Lecture 8

In which we use the quantum Fourier transform to solve the period-finding problem.

1 The Period Finding Problem

Let $f : \{0, \ldots, M-1\} \to \{0, \ldots, M-1\}$ be a periodic function of period r, meaning that $\forall x \in \{0, \ldots, M-r-1\}$ we have that $f(x) = f(x+r)$ and the values $f(x), f(x+1), \ldots, f(x+r-1)$ are all distinct. Suppose also that $M = 2^m$ is a power of 2 and that $r \leq \sqrt{M}/2$.

Today we will describe a quantum algorithm that finds r in time polynomial in m and in the size of a classical circuit computing f.

The importance of the period-finding problem is that, as we will see next time, the integer factoring problem reduces to it, and so the quantum polynomial time for period-finding yields a quantum polynomial time algorithm for integer factoring.

2 The Algorithm

The algorithm is essentially identical to the algorithm for the "Boolean period-finding problem" discussed in the last lecture.

1. **Create the quantum state $\frac{1}{\sqrt{M}} \sum_x |x\rangle|f(x)\rangle$**

   Let U_f be a unitary transformation on $\ell = m + m + O(S)$ bits that maps $|x\rangle|0\cdots0\rangle|0\cdots0\rangle$ to $|x\rangle|f(x)\rangle|0\cdots0\rangle$, where S is the size of a classical circuit that computes f. We construct a circuit over ℓ qubits that first applies Hadamard gates to each of the first m qubits. After these operations, starting from the input $|0^\ell\rangle$ we get $\frac{1}{\sqrt{M}}|x\rangle|0^{\ell-m}\rangle$. Then we apply U_f, which gives us the state $\frac{1}{\sqrt{M}}|x\rangle|f(x)\rangle|0^{\ell-2m}\rangle$. From this point on, we ignore the last $\ell - 2m$ wires.

2. **Measure the last m bits of the state**

   The outcome of this measurement will be a possible output y of $f(\cdot)$. Let us call x_0 the smallest input such that $f(x_0) = y$. For such an outcome, the residual state will be

$$\frac{1}{\sqrt{\left\lceil \frac{M}{r} \right\rceil}} \sum_{t=0}^{\left\lceil \frac{M}{r} \right\rceil - 1} |x_0 + tr\rangle |f(x_0)\rangle$$

   where $[M/r]$ stands for $\lfloor M/r \rfloor$ or for $\lceil M/r \rceil$ depending on x_0.

   From this point on we ignore the last n bits of the state because they have been fixed by the measurement.

3. **Apply the Fourier transform to the first m bits**

   The state becomes

$$\frac{1}{\sqrt{M}} \frac{1}{\sqrt{\left\lceil \frac{M}{r} \right\rceil}} \sum_s \sum_{t=0}^{\left\lceil \frac{M}{r} \right\rceil - 1} \omega^{(x_0 + tr) \cdot s} |s\rangle$$

   where $\omega = e^{-2\pi i/M}$.

4. **Measure the first m bits.**

   The measurement will give us an integer s with probability

$$\frac{1}{M} \cdot \frac{1}{\left\lceil \frac{M}{r} \right\rceil} |\omega^{x_0 s}|^2 \left| \sum_{t=0}^{\left\lceil \frac{M}{r} \right\rceil - 1} \omega^{(x_0 + tr) \cdot s} \right|^2$$

$$= \frac{1}{M} \cdot \frac{1}{\left\lceil \frac{M}{r} \right\rceil} \left| \sum_{t=0}^{\left\lceil \frac{M}{r} \right\rceil - 1} \omega^{trs} \right|^2$$

We will now discuss how to use the measurement done in step (4) in order to estimate r. The point will be that, with noticeably high probability, s/M will be close to k/r for a random k, and this information will be sufficient to identify r, after executing the algorithm a few times in order to obtain multiple samples. The key to the analysis is to understand the probability distribution of outcomes of the measurement in step (4). The analysis is simpler in the special case in which r divides M, so we begin with this special case.

3 If M is a Multiple of r

Suppose that $q := M/r$ is an integer, and let us call a value s "good" if s is a multiple of q. Note that there are exactly r good values of s, namely $0, M/r, 2M/r, \ldots, M-r$.

If s is good, then, for every t, trs is a multiple of M, and so $\omega^{trs} = 1$, and the probability that s is sampled is $1/r$, and so the good values of s contain all the probability mass of the distribution..

This means that in step 4 we sample a number s which is uniformly distributed in $\{0, M/r, 2M/r, \ldots, M-r\}$, and the rational number s/M, which we can compute after sampling s, is of the form k/r for a random $k \in \{0, \ldots, r-1\}$. After simplifying the fraction s/M, we get coprime integers a, b such that $\frac{s}{M} = \frac{a}{b}$; if k and r are coprime, then $r = b$, otherwise b is a divisor of r.

If we execute the algorithm twice, we get two numbers s_1, s_2 such that $s_i/M = k_i/r$ for random k_1, k_2. If we compute the simplified fractions $\frac{a_i}{b_i} = \frac{s_i}{M}$, then each b_i is either r or a divisor of r and, more precisely, we have $b_i = r/\gcd(k_i, r)$. Now, if $\gcd(k_1, r)$ and $\gcd(k_2, r)$ are coprime, then $r = \mathrm{lcm}(b_1, b_2)$.

This gives us a quantum algorithm that computes r and whose error probability is the probability that picking two random number $k_1, k_2 \in \{0, \ldots, r-1\}$ we have that r, k_1, k_2 all share a common factor. The probability that this happens is at most the probability that k_1, k_2 share a common factor, which is at most

$$\sum_{p \text{ prime}} \mathbb{P}[k_1 \text{ multiple of } p \wedge k_2 \text{ multiple of } p \leq \sum_{p \text{ prime}} \frac{1}{p^2} < \sum_{n \geq 2} \frac{1}{n} = \frac{\pi^2}{6} - 1 < .65$$

So we have at least a probability of about $1/3$ of finding the correct r, and this can be boosted to be arbitrarily close to 1 by repeating the algorithm several times.

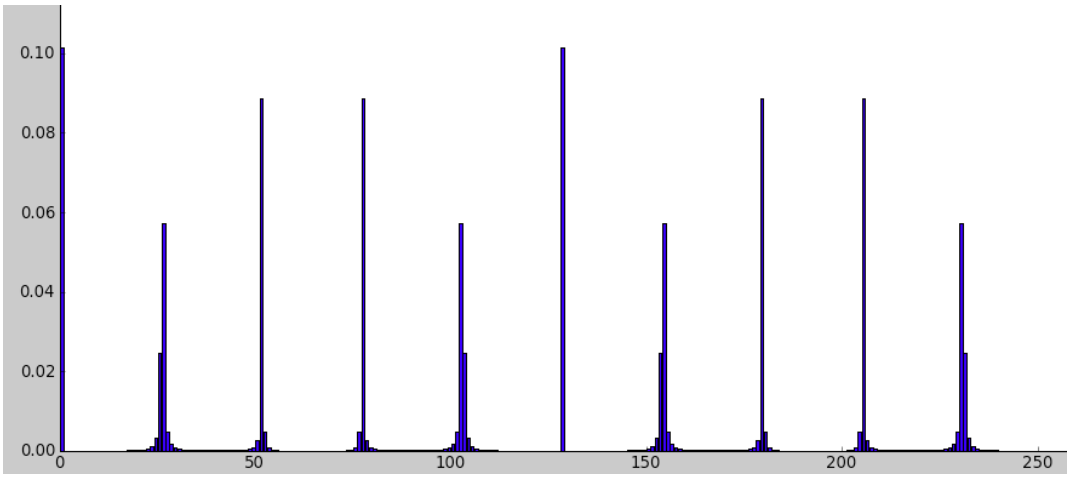
(A final note: if we repeat the algorithm several times, we collect several values $r_1, \ldots, r_a$ such that, with high probability, at least one of them is the correct period. How do we find the correct period out of this list? First of all we check for each r_i whether, say, $f(0) = f(r_i)$, and if not we remove it from the list. The remaining values are either the correct period or multiples of the correct period. We then output the smallest value of the remaining ones.)

4 The General Case

Suppose now that, as will usually be the case in the application to factoring, r does not divide M. For the general case, we will develop an "approximate version" of the argument of the previous section. We will define a value of s to be good if

sr is approximately a multiple of M, we will show that there are approximately r good values of s, each with probability approximately approximately $1/r$, so that the measurement at step (4) will give us with good probability a value of s such that s/M is close to a multiple of $1/r$, from which we will be able to get a divisor of r, and then, by repeating the algorithm several times, the actual value of r.

Before realizing the above plan, let us consider a concrete example, to get a sense of what we may hope for. Suppose that $M = 2^8 = 256$, that $r = 10$, and that, at step (2), we measured a value y equal to $f(3)$. Then, at step (4), the probability of the various values of s are as in the graph below:



If $r = 10$ had been a divisor of M, we would expect 10 values of s to have probability .1 each, those being the multiples of $M/10$, and all the other values to have probability 0.

In the above example, $s = 0$ and $s = 128$, which are the only integer multiples of $M/10 = 25.6$ have the highest probability, which is about .1015. The values $s = 25$ and $s = 26$ are both close to $M/10$, but neither is very close, and their probabilities are .0246 and .0571 respectively. The value $s = 51$ is quite close to $2 \cdot M/10 = 51.2$, and its probability is .08852, much closer to the 10% probability of the ideal case. The next near-multiple is $s = 77 \approx 3 \cdot M/10 = 76.8$, which also has probability .08852. The next probability spike is near $4 \cdot M/10 = 102.4$ and it is split between $s = 102$ and $s = 103$, which have probability .0571 and .0246, respectively, and so on.

A couple of observations that follow from the above calculations are that (1) the probability of sampling a value s depends entirely on the difference between s and the closest multiple of M/r or, equivalently, it depends exclusively on $sr \bmod M$, a fact that is easy to verify rigorously, and that (2) the values of s that differ by less than $1/2$ from the closest multiple of M/r have an overall probability of more than 80% (in fact, nearly 90%) of being sampled. We will give a rigorous proof of a weaker bound that holds in general.

Definition 1 (Good s) *We say that a value of s is good if*

$$-\frac{r}{2} \leq sr \bmod M \leq r2$$

Lemma 2 *Every good s has probability at least $\frac{1}{8r} - o(1/r)$ of being sampled.*

Proof: We need to estimate

$$\mathbb{P}[\text{outcome } s] = \frac{1}{M} \cdot \frac{1}{\left[\frac{M}{r}\right]} \left| \sum_{t=0}^{\left[\frac{M}{r}\right]-1} \omega^{trs} \right|^2 = \frac{1}{M} \cdot \frac{1}{\left[\frac{M}{r}\right]} \left| \sum_{t=0}^{\left[\frac{M}{r}\right]-1} e^{2\pi i \frac{1}{M} trs} \right|^2$$

(Recall that $\omega = e^{-2\pi i/M}$, and that conjugation does not affect the magnitude of a complex number.)

Let us consider the case $0 \leq sr \bmod M \leq \frac{r}{2}$, and the other case will be analogous. Call $g := (sr \bmod M)/M$. Then

$$\sum_{t=0}^{\left[\frac{M}{r}\right]-1} e^{2\pi i \frac{1}{M} trs} = \sum_{t=0}^{\left[\frac{M}{r}\right]-1} e^{2\pi i g t}$$

So we are summing $e^{\theta i}$ for values of θ that range from 0 to $2\pi g \frac{M}{r} \leq \pi$ in $[M/r]$ equal increments. Thinking of complex numbers as two-dimensional vectors, we are summing $[M/r]$ equally spaced unit vectors within an angle of $\leq \pi$. This means that at least half of the vectors in the sum form an angle of $\leq \pi/4$ with the sum vector, and each of those contributes at least $1/\sqrt{2}$ to the sum vector. This means that the magnitude (length) of the sum is at least $\frac{1}{2}\frac{1}{\sqrt{2}} \cdot \left[\frac{M}{r}\right]$ and the overall probability of s is at least

$$\frac{1}{M} \cdot \frac{1}{\left[\frac{M}{r}\right]} \cdot \frac{1}{8} \cdot \left[\frac{M}{r}\right]^2 \geq \frac{1}{8r} \cdot (1 - o(1))$$

□

This seems considerably less than the probabilities we saw in the example above, where the good s had probability at least .0571, which was $.571/r$. Indeed the lower bound could be improved to $\frac{4}{\pi^2} \cdot \frac{1}{r} > .4052\frac{1}{r}$ by proving the following

Claim 3 *For every good s*

$$\left| \sum_{t=0}^{\left[\frac{M}{r}\right]-1} e^{2\pi i \frac{1}{M} str} \right|^2 \geq (1 - o(1)) \frac{M^2}{r^2} \frac{4}{\pi^2}$$

PROOF:[Sketch] Again we will only look at the case $0 \leq sr \bmod M \leq \frac{r}{2}$. Ignoring the difference between $[M/r]$ and M/r (which will be accounted for in the $o(1)$ error term, we want to study

$$\left| \frac{M}{r} \sum_{t=0}^{\frac{M}{r}-1} e^{2\pi i \frac{1}{M} srt} \right|^2$$

which we can think of as

$$\left| \mathop{\mathbb{E}}_{t \sim \{0,\ldots M/r-1\}} e^{2\pi i \frac{1}{M} srt} \right|^2$$

and if we call $\theta_{\max} := 2\pi \frac{1}{M}(sr \bmod M)\frac{M}{r}$ (note that $0 \leq \theta_{\max} \leq \pi$) we can approximate the discrete set $\{2\pi \frac{1}{M} srt : t = 0, \ldots M/r - 1\}$ with the continuous interval $[0, \theta_{\max}]$, so that, up to an error that is accounted for in the $o(1)$, we have to bound the expectation

$$\left| \mathop{\mathbb{E}}_{\theta \sim [0,\theta_{\max}]} e^{i\theta} \right|^2 = \left| \frac{1}{\theta_{\max}} \int_0^{\theta_{\max}} e^{i\theta} d\theta \right|^2 = \frac{1}{\theta_{\max}^2}(2 - 2\cos\theta_{\max})$$

The function in the right-hand side of the above equation is monotone decreasing for $0 \leq \theta_{\max} \leq \pi$, and so its minimum is achieved at $\theta_{\max} = \pi$, and it is $4/\pi^2$. □

Lemma 4 *There are at least r good s*

PROOF: For every $k \in \{0, \ldots, r-1\}$, there must be an integer s_k in the interval $[kM/r - 1/2, kM/r + 1/2]$, and such an integer satisfies $-r/2 \leq s_k r \bmod M \leq r/2$. Furthermore, the integers s_k are all distinct because they belong to disjoint intervals. □

Suppose now that we have measured a good s at step (4). Then, for some k, we have $|sr - kM| \leq r/2$, that is

$$\left| \frac{s}{M} - \frac{k}{r} \right| \leq \frac{1}{2M}$$

Now we want to see how to use such an s to find the *exact* ratio k/r. First of all, let us see that this is possible in principle.

Fact 5 *If $\frac{a}{b}$ and $\frac{a'}{b'}$ are two different rationals such that $b, b' \leq D$, then*

$$\left| \frac{a}{b} - \frac{a'}{b'} \right| \geq \frac{1}{D^2}$$

PROOF: The difference is a multiple of bb'. $\square$

This means that if we have a real number ρ, and a bound D, then there can be at most one rational number a/b with $b \leq D$ such that $|\rho - a/b| < 1/2D^2$. Interestingly, if such a number exists, it can be found efficiently.

Fact 6 (Continued Fraction Algorithm) *Given a real number ρ and a bound D, there is a $O((\log D)^{O(1)})$ time algorithm that finds the unique rational number a/b with $b \leq D$ such that $|\rho - a/b| \leq 1/2D^2$, if such a number exists.*

Putting it all together, the measurement at step (4) gives us, with $\Omega(1)$ probability, a good s, and from such an s, using the fact that, for some k, $|s/M - k/r| \leq 1/2M$ and that $r < \sqrt{M}$, we can use the continued fraction algorithm to find coprime a, b such that $a/b = k/r$, which is the same outcome that we had reached in the previous section. Furthermore, each k has probability $\Omega(1/r)$, and so there is $\Omega(1)$ probability that, repeating the algorithm twice, we obtain $a_1/b_1 = k_1/r$ and $a_2/b_2 = k_2/r$ where k_1 and k_2 are coprime, so that $\text{lcm}(b_1, b_2) = r$.