Chapter 25 Quantum Techniques in Cryptography

Quantum computing is a new area of research that has only recently started to blossom. Quantum computing and quantum cryptography were born out of the study of how quantum mechanical principles might be used in performing computations. The Nobel Laureate Richard Feynman observed in 1982 that certain quantum mechanical phenomena could not be simulated efficiently on a classical computer. He suggested that the situation could perhaps be reversed by using quantum mechanics to do computations that are impossible on classical computers. Feynman didn't present any examples of such devices, and only recently has there been progress in constructing even small versions.

In 1994 the field of quantum computing had a significant breakthrough when Peter Shor of AT&T Research Labs introduced a quantum algorithm that can factor integers in (probabilistic) polynomial time (if a suitable quantum computer is ever built). This was a dramatic breakthrough as it presented one of the first examples of a scenario in which quantum techniques might significantly outperform classical computing techniques.

In this chapter we introduce a couple of examples from the area of quantum computing and quantum cryptography. By no means is this chapter a thorough treatment of this young field, for even as we write this chapter significant breakthroughs are being made at NIST and other places, and the field likely will continue to advance rapidly.

There are many books and expository articles being written on quantum computing. One readable account is

[Rieffel-Polak].

25.1 A Quantum Experiment

Quantum mechanics is a difficult subject to explain to nonphysicists since it deals with concepts where our everyday experiences aren't applicable. In particular, the scale at which quantum mechanical phenomena take place is on the atomic level, which is something that can't be observed without special equipment. There are a few examples, however, that are accessible to us, and we now present one such example and use it to develop the mathematical formulation needed to describe some quantum computing protocols.

Since quantum mechanics is a particle-level physics, we need particles that we are able to observe. Photons are the particles that make up light and are therefore observable (similar demonstrations using other particles, such as electrons, can be performed but require more sophisticated equipment).

In order to understand this experiment better, we recommend that you try it at home. Start with a light source and three Polaroid® filters from a camera supply store or three lenses from Polaroid sunglasses.

Label the three filters A, B, and C. Rotate them so that they have the following polarizations: horizontal, 45°, and vertically, respectively (we will explain polarization in more detail after the experiment). Shine the light at the wall and insert filter A between the light source and the wall as in Figure 25.1. The photons coming out of the filter will have horizontal polarization. Now insert filter C as in Figure 25.2. Since filter C has vertical polarization, it filters out all of the horizontally polarized photons from filter A. Notice that no light arrives at the wall after this step, the two filters have removed all of the

light components. Now for the final (and most bizarre) step, insert filter B in between filter A and C. You should observe that there is now light arriving at the wall, as depicted in Figure 25.3. This is puzzling, since filter A and C were enough to remove all of the light, yet the addition of a third filter allows for light to reach the wall.

Figure 25.1 The Photon Experiment with Only Filter A Inserted

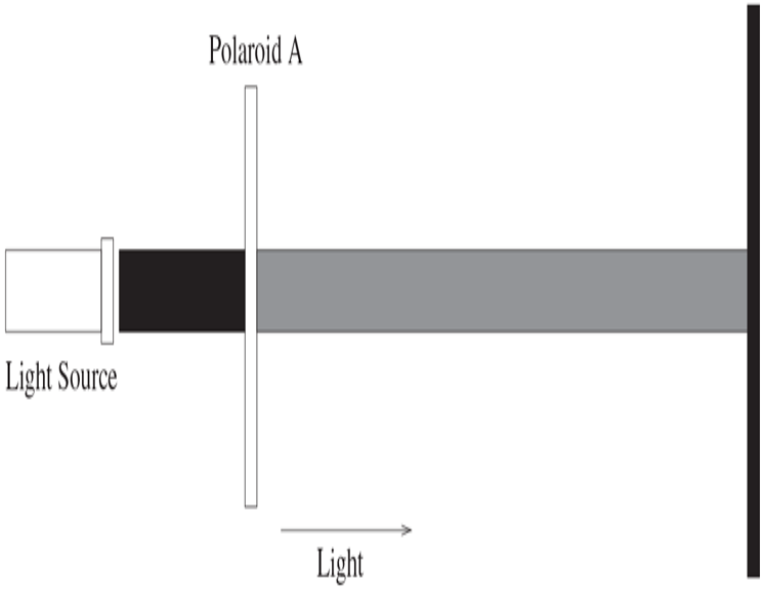

Figure 25.1 Full Alternative Text

Figure 25.2 The Photon Experiment with Filters A and C Inserted

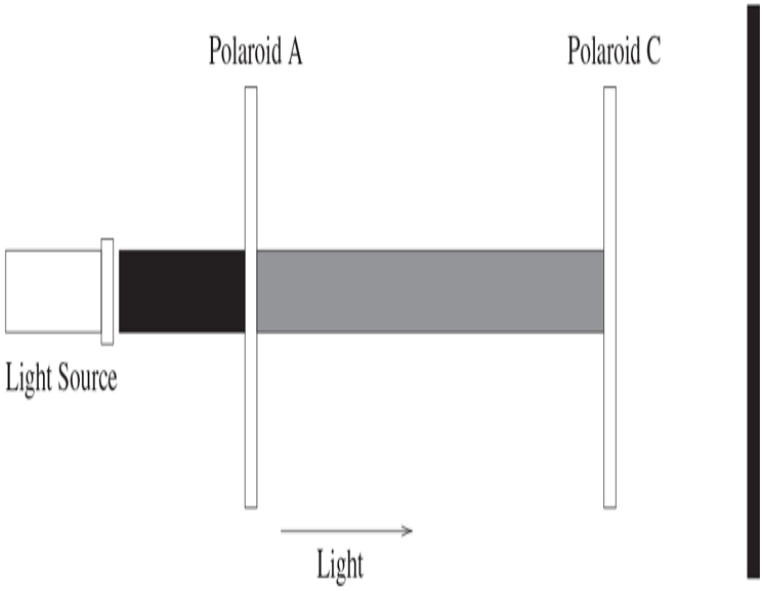

Figure 25.2 Full Alternative Text

Figure 25.3 The Photon Experiment after All Filters Have Been Inserted

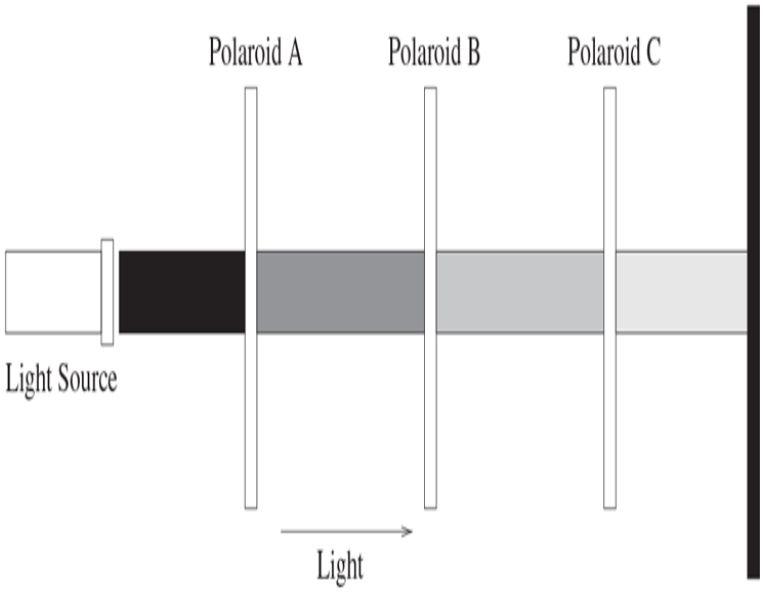

Figure 25.3 Full Alternative Text

In order to explain this demonstration, we need to discuss the concept of polarization of light.

Light is an example of an electromagnetic wave, meaning that it consists of an electric field that travels orthogonally to a corresponding magnetic field. In order to visualize this, consider the light traveling along the x-axis. Now imagine, for example, that the electric field is a wavelike function that lies in the xz-plane. Then the corresponding magnetic field would be a wavelike function in the xy-plane. For such a scenario, the light is referred to as vertically polarized. In general, polarization refers to the direction in which the electric field lies. There is no constraint on this direction.

We will represent a photon's polarization by a unit vector in the two-dimensional complex vector space (however, for our present purposes, real numbers suffice). This vector space has a dot product given by $(a, b) \cdot (c, d) = a\bar{c} + b\overline{d}$, where $\bar{c}$ and $\overline{d}$ denote the complex conjugates of c and d. The square of the length of a vector (a, b) is then $(a, b) \cdot (a, b) = |a|^2 + |b|^2$. Choose a basis, which we shall denote $|\uparrow\rangle$ and $|\rightarrow\rangle$, for this vector space. We are choosing to use the ket (the second half of "bracket") notation from physics to represent vectors. We can think of $|\uparrow\rangle$ as being the vertical direction and $|\rightarrow\rangle$ as being horizontal. Therefore, an arbitrary polarization may be represented as $a|\uparrow\rangle + b|\rightarrow\rangle$, where a and b are complex numbers. Since we are working with unit vectors, the following property holds: $|a|^2 + |b|^2 = 1$. We could just have well chosen a different orthogonal basis, for example, one corresponding to a 45° rotation: $|\nwarrow\rangle$ and $|\nearrow\rangle$.

The Polaroid filters perform a measurement of the polarity of the photon. There are two possible outcomes: Either the photon is aligned with the filter, or it is perpendicular to the direction of the filter. If the vector $a|\uparrow\rangle + b|\rightarrow\rangle$ is measured by a vertical filter, then the probability that the photon has vertical polarity after passing through the filter is $|a|^2$. The probability that it

is measured as having horizontal polarity, and therefore not pass through, is $|b|^2$.

Similarly, suppose we measure a vertically aligned photon with respect to a 45° filter. Since

$$|\uparrow\rangle = \frac{1}{\sqrt{2}}|\nwarrow\rangle + \frac{1}{\sqrt{2}}|\nearrow\rangle,$$

the probability that the photon passes through the filter (which means that it is measured as being aligned at 45°) is $(1/\sqrt{2})^2 = 1/2$. Similarly, the probability that it doesn't pass through the filter (which means that it is measured at -45°) is also $1/2$.

One of the basic principles of quantum mechanics is that such a measurement forces the photon into a definite state. After being measured, the state of the photon will be changed to the result of the measurement. Therefore, if we measured the state of $a|\uparrow\rangle + b|\rightarrow\rangle$ as $|\rightarrow\rangle$, then, from that moment on, the photon will have the state $|\rightarrow\rangle$. If we then measure this photon with a $|\rightarrow\rangle$ filter, we will always observe that the photon is in the $|\rightarrow\rangle$ state; however, if we measure with a $|\uparrow\rangle$ filter, we will never observe that the photon is in the $|\uparrow\rangle$ state.

Let's now explain the interpretation of the experiment. The original light was emitted with random polarization, so only half of the photons being emitted will pass through the $|\rightarrow\rangle$ filter, and these photons will have their state changed to $|\rightarrow\rangle$. The remaining half will be absorbed or reflected and will be changed to $|\uparrow\rangle$. When we place the vertical filter after the horizontal filter, the photons that hit it, which are in state $|\rightarrow\rangle$, will be stopped.

When we insert filter B in the middle, it corresponds to measuring with respect to $|\nearrow\rangle$, and hence those photons that had $|\rightarrow\rangle$ polarity will come out having $|\nearrow\rangle$ polarity with probability $1/2$. Therefore, there has been a $4:1$

reduction in the amount of photons passing through up to filter B. Now the $|\nearrow\rangle$ photons pass through the $|\uparrow\rangle$ filter with probability $1/2$ also, and so the total intensity of light arriving at the wall is $1/8$th the original intensity.

25.2 Quantum Key Distribution

Now that we have set up some of the ideas behind quantum mechanics, we can use them to describe a technique for distributing bits through a quantum channel. These bits can be used to establish a key that can be used for communicating across a classical channel, or any other shared secret.

We begin by describing a quantum bit. Start with a two-dimensional complex vector space. Choose a pair of orthogonal vectors of length 1; call them $|0\rangle$ and $|1\rangle$. For example, these two vectors could be either of the two pairs of orthogonal vectors used in the previous section. A **quantum bit**, also known as a **qubit**, is a unit vector in this vector space. For the purposes of the present discussion, we can think of a qubit as a polarized photon. We have chosen $|0\rangle$ and $|1\rangle$ as notation to conveniently represent the 0 and 1 bits, respectively. The other qubits are linear combinations of these two bits.

Since a qubit is a unit vector, it can be represented as $a|0\rangle + b|1\rangle$, where a and b are complex numbers such that $|a|^2 + |b|^2 = 1$. Just as in the case for photons from the preceding section, we can measure this qubit with respect to the basis $|0\rangle$, $|1\rangle$. The probability that we observe it in the $|0\rangle$ state is $|a|^2$.

Let us now examine how Alice and Bob can communicate with each other in order to establish a message. They will need two things: a quantum channel and a classical channel. A quantum channel is one through which they can exchange polarized photons that are isolated from interactions with the environment (that is, the environment doesn't alter the photons). The classical

channel will be used to send ordinary messages to each other. We assume that the evil observer Eve can observe what is being sent on the classical channel and that she can observe and resend photons on the quantum channel.

Alice starts the establishment of a message by sending a sequence of bits to Bob. They are encoded using a randomly chosen basis for each bit as follows. There are two bases: $B_1 = \{|\uparrow\rangle, |\rightarrow\rangle\}$ and $B_2 = \{|\nwarrow\rangle, |\nearrow\rangle\}$. If Alice chooses B_1, then she encodes 0 as $|\uparrow\rangle$ and 1 as $|\rightarrow\rangle$, while if she chooses B_2 then she encodes 0 and 1 using the two elements of B_2.

Each time Alice sends a photon, Bob randomly chooses to measure with respect to either basis B_1 or B_2. Therefore, for each photon, he obtains an element of that choice of basis as the result of his measurement. Bob records the measurements he has made and keeps them secret. He then tells Alice the basis with which he measured each photon. Alice responds to Bob by telling him which bases were the correct bases for the polarity of the photons that she sent. They keep the bits that used the same bases and discard the other bits. Since two bases were used, Alice and Bob will agree on roughly half of the amount of bits that Alice sent. They can then use these bits as the key for a conventional cryptographic system.

Example

Suppose Alice wants to send the bits $0, 1, 1, 1, 0, 0, 1, 0$. She randomly chooses the bases $B_1, B_2, B_1, B_1, B_2, B_2, B_1, B_2$. Therefore, she sends the qubits (photons)

$$|\uparrow\rangle, |\nearrow\rangle, |\rightarrow\rangle, |\rightarrow\rangle, |\nwarrow\rangle, |\nwarrow\rangle, |\rightarrow\rangle, |\nwarrow\rangle$$

to Bob. He chooses the bases B_2, B_2, B_2, B_1, B_2, B_1, B_1, B_2. He measures the qubits that Alice sent and also tells Alice which bases he used. Alice tells him that the second, fourth, fifth, seventh, and eighth match her choices. These yielded measurements

$$|\nearrow\rangle, |\rightarrow\rangle, |\nwarrow\rangle, |\rightarrow\rangle, |\nwarrow\rangle$$

for Bob, and they correspond to the bits 1, 1, 0, 1, 0. Therefore, both Alice and Bob have the same string 1, 1, 0, 1, 0. They use 11010 as a key for future communication (for example, if they obtained a longer string, they could use the first 128 characters for an AES key).

The security behind quantum key distribution is based upon the laws of quantum mechanics and the fundamental principle that following a measurement of a particle, that particle's state will be altered. Since an eavesdropper Eve must perform measurements in order to observe the photon transmissions between Alice and Bob, Eve will introduce errors in the data that Alice and Bob agreed upon.

Let's see how this happens. Suppose Eve measures the states of the photons transmitted by Alice and allows these measured photons to proceed onto Bob. Since these photons were measured by Eve, they will have the state that Eve observed. Eve will use the wrong basis half of the time when performing the measurement. When Bob performs his measurements, if he uses the correct basis there will be a 25 chance that he will have measured the wrong value.

Let's examine this last statement in more detail. Suppose that Alice sends a photon corresponding to $|\rightarrow\rangle$ and that Bob uses the same basis B_1 as Alice. If Eve uses B_1, then the photon is passed through correctly and then Bob measures the photon correctly. However, if Eve used B_2,

then she will measure $|\nearrow\rangle$ and $|\nwarrow\rangle$ equally likely. The photons that pass to Bob will have one of these orientations and he will therefore half the time measure them correctly as $|\rightarrow\rangle$ and half the time incorrectly. Combining the two possible choices of basis that Eve has causes Bob to have a 25 chance of measuring the incorrect value.

Thus, any eavesdropping introduces a higher error rate in the communication between Alice and Bob. If Alice and Bob test their data for discrepancies over the conventional channel (for example, they could send parity bits), they will detect any eavesdropping.

Actual implementations of this technique have been used to establish keys over distances of more than 100 km using conventional fiber optical cables.

25.3 Shor's Algorithm

Quantum computers are not yet a reality. The current versions can only handle a few qubits. But, if the great technical problems can be overcome and large quantum computers are built, the effect on cryptography will be enormous. In this section we give a brief glimpse at how a quantum computer could factor large integers, using an algorithm developed by Peter Shor. We avoid discussing quantum mechanics and ask the reader to believe that a quantum computer should be able to do all the operations we describe, and do them quickly. For more details, see, for example, [Ekert-Josza] or [Rieffel-Polak].

What is a quantum computer and what does it do? First, let's look at what a classical computer does. It takes a binary input, for example, 100010, and gives a binary output, perhaps 0101. If it has several inputs, it has to work on them individually. A quantum computer takes as input a certain number of qubits and outputs some qubits. The main difference is that the input and output qubits can be linear combinations of certain basic states. The quantum computer operates on all basic states in this linear combination simultaneously. In effect, a quantum computer is a massively parallel machine.

For example, think of the basic state $|100\rangle$ as representing three particles, the first in orientation 1 and the last two in orientation 0 (with respect to some basis that will implicitly be fixed throughout the discussion). The quantum computer can take $|100\rangle$ and produce some output. However, it can also take as input a normalized (that is, of length 1) linear combination of basic quantum states such as

$$\frac{1}{\sqrt{3}}\big(|100\rangle + |011\rangle + |110\rangle\big)$$

and produce an output just as quickly as it did when working with a basic state. After all, the computer could not know whether a quantum state is one of the basic states, or a linear combination of them, without making a measurement. But such a measurement would alter the input. It is this ability to work with a linear combination of states simultaneously that makes a quantum computer potentially very powerful.

Suppose we have a function $f(x)$ that can be evaluated for an input x by a classical computer. The classical computer asks for an input and produces an output. A quantum computer, on the other hand, can accept as input a sum

$$\frac{1}{C}\sum_x |x\rangle$$

(C is a normalization factor) of all possible input states and produce the output

$$\frac{1}{C}\sum_x |x,\, f(x)\rangle,$$

where $|x,\, f(x)\rangle$ is a longer sequence of qubits, representing both x and the value of $f(x)$. (*Technical point:* It might be notationally better to input $(1/C)\sum |x,\, 00\cdots\rangle$ in order to have some particles to change to $f(x)$. For simplicity, we will not do this.) So we can obtain a list of all the values of $f(x)$. This looks great, but there is a problem. If you make a measurement, you force the quantum state into the result of the measurement. You get $|x_0,\, f(x_0)\rangle$ for some randomly chosen x_0, and the other states in the output are destroyed. So, if you are going to look at the list of values of $f(x)$, you'd better do it carefully, since you get only one chance. In particular, you probably want to apply some transformation to the output in order to put it into a more desirable form. The skill in programming a quantum computer is in designing the computation so that the outputs you want to examine appear with much

higher probability than the others. This is what is done in Shor's factorization algorithm.

25.3.1 Factoring

We want to factor n. The strategy is as follows. Recall that if we can find (nontrivial) a and r with $a^r \equiv 1 \pmod{n}$, then we have a good chance of factoring n (see the factorization method in Subsection 9.4.1). Choose a random a and consider the sequence $1, a, a^2, a^3, \ldots \pmod{n}$. If $a^r \equiv 1 \pmod{n}$, then this sequence will repeat every r terms since $a^{j+r} \equiv a^j a^r \equiv a^j \pmod{n}$. If we can measure the period of this sequence (or a multiple of the period), we will have an r such that $a^r \equiv 1 \pmod{n}$. We therefore want to design our quantum computer so that when we make a measurement on the output, we'll have a high chance of obtaining the period.

25.3.2 The Discrete Fourier Transform

We need a technique for finding the period of a periodic sequence. Classically, Fourier transforms can be used for this purpose, and they can be used in the present situation, too. Suppose we have a sequence

$$a_0, a_1, \ldots, a_{2^m-1}$$

of length 2^m, for some integer m. Define the Fourier transform to be

$$F(x) = \frac{1}{\sqrt{2^m}} \sum_{c=0}^{2^m-1} e^{\frac{2\pi i c x}{2^m}} a_c,$$

where $0 \leq x < 2^m$.

For example, consider the sequence

$$1, 3, 7, 2, 1, 3, 7, 2$$

of length 8 and period 4. The length divided by the period is the frequency, namely 2, which is how many times the sequence repeats. The Fourier transform takes the values

$$\begin{aligned} F(0) = 26/\sqrt{8}, \qquad F(2) = (-12+2i)/\sqrt{8}, \\ F(4) = 6/\sqrt{8}, \qquad F(6) = (-12-2i)/\sqrt{8}, \\ F(1) = F(3) = F(5) = F(7) = 0. \end{aligned}$$

For example, letting $\zeta = e^{2\pi i/8}$, we find that

$$\sqrt{8}F(1) = 1 + 3\zeta + 7\zeta^2 + 2\zeta^3 + \zeta^4 + 3\zeta^5 + 7\zeta^6 + 2\zeta^7.$$

Since $\zeta^4 = -1$, the terms cancel and we obtain $F(1) = 0$. The nonzero values of F occur at multiples of 2, which is the frequency.

Let's consider another example: $2, 1, 2, 1, 2, 1, 2, 1$. The Fourier transform is

$$\begin{aligned} F(0) = 12/\sqrt{8}, \quad F(4) = 4/\sqrt{8}, \\ F(1) = F(2) = F(3) = F(5) = F(6) = F(7) = 0. \end{aligned}$$

Here the nonzero values of F are again at the multiples of the frequency.

In general, if the period is a divisor of 2^m, then all the nonzero values of F will occur at multiples of the frequency (however, a multiple of the frequency could still yield 0). See Exercise 2.

Suppose now that the period isn't a divisor of 2^m. Let's look at an example. Consider the sequence $1, 0, 0, 1, 0, 0, 1, 0$. It has length 8 and almost has period 3 and frequency 3, but we stopped the sequence before it had a chance to complete the last period. In Figure 25.4, we graph the absolute value of its Fourier transform (these are real numbers, hence easier to graph than the complex values of the Fourier transform). Note that there are peaks at 0, 3, and 5. If we continued $F(x)$

to larger values of x we would get peaks at $8, 11, 13, 16, \ldots$. The peaks are spaced at an average distance of $8/3$. Dividing the length of the sequence by the average distance yields a period of $8/(8/3) = 3$, which agrees with our intuition.

Figure 25.4 The Absolute Value of a Discrete Fourier Transform

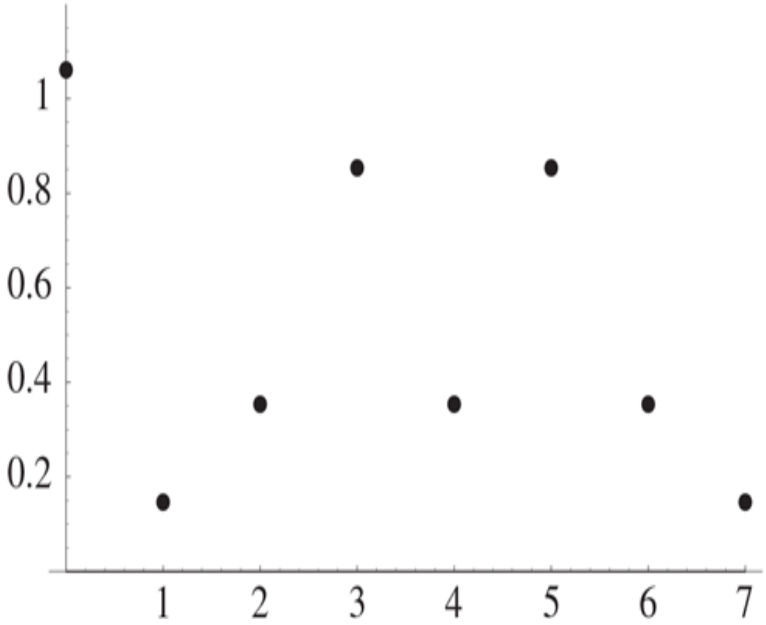

Figure 25.4 Full Alternative Text

The fact that there is a peak at 0 is not very surprising. The formula for the Fourier transform shows that the value at 0 is simply the sum of the elements in the sequence divided by the square root of the length of the sequence.

Let's look at one more example: 1, 0, 0, 0, 0, 1, 0, 0, 0, 0 1, 0, 0, 0, 0, 1. This sequence has 16 terms. Our intuition might say that the period is around 5 and the frequency is slightly more than 3. Figure 25.5 shows the graph of the absolute value of its Fourier transform. Again, the

peaks are spaced around 3 apart, so we can say that the frequency is around 3. The period of the original sequence is therefore around 5, which agrees with our intuition.

Figure 25.5 The Absolute Value of a Discrete Fourier Transform

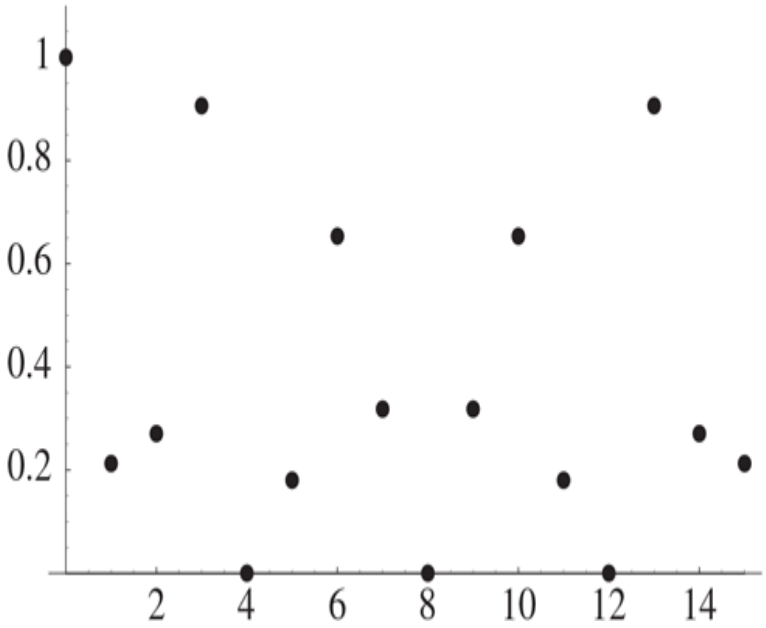

Figure 25.5 Full Alternative Text

In the first two examples, the period was a divisor of the length (namely, 8) of the sequence. We obtained nonzero values of the Fourier transform only at multiples of the frequency. In these last two examples, the period was not a divisor of the length (8 or 16) of the sequence. This introduced some “noise” into the situation. We had peaks at approximate multiples of the frequency and values close to 0 away from these peaks.

The conclusion is that the peaks of the Fourier transform occur approximately at multiples of the frequency, and

the period is approximately the number of peaks. This will be useful in Shor's algorithm.

25.3.3 Shor's Algorithm

Choose m so that $n^2 \leq 2^m < 2n^2$. We start with m qubits, all in state 0:

$$|000000000\rangle.$$

As in the previous section, by changing axes, we can transform the first bit to a linear combination of $|0\rangle$ and $|1\rangle$, which gives us

$$\frac{1}{\sqrt{2}}\big(|000000000\rangle + |100000000\rangle\big).$$

We then successively do a similar transformation to the second bit, the third bit, up through the mth bit, to obtain the quantum state

$$\frac{1}{\sqrt{2^m}}\big(|000000000\rangle + |00000001\rangle + |000000010\rangle + \cdots + |111111111\rangle\big).$$

Thus all possible states of the m qubits are superimposed in this sum. For simplicity of notation, we replace each string of 0s and 1s with its decimal equivalent, so we write

$$\frac{1}{\sqrt{2^m}}\big(|0\rangle + |1\rangle + |2\rangle + \cdots + |2^m - 1\rangle\big).$$

Choose a random number a with $1 < a < n$. We may assume $\gcd\,(a, n) = 1$; otherwise, we have a factor of n. The quantum computer computes the function $f(x) = a^x \pmod n$ for this quantum state to obtain

$$\frac{1}{\sqrt{2^m}}\big(|0, a^0\rangle + |1, a^1\rangle + |2, a^2\rangle + \cdots + |2^m - 1, a^{2^m-1}\rangle\big)$$

(for ease of notation, a^x is used to denote $a^x \pmod n$). This gives a list of all the values of a^x. However, so far we are not any better off than with a classical computer.

If we measure the state of the system, we obtain a basic state $|x_0,\, a^{x_0}\rangle$ for some randomly chosen x_0. We cannot even specify which x_0 we want to use. Moreover, the system is forced into this state, obliterating all the other values of a^x that have been computed. Therefore, we do not want to measure the whole system. Instead, we measure the value of the second half. Each basic piece of the system is of the form $|x,\, a^x\rangle$, where x represents m bits and a^x is represented by $m/2$ bits (since $a^x \pmod n < n < 2^{m/2}$). If we measure these last $m/2$ bits, we obtain some number $u \pmod n$, and the whole system is forced into a combination of those states of the form $|x,\, u\rangle$ with $a^x \equiv u \pmod n$:

$$\frac{1}{C} \sum_{\substack{0 \le x \le 2^m \\ a^x \equiv u \pmod n}} |x,\, u\rangle,$$

where C is whatever factor is needed to make the vector have length 1 (in fact, C is the square root of the number of terms in the sum).

Example

At this point, it is probably worthwhile to have an example. Let $n = 21$. (This example might seem simple, but it is the largest that quantum computers using Shor's algorithm can currently handle. Other algorithms are being developed that can go somewhat farther.) Since $21^2 < 2^9 < 2 \cdot 21^2$, we have $m = 9$. Let's choose $a = 11$, so we compute the values of $11^x \pmod{21}$ to obtain

$$\begin{aligned}\frac{1}{\sqrt{512}} \big(&|0,\, 1\rangle + |1,\, 11\rangle + |2,\, 16\rangle + |3,\, 8\rangle + |4,\, 4\rangle + |5,\, 2\rangle + |6,\, 1\rangle + |7,\, 11\rangle + \\ &|8,\, 16\rangle + |9,\, 8\rangle + |10,\, 4\rangle + |11,\, 2\rangle + |12,\, 1\rangle + |13,\, 11\rangle + |14,\, 16\rangle + \\ &|15,\, 8\rangle + |16,\, 4\rangle + |17,\, 2\rangle + |18,\, 1\rangle + |19,\, 11\rangle + |20,\, 16\rangle + \cdots \\ &+|508,\, 4\rangle + |509,\, 2\rangle + |510,\, 1\rangle + |511,\, 11\rangle\big).\end{aligned}$$

Suppose we measure the second part and obtain 2. This means we have extracted all the terms of the form $|x,\ 2\rangle$

to obtain

$$\frac{1}{\sqrt{85}}\big(|5,\ 2\rangle + |11,\ 2\rangle + |17,\ 2\rangle + |23,\ 2\rangle + \cdots + |497,\ 2\rangle + |503,\ 2\rangle + |509,\ 2\rangle\big).$$

For notational convenience, and since it will no longer be needed, we drop the second part to obtain

$$\frac{1}{\sqrt{85}}\big(|5\rangle + |11\rangle + |17\rangle + |23\rangle + \cdots + |497\rangle + |503\rangle + |509\rangle\big).$$

If we now measured this system, we would simply obtain a number x such that $11^x \equiv 2 \pmod{21}$. This would not be useful.

Suppose we could take two measurements. Then we would have two numbers x and y with $11^x \equiv 11^y \pmod{21}$. This would yield $11^{x-y} \equiv 1 \pmod{21}$. By the factorization method of Subsection 9.4.1, this would give us a good chance of being able to factor 21. However, we cannot take two independent measurements. The first measurement puts the system into the output state, so the second measurement would simply give the same answer as the first.

Not all is lost. Note that in our example, the numbers in our state are periodic with period 6. In general, the values of $a^x \pmod{n}$ are periodic with period r, with $a^r \equiv 1 \pmod{n}$. So suppose we are able to make a measurement that yields the period. We then have a situation where $a^r \equiv 1 \pmod{n}$, so we can hope to factor n by the method from Subsection 9.4.1 mentioned above.

The **quantum Fourier transform** is exactly the tool we need. It measures frequencies, which can be used to find the period. If r happens to be a divisor of 2^m, then the frequencies we obtain are multiples of a fundamental frequency f_0, and $rf_0 = 2^m$. In general, r is not a divisor of 2^m, so there will be some dominant

frequencies, and they will be approximate multiples of a fundamental frequency f_0 with $rf_0 \approx 2^m$. This will be seen in the analysis of our example and in Figure 25.6.

Figure 25.6 The Absolute Value of $g(c)$

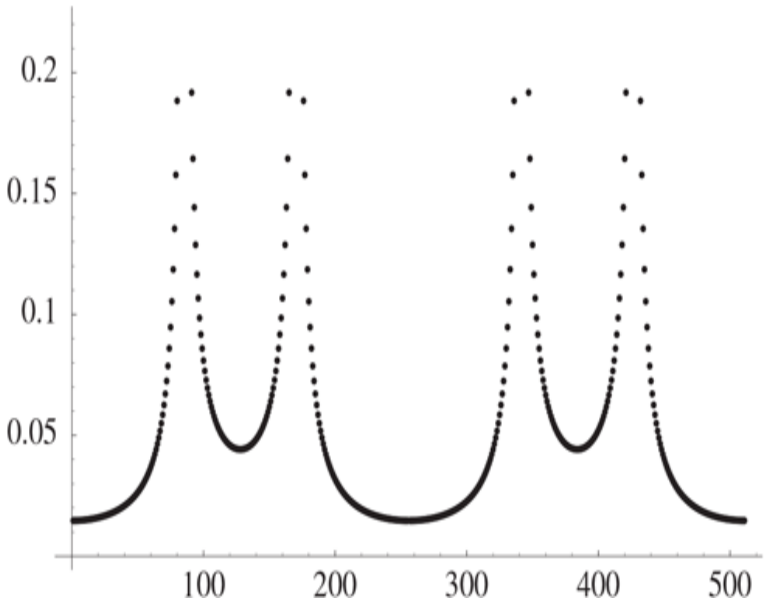

Figure 25.6 Full Alternative Text

The quantum Fourier transform is defined on a basic state $|x\rangle$ (with $0 \leq x < 2^m$) by

$$QFT(|x\rangle) = \frac{1}{\sqrt{2^m}} \sum_{c=0}^{2^m-1} e^{\frac{2\pi icx}{2^m}} |c\rangle.$$

It extends to a linear combination of states by linearity:

$$QFT\big(a_1|x_1\rangle + \cdots + a_t|x_t\rangle\big) = a_1 QFT(|x_1\rangle) + \cdots + a_t QFT(|x_t\rangle).$$

We can therefore apply $QF <$ to our quantum state.

In our example, we compute

$$QFT < \left(\frac{1}{\sqrt{85}} \left(|5\rangle + |11\rangle + |17\rangle + |23\rangle + \cdots + |497\rangle + |503\rangle + |509\rangle \right) \right)$$

and obtain a sum

$$\frac{1}{\sqrt{85}} \sum_{c=0}^{511} g(c) \, |c\rangle$$

for some numbers $g(c)$.

The number $g(c)$ is given by

$$g(c) = \tfrac{1}{\sqrt{512}} \sum_{\substack{0 \le x < 512 \\ x \equiv 5 \pmod 6}} e^{\frac{2\pi i c x}{512}},$$

which is the discrete Fourier transform of the sequence

$$0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, \ \ldots, 0, 0, 0, 0, 0, 1, 0, 0.$$

Therefore, the peaks of the graph of the absolute value of g should correspond to the frequency of the sequence, which should be around $512/6 \approx 85$. The graph in Figure 25.6 is a plot of $|g|$.

There are sharp peaks at $c = 0, 85, 171, 256, 341, 427$ (the ones at 0 and 256 do not show up on the graph since they are centered at one value; see below). These are the dominant frequencies mentioned previously. The values of g near the peak at $c = 341$ are

338	339	340	341	342	343	344	345
0.305	0.439	0.773	3.111	1.567	0.631	0.398	0.291

The behavior near $c = 85, 171$, and 427 is similar. At $c = 0$ and 256, we have $g(0) = 3.756$, while all the nearby values of c have $g(c) \approx 0.015$.

The peaks are approximately at multiples of the fundamental frequency $f_0 = 85$. Of course, we don't really know this yet, since we haven't made any measurements.

Now we measure the quantum state of this Fourier transform. Recall that if we start with a linear combination of states $a_1|x_1\rangle + \cdots + a_<|x_<\rangle$ normalized such that $\sum |a_j|^2 = 1$, then the probability of of obtaining $|x_k\rangle$ is $|a_k|^2$. More generally, if we don't assume $\sum |a_j|^2 = 1$, the probability is

$$|a_k|^2 / \sum |a_j|^2.$$

In our example,

$$3.111^2 / \sum |a_j|^2 \approx .114,$$

so if we sample the Fourier transform, the probability is around $4 \times .114 = .456$ that we obtain one of $c = 85$, 171, 341, 427. Let's suppose this is the case; say we get $c = 427$. We know, or at least expect, that 427 is approximately a multiple of the frequency f_0 that we're looking for:

$$427 \approx j f_0$$

for some j. Since $r f_0 \approx 2^m = 512$, we divide to obtain

$$\frac{427}{512} \approx \frac{j}{r}.$$

Note that $427/512 \approx .834 \approx 5/6$. Since we must have $r \leq \phi(21) < 21$, a reasonable guess is that $r = 6$ (see the following discussion of continued fractions).

In general, Shor showed that there is a high chance of obtaining a value of $c/2^m$ with

$$\left|\frac{c}{2^m} - \frac{j}{r}\right| < \frac{1}{2^{m+1}} < \frac{1}{2n^2},$$

for some j. The method of continued fractions will find the unique (see Exercise 3) value of j/r with $r < n$ satisfying this inequality.

In our example, we take $r = 6$ and check that $a^r = 11^6 \equiv 1 \pmod{21}$.

We want to use the factorization method of Subsection 9.4.1 to factor 21. Recall that this method writes $r = 2^k m$ with m odd, and then computes $b_0 \equiv a^m \pmod{n}$. We then successively square b_0 to get b_1, b_2, ..., until we reach $1 \pmod{n}$. If b_u is the last $b_i \not\equiv 1 \pmod{n}$, we compute $\gcd(b_u - 1, n)$ to get a factor (possibly trivial) of n.

In our example, we write $6 = 2 \cdot 3$ (a power of 2 times an odd number) and compute (in the notation of Subsection 9.4.1)

$$\begin{aligned} b_0 &\equiv 11^3 \equiv 8 \pmod{21} \\ b_1 &\equiv 11^6 \equiv 1 \pmod{21} \end{aligned}$$

$$\gcd(b_0 - 1,\ 21) = \gcd(7,\ 21) = 7,$$

so we obtain $21 = 7 \cdot 3$.

In general, once we have a candidate for r, we check that $a^r \equiv 1 \pmod{n}$. If not, we were unlucky, so we start over with a new a and form a new sequence of quantum states. If $a^r \equiv 1 \pmod{n}$, then we use the factorization method from Subsection 9.4.1. If this fails to factor n, start over with a new a. It is very likely that, in a few attempts, a factorization of n will be found.

We now say more about continued fractions. In Chapter 3, we outlined the method of continued fractions for finding rational numbers with small denominator that approximate real numbers. Let's apply the procedure to the real number $427/512$. We have

$$\frac{427}{512} = 0 + \cfrac{1}{1 + \cfrac{1}{5 + \cfrac{1}{42 + \cfrac{1}{2}}}}.$$

This yields the approximating rational numbers

$$0, \quad 1, \quad \frac{5}{6}, \quad \frac{211}{253}, \quad \frac{427}{512}.$$

Since we know the period in our example is less than $n = 21$, the best guess is the last denominator less than n, namely $r = 6$.

In general, we compute the continued fraction expansion of $c/2^m$, where c is the result of the measurement. Then we compute the approximations, as before. The last denominator less than n is the candidate for r.

25.3.4 Final Words

The capabilities of quantum computers and quantum algorithms are of significant importance to economic and government institutions. Many secrets are protected by cryptographic protocols. Quantum cryptography's potential for breaking these secrets as well as its potential for protecting future secrets has caused this new research field to grow rapidly over the past few years.

Although the first full-scale quantum computer is probably many years off, and there are still many who are skeptical of its possibility, quantum cryptography has already succeeded in transmitting secure messages over a distances of more than 100 km, and quantum computers have been built that can handle a (very) small number of qubits. Quantum computation and cryptography have already changed the manner in which computer scientists and engineers perceive the capabilities and limits of the computer. Quantum computing has rapidly become a popular interdisciplinary research area and promises to offer many exciting new results in the future.

25.4 Exercises

1. Consider the sequence $2^0, 2^1, 2^2, \ldots \pmod{15}$.

    a. What is the period of this sequence?

    b. Suppose you want to use Shor's algorithm to factor $n = 15$. What value of m would you take?

    c. Suppose the measurement in Shor's algorithm yields $c = 192$. What value do you obtain for r? Does this agree with part (a)?

    d. Use the value of r from part (c) to factor 15.

2. a. Let $0 < s \leq m$. Fix an integer c_0 with $0 \leq c_0 < 2^s$. Show that

$$\sum_{\substack{0 \leq c < 2^m \\ c \equiv c_0 \,(\mathrm{mod}\, 2^s)}} e^{\frac{2\pi i c x}{2^m}} = 0$$

    if $x \not\equiv 0 \,(\mathrm{mod}\, 2^{m-s})$ and $= 2^{m-s} e^{2\pi i x c_0 / 2^m}$ if $x \equiv 0 \,(\mathrm{mod}\, 2^{m-s})$. (Hint: Write $c = c_0 + j2^s$ with $0 \leq j < 2^{m-s}$, factor $e^{2\pi i x c_0/2^m}$ off the sum, and recognize what's left as a geometric sum.)

    b. Suppose $a_0, a_1, \ldots, a_{2^m-1}$ is a sequence of length 2^m such that $a_k = a_{k+j2^s}$ for all j, k. Show that the Fourier transform $F(x)$ of this sequence is 0 whenever $x \not\equiv 0 \,(\mathrm{mod}\, 2^{m-s})$.

    This shows that if the period of a sequence is a divisor of 2^m then all the nonzero values of F occur at multiples of the frequency (namely, 2^{m-s}).

3. a. Suppose j/r and j_1/r_1 are two distinct rational numbers, with $0 < r < n$ and $0 < r_1 < n$. Show that

$$\left|\frac{j_1}{r_1} - \frac{j}{r}\right| g > \frac{1}{n^2}.$$

    b. Suppose, as in Shor's algorithm, that we have

$$\left|\frac{c}{2^m} - \frac{j}{r}\right| < \frac{1}{2n^2} \text{ and } \left|\frac{c}{2^m} - \frac{j_1}{r_1}\right| < \frac{1}{2n^2}.$$

    Show that $j/r = j_1/r_1$.