Factoring Using Quantum Computers

KG

October 24, 2019

Contents

1 Introduction

In this note, we shall describe a machine for quickly factoring an RSA modulus n, which is the product of two large primes p and q. Let $g \in \mathbb{Z}$ be relatively prime to n.

Consider the function $f : \mathbb{Z} \to \mathbb{Z}$ given by $a \mapsto g^a \bmod n$. This function is *periodic*, and the period is the same as the multiplicative order of $g + \langle n \rangle$ in $\mathbb{Z}_n^*$. The period of this function therefore divides $(p-1)(q-1)$, and for most n and g the period will be close to $(p-1)(q-1)$.

We have previously seen that we can easily factor n if we know a multiple of $(p-1)(q-1)$. This means that if we can find the a multiple of the period of such a function, then by adapting our previous results, we can factor n.

So to meet our goal of factoring, it suffices to construct a machine that is capable of finding (a multiple of) the period of a periodic function $f : \mathbb{Z} \to \mathbb{Z}$. In the following, r will be the period of f.

2 Background

2.1 Rational Approximations

A *rational approximation* of a real number is a fraction that is close to the real number. Continued fractions are a useful tool to find rational approximations.

T **Fact 1.** *Let a, b, c, d be integers such that $|a/b - c/d| \leq 1/(2d^2)$. Then c/d is a convergent for the continued fraction expansion of a/b, and this convergent can be found using the Euclidian algorithm.*

Let N be an integer satisfying $N > n^2 > r^2$. Suppose we also have an integer k that is close to a multiple of the fraction N/r, that is, satisfying

$$k = l\frac{N}{r} + \delta_k \qquad \text{where } l, N, r \in \mathbb{Z} \text{ and } |\delta_k| \leq 1/2. \tag{1}$$

Then

$$\left| \frac{k}{N} - \frac{l}{r} \right| \leq \frac{1}{2N}.$$

This means that l/r is a rational approximation of the rational number k/N. This rational approximation can be found quickly using the Euclidian algorithm, but note that as there may be factors in common between l and r, we may not find r exactly.

This means that if we know N and find k satisfying (1), then we can often find a rational approximation l/r, which will give us r (up to cancellation of common factors).

E *Example* 1. Consider the rational number $k/N = 1365/2048$. In a procedure which is essentially the Euclidian algorithm, we get

$$\frac{1365}{2048} = \frac{1}{1 + \frac{683}{1365}} \qquad \frac{683}{1365} = \frac{1}{1 + \frac{682}{683}} \qquad \frac{682}{683} = \frac{1}{1 + \frac{1}{682}}$$

from which we get the continued fraction

$$\frac{1365}{2048} = \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \frac{1}{682}}}}.$$

We get the convergents

$$0, \frac{1}{1}, \frac{1}{2}, \frac{2}{3}, \frac{1365}{2048}$$

where we see that $2/3$ is a good rational approximation. Indeed, so would $4/6$, $8/12$, etc. also be.

E *Example* 2. Consider the real number $k/N = 1195/2048$. Following the procedure from the previous example we get

$$\frac{1195}{2048} = \frac{1}{1+\frac{853}{1195}} \qquad \frac{853}{1195} = \frac{1}{1+\frac{342}{853}} \qquad \frac{342}{853} = \frac{1}{2+\frac{169}{342}}$$
$$\frac{169}{342} = \frac{1}{2+\frac{4}{169}} \qquad \frac{4}{169} = \frac{1}{42+\frac{1}{4}}$$

from which we get the continued fraction

$$\frac{1195}{2048} = \cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{2+\frac{1}{42+\frac{1}{4}}}}}}.$$

We get the convergents

$$0, \frac{1}{1}, \frac{1}{2}, \frac{3}{5}, \frac{7}{12}, \frac{297}{509}, \frac{1195}{2048}$$

from which we see that 7/12 and 297/509 are good rational approximations.

2.2 Discrete Fourier Transform

Let $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_{N-1}) \in \mathbb{C}^N$. Then a normalized *discrete Fourier transform* of $\boldsymbol{\alpha}$ is $\boldsymbol{\beta} \in \mathbb{C}^N$ given by

$$\beta_k = \frac{1}{\sqrt{N}} \sum_j \alpha_j \exp\left(2\pi i \frac{kj}{N}\right), \qquad 0 \le k < N.$$

Since it is normalized, it preserves the norm of vectors, so $\|\boldsymbol{\alpha}\| = \|\boldsymbol{\beta}\|$. In fact, it is a linear map given by a unitary matrix (a normalized Vandermonde matrix).

E *Exercise* 1. Let $\omega = \exp(2\pi i/N) \in \mathbb{C}$, which is a primitive Nth root of unity. Let $V = (v_{kj})$ be the Vandermonde matrix with $v_{kj} = \omega^{kj}$, $0 \le k, j < N$.

1. Show that the $\overline{\omega^j} = \omega^{N-j}$, where $\bar{z}$ denotes complex conjugation.
2. Show that $\sum_{j=0}^{N-1} (\omega^k)^j = 0$ for any $k \not\equiv 0 \pmod{N}$.
3. Show that $VV^* = NI$, where V^* denotes the conjugate transpose of V.
4. Show that when $\boldsymbol{\beta}$ is the discrete Fourier transform of $\boldsymbol{\alpha}$, then

$$\boldsymbol{\beta} = \frac{1}{\sqrt{N}} V \boldsymbol{\alpha}.$$

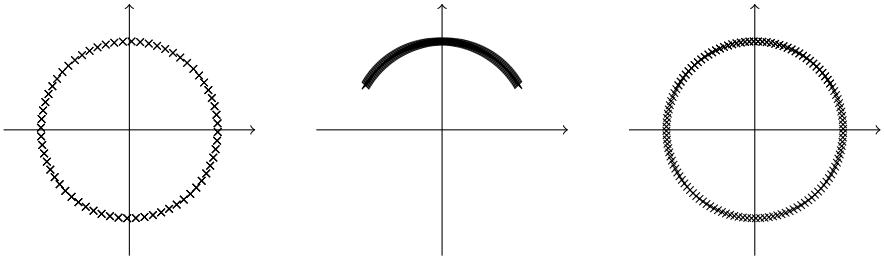

Figure 1: Plot of the terms in the final sum in (3) for $N = 2048$, $r = 12$ and $t_0 = 7$, and for $k = 1192, 1195, 1196$.

Exercise 2. Let U be an $N \times N$ unitary matrix. Show that for any vector $\boldsymbol{\alpha} \in \mathbb{C}^N$,

$$\|U\boldsymbol{\alpha}\| = \|\boldsymbol{\alpha}\|.$$

We want to study the Fourier coefficients of a very special complex vector. Let t_0 be an integer such that $0 \leq t_0 < r$, and let m be minimal such that $mr + t_0 \geq N$. Let $\boldsymbol{\alpha} \in \mathbb{C}^N$ be given by

$$\alpha_k = \begin{cases} \frac{1}{\sqrt{m}} & k = t_0 + jr, \text{ and} \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Note that $\|\boldsymbol{\alpha}\| = 1$ and that $1 - mr/N = 1 - (N - t_0)/N \leq r/N \leq 1/r$.

Applying the discrete Fourier transform gives us

$$\begin{aligned} \beta_k &= \frac{1}{\sqrt{N}} \sum_{j=0}^{m-1} \frac{1}{\sqrt{m}} \exp\left(2\pi i \frac{k(t_0 + jr)}{N}\right) \\ &= \frac{1}{\sqrt{N}} \frac{1}{\sqrt{m}} \exp\left(2\pi i \frac{kt_0}{N}\right) \sum_{j=0}^{m-1} \exp\left(2\pi i \frac{kjr}{N}\right). \end{aligned} \tag{3}$$

Example 3. Consider the vector $\boldsymbol{\alpha}$ from (2) for $N = 2048$, $r = 12$ and $t_0 = 7$. The terms of the sum in the expression for β_k from (3) has been plotted in the complex plane for $k = 1192, 1195, 1196$ in Figure 1.

Since the Fourier coefficient β_k is the sum of all the plotted complex numbers, it is easy to imagine that the sum of these complex numbers will tend to cancel out in the left and right cases, while for the middle case with $k = 1195$ there will be no cancellation. In other words, the amplitude of β_{1192} and β_{1196} will be small, while β_{1195} will be large.

We shall only care about the absolute value, which means that we shall only care about the sum factor. Furthermore, we shall only care about those k that are close to a multiple of the fraction N/r, that is those k satisfying (1). For such a k, our sum factor

becomes

$$\begin{aligned}\sum_{j=0}^{m-1}\exp\left(2\pi i\frac{kjr}{N}\right) &= \sum_{j=0}^{m-1}\exp\left(2\pi i\frac{jr}{N}\left(l\frac{N}{r}+\delta_k\right)\right)\\ &= \sum_{j=0}^{m-1}\exp(2\pi i\underbrace{jl}_{\in\mathbb{Z}})\exp\left(2\pi i\frac{jr}{N}\delta_k\right)\\ &= \sum_{j=0}^{m-1}\exp\left(2\pi i\frac{jr}{N}\delta_k\right) = \sum_{j=0}^{m-1}\left(\exp\left(2\pi i\frac{r}{N}\delta_k\right)\right)^j\\ &= \frac{1-\exp\left(2\pi i\frac{r}{N}\delta_k m\right)}{1-\exp\left(2\pi i\frac{r}{N}\delta_k\right)} = S.\end{aligned}$$

Exercise 3. Let $z\in\mathbb{R}$.

1. Prove that $|1-\exp(2\pi iz)|^2 = 4\sin^2(\pi z)$.
2. If $0\le|z|\le\pi/2$, prove that
$$\frac{\sin z}{z}\ge\frac{2}{\pi}.$$

Recall that when z is a very small real number, $|\sin z|\le|z|$, while for $|z|<\pi/2$ we know that $|\sin z|$ is increasing. Since N is large relative to r, r/N will be small. Furthermore, $(m-1)r+t_0\le N<mr+t_0$, which means that mr/N is slightly smaller than 1. Using Exercise 3 we get

$$|S|^2 = \frac{\sin^2\left(\pi\frac{r}{N}\delta_k m\right)}{\sin^2\left(\pi\frac{r}{N}\delta_k\right)} \ge \frac{\sin^2\left(\pi\delta_k\right)}{\sin^2\left(\pi\frac{r}{N}\delta_k\right)} \ge \frac{\sin^2(\pi\delta_k)}{(\pi\delta_k r/N)^2} = \frac{N^2}{r^2}\left(\frac{\sin\pi\delta_k}{\pi\delta_k}\right)^2 \ge \frac{N^2}{r^2}\left(\frac{2}{\pi}\right)^2.$$

This means that for the indexes k satisfying (1), the discrete Fourier coefficient satisfies

$$|\beta_k|^2 \ge \frac{1}{N}\frac{1}{m}\frac{N^2}{r^2}\frac{4}{\pi^2} = \frac{1}{mr}\frac{N}{r}\frac{4}{\pi^2} \ge \frac{1}{mr}\frac{mr}{N}\frac{N}{r}\frac{4}{\pi^2} = \frac{1}{r}\frac{4}{\pi^2}.$$

Next, we want to sum these values for all k satisfying (1), and since there are at least $r-1$ such k, we get an approximate lower bound

$$(r-1)\frac{1}{r}\frac{4}{\pi^2}\approx\frac{4}{\pi^2}. \tag{4}$$

This result essentially says that when we sum the squared absolute values of all the Fourier coefficients, those corresponding to values of k satisfying (1) constitute a relatively large fraction of the total sum.

Example 4. Figure 2 shows a plot of the amplitude of the discrete Fourier transform of a function of the form (2). The locations of the amplitude peaks for the discrete fourier transform are given in Table 1.

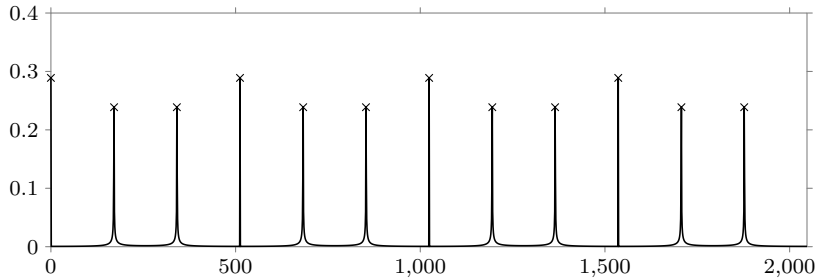

Figure 2: The amplitude of the discrete Fourier transform of a function of the form (2) with period $r = 12$ and $N = 2048$.

Table 1: The locations of the maximal amplitudes from Figure 2, with reference to (1).

k	$\lfloor k/(N/r) \rceil$	$\lvert\delta\rvert$	rat. approx.
0	0.0	0.0	-
171	1.0	0.002	1/12
341	2.0	0.002	1/6
512	3.0	0.0	1/4
683	4.0	0.002	1/3
853	5.0	0.002	5/12
1024	6.0	0.0	1/2
1195	7.0	0.002	7/12
1365	8.0	0.002	2/3
1536	9.0	0.0	3/4
1707	10.0	0.002	5/6
1877	11.0	0.002	11/12

The cumulative probability of these values is approximately 0.79. We see that for only a few of these values will the rational approximation of k/N give us the period $r = 12$, but for all except $k = 0$, it is within a small multiple. (The explanation is that the multiple l and the period r have common factors, which cancel.)

Note, however, that we cannot compute the discrete Fourier transform, since N is too large, and even if we could, we could not easily find out which values of k are included in the sum. So the discrete Fourier transform does not immediately help us find a k satisfying (1).

However, if we could arrange some probabilistic process in such a way that it will sample an integer from $\{0, 1, \ldots, N-1\}$ with probability according to the squared absolute value of the Fourier coefficients, then the above results says that it is reasonably likely that we will sample a k satisfying (1).

3 Quantum Computation

A *quantum bit* (often called *qubit*) is a physical system with two states. We interpret the two states as 0 and 1. A quantum system can be in a *superposition* of its two states, which is described by two complex numbers α and β such that $|\alpha|^2 + |\beta|^2 = 1$. The two values $|\alpha|^2$ and $|\beta|^2$ are the probabilities of measuring 0 and 1, respectively. We write the state as

$$\alpha|0\rangle + \beta|1\rangle.$$

A physical system can have more than two states, of course. If we number the states from 0 and to $N-1$, a superposition of states is described a complex vector $\boldsymbol{\alpha}$ with norm 1 and we write the state as

$$\sum_{j=0}^{N-1} \alpha_j |j\rangle.$$

Two physical systems (with N and N' states, respectively) can be *entangled*, which means that the outcome of measuring one of the systems will influence the outcome of measuring the other system. We write the system state as

$$\sum_{k=0}^{N-1} \sum_{j=0}^{N'-1} \alpha_{kj} |k\rangle|j\rangle.$$

If we measure the second system, we will get the outcome j_0 with probability

$$\sum_{k=0}^{N-1} |\alpha_{kj_0}|^2$$

and if j_0 was the outcome, we will get the quantum state

$$\sum_{k=0}^{N-1} \frac{\alpha_{kj_0}}{\sqrt{\sum_{k=0}^{N-1} |\alpha_{kj_0}|^2}} |k\rangle|j_0\rangle.$$

When later measuring the first system, the probability of outcome k_0 will be

$$\frac{|\alpha_{k_0 j_0}|^2}{\sum_{k=0}^{N-1} |\alpha_{kj_0}|^2}.$$

Multiple quantum bits can be entangled, which means that the outcome of measuring one bit may influence the outcome of measuring the other bits. In this case, a system of l qubits will have $N = 2^l$ states. A system of multiple entangled qubits is called a *quantum register*.

3.1 Computing on Quantum Registers

Computations on quantum bits can be done using *quantum gates*. Any such computation can be described by an invertible linear map that preserves the norm, that is, a unitary matrix U. That is, the transition

$$\sum_{j=0}^{N-1} \alpha_j |j\rangle \qquad \stackrel{U}{\longmapsto} \qquad \sum_{j=0}^{N-1} \beta_j |j\rangle$$

is given by by $\boldsymbol{\beta} = U\boldsymbol{\alpha}$.

Example 5. Consider a system with two states. We want to compute the function $f(z) = 1 - z$, the negation function, on this state. The unitary matrix

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

will compute exactly this function, mapping the quantum state (α_0, α_1) to (α_1, α_0).

Example 6. Consider a system of three qubits. We want to compute the function $f(z_1, z_2, z_3) = (z_1, z_2, z_1 z_2 + z_3 \bmod 2)$, that is, negate z_3 if and only if both z_1 and z_2 are set. If we consider the natural basis $|000\rangle, |001\rangle, \ldots |110\rangle, |111\rangle$, the matrix

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

will compute exactly this function, swapping the quantum states 110 and 111, while leaving all other states unchanged. (It can be shown that this gate, the Toffoli gate, is universal for classical computations. Which means that this gate is sufficient in order to do any classical computation.)

Example 7. Consider a system with two states. The gate given by the unitary matrix

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

defines an operation that maps the state $|0\rangle$ to a superposition of $|0\rangle$ and $|1\rangle$, where each outcome will be equally likely. (Such gates can be used to create superpositions.)

Exercise 4. Show that the above three matrices are unitary.

Since unitary matrices are invertible, computations on quantum bits must also be invertible. But with a bit of care and effort any classical computation can be repeated on quantum bits.

In particular, we can compute the function f discussed above, even though it is not invertible. Typically, we will have two quantum registers and map $|k\rangle|j\rangle$ to $|k\rangle|j+f(k)\rangle$, which is an invertible computation, at least if we use modular addition.

If we compute on a superposition, our result will be a superposition of the function values. In the example above

$$\sum_{k=0}^{N-1}\sum_{j=0}^{N'-1}\alpha_{kj}|k\rangle|j\rangle \qquad \longmapsto \qquad \sum_{k=0}^{N-1}\sum_{j=0}^{N'-1}\alpha_{kj}|k\rangle|j+f(k)\rangle.$$

3.2 Quantum Fourier Transform

Example 6 and the discussion in the previous section show that any classical computation can be replicated on a quantum computer. However, Example 7 shows that there are operations on quantum bits that do not correspond directly to classical computations. One such operation that will be important for us is the quantum Fourier transform.

The quantum Fourier transform on a system with N states is defined by

$$\sum_{j=0}^{N-1}\alpha_j|j\rangle \qquad \overset{QFT}{\longmapsto} \qquad \sum_{j=0}^{N-1}\beta_j|j\rangle$$

where

$$\beta_k = \frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}\alpha_j \exp(2\pi ijk/N).$$

This is of course exactly the same as our normalized discrete Fourier transform.

Unlike the discrete Fourier transform, it turns out that a quantum Fourier transform of a quantum system made up of l quantum bits (with $N = 2^l$) can be implemented using about l^2 elementary quantum gates. The reason for this is that the classical discrete Fourier transform must be computed using operations that involve all the complex numbers $\alpha_0, \alpha_1, \ldots$. The quantum Fourier transform is computed by manupulating the l quantum bits, which only indirectly manipulates the quantum probabilities $\alpha_0, \alpha_1, \ldots$.

4 Factoring Using a Quantum Computer

Our goal is now to set up a quantum system with an amplitude distribution like (2). We can then apply the quantum Fourier transform, which will result in a quantum system where we know that the likelihood of measuring a k satisfying (1) is at least $4/\pi^2$. And if we measure such a k, computing a rational approximation will give us the period of the function f, which will allow us to factor n.

We begin with a two entangled quantum registers, one with $\log_2 N$ qubits and the other with $\lceil \log_2 n \rceil$ qubits. The first register should contain a superposition of the integers $0, 1, 2, \ldots, N-1$, all with the same amplitude, while the second register should contain 0. We have the state

$$\sum_{j=0}^{N-1} \frac{1}{\sqrt{N}} |j\rangle|0\rangle.$$

Using a quantum circuit, we compute f on the first register, storing the result in the second register. Since any classical computation can be done on a quantum computer, computing f is easy. Our two quantum registers now contain a superposition of $(k, f(k))$ for $k = 0, 1, \ldots, N-1$, all being equally likely, and we have the new state

$$\sum_{j=0}^{N-1} \frac{1}{\sqrt{N}} |j\rangle|f(j)\rangle.$$

Next, we measure the second register. Suppose we measure the value s. Since we have not yet measured the first register, we do not know what it contains, but since it was entangled with the second register, every value we could possible measure must be consistent with s. Suppose t_0 is the smallest non-negative integer such that $f(t_0) = s$. Then the only values we can measure in the first register are the integers $t_0 + jr$.

Since we had a uniform probability before we measured the second register, we must have a uniform probability after measuring. We now have m possible states, so we have the state

$$\sum_{j=0}^{m-1} \frac{1}{\sqrt{m}} |t_0 + jr\rangle|s\rangle.$$

In other words, if we ignore the second register which is constant, our first register now has exactly the amplitudes given by (2).

We then use a second quantum circuit to compute the Quantum Fourier Transform on the first register. And then we measure the first register. As discussed, we will with significant probability measure a k satisfying (1), which will allow us to factor n.

Example 8. Consider $n = 35$. We choose $g = 2$ and want to compute (something close to) the period of $a \mapsto 2^a \bmod 35$. To this end, we have a quantum computer with two registers, the first of which has 2048 qubits.

We initialise the quantum registers with a superposition of all possible values in the first register and zero in the second register. We apply the exponentiation function and measure the function value in the second register. Suppose we get the answer 23, which implies that the amplitude of the collapsed first register corresponds to that from Example 3.

We then compute the quantum Fourier transform of the first register, and the end result is the amplitude shown in Example 4.

Then we measure the first register. Suppose we get the answer 1195. Rational approximation as in Example 2 gives us a period of 12.

We now compute

$$2^3 \equiv 8 \pmod{35}$$

and we immediately get that $\gcd(8-1, 35) = 7$.

(Even if we had measured 1365 and gotten the wrong period from the rational approximation in Example 1, we would still have factored 35. But that would have been just lucky.)

As of today, quantum computers with a few qubits capable of running this algorithm have been demonstrated. It is unclear if we will ever be able to build larger quantum computers, but for applications requiring long-term security, quantum computers will at least constitute a source of uncertainty for a long time.