Notes 8.370/18.435 Fall 2022

Lecture 22 Prof. Peter Shor

In this lecture, we talked about the number theory that we needed for the factoring algorithm..

1 The Euclidean Algorithm and the Extended Euclidean Algorithm

Let's recall how we found the factors of N. To make the exposition easier, we will assume that N is a product of two primes, $N = PQ$ in these notes, but the factoring algorithm works fine in the general case when more than two primes divide N.

Recall that in order to factor, we found the period of the sequence

$$g, g^2 \pmod N, g^3 \pmod N, \ldots.$$

If the period of this sequence is r, then we must have $g^r \equiv 1 \pmod N$. Why? Because there are at most $N-2$ different values in this sequence, we must have $g^a = g^{r+a} \pmod N$ for some a. But then, multiplying by $g^{-a} \pmod N$, we get $1 = g^r \pmod N$.

Now that we have $g^r \equiv 1 \pmod N$, if r is even, we can factor this expression to get

$$(g^{r/2} - 1)(g^{r/2+1}) \equiv 0 \pmod N;$$

We have two numbers multiplying to a multiple of N. If neither of them is a multiple of N, then we have P must divide one of the numbers and Q the other. Let's try to factor 33 in this way. Suppose we take $g = 2$. We see that $2^{10} \equiv 1 \pmod{33}$, so we get $(2^5 - 1)(2^5 + 1) \equiv 0 \pmod 3$. Unfortunately, this doesn't give us a factor, because $2^5 + 1 \equiv 33$.

So let's take $g = 5$. We then have (again) that $5^{10} \equiv 1 \pmod N$, so $(5^5 - 1)(5^5 + 1) \equiv 0 \pmod{33}$. We an compute that $5^5 \pmod{33} = 23$, so this gives $(23 - 1)(23 + 1) \equiv 0 \pmod{33}$. And this time it worked! 22 contains the factor 11 and 24 contains the factor 3. How do we recover 3 from 24 and 33. We use the Euclidean algorithm for finding the greatest common divisor of two numbers. How do we implement this? One standard way is to put the two numbers we start with in a row, with the larger first. We then repeatedly move the number in the right column to the left column, and replace the number in the right column by the remainder we get when dividing these two numbers. For example, to find gcd$(24, 9)$,

$$\begin{array}{rr} 33 & 24 \\ 24 & 9 \\ 9 & 6 \\ 6 & 3 \end{array}$$

Here, in the first step, we divide 33 by 24, and get remainder 9. In the second step, we divide 24 by 9 and get remainder 6, and so on. If something divides both of the

first numbers, it will divide all the other numbers in our array. We keep decreasing the size of the numbers, so eventually we will reach the greatest common divisor of the numbers.

Now, let's look again at our algorithm. We needed to find the period of the unitary map $|y\rangle \to |gy \pmod N\rangle$. How can we implement this map.

Recall from our discussion of reversible classical computation, that if we have a classical circuit taking y to $gy \pmod N$ and a circuit take $gy \pmod N$ to y, we can find a reversible circuit whose input is y (along with some workbits whose initial values are 0 and whose output is $gy \pmod N$, where the values of the workbits have been returned to 0. Finding a circuit that takes y to $gy \pmod N$ is easy—it's just multiplication. But how do we go the other way? What we need to do is find $g^{-1} \pmod N$ and then use a circuit for multiplication that takes x to $g^{-1}x \pmod N$. So the only hard part of this is finding $g^{-1} \pmod N$. For this, we use something called the extended Euclidean algorithm.

As an example, let's find $5^{-1} \pmod{33}$. The first thing we do is use the Euclidean algorithm to find the greatest common divisor of 5 and 33. Recall that 5^{-1} only exists if this $\gcd(5, 33) = 1$. What we do is divide 33 by 5 to get the remainder 3, and then repeat with these two numbers—we divide 5 by 3 to get the remainder 2:

$$\begin{array}{lll} 33 & 5 & 33 - 6 \cdot 5 = 3 \\ 5 & 3 & 5 - 1 \cdot 3 = 2 \\ 3 & 2 & 3 - 1 \cdot 2 = 1 \\ 2 & 1 & \end{array}$$

Our next goal is to find two integers s and t such that $s \cdot 3 + t \cdot 5 = 1$. What we do is start from the last row and work backwards. In the second to last row, we have that $1 \cdot 3 - 1 \cdot 2 = 1$. What we do is plug in the expression for 2 in the second last row of this array to get So $1 \cdot 3 - 1 \cdot (5 - 3) = 1$, and simplifying this gives $2 \cdot 3 - 1 \cdot 5 = 1$. Now, we plug in the expression for 3 in the first row of our array, giving $2 \cdot (33 - 6 \cdot 5) - 1 \cdot 5 = 1$. This simplifies to $2 \cdot 33 - 13 \cdot 5 = 1$. But from this expression, we can find the inverse of $5 \pmod{33}$. Since the first term is a multiple of 33, we have $-13 \cdot 5 = 1 \pmod N$, which gives $5^{-1} = 33 - 13 = 20$.

In general, to implement the extended Euclidean algorithm, we start at the last row given by the Euclidean algorithm and work backwards. Let's say the first row of the Euclidean algorithm is r_1, r_2, the second row r_2, r_3, and so forth. For each row, we get an equation $r_j - q_j r_{j+1} = r_{j+2}$. Now, lets say we have found two integers s and t such that

$$sr_{j+1} + tr_{j+2} = 1$$

We plug in our formula for r_{j+2} into this equation to get

$$sr_{j+1} + t(r_j - q_j r_{j+1}) = 1.$$

Simplifying this will give us s' and t' so that $s'r_j + t'r_{j+1} = 1$. when we reach the top row, we have $sr_1 + tr_2 = 1$. This means that $t = r_2^{-1} \pmod{r_1}$.

2 Continued Fractions

First, we're going to go through an example to show how the continued fraction algorithm works. After that, we will prove some properties of it.

Let's use 33 as the example number we want to factor. We first pick a number g and find the period of $g^x \pmod{33}$. Recall that if we chose $g = 5$, the period was 10. We will need to choose $L \approx 2 \log N$ in the phase estimation algorithm. What the phase estimation algorithm finds is an estimate of the eigenvalue of $e^{2\pi i k/10}$ for some k, let's say $k = 3$. The phase estimation algorithm then returns a number of the form $d/2^L$that is close to $3/10$. Let's choose $2^L = 2048$. Then the approximation for $3/10$ would have a denominator of 2048. Let's say this approximation is $615/2048$. How do we recover $3/10$ from $615/2048$?

What we do is use *continued fractions*. A continued fraction is a number of the form

$$\cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cfrac{1}{\ddots}}}}}$$

Taking the continued fraction of $615/2048$, we find that

$$a_1, a_2, a_3, a_4, a_5, \ldots = 3, 3, 33, 1, 5.$$

How did we find this? We start by dividing 2048 by 615, and get 3, with remainder 203 This shows us that

$$\frac{615}{2048} = \cfrac{1}{3 + \cfrac{203}{615}}$$

and so forth.

We now proceed by finding the continued fraction of $\frac{203}{615}$. Since $615 = 3 \cdot 203 + 9$, this gives

$$\cfrac{1}{3 + \cfrac{203}{615}} = \cfrac{1}{3 + \cfrac{1}{3 + \cfrac{6}{203}}}$$

Continuing this process, we get

$$\begin{aligned}
\frac{615}{2048} &= \frac{1}{3+\dfrac{203}{615}} \\
&= \frac{1}{3+\dfrac{1}{3+\dfrac{6}{203}}} \\
&= \frac{1}{3+\dfrac{1}{3+\dfrac{1}{33+\dfrac{5}{6}}}} \\
&= \frac{1}{3+\dfrac{1}{3+\dfrac{1}{33+\dfrac{1}{1+\dfrac{1}{5}}}}}
\end{aligned}$$

The property of continued fractions that we will be using is that all the best approximations of a real number R by a rational number are the convergents of the continued fractions for the number. What are the convergents? With the example above, the first few convergents are:

$$\begin{aligned}
\frac{1}{3} &= \frac{1}{3} = 0.3333, \\
\frac{1}{3+\dfrac{1}{3}} &= \frac{3}{10} = 0.3, \\
\frac{1}{3+\dfrac{1}{3+\dfrac{1}{33}}} &= \frac{100}{333} = 0.3003003,
\end{aligned}$$

You can see that these values keep getting closer to $615/2048 = .300293$. The third convergent has a denominator of 333, which is clearly too large when we're factoring 33, so the right convergent to choose is the second one, which has 10 as the denominator, and which gives us the correct factorization.

The remaining thing to do is to show that all the close approximations to a fraction are convergents of its continued fraction. We will show:

Theorem 1 *if* $|R - \frac{p}{q}| \leq \frac{1}{2q^2}$*, then* $\frac{p}{q}$ *is one of the convergents of* r*.*

How do we show this theorem? We will first prove a lemma:

Lemma 1 *Suppose* $\frac{p}{q} < R < \frac{p'}{q'}$ *and* $pq' = 1 + p'q$ *With these conditions if* $q < q'$*, then* $\frac{p}{q}$ *is one of the convergents of R, and if* $q' < q$*, then* $\frac{p'}{q'}$ *is one of the convergents of R.*

Let's take as an example $R = 615/2048 \approx .30030293$. We have

$$\frac{3}{10} = 0.3 < 615/2048 = 0.30030293 < \frac{10}{33} = 0.30303.$$

We can easily check that $3 \cdot 33 + 1 = 10 \cdot 10$ This shows that $\frac{3}{10}$ is a convergent. ($\frac{33}{100}$ is not, although it is something called a *semiconvergent*).

Proof of Lemma:

First, let's look at the continued fractions for $\frac{p}{q}$ and $\frac{p'}{q'}$. I claim that they cannot be of the forms

$$\frac{p}{q} = \frac{1}{a + \dfrac{1}{b + \ldots}}$$

and

$$\frac{p'}{q'} = \frac{1}{a' + \dfrac{1}{b' + \ldots}}$$

with $a > a'$. Suppose they were. Then the fraction $\dfrac{1}{a}$ would be between $\frac{p}{q}$ and $\frac{p'}{q'}$, and would have a lower denominator than either, and it would be impossible for $\frac{p'}{q'} - \frac{p}{q} = \frac{1}{qq'}$. Thus, p' and q' both must start with $a = a'$. Now, let's consider the continued fractions

$$\frac{p}{q} = \frac{1}{a + \frac{b}{c}} \quad \text{and} \quad \frac{p'}{q'} = \frac{1}{a + \frac{b'}{c'}}$$

We will show $p'q - pq' = 1$ if and only if $b'c - bc' = 1$.

First, we calculate

$$\frac{p}{q} = \frac{c}{ac + b} \quad \text{and} \quad \frac{p'}{q'} = \frac{c'}{ac' + b'}$$

so

$$p'q - pq' = c(ac' + b') - c'(ac + b) = b'c - bc'.$$

What this shows is that if the two continued fractions

$$\frac{p}{q} = \frac{1}{a + \dfrac{1}{b + \dfrac{1}{c + \ldots}}} \quad \text{and} \quad \frac{p'}{q'} = \frac{1}{a' + \dfrac{1}{b' + \dfrac{1}{c' + \ldots}}},$$

satisfy $|\frac{p'}{q'} - \frac{p}{q}| = \frac{1}{qq'}$, then the continued fractions

$$\frac{r}{s} = \cfrac{1}{b + \cfrac{1}{c + \dots}} \quad \text{and} \quad \frac{r'}{s'} = \cfrac{1}{b' + \cfrac{1}{c' + \dots}},$$

satisfy $|\frac{r}{s} - \frac{r'}{s'}| = \frac{1}{ss'}$. We can in this way keep removing the first terms of the continued fractions and preserve the relation between the remaining terms. When can this process end? It can only end when one of the two continued fractions has been reduced to the form $\frac{1}{a}$. At this point, the other continued fraction must look like

$$\cfrac{1}{a + \cfrac{1}{b + \dots}},$$

so the first continued fraction is a convergent of the second one. And since R is sandwiched between them, the first continued fraction must also be a convergent of R.

We now use the lemma to prove the theorem. Suppose that $\frac{p}{q} < R$ (the case of $\frac{p'}{q'} > R$ is completely analogous) and that $R - \frac{p}{q} < \frac{1}{2q^2}$. Now, $\frac{p}{q}$ must be the closest fraction to R with denominator at most q, because the closest two fractions with denominator less than or equal to q can be to each other is $\frac{1}{q(q-1)}$. There must also be a smallest fraction larger than $\frac{p}{q}$ with denominator at most q. Call this fraction $\frac{p'}{q'}$. Because there are no fractions between $\frac{p}{q}$ and $\frac{p'}{q'}$ with denominator at most q, we must have $pq' + 1 = p'q$. And we must have

$$\frac{p}{q} < R < \frac{p'}{q'}$$

Let's consider the fraction $\frac{p+p'}{q+q'}$. We have

$$\begin{aligned} \frac{p+p'}{q+q'} - \frac{p}{q} &= \frac{q(p+p') - p(q+q')}{q(q+q')} \\ &= \frac{p'q - q'p}{q(q+q')} \\ &= \frac{1}{q(q+q')} > \frac{1}{2q^2} \end{aligned}$$

This is larger than the distance between $\frac{p}{q}$ and R, so R must be between $\frac{p}{q}$ and $\frac{p+p'}{q+q'}$. And clearly the denominator q is less than the denominator $q + q'$. This shows that $\frac{p}{q}$ satisfies the conditions of the Lemma to be a convergent of R, and we have proved Theorem 1.

3 The Chinese Remainder Theorem

Recall that I said that the probability of finding an r such that $\gcd(a^{r/2} \pm 1, N)$ gave you a factor was at least $\frac{1}{2}$. Why is this true? You need the Chinese remainder theorem to prove this.

What is the Chinese remainder theorem? It says that if you have a product $N = PQ$, and if P and Q sre relatively prime, then there is a one-to-one correspondence between numbers modulo N and pairs of numbers modulo P and Q. That is, we have a correspondence between

$$x \bmod N \longleftrightarrow \big(x \bmod P, x \bmod Q\big)$$

Let's take $77 = 7 \cdot 11$ as an example. Suppose we have the number $60 \bmod 77$. We want the pair (x, y) corresponding with 77. We find this pair by finding the remainder when 60 is divided by 7 and 11, respectively. Thus $53 \leftrightarrow (4, 6)$. There is a polynomial-time algorithm to go the other way; that is, from the pair $(2, 8)$, it is possible to find $30 \bmod 77$, but we won't cover this calculation in these notes.

How does r for some a depend on $a \bmod P$ and $a \bmod Q$? It turns out that 3 is a multiplicative generator modulo 7 and 2 is a multiplicative generator mod 11. Let's use this to make a table of the numbers modulo 7 and 11. For a number x, we let $r_7(x)$ and r_{11} be the smallest power to which we have to raise x to get 1.

power of 3	$x \bmod 7$	r_7	power of 2	$x \bmod 11$	r_{11}
3^1	3	6	2^1	2	10
3^2	2	3	2^2	4	5
3^3	6	2	2^3	8	10
3^4	4	3	2^4	5	5
3^5	5	6	2^5	10	2
3^6	1	1	2^6	9	5
			2^7	7	10
			2^8	3	5
			2^9	6	10
			2^{10}	1	1

You can see from a little thought that if g is a generator of the multiplicative group mod P, then for $a = g^x$, $r_P = P - 1$ if and only if $\gcd(x, P - 1) = 1$.

How do we combine the r_7 and r_{11} to get r_{77}. For a number to be 1 mod 77, it has to be 1 both mod 7 and mod 11. Thus, we need to take the least common multiple (lcm) of r_7 and r_{11}. For example, let's look at the number $53 \longleftrightarrow (4, 6)$. From the table, we need 4^3 to make it 1 mod 11 and 6^{10} to make it 1 mod 11. Thus, if we choose $a = 53$, we get $r = \mathrm{lcm}(3, 10) = 30$. So what is 53^{15}. It corresponds to $4^{15} \equiv (4^3)^5 \equiv 1 \pmod 7$ and $6^{15} \equiv 6^{10}6^5 \equiv 6^5 \equiv -1 \pmod{11}$. Thus, $53^{15} \neq \pm 1$, and will it give us a factor.

If we had chosen $a = 30 \longleftrightarrow (2, 8)$, we would get $r_7 = 3$ and $r_{11} = 5$. Then, $r = 15$ so it is not even, and thus it doesn't work.

How about $a = 6 \longrightarrow (6, 6)$? Then we get $r_7 = 2$ and $r_{11} = 10$. Thus, $r = 10$. We can see from the table that $6^{10} \equiv 1 \bmod 11$, so $6^5 \bmod 11$ is a square root of 1. There is only one square root of 1 modulo any prime P, so $6^5 \equiv 1 \bmod 11$. Similarlly, $6^5 \equiv -1 \bmod 7$, so $6^5 \equiv -1 \bmod 77$, and we don't find a factor.

What are the conditions for giving a factor for a general $N = PQ$, with P and Q prime? r cannot be odd, and $a^{r/2}$ cannot be -1. We have r is odd if and only if $r_P(a)$

and $r_Q(a)$ are odd, so 2 doesn't divide either $r_P(a)$ and $r_Q(a)$. Put another way, the largest power of 2 dividing $r_P(a)$ (and $r_Q(a)$) is 0.

Now, if $a^{r/2} = -1$ modulo both P and Q, it must be the case that the same largest power of 2 divides both r_P and r_Q. To see this, consider an example. if $r_P(a) = 4s$ and $r_Q(a) = 2t$, where s and t are odd, then $r = \mathrm{lcm}(r_P, r_Q) = 4\mathrm{lcm}(s,t)$. And $a^{2\mathrm{lcm}(r_P,r_Q} \equiv -1 \bmod P$ but $a^{2\mathrm{lcm}(r_P,r_Q} \equiv 1 \bmod P$. So an a will result in a factor if and only if the largest powers of 2 dividing r_P and r_Q are different.

Now consider an arbitrary P and a generator g for it. For $a = g^x \bmod P$, if x is odd, then the largest power of 2 dividing $P-1$ is the largest power of two dividing $r_P(a)$, and if g is even, the largest power of 2 dividing $r_P(a)$ is smaller than the largest power of 2 dividing $P-1$. Thus, if a is chosen at random, the largest powers of 2 dividing $r_P(a)$ and $r_Q(a)$ are the same with probability at most $\frac{1}{2}$, and we see that a random a works with probability at least $\frac{1}{2}$.