MAT 4930: Quantum Algorithms and Complexity **Spring 2021**

Lecture 11: Shor's factoring algorithm

Lecturer: Jean-François Biasse *TA: Robert Hart*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

11.1 Order finding in $(\mathbb{Z}/N\mathbb{Z})^*$

As we previously saw, the factorization of an RSA integer N can be reduced to the search for the order of a random element $a \in \{0, \dots, N-1\}$ modulo N (and if a is not coprime to N, then there is no need for a quantum computer). In this section, we show how this could be done if the order of the group $(\mathbb{Z}/N\mathbb{Z})^*$ were known. This is obviously not the case when we do not know the factorization of N. Indeed, the only known method to compute $\phi(N) = |(\mathbb{Z}/N\mathbb{Z})^*| = (p-1)(q-1)$ is to compute p and q, i.e. to factor N. If we do that, then we are no longer interested in the order of a in $(\mathbb{Z}/N\mathbb{Z})^*$. However, this hypothetical search of the order of a modulo N shows some of the essential ingredients of Shor's algorithm which we formally describe in Section 11.2.

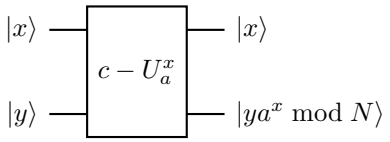
We assume here that we have a Quantum Fourier Transform modulo $\phi(N)$. In the actual Shor's algorithm, we will work modulo 2^n for n a large enough bound on the bit size of N. This means that we have an implementation of

$$|x\rangle \longmapsto \frac{1}{\sqrt{\phi(N)}} \sum_{y=0}^{\phi(N)-1} e^{\frac{2i\pi xy}{\phi(N)}} |y\rangle .$$

We assume $a \in \{0, \dots, \phi(N)-1\}$ coprime to n was drawn uniformly at random. Its (unknown) order r satisfies $r \mid \phi(N)$. Now we additionally assume that we can create a uniform superposition of elements in $(\mathbb{Z}/N\mathbb{Z})^*$:

$$|\psi\rangle := \frac{1}{\sqrt{\phi(N)}} \sum_{x=0}^{\phi(N)-1} |x\rangle .$$

When working modulo 2^n, this operation is simply done by applying $H^{\otimes n}$ to the input state $|0\rangle^{\otimes n}$, but when $\phi(N)$ is not a power of 2, this is potentially a little more complicated. Then we remeber that we can efficiently implement the a-to-the-power-x gate



On input $|\psi\rangle \otimes |1 \bmod N\rangle$, this circuit yields the state

$$|\psi_0\rangle = \frac{1}{\sqrt{\phi(N)}} \sum_{x=0}^{\phi(N)-1} |x\rangle\, |a^x \bmod N\rangle .$$

As a has order $r \mid \phi(N)$, the possible values for $a^x \bmod N$ for $x \in \{0, \ldots, N-1\}$ are

$$1, a, a^2, \ldots, a^{r-1}, 1, a^2, \ldots, a^{r-1}, \ldots\ldots, 1, a, a^2, \ldots, a^{r-1}.$$

The sequence $1, a, a^2, \ldots, a^{r-1}$ is repeated $\phi(N)/r$ times. Therefore, the state can be re-written as

$$|\psi_0\rangle = \sum_{b=0}^{r-1} \left(\frac{1}{\sqrt{\phi(N)}} \sum_{z=0}^{\phi(N)/r-1} |zr+b\rangle \right) \left|a^b \bmod N\right\rangle.$$

Then, we measure the second register to obtain a^b for some $b \in \{0, \ldots, r-1\}$. This leaves the system in the following state

$$|\psi_1\rangle := \frac{1}{\sqrt{\phi(N)/r}} \sum_{z=0}^{\phi(N)/r-1} |zr+b\rangle \left|a^b \bmod N\right\rangle.$$

We can discard the second register, and apply to the first register the inverse of the QFT modulo $\phi(N)$ that does

$$\mathrm{QFT}_{\phi(N)}^{-1} : |x\rangle \longmapsto \frac{1}{\sqrt{\phi(N)}} \sum_{y=0}^{\phi(N)-1} e^{-\frac{2i\pi xy}{\phi(N)}} |y\rangle.$$

We thus obtain

$$\begin{aligned}
\frac{\mathrm{QFT}_{\phi(N)}^{-1}}{\sqrt{\phi(N)/r}} \sum_{z=0}^{\phi(N)/r-1} |zr+b\rangle &= \frac{1}{\sqrt{\phi(N)}\sqrt{\phi(N)/r}} \sum_{z=0}^{\phi(N)/r-1} \sum_{y=0}^{\phi(N)-1} e^{-\frac{2i\pi(zr+b)y}{\phi(N)}} |y\rangle \\
&= \frac{1}{\sqrt{\phi(N)}\sqrt{\phi(N)/r}} \sum_{y=0}^{\phi(N)-1} \sum_{z=0}^{\phi(N)/r-1} e^{-\frac{2i\pi(zr+b)y}{\phi(N)}} |y\rangle \\
&= \frac{1}{\sqrt{\phi(N)}\sqrt{\phi(N)/r}} \sum_{y=0}^{\phi(N)-1} e^{-\frac{2i\pi by}{\phi(N)}} \left(\sum_{z=0}^{\phi(N)/r-1} e^{-\frac{2i\pi zry}{\phi(N)}} \right) |y\rangle \\
&= \frac{1}{\sqrt{\phi(N)}\sqrt{\phi(N)/r}} \sum_{y=0}^{\phi(N)-1} e^{-\frac{2i\pi by}{\phi(N)}} \left(\sum_{z=0}^{\phi(N)/r-1} \zeta^z \right) |y\rangle,
\end{aligned}$$

where $\zeta := e^{-\frac{2i\pi y}{\phi(N)/r}}$. When $y \not\equiv 0 \bmod \phi(N)/r$, we have

$$\sum_{z=0}^{\phi(N)/r-1} \zeta^z = \frac{1-\zeta^{\phi(N)/r}}{1-\zeta} = 0.$$

On the other hand, if $y \equiv 0 \bmod \phi(N)/r$, then let $0 \le k < r$ such that $y = k\frac{\phi(N)}{r}$. We have

$$\sum_{z=0}^{\phi(N)/r-1} \zeta^z = \frac{\phi(N)}{r} \quad \text{and} \quad e^{-\frac{2i\pi by}{\phi(N)}} = e^{-\frac{2i\pi bk}{r}}.$$

This means that our state can be re-written as

$$\frac{1}{\sqrt{\phi(N)}\sqrt{\phi(N)/r}} \sum_{y=0}^{\phi(N)-1} e^{-\frac{2i\pi by}{\phi(N)}} \left(\sum_{z=0}^{\phi(N)/r-1} \zeta^z \right) |y\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-\frac{2i\pi bk}{r}} |k\phi(N)/r\rangle.$$

Hence, a measurement of the system yields $y = k\phi(N)/r$ for $k \in \{0, \ldots, r-1\}$ distributed uniformly at random. Such a y satisfies $\frac{y}{\phi(N)} = \frac{k}{r}$. Whenever k is coprime with r, this fraction yields r directly.

Proposition 11.1 (Probability of success) *With the above notations, the probability p of successfully recovering r satisfies*

$$p \in \Omega\left(\frac{1}{\log\log(r)}\right) \subseteq \Omega\left(\frac{1}{\log\log(N)}\right),$$

where $\Omega(f(n))$ denotes the set of functions g such that there is a constant $C > 0$, and n_0 with $\forall n > n_0$, $g(n) > Cf(n)$.

Proof: The order finding algorithm returns $\frac{k}{r}$ for $k \in \{0, \ldots, r-1\}$ distributed uniformly at random. There are $\phi(r)$ elements in $\{0, \ldots, r-1\}$ that are coprime to r, and thefore, the probability of success is $\phi(r)/r$. The Euler function ϕ fluctuates a lot, but some very pessimistic bounds are known, including the fact that there is a constant $C > 0$ such that $\phi(r)/r > C/\log\log(r)$. The second inequality in the result follows from the fact that $r \leq \phi(N) \leq N$. ■

11.2 Quantum factoring

Quantum factoring reduces to finding the order of an element modulo N, but with the essential restriction that we do not know the order of the multiplicative group $(\mathbb{Z}/N\mathbb{Z})^*$. This complicates things a little since all we can do is work modulo 2^n for a large enough n. Then instead of obtaining $\frac{y}{\phi(N)} = \frac{k}{r}$ for some k, we get $\frac{y}{2^n} \approx \frac{k}{r}$, and if this approximation is good enough, then $\frac{k}{r}$ is one of the convergents of the continued fraction expansion of $\frac{y}{2^n}$. So the main difficulty is to prove that we measure y such that $\frac{y}{2^n}$ is close enough to a fraction of the form $\frac{k}{r}$ with good probability.

Interestingly, the rest of the algorithm remains pretty much similar to what we have seen in the previous section. We recap the main steps in order to properly introduce additional notations necessary due to the fact that we do not know $\phi(N)$. Using $H^{\otimes n}$ and $c - U_a^x$, we produce the state

$$|\psi_0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \, |a^x \bmod N\rangle \,.$$

This state can be re-written as

$$|\psi_0\rangle = \sum_{b=0}^{r-1} \left(\frac{1}{\sqrt{2^n}} \sum_{z=0}^{m_b-1} |zr+b\rangle \right) \left|a^b \bmod N\right\rangle,$$

where $m_b := |\{z, \mid 0 \leq zr + b < 2^n\}| = \left\lfloor \frac{2^n - b - 1}{r} \right\rfloor$ is the largest integer such that $(m_b - 1)r + b \leq 2^n - 1$. Then, we measure the second register (and discard it to obtain a^b for some $b \in \{0, \ldots, r-1\}$. This leaves the system in the following state

$$|\psi_1\rangle := \frac{1}{\sqrt{m_b}} \sum_{z=0}^{m_b-1} |zr+b\rangle \left|a^b \bmod N\right\rangle.$$

Then we apply $\mathrm{QFT}_{2^n}^{-1}$ to the first register that is in the state $\frac{1}{\sqrt{m_b}}\sum_{z=0}^{m_b-1}|zr+b\rangle$, and thus obtain

$$\begin{aligned}
\frac{\mathrm{QFT}_{2^n}^{-1}}{\sqrt{m_b}}\sum_{z=0}^{m_b-1}|zr+b\rangle &= \frac{1}{\sqrt{2^n}\sqrt{m_b}}\sum_{z=0}^{m_b-1}\sum_{y=0}^{2^n-1}e^{-\frac{2i\pi(zr+b)y}{2^n}}|y\rangle \\
&= \frac{1}{\sqrt{2^n}\sqrt{m_b}}\sum_{y=0}^{2^n-1}\sum_{z=0}^{m_b-1}e^{-\frac{2i\pi(zr+b)y}{2^n}}|y\rangle \\
&= \frac{1}{\sqrt{2^n}\sqrt{m_b}}\sum_{y=0}^{2^n-1}e^{-\frac{2i\pi by}{2^n}}\left(\sum_{z=0}^{m_b-1}e^{-\frac{2i\pi zry}{2^n}}\right)|y\rangle \\
&= \frac{1}{\sqrt{2^n}\sqrt{m_b}}\sum_{y=0}^{2^n-1}e^{-\frac{2i\pi by}{2^n}}\left(\sum_{z=0}^{m_b-1}\zeta^z\right)|y\rangle\,,
\end{aligned}$$

where $\zeta := e^{-\frac{2i\pi ry}{2^n}}$. The probability of measuring a given y is

$$\frac{1}{m_b 2^n}\left|\sum_{z=0}^{m_b-1}\zeta^z\right|^2 = \frac{1}{m_b 2^n}\left|\frac{\zeta^{m_b}-1}{\zeta-1}\right|^2 = \frac{1}{m_b 2^n}\left|\frac{\zeta^{m_b/2}-\zeta^{m_b/2}}{\zeta^{1/2}-\zeta^{1/2}}\right|^2 = \frac{1}{m_b 2^n}\left|\frac{\sin\left(\frac{\pi m_b ry}{2^n}\right)}{\sin\left(\frac{\pi ry}{2^n}\right)}\right|^2$$

Lemma 11.2 *The probability of drawing y such that*

$$\left|\frac{x}{2^n}-\frac{k}{r}\right| \leq \frac{1}{2m_b r}$$

for some integer k is at least $\frac{m_b}{2^n}\frac{4}{\pi^2}$.

Proof: We begin this proof by showing that if $|\theta| \leq \frac{\pi}{2M}$, then $\frac{1}{M^2}\frac{\sin^2(M\theta)}{\sin^2(\theta)} \geq \frac{4}{\pi^2}$ for any M. This is a direct consequence of the fact that when $|x| \leq \frac{\pi}{2}$, we have $\frac{2}{\pi} \leq \frac{\sin(x)}{x}$ (proof by direct analysis of the variations of the function of the variable x). When used with $x = M\theta$ this yields

$$\frac{1}{M^2}\frac{\sin^2(M\theta)}{\sin^2(\theta)} \geq \frac{1}{M^2}\frac{4M^2\theta^2}{\pi^2}\frac{1}{\sin^2(\theta)} \geq \frac{1}{M^2}\frac{4M^2\theta^2}{\pi^2}\frac{1}{\theta^2} = \frac{4}{\pi^2}.$$

Now we notice that for all $k \in \mathbb{Z}$,

$$\sin^2\left(\frac{\pi ry}{2^n}\right) = \sin^2\left(\frac{\pi ry}{2^n}-k\pi\right) = \sin^2\left(\pi r\left(\frac{y}{2^n}-\frac{k}{r}\right)\right).$$

We define $\theta := \pi r\left(\frac{y}{2^n}-\frac{k}{r}\right)$ and $M := m_b$. Since $\left|\frac{x}{2^n}-\frac{k}{r}\right| \leq \frac{1}{2m_b r}$, we have $|\theta| \leq \frac{\pi}{2M}$, and therefore the probability of measuring a given y satisfies

$$\frac{1}{m_b 2^n}\left|\frac{\sin\left(\frac{\pi m_b ry}{2^n}\right)}{\sin\left(\frac{\pi ry}{2^n}\right)}\right|^2 = \frac{1}{m_b 2^n}\frac{\sin^2(M\theta)}{\sin^2(\theta)} \geq \frac{1}{m_b 2^n}\frac{4M^2}{\pi^2} = \frac{1}{m_b 2^n}\frac{4m_b^2}{\pi^2} = \frac{m_b}{2^n}\frac{4}{\pi^2}.$$

■

The previous lemma shows us that if m_b is large enough, then $\frac{x}{2^n}$ will be close to $\frac{k}{r}$, which will force the latter to appear as a convergent in the continued fraction expansion of $\frac{x}{2^n}$.

Lemma 11.3 *If we choose n such that $2^n \geq 2r^2$, then necessarily $m_b \geq r$.*

Proof: By definition, we have that $m_b := |\{z, \mid 0 \leq zr + b < 2^n\}| = \left\lfloor \frac{2^n - b - 1}{r} \right\rfloor$. It can be re-written as

- $m_b = \frac{2^n - (2^n \bmod r)}{r} + 1$ if $0 \leq b < 2^n \bmod r$,
- $m_b = \frac{2^n - (2^n \bmod r)}{r}$ if $2^n \bmod r \leq b < r$.

In each case, if $2^n \geq 2r^2$, $m_b \geq \frac{2r^2 - r}{r} \geq r$. ■

Proposition 11.4 (Probability of success) *If $2^n \geq 2r^2$, then the probability p of success of Shor's algorithm satisfies*

$$p \in \Omega\left(\frac{1}{\log\log(r)}\right) \subseteq \Omega\left(\frac{1}{\log\log(N)}\right).$$

Proof: We saw from the previous lemmas that whenever $\left|\frac{x}{2^n} - \frac{k}{r}\right| \leq \frac{1}{2m_b r}$, the probability of drawing x is at least $\frac{m_b}{2^n}\frac{4}{\pi^2}$. When $2^n \geq 2r^2$, we addtionally have that $\frac{1}{2m_b r} \leq \frac{1}{2r^2}$, which means that $\frac{k}{r}$ appears in the list of convergents of $\frac{x}{2^n}$.

As we saw when finding orders in $(\mathbb{Z}/N\mathbb{Z})^*$, there are $\phi(r)$ potential $k \in \{0, \ldots, r-1\}$ that are coprime to r (meaning that k/r yields r). For each of these k, there is a $y \in \{0, 2^n - 1\}$ such that $\left|\frac{x}{2^n} - \frac{k}{r}\right| \leq \frac{1}{2^{n+1}}$. Therefore, the probability of drawing a y in the desired range is at least $\phi(r)\frac{m_b}{2^n}\frac{4}{\pi^2}$. Now we have

$$\frac{m_b}{2^n} \geq \frac{1}{t}\frac{2^n = (2^n \bmod r)}{2^n} \geq \frac{1}{r}\frac{2^n - r}{2^n} \geq \frac{1}{r}\frac{2^n - 2^{(n-1)/2}}{2^n} \geq \frac{1}{r}\left(1 - \frac{1}{2^{(n+1)/2}}\right) \geq \frac{2}{r}.$$

This means that we draw a y in the desired range with probability at least $\frac{\phi(r)}{r}\frac{2}{\pi^2}$. As we previously saw that $\frac{\phi(r)}{r} \in \Omega(1/\log\log(r))$, we can conclude that the probability of success is in $\Omega(1/\log\log(r))$ as well. ■

So far, we have define "success" as the measurement of y such that the continued fraction expansion of $\frac{y}{2^n}$ contains $\frac{k}{r}$ for k coprime to r. We need however to be able to recognize the correct convergent.

Proposition 11.5 (Recognizing the right convergent) *We have the following information about the right convergent:*

1. *There can be at most one convergent $\frac{a_i}{b_i} \neq \frac{x}{2^n}$ satisfying $\left|\frac{x}{2^n} - \frac{a_i}{b_i}\right| \leq \frac{1}{2r^2}$ and $b_i \leq r$.*
2. *If $2^n \geq 2r^2$ and $\left|\frac{x}{2^n} - \frac{k}{r}\right| \leq \frac{1}{2^n}$, then $\frac{a_i}{b_i} = \frac{k}{r}$ will be the only convergent of $\frac{x}{2^n}$ with $b_i \leq 2^{\frac{(n-1)}{2}}$ and $\left|\frac{x}{2^n} - \frac{a_i}{b_i}\right| \leq \frac{1}{2^n}$.*

Proof: 1. Assume that we have $\frac{a_i}{b_i}$ and $\frac{a_j}{b_j}$ such that $\left|\frac{x}{2^n} - \frac{a_i}{b_i}\right| \leq \frac{1}{2r^2}$ and $\left|\frac{x}{2^n} - \frac{a_j}{b_j}\right| \leq \frac{1}{2r^2}$ with $b_i, b_i' \leq r$. Then by triangular inequality, $\left|\frac{a_i}{b_i} - \frac{a_j}{b_j}\right| \leq \frac{1}{r^2}$. Without loss of generality, we can assume $i < j$ and thus $b_i b_j < b_j^2 \leq r^2$. This implies

$$\left|\frac{a_i}{b_i} - \frac{a_j}{b_j}\right| = \left|\frac{b_j a_i - a_j b_i}{b_i b_j}\right| > \left|\frac{b_j a_i - a_j b_i}{r^2}\right|.$$

Since $\left|\frac{a_i}{b_i} - \frac{a_j}{b_j}\right| \leq \frac{1}{r^2}$, the only possibility is $b_j a_i - a_j b_i = 0$, i.e. $\frac{a_i}{b_i} = \frac{a_j}{b_j}$.

2. The proof is similar to that of 1. with $2r^2$ replaced by 2^n. ■

In case 1. the issue is that we do not know r, therefore we could see multiple candidates for the right convergent. From the proof of Proposition 11.4, we know that we are in the case 2. Therefore, for each convergent $\frac{a_i}{b_i}$, we can check how close it is from $\frac{y}{2^n}$, and whether b_i exceeds $2^{(n-1)/2}$. If we reach $b_j > 2^{(n-1)/2}$ without having found a convergent close enough to $\frac{y}{2^n}$, then we declare a failure and try the algorithm again.