

named after the mathematician Jacques Hadamard for historical reasons.

It has this effect: [reads equation(s)] The Hadamard operation is like the NOT operation, a square root of the unit operation. It's its own inverse. But unlike NOT, it doesn't have a classical analog. By the way, the square root of 2 factor is just there to normalize the states. Now, here's the definition of the controlled NOT gate in the Schrödinger picture. Look how simple it is: it evolves a system of two qubits in the eigenstate x comma y of the observables \hat{Z}_2 and \hat{Z}_1 into the state x comma xy . The ket x comma y is just another way of writing the tensor product ket x ket y . And again, compare this with the Heisenberg picture definition of the controlled NOT gate.

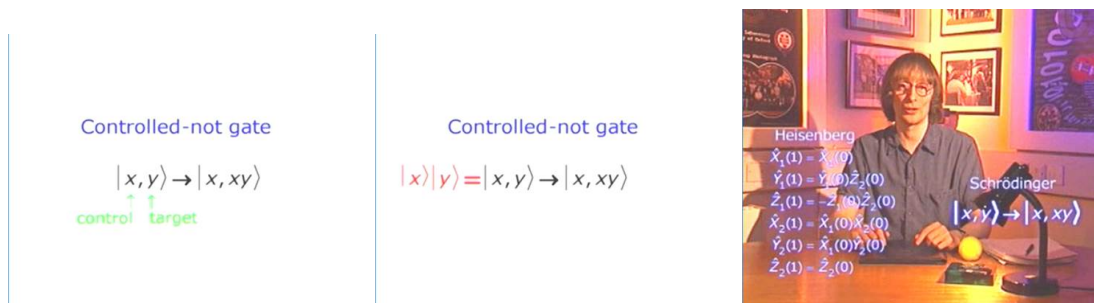


Table 5.16: Now, here's the definition of [...]

The operation of the oracle is as follows: [reads equations]

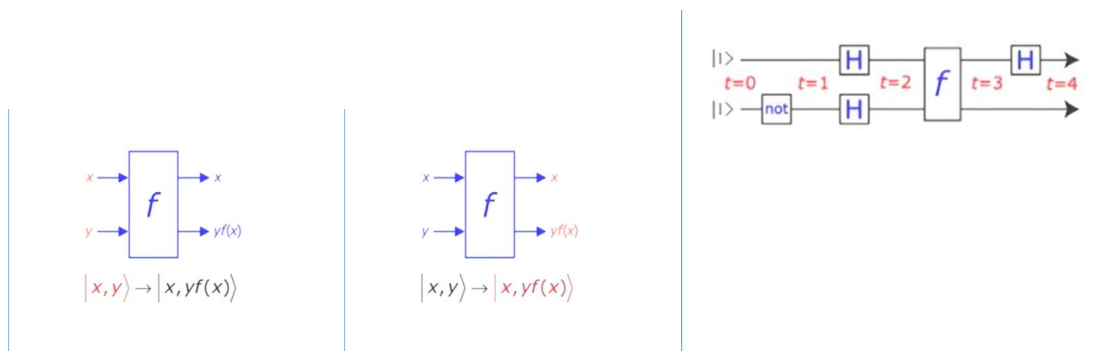


Table 5.17: The quantum network that solves our problem [...]

Now, for the algorithm. Our two qubits start in the state where they're both sharp with the value +1. We could think of that as the 'blank state' of our two-qubit computer memory. By the way, when analyzing algorithms in general, the algorithm should always start with a state where all the qubits that are not given as part of the task are blank. That way your analysis will automatically account for the resources required for any other initialization you might want

to do. The quantum network that solves our problem contains a NOT gate, three Hadamard gates, and the oracle. At time zero, the beginning of the computation, the state of each qubit is the eigenvalue +1 eigenstate of its \hat{Z} -observable. So the overall state at time zero is the tensor product of those, which we can write “1 comma 1”. Then, the second qubit encounters a NOT gate which flips its \hat{Z} observable to -1 . Next, both qubits pass through Hadamard gates. And we can just substitute from the definition of the Hadamard gate what the state will be at time 2: [reads equation(s)] Then, the oracle acts on the qubits—once. But neither of the \hat{Z} -observables in the input is now sharp, so we are presenting the oracle with four different inputs in different universes.

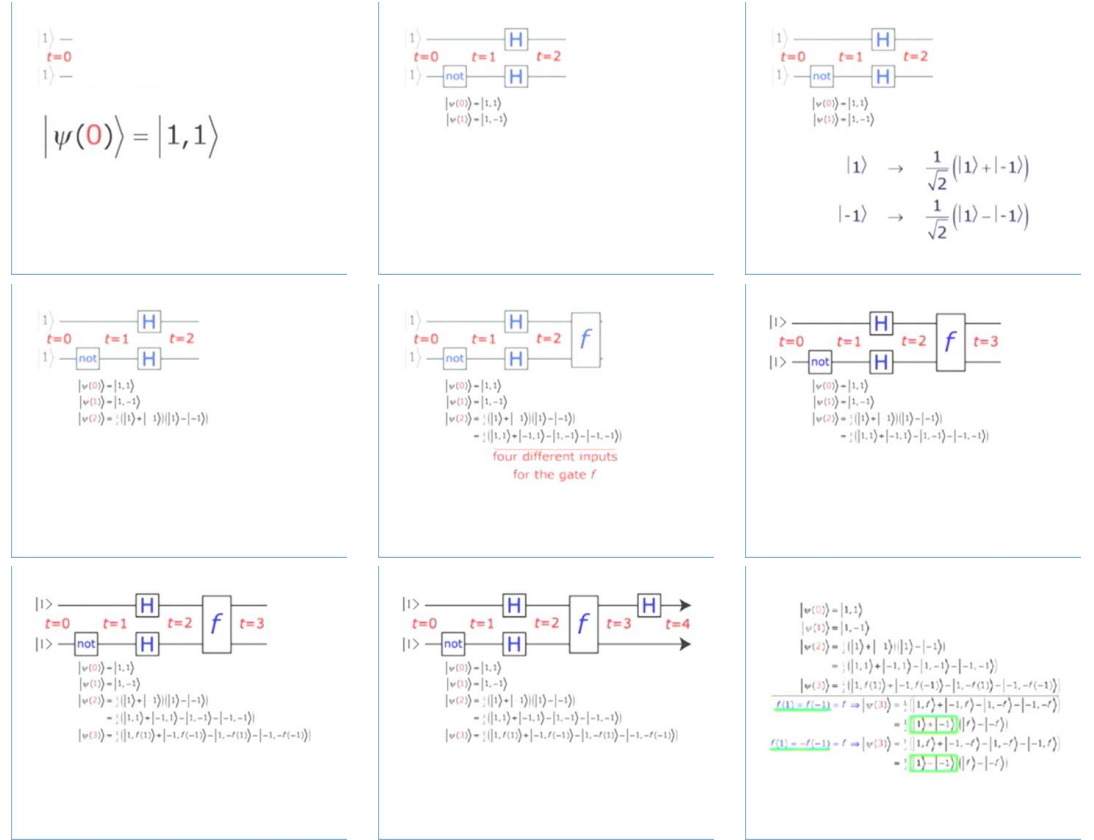


Table 5.18: At time zero [...] Then, the second qubit [...]

Next, well, the oracle will in general take a long time to run but the running time is a constant independent of the input, and we are not interested in what it is at the moment. So let's just call the time when its complete ‘time 3’. And again, we can fill in what ψ of 3 will be. At this point, the computation isn't quite finished yet, we still have that final Hadamard gate to go. But let's just

look at this state. Suppose that $f(1)$ does equal $f(-1)$, so we can call them both f . Then the state at time 3 will take this form. Which factorizes into a tensor product. So we see that the qubits are still individually in pure states, and the first cubit is in a state proportional to ket 1 plus ket -1 . Okay, that's if $f(1)$ equals $f(-1)$. If they're unequal then $f(1)$ must equal $-f(-1)$, which we can call ' f ' again. And again, the qubits are each in a pure state. This time qubit 1 is in a state proportional to ket 1 minus ket -1 . If we can distinguish those two states of cubit 1, we shall have determined whether $f(1)$ and $f(-1)$ are the same or different. And that's just what the final Hadamard gate does, because Hadamard acting on [reads equation]:

Hadamard Gate, H

$$H|1\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |-1\rangle)$$
$$H|-1\rangle = \frac{1}{\sqrt{2}}(|1\rangle - |-1\rangle)$$

hence: $H=H^{-1}$, and so:

$$H\frac{1}{\sqrt{2}}(|1\rangle + |-1\rangle) = |1\rangle$$
$$H\frac{1}{\sqrt{2}}(|1\rangle - |-1\rangle) = |-1\rangle$$

$$H\frac{1}{\sqrt{2}}(|1\rangle + |-1\rangle) = |1\rangle$$
$$H\frac{1}{\sqrt{2}}(|1\rangle - |-1\rangle) = |-1\rangle$$
$$f(1) = f(-1) = f \Rightarrow |\psi(3)\rangle = \frac{1}{2}(|1, f\rangle + |-1, f\rangle - |1, -f\rangle - |-1, -f\rangle)$$
$$= \frac{1}{2}(|1\rangle - |-1\rangle)(|f\rangle - |-f\rangle)$$
$$f(1) = -f(-1) = f \Rightarrow |\psi(3)\rangle = \frac{1}{2}(|1, f\rangle + |-1, -f\rangle - |1, -f\rangle - |-1, f\rangle)$$
$$= \frac{1}{2}(|1\rangle - |-1\rangle)(|f\rangle + |-f\rangle)$$
$$|\psi(4)\rangle = \frac{1}{\sqrt{2}}|f(1) \neq f(-1)\rangle(|f\rangle - |-f\rangle)$$

Table 5.19: Although these computations were performed in [...]

And we can summarize that as: at time 4, the state is proportional to ket $f(1)f(-1)$ times some state of cubit number 2. So the \hat{Z} -observable of qubit 1 now contains information that depends logically on the outcomes of the computations of both $f(1)$ and $f(-1)$. And it's sharp. Although those computations were performed in different universes between times 2 and 3, the interference phenomenon effected by the Hadamard gate between times 3 and 4 combine those values, and caused the answer to appear in qubit 1 in all the universes. \hat{Z}_1 is sharp at time 4.

31:03



Table 5.20: The version I've showed you [...]

The problem solved by this algorithm has come to be known as the Deutsch

problem. And the algorithm I've just described as the Deutsch algorithm. I should say that that gives me slightly too much credit—the algorithm that I originally proposed was somewhat less elegant and significantly less powerful than this one. I refer you to the work examples to see in what way it was less powerful. The version I've shown you was published in 1998 by Richard Cleve, Art Ekert, Kiara Macchiavello, and Mike Mosca. The way this algorithm works is typical of how quantum algorithms work in general—at least typical of those that have been discovered so far.

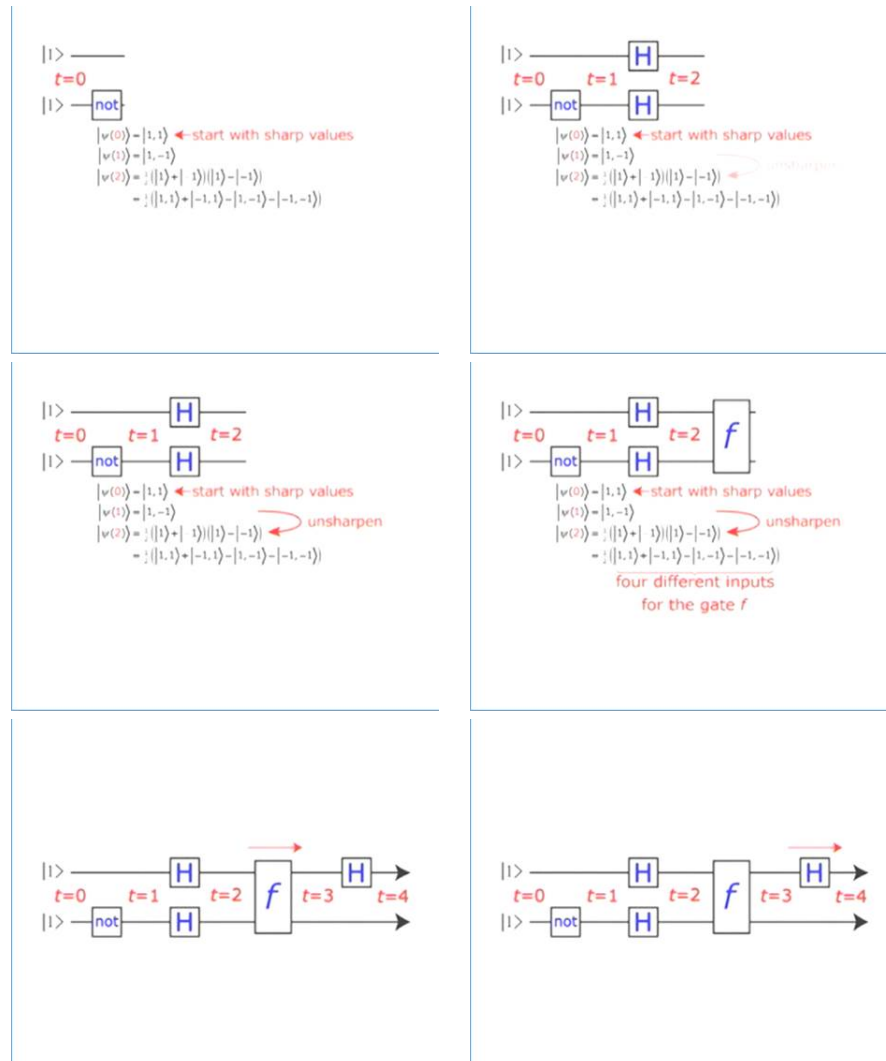


Table 5.21: [...] means that these computations are interference phenomena.

Namely, we start out with the state of the system being an element of the computational basis. That is to say we prepare the \hat{Z} observables with particular initial values. And the first thing we do is unsharpen those observables, so that they contain many possible values. In our case, all four possible values.

That unsharpening is a quantum operation. Then, we perform a classical reversible computation using coherent quantum components. In our case, that's the computation of f by the oracle. But this computation is being done with different inputs in different universes, giving different outputs too, and then these outputs are somehow combined using another quantum operation which gives a sharp answer in this case, and more generally as sharp as possible an answer as the output. And that means that these computations are interference phenomena. Observables which are sharp become unsharp, and then sharp again. An algorithm which performs multiple classical computations on unsharp inputs and then combines them is said to be using 'quantum parallelism'. That's because, as I've shown, the process is reminiscent of classical parallel computation, with the difference that you don't need a second copy of the oracle. You use the parallel universe counterparts of the one you're given.



Table 5.22: [...] That unsharpening is a quantum operation.

Okay, so by using this quantum algorithm, we've gained a factor of 2 in speed.

But the real significance of the existence of this algorithm is not its speed, it's the fact of being able to manage with just one evaluation of f to obtain information that depends logically on both values. It's the fact that during that function evaluation, something is going on that cannot be analyzed as the sequence of states in which each computational observable has a single value.

This is the characteristic of this new mode of information processing, which is not the implementation of any classical algorithm, and performs a task that no classical algorithm can perform. The ability to perform this particular task is unlikely to have any practical application though. Well, in some very contrived circumstances it just might—say you have a time limit by which you have to perform a very important computation that consists of evaluating $f(1)f(-1)$ for some complex algorithm f . As I said, the real significance is theoretical.

But next time, I'll describe an amazing quantum algorithm, again using quantum parallelism, that is both theoretically interesting and likely to be useful in a wide range of practical applications.

35:38



Table 5.23: [...] both theoretically interesting and likely to be useful [...]

Chapter 6

Grover's Search Algorithm

Today I am going to describe a remarkable quantum algorithm which was invented by Lov Grover in 1996. It's a search algorithm suitable for a very broad category of computational tasks known as algorithmic searching. Algorithmic searching is just exhaustive searching for a number with a given mathematical property. It's a programmer's last resort in addressing computational tasks where there's a quick way of verifying that a given number is the answer once you have it, but no easy way of constructing the answer. In other words, the task is: we're given a criterion for whether a number is the answer or not. We're given it in the form of an algorithm F that delivers an output $+1$ or -1 where -1 means that the input meets the criterion and $+1$ that it doesn't. And we have to find a value t , the target value, such that $F(t) = -1$.

<p>Algorithmic Searching</p> <p>Given an algorithm F, find a number t such that $F(t) = -1$</p>	
---	--

Table 6.1: It's a programmer's last resort [...]

Suppose there are N possible values that have to be searched and suppose for simplicity that N is exactly 2 to the L for some integer L , so that we can encode each possible value uniquely in a register of L qubits. By the way, note that the conventional representation for binary numbers stored in qubits is that the eigenvalue $+1$ state represents the binary digit 0, and the -1 state represents the digit 1 (see pic above). Anyway, suppose that there's exactly one value that meets the criterion in the range that we want to search. that's the target value t with $F(t) = -1$. For all other values x not equal to t , $F(x) = +1$ and we're given a classical reversible algorithm for evaluating F on an arbitrary L

bit input. Like last time, we're given this in the form of an oracle. It's some sort of computer operating coherently on a certain number of qubits that we can give it as input, and then get out again after a fixed period. But we're not allowed to look inside it. Now that models the fact that we're only doing this search in the first place because we've exhausted all ways of simplifying the problem analytically, and we're looking for an exhaustive search algorithm, one that doesn't depend on our knowing anything about the structure of the function F . So, the oracle computes F on an L bit input and gives us a 1 bit output—whether it meets the criterion or not. But again, because of reversibility it must actually have the same number of outputs as inputs so it must operate something like this.

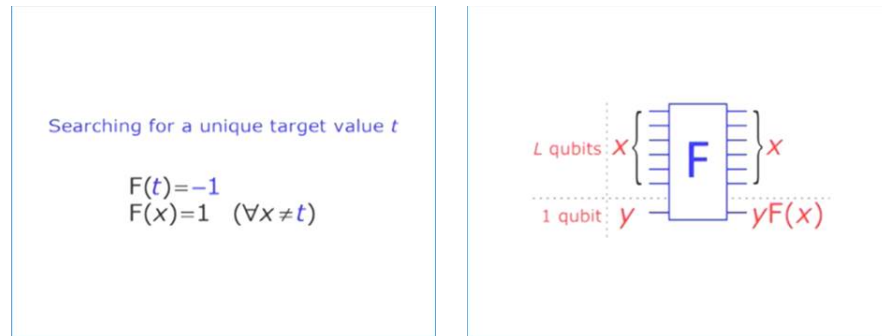


Table 6.2: Like last time, we're given this in the form of an [...]

For inputs x and y , where x can take one of 2 to the L values and y can take one of 2 values $+$ and -1 , it gives an output x and $yF(x)$. So the oracle as a whole operates on $L + 1$ cubits. And it may also contain an unknown number of internal qubits that we never see. How many evaluations of F —how many invocations of the oracle—are we going to need to find the unique value of t such that $F(t) = -1$? Well, by similar argument to last time, it's clear that we may need to invoke the oracle as many as $N - 1$ times. It's $N - 1$ rather than N because if the conditions of the task say that there's exactly one value meeting the criterion, and we've checked all but one possible values, then we know that the answer is the remaining value. But that's the only break we get. We could check the values randomly and then on average we'd probably need about $\frac{N}{2}$ oracle invocations—probably. Using quantum computation in which the oracle receives different inputs in different universes, we can do a great deal better than that. And that's what Grover's algorithm does.

06:05

Grover's algorithm uses three simple subroutines, so I'll describe those separately first and then put them together. The first subroutine involves the Hadamard gate that I introduced last time. It consists of applying a Hadamard gate to each of L qubits. In our case, it's the L cubits in our L qubit register, like this. I'll call this operation "bold face H" to distinguish it from a single qubit Hadamard gate. The algorithm begins, as I advocated last time, with all

L of these qubits in their blank initial state, where they all have sharp values $+1$. And I'll call that state ket 0, to state if L qubits all in state $+1$. The effect of H on the blank state is H on plus, plus, plus and so on, which equals 1 over root 2 to the L times ket plus plus ket minus times ket plus plus ket minus and so on, L factors in the tensor product. This state of the L qubits plays an important role in the analysis of this algorithm, so let me give it a name: μ .

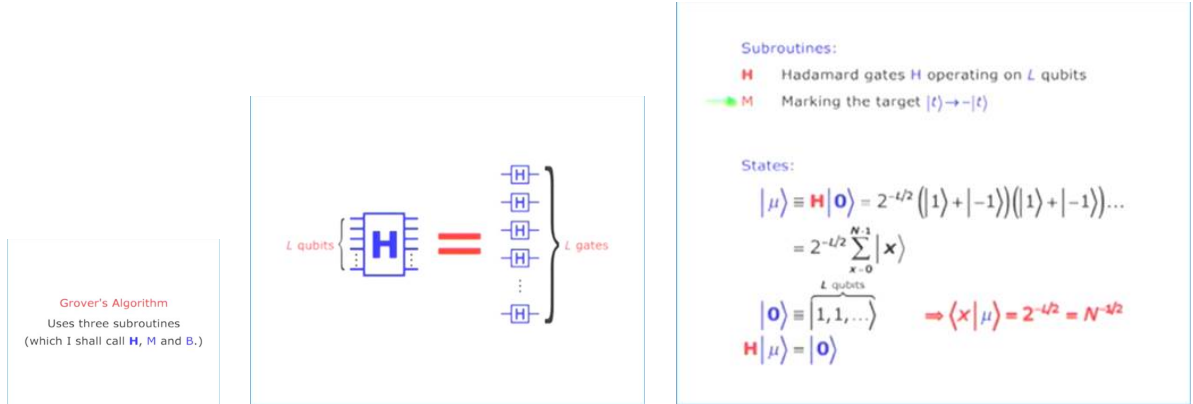


Table 6.3: Grover's algorithm uses three simple subroutines [...]

So, H acting on ket 0 equals ket μ . And since the Hadamard gate is its own inverse, it must also be true that H acting on ket μ equals ket 0. If we multiply out μ , we see that it consists of a superposition of states with all N , or 2 to the L , possible sequences of plus and minus values—representing the 2 to the L possible values of x . Hence, if ket x is any computation basis state, the scalar product of x with μ is 1 over root N . Okay, the next subroutine involves the oracle, with its L qubits to hold the argument of F and its one auxiliary qubit. I'll call this the marking subroutine M —you'll see why in a moment.

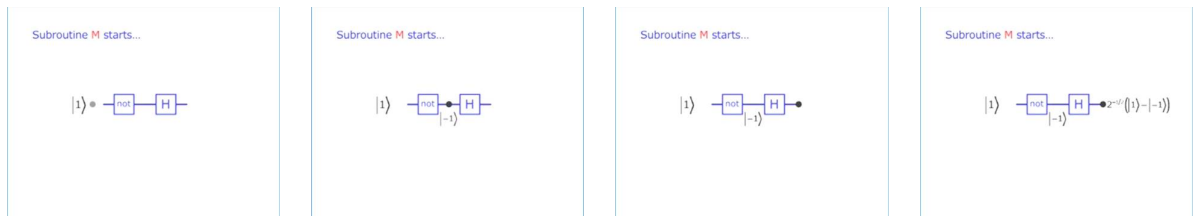


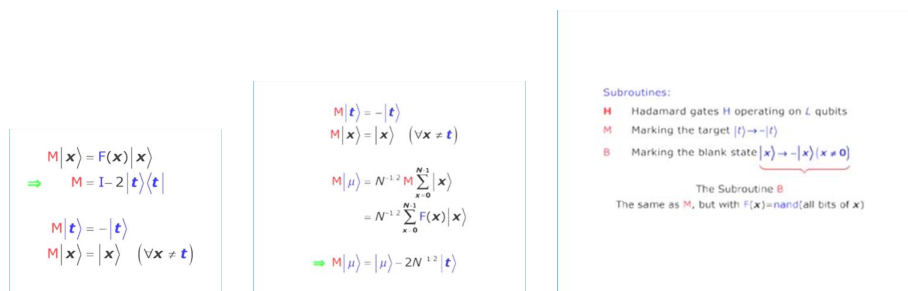
Table 6.4: I'll call this the marking subroutine M [...]

It starts with the auxiliary bit being put through a NOT gate and then a Hadamard gate. That's just to prepare it in the state 1 over root 2 ket plus minus ket minus. With that as the input for the auxiliary qubit, the effect of the oracle on the sharp state of the first L qubits with the value x is—well as

Figure 1 illustrates the implementation of a quantum algorithm through a sequence of six diagrams, each representing a step in the circuit's execution. The diagrams are arranged in two rows of three.

- Diagram 1 (Top Left):** Labeled "Subroutine M". It shows a quantum circuit with a NOT gate, a Hadamard gate, and a phase shift $z^{-1/2}(|1\rangle - |-1\rangle)$ before a unitary F . The output is a state $|x\rangle$.
- Diagram 2 (Top Middle):** Labeled "Subroutine M". It shows the same circuit as Diagram 1, but the phase shift is now $z^{-1/2}(|1\rangle - |-1\rangle)$.
- Diagram 3 (Top Right):** Labeled "Subroutine M". It shows the same circuit as Diagram 1, but the phase shift is now $2^{-1/2} |x\rangle(|1\rangle - |-1\rangle)$.
- Diagram 4 (Bottom Left):** Labeled "Subroutine M". It shows the same circuit as Diagram 1, but the phase shift is now $2^{-1/2} |x\rangle(|1\rangle - |-1\rangle)$.
- Diagram 5 (Bottom Middle):** Labeled "Subroutine M". It shows the same circuit as Diagram 1, but the phase shift is now $2^{-1/2} F(x) |x\rangle(|1\rangle - |-1\rangle)$.
- Diagram 6 (Bottom Right):** Labeled "Subroutine M". It shows the same circuit as Diagram 1, but the phase shift is now $2^{-1/2} F(x) |x\rangle(|1\rangle - |-1\rangle)$.

So we don't have to mention the auxiliary qubit explicitly when we analyze algorithms involving M , just as we don't explicitly mention the unseen qubits in the oracle when it computes F . The net effect of M on the L qubits holding x can be summarized as: M acting on x is $F(x)$ times x , explicitly the unitary matrix that is M is unit matrix minus 2 times ket t bra t .



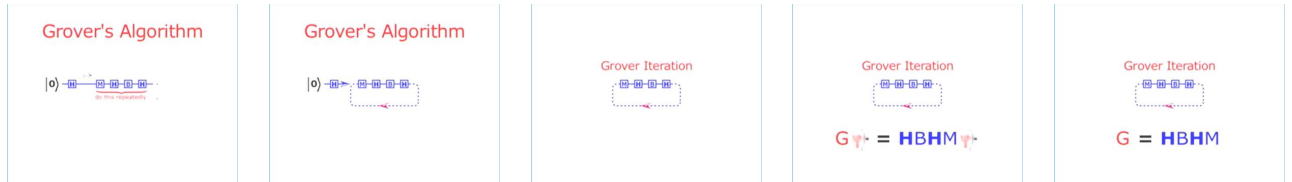
That's because the computation basis states are orthonormal, so M operating on t is minus ket t and M operating on any other computation basis state leaves the state unchanged. Hence also, the effect of M on the state μ is $M\mu$

equals 1 over root N times M [on] the sum of all possible values of x of ket x , which equals 1 over root N times the sum of $F(x)$ times x . Now, for all but one value of x , $F(x)$ is one, only $F(t)$ is minus one. Now you can see why I called this the “marking operation”—it “marks” the target term in the superposition by changing its sign. And we can write the effect of M in terms of states I’ve already named like this: $M\mu$ equals μ minus 2 over root N ket t . The third subroutine, which I’ll call B , is just marking the blank state. Or for convenience actually we are going to mark everything but the blank state. So, this operation is like M , except that instead of the oracle it uses a network that performs the NAND operation on all L inputs, and exclusive OR’s the result into the auxiliary qubit. That means that for computation basis states inputs the auxiliary qubit is flipped unless the first L inputs are all +1. And we’ll use the same trick of using a Hadamard gate to prepare the auxiliary qubit in the state 1 over root 2 ket plus minus ket minus, giving us an L qubit operation B . The effect of B is: [that] B on the blank state zero equals plus the blank state, and B on any other computation basis state is minus that state. So B is the unitary matrix 2 ket 0 bra 0 minus the unit matrix.

<p>The Subroutine B (The same as M, but with $F(x) = \text{nand}(\text{all bits of } x)$)</p> $B 0\rangle = + 0\rangle$ $B x\rangle = - x\rangle$ $B = 2 0\rangle\langle 0 - I$	<p>Ingredients of Grover's Algorithm</p> <ul style="list-style-type: none"> H Hadamard gates H operating on L qubits M Marking the target $t\rangle \rightarrow - t\rangle$ B Marking the blank state $x\rangle \rightarrow - x\rangle (x \neq 0)$ $ 0\rangle \equiv \frac{1}{\sqrt{2^L}} \sum_{x=0}^{2^L-1} x\rangle$ $H 0\rangle = \mu\rangle = \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2^L}} \sum_{x=0}^{2^L-1} x\rangle \right) \left(\frac{1}{\sqrt{2}} \sum_{x=0}^{2^L-1} x\rangle \right) \dots$ $= N^{-1/2} \sum_{x=0}^{2^L-1} x\rangle$ $H \mu\rangle = 0\rangle$	<p>Ingredients of Grover's Algorithm</p> <ul style="list-style-type: none"> H Hadamard gates H operating on L qubits M Marking the target $t\rangle \rightarrow - t\rangle$ B Marking the blank state $x\rangle \rightarrow - x\rangle (x \neq 0)$ $ 0\rangle \equiv \frac{1}{\sqrt{2^L}} \sum_{x=0}^{2^L-1} x\rangle$ $H 0\rangle = \mu\rangle = \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2^L}} \sum_{x=0}^{2^L-1} x\rangle \right) \left(\frac{1}{\sqrt{2}} \sum_{x=0}^{2^L-1} x\rangle \right) \dots$ $= N^{-1/2} \sum_{x=0}^{2^L-1} x\rangle$ $H \mu\rangle = 0\rangle$
---	--	--

Table 6.7: So B is the unitary matrix [...]

Those are the ingredients of Grover’s algorithm. All of them are operations on L cubits. There’s the operation H which transforms the blank state to the grand superposition μ , and viceversa; M which marks the target state; and B , which marks all computation basis states except the blank state.



Grover’s algorithm consists of starting with a blank initial state, perform H which makes the state μ , and then perform a certain number of iterations of the sequence: M , then H , then B , then H again. So the net effect of such an iteration, called the Grover iteration, on an arbitrary state, is given by

the unitary matrix $HBHM$. They're in reverse order to the order they're performed in because it's the right-most factor that hits the ket first. We'll see in a moment how many of these Grover iterations you have to do. That's it.

16:30

Now, what does all that do and why does it work? There's a beautiful geometric interpretation of what it's doing. Let's take a look at a 2 dimensional plane through the 2 to the L dimension Hilbert space of these L qubits. The plane is defined as the one containing both the target state t and the superposition μ . Now, t and μ are very nearly orthogonal for large N , but never quite orthogonal, their scalar product is 1 over root N . Let's define a state that is orthogonal to t here. Call it ϕ . And now consider a family of states all lying in this plane. $\cos \theta$ times ϕ plus $\sin \theta$ times t , where θ is a real parameter.

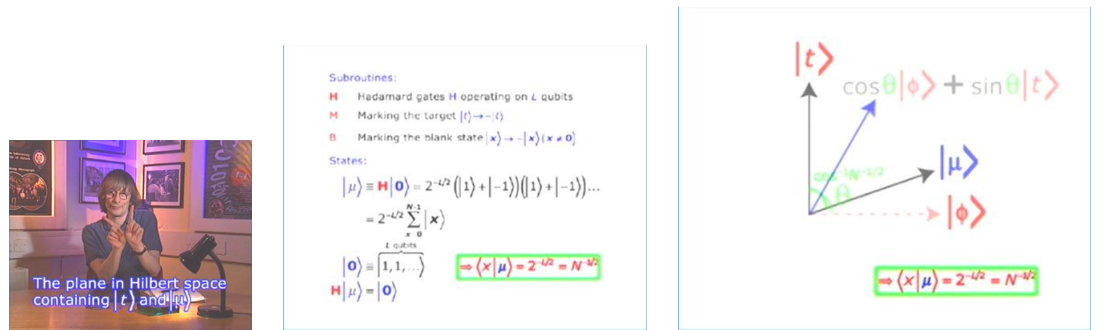
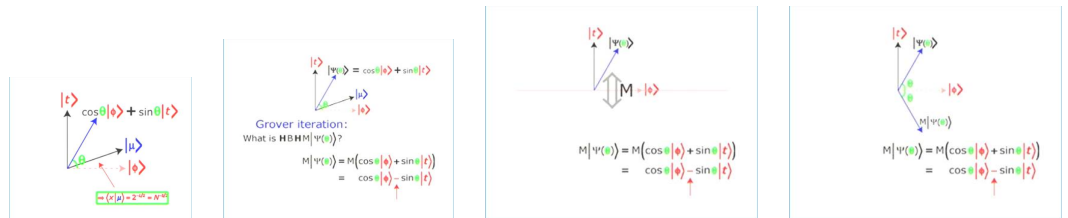


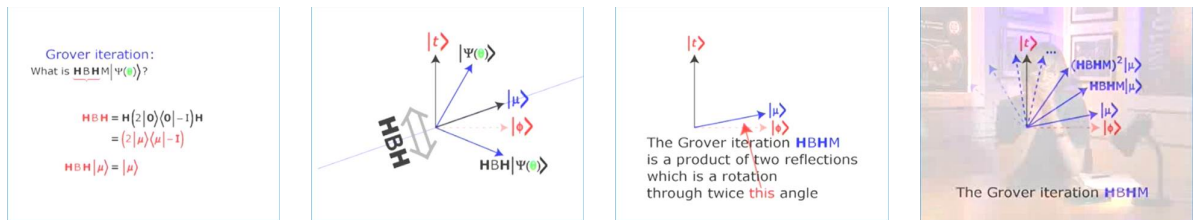
Table 6.8: So B is the unitary matrix [...]

The state that we want to end up in, namely ket t , is in this parameterized family with θ equals π by 2. And at the moment after our initial Hadamard operation, just before we start the iterations, we're in the state μ which is also of this form but with θ equals arcsin of 1 over root N . For a general state of this form, what's the effect of one Grover iteration?

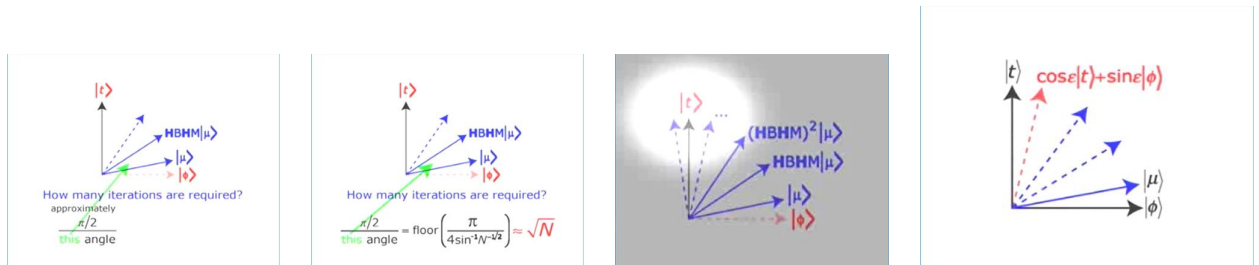


Well, the Grover iteration begins with an M operation and we know what the effect of M is. It changes the sign of the coefficient of the target ket t , and leaves all other computation basis states unchanged. Hence the effect of M on the general state in this plane is to reflect it in the horizontal axis. After M , we do HBH . And we can simplify that if we just substitute what H and B are as unitary matrices. We get HBH equals 2 ket μ bra μ minus 1, which has no effect on the state μ but changes the sign of any state perpendicular to μ .

In other words, the effect of HBH on an arbitrary state in this plane is to reflect it in the line μ . The product of these two reflections, namely the Grover iteration as a whole, is a rotation. You can easily prove that it's a rotation through an angle $2 \arcsin 1/\sqrt{N}$; exactly twice this angle. And in the anti-clockwise direction—in other words, towards t . So each Grover iteration rotates the state a little closer to our target, until we go past it and get further away from it again. So it's vital to choose the right number of iterations.

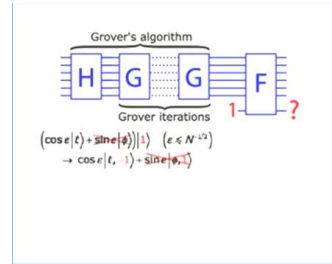
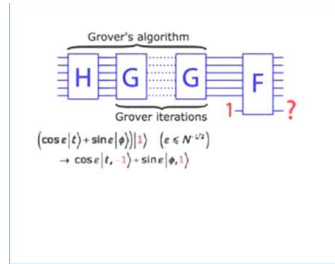


The right number will be: π over 2 divided by this angle that the rotation rotates the state by, except that because we start the rotation at half that angle away from the horizontal, that means that the state will be closest to the target t after the integer part of that number of operations, namely floor of π over $4 \arcsin 1/\sqrt{N}$. Roughly speaking, that's a constant times the square root of N . So, Grover's algorithm completes its search using about the square root of the number of oracle calls required using classical computation. That's a tremendous saving. You can search a million possibilities in only about a thousand iterations. A trillion possibilities in only a million iterations.



Now, just a small detail. The state at the end won't have exactly hit the target state. Except when N equals 4, by the way, you might like to check that interesting special case where a single oracle call is actually enough to do the whole search. But for all other values, the output of Grover's algorithm is not a computation basis state at all, it's a state of the L qubits that's close to but not equal to the target state. Does that matter? It doesn't, because, well, suppose Grover's algorithm delivers a state $\cos \epsilon |t\rangle + \sin \epsilon |\phi\rangle$. From the argument I've just given, we can expect epsilon to be less than about $1/\sqrt{N}$ which is very small. Well, suppose we then measure the output value of the L qubits and check whether it's correct—whether it's the value such that

$F(t) = -1$. We can do that very easily by using the oracle one more time, this time with the auxiliary qubit initially blank with the sharp value $+1$ like this, and then running the oracle will flip the sign of the component that goes with ket t and leave the other one alone.



Grover's algorithm

Grover iterations

$$(\cos \epsilon |t\rangle + \sin \epsilon |\phi, 1\rangle) \quad (\epsilon \leq N^{-1/2})$$

$$\rightarrow \cos \epsilon |t, -1\rangle + \sin \epsilon |\phi, 1\rangle$$

$$\Rightarrow \rho = (\cos \epsilon |t, -1\rangle + \sin \epsilon |\phi, 1\rangle)(\cos \epsilon \langle t, -1| + \sin \epsilon \langle \phi, 1|)$$

$$\langle \hat{Z} \rangle = \text{Tr } \rho \hat{Z}$$

$$= (\cos \epsilon \langle t, -1| + \sin \epsilon \langle \phi, 1|) \hat{Z} (\cos \epsilon |t, -1\rangle + \sin \epsilon |\phi, 1\rangle)$$

$$= -\cos^2 \epsilon + \sin^2 \epsilon$$

$$= -1 + O(1/N)$$

What is then the expectation value of the \hat{Z} component of that auxiliary qubit?

If the output had been sharply at the target value, then \hat{Z} for the auxiliary qubit would also be sharp with value -1 . If its expectation value is anything above that, that means that the answer was wrong in some universes—and we can work out how many. The expectation value of \hat{Z} is Trace of $\rho \hat{Z}$. And ρ for this pure state is this [equation appears on screen]. The expectation value of \hat{Z} is this [screen gets updated with new equation]. So we get the expectation value minus cos squared epsilon + sin squared epsilon, which is -1 plus the order of epsilon squared. So, the probability of seeing a wrong answer—that is, the proportion of universes in which the answer will be wrong—is of order 1 over N . Very small. In practice, many other small errors will be introduced by physical factors such as noise, or the fact that the computer doesn't perfectly match its ideal design. But it's in the nature of algorithmic searching that we can't be lead astray by such errors. We check the result and in the rare cases when it's wrong we just run the algorithm again.

Grover's algorithm also works in cases where there are M values t_1, t_2, \dots, t_M , each satisfying $F(t_i) = -1$.

It then requires about $\sqrt{N/M}$ oracle invocations.



You'll see in the worked examples that Grover's algorithm can also be used in cases where more than 1 value satisfies the target criterion, and that it speeds up such searches by the same factor. It has been proved that the

algorithm is optimal—that is, no algorithm quantum or classical can ever do an exhaustive algorithmic search than Grover’s. Though bear in mind that the proof of this applies only to the oracle version of the searching task, which is an idealization. In real life algorithmic searches, we often do know something about the structure of the function we’re investigating. And this knowledge can be used to speed up searching. The possibility is still open that in some cases it speeds up quantum searching by more than classical searching. In other cases though, it offsets the benefit of quantum searching altogether. I once mentioned in a popular article about quantum computers that one day, Grover’s algorithm will be used to make super powerful chess playing machines. My thought was that since the best classical chess playing algorithms work by exhaustively searching all possible continuations of the game from a given position, Grover’s algorithm would allow one to search the square of the number of possibilities in given number of steps—a huge improvement. But no, as was pointed out by Richard Cleve, the nature of that kind of search which has a tree like structure is that most of the work of computing any of the final positions is shared with many other final positions. And under those circumstances the advantage of Grover’s algorithm over conventional tree searching disappears. And the same seems to be true of game-playing in general. Though I should add that that doesn’t rule out the possibility that there may be other quantum algorithms for playing certain games.



Grover’s paper in which he first published his algorithm was called “a fast quantum mechanical algorithm for database search”. That’s a slightly misleading name because the algorithm is probably at its least useful in searching databases. First of all, if your database is a custom-built physical object like an oracle, containing essentially a read-only memory of N bits or N qubits, then it would surely be easier just to store the single value t and it could just tell you that value on request and you wouldn’t have to do any searching. On the other hand, if the database is not known in advance, or if the query you’re going to make is not known in advance—suppose the data was some recently gathered data from the search for extra-terrestrial intelligence and you want to search it for a given pattern—well, then the act of reading the data into a

quantum computer memory or imprinting it permanently on an oracle is itself a task that takes of the order of N operations. So, Grover's algorithm would only speed up database searching by perhaps a constant factor. And even that factor might be swamped by the extra technological difficulty of doing coherent quantum computations as opposed to incoherent classical ones. Or it might not. But in any case, once there are quantum computers, Grover's algorithm will be unrivaled when it comes to algorithmic searching which includes a wide variety of very important tasks. There's a large number of tasks, some of which currently take up a lot of computer time, where there's at present no known alternative but the hard slog of try $F(1)$, try $F(2)$, try $F(3)$ and so on. And where having tried any number of such guesses gives you little or no useful information about where the target may be among the remaining possibilities. Perhaps the archetypal case of this is "crypt-analysis"—the science of reading encrypted messages. One of the basic computational tasks of cryptanalysis comes up when we know the cryptographic system that was used by the writer of an encrypted message, but we don't know the key that they used for that session. We know the decoding algorithm but not the key that would make it decode the cipher text correctly. Say we have an encrypted message and ciphertext c , and we know that the original plain text was written in English, say. To solve that problem by algorithmic search means to try one value of k after another until we find one for which the decoding algorithm yields a message that seems to be an English text.

A message m has been encrypted to form a cyphertext c .

A basic cryptanalytic task: Find m .

We know c and the decoding algorithm D where

$$m = D_k(c)$$

but we don't know the key k .

If all else fails, we try all possible values of k until we find one for which $D_k(c)$ makes sense as a message.



The Complexity Class NP

A computational task with input parameter N is in NP if there exists an 'efficient' classical algorithm for recognising a solution.

'Efficient' means requiring at most a power of $\log N$ computational steps

More generally, the problems for which algorithmic search is useful are in the complexity class called NP. Basically, problems where it's easy to recognize a solution once you have it. And they especially include the important subset of NP called NP-complete, which includes problems such as the traveling salesman problem. Grover's algorithm will speed up the solution of such problems by a factor of root N . That doesn't make such problems tractable on a quantum computer in the language of complexity theorists, because the technical definition of a tractable task is one which for very large N requires only some power of $\log N$ steps. But if you're a programmer contemplating an algorithmic search of a trillion possibilities to solve some vital design problem

or whatever, it will be a great help to have a great of doing that that uses only a million function calls instead of a trillion. With cryptanalysis, the numbers are much larger still. And in general, the harder the task is the greater the advantage Grover's algorithm will confer.

Let me try to give a feel for the numbers involved. Imagine some future quantum computer that's performing Grover's algorithm. Say it has a quantum processor capable of performing 100 million calls of a criterion F per second. And suppose its engaged in a particularly arduous search through a space of 10 to the 30 possible solutions looking for the one with $F(t) = -1$. How long will that take? Well, it will require about 10 to the 15 calls of F , which at a hundred million calls per second, will take about 10 to the 7 seconds, or about 4 months. A classical computer with the same processor speed would require about 5 times 10 to the 29 calls of F which would take it 5 times 10 to the 21 seconds, or very nearly the age of the universe.

That's the size of the practical benefit of the quantum search algorithm over the classical. But the theoretical implication is even more mind boggling. Because we know how Grover's algorithm works. It isn't just doing 10 to the 15 evaluations of F in parallel, which you could mimic classically, by say, making all the computers on Earth work on nothing but this problem. The quantum processor is doing 10 to the 30 possible computations of F in parallel every time it's invoked. If all the silicon in the whole of our planet were made into microchips of one cubit millimeter each, and they were all set performing different computations, there would still be fewer distinct computations going on in that giant parallel computer than there would be in a single quantum processor performing Grover's algorithm.

That's a measure of the complexity of structure and process that exists in ordinary matter just beyond our perception because quantum processes are of course going on all the time everywhere. In a quantum computer some small part of that complexity is put to good use.



Table 6.9: Our intention was/is to promote the videos (sponsored by Quiprocone Network of Excellence and Hewlett-Packard) by capturing their essence (content plus style) in a format that is friendlier to browsing and as a reference. (Lectures available at http://www.quiprocone.org/Protected/DD_lectures.htm)