Quantum Fourier transform and Shor's algorithm

August 6, 2018

1 The discrete Fourier transform

Let us start with a short summary of the definition and some of the properties of the classical discrete Fourier transform. Given an integer N, the discrete Fourier transform is usually defined to be a mapping from the space of complex sequences with N elements to itself. To simplify our notation a bit, we will denote the i-th element of a sequence of N complex numbers as $x[i]$ instead of x_i and let the index i start at zero, so that such a sequence is given by the N complex numbers $x[0], x[1], \ldots, x[N-1]$. We combine these numbers into a vector $x \in \mathbb{C}^N$.

Given such a sequence x, the discrete Fourier transform of x is defined to be the sequence X with elements

$$X[k] = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{-\frac{2\pi i}{N} jk} = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \eta^{-jk}$$

where we denote by $\eta = e^{\frac{2\pi i}{N}}$ the standard N-th root of unity. Mapping vectors to vectors, we can think of the Fourier transform as a mapping

$$\mathcal{F} \colon \mathbb{C}^N \to \mathbb{C}^N$$

which is clearly linear. We will now show that this mapping is in fact unitary. For this purpose, consider the vectors u_k which are defined as follows.

$$u_k[i] = \frac{1}{\sqrt{N}} \eta^{ik}$$

Clearly, the complex conjugates of these vectors are the image of the standard basis under the discrete Fourier transform. In fact, if we consider the vector e_k defined by $e_k[j] = \delta_{jk}$, then the j-th component of the Fourier transform of this vector is

$$\begin{aligned} \mathcal{F}(e_k)[j] &= \frac{1}{\sqrt{N}} \sum_{s=0}^{N-1} e_k[s] \eta^{-sj} \\ &= \frac{1}{\sqrt{N}} \sum_{s=0}^{N-1} \delta_{sk} \eta^{-sj} \\ &= \frac{1}{\sqrt{N}} \eta^{jk} = u_k[j]^* \end{aligned}$$

Let us now show that the vectors u_k form an orthonormal basis of $\mathbb{C}^N$. For given k and k', we can compute the hermitian product

$$\begin{aligned}\langle u_k, u_{k'}\rangle &= \sum_{s=0}^{N-1} u_k[s]u_{k'}^*[s] \\ &= \frac{1}{N}\sum_{s=0}^{N-1} \eta^{-sk}\eta^{sk'} = \frac{1}{N}\sum_{s=0}^{N-1}\eta^{s(k'-k)}\end{aligned}$$

Now this is a geometric series with $q = \eta^{k-k'}$. If $k = k'$, then $q = 1$ and the series sums up to N, so that the hermitian product is one. If, however, $k \neq k'$, then the sum gives

$$\frac{1-q^N}{1-q}$$

according to the summation formula for a geometric series. However, as η is an N-th root of unity, $q^N = 1$ and the series sums up to zero. This proves our claim that the u_k form an orthonormal basis and at the same time demonstrates that the mapping given by the Fourier transform $\mathcal{F}$ is a unitary matrix. Also note that the Fourier coefficient $X[k]$ can be written as

$$X[k] = \sum_s x[s]u_k^*[s] = \langle x, u_k\rangle$$

Using the properties of the vectors u_k, it is also easy to find a formula for the inverse of the Fourier transform. In fact, given a vector x, we can expand x in the basis u_k and obtain

$$\begin{aligned}x &= \sum_k \langle x, u_k\rangle u_k \\ &= \sum_k X[k]u_k \\ &= \sum_k X[k]\sum_s u_k[s]e_s = \frac{1}{\sqrt{N}}\sum_k(\sum_s X[s]\eta^{sk})e_k\end{aligned}$$

where we have renamed the indices in the last line. From this result, we can immediately read off that the formula for the cofficients $x[k]$ in terms of the $X[k]$, i.e. the formula for the inverse Fourier transform, is given by

$$x[k] = \frac{1}{\sqrt{N}}\sum_s X[s]\eta^{sk}$$

This is very similar to the formula for the Fourier transform itself, up to the sign in the exponent. It is worth mentioning that some authors use this as the definition of the Fourier transform, not its inverse, but it should be clear that this is merely a convention.

2 Quantum Fourier transform

Let us now jump right back into the world of quantum computing. Imagine that we have a quantum computer with n qubits. The states of this quantum computer are then described by rays in a Hilbert space with $N = 2^n$ dimensions, namely the n-fold tensor product of the one-qubit Hilbert space. With respect to the usual standard basis labeled by the vectors $|x\rangle$, with x ranging from 0 to N-1, we can then consider any vector as a sequence, using the identification

$$x = \sum_k x[k]|x\rangle$$

To this sequence, we can apply the Fourier transform. This will give us a unitary transformation $\mathcal{F}$, described by

$$\mathcal{F}(\sum_k x[k]|k\rangle) = \sum_k X[k]|k\rangle = \frac{1}{\sqrt{N}} \sum_{s,k} x[s]\eta^{-sk}|k\rangle$$

As any unitary transformation, this transformation can be realized by a quantum circuit. In fact, one can show (see [2] and the references therein) that this can be done with a number of quantum gates that scales as $\mathrm{O}(n^2)$.

3 Period finding

Let us now see how the quantum Fourier transform can be applied to find the period of a function - a discovery which is at the heart of Shor's algorithm.

Suppose we are given a large number M and an element x in $\mathbb{Z}_M$, or, in other words, an integer $x \in \{0, 1, \ldots, M-1\}$. If x and M are co-prime, we know that x is a unit modulo m, thus a certain power of x will be one modulo M. The purpose of the period-finding algorithm is to find the smallest number r such that

$$x^r \equiv 1 \mod M$$

We call this number the *period* of x. To find the value of the period, we first find a power of two, i.e. some $N = 2^n$, such that $M^2 \leq N \leq 2M^2$ (in other words, n is the smallest number such that M^2 can be represented as an n-bit number). Next, we prepare a register of n+n qubits in the usual superposition

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

which can be done by starting with the fiducial state that is zero in all qubits and then applying the Hadamard-Walsh operator.

In the next step called *modular exponentation*, we apply the operator U_f to this superposition, where f is the function $f(a) = x^a$. We know on general grounds that this can be done, and Shor demonstrates in [2] that this can be done with $\mathrm{O}(n^3)$ gates and O(n) qubits and refers to further results improving this to $\mathrm{O}(n^2 \log n \log\log n)$ gates at the cost of increasing the required space slightly.

After applying this step to our quantum state, we end up with the state $|\psi\rangle$ given by the expression

$$|\psi\rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} |k\rangle |x^k \mod M\rangle$$

Now there are different ways to proceed. As in the expositions given in [3] and [4], we add a measurement step to the algorithm that simplifies the arguments a bit (but is actually not needed and not done in [2])). In fact, at this point of the algorithm, we perform a measurement of the second register. Let us call the measured value y. As measuring amounts to the projection onto the eigenstates, only those components of the state $|\psi\rangle$ survice the measurement for which, in the expression above, $x^k \equiv y \mod M$. Let us denote the smallest such number by x_0. Then, after the measurement, our state will be

$$\frac{1}{\sqrt{A}} \sum_{j=0}^{A-1} |x_0 + jr\rangle |y\rangle$$

where A is the largest number such that $x_0 + (A-1)r < N$. In particular, it follows that from $x_0 \geq 0$ and $x_0 \leq r$ that $(A-1)r < N$ and $(A+1)r > N$. At this point, we ignore the second register and focus on the first one and apply the inverse of a quantum Fourier transform to it. The resulting state will be

$$\frac{1}{\sqrt{AN}} \sum_{j=0}^{A-1} \sum_{s=0}^{N-1} \eta^{(x_0+jr)s} |s\rangle$$

Let us now reorganize this term a bit. First, we can collect all terms that contribute to the amplitude for a given value of s and find that the coefficient of $|s\rangle$ is given by

$$\frac{1}{\sqrt{AN}} \eta^{x_0 s} \left[\sum_{j=0}^{A-1} \eta^{jrs} \right]$$

Thus, if we now perform a measurement of the first register, we obtain the value s with the probability

$$P(s) = \frac{1}{AN} \left| \sum_{j=0}^{A-1} \eta^{jrs} \right|^2$$

Now, the sum can again be simplified by recognizing it as a geometric series with coefficient $q = \eta^{rs}$. If $q = 1$, which will happen exactly if s is a multiple of $\frac{N}{r}$, we find that the probability is $\frac{A}{N}$. We now claim that in general, the probability distribution given by $P(s)$ is peaked at those values of s for which $\frac{sr}{N}$ is close to an integer.

We have already seen that in case this number is exactly an integer, which happens if and only if $q = 1$, the probability is $\frac{A}{N}$. For the general case, we can therefore assume that q is not one and sum the geometric series to obtain

$$P(s) = \frac{A}{N} \left| \frac{1}{A} \frac{1 - q^A}{1 - q} \right|^2$$

We want to show that this distribution has a peak in the range of those s where $s \cdot r$ is close to a multiple of N. If we denote, for a given integer x, by $\{x\}_N$ the residual modulo N in the range between $-\frac{N}{2}$ and $\frac{N}{2}$, we can express this condition more formal as

$$-\frac{r}{2} \leq \{sr\}_N \leq \frac{r}{2}$$

How many values of s do exist that fulfill this condition? We claim that there are precisely r such values. In fact, let us follow the argument in [4] and draw the grids of integers spanned by the numbers r and N, as done in figure 1.

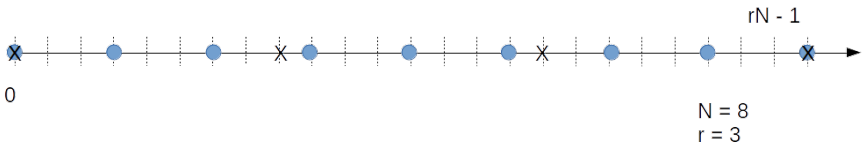

Figure 1

Here the blue bubbles mark multiples or r, i.e. points of the form sr, and the crosses mark multiples of N, i.e. points of the form kN. Now each such multiple of N is located in a stripe between two consecutive blue bubbles, and one of the bubbles will be closer than $\frac{r}{2}$ to it. Thus we can select, for each kN, a number rs such that $|rs - kN| \leq \frac{r}{2}$. Now we let k run over the r values from 0 to r - 1. This will select r corresponding values of s, and these are exactly the values for which our condition holds (if you wanted to formally prove that no s is selected twice using this procedure, you would once more need that $r \ll N$).

We will now estimate the probability P(s) in this range, using the argument given in [4]. The argument is based on the elementary geometrical observations - for which we will give proofs in the appendix - that

$$|1 - e^{i\Phi}| \geq \frac{2|\Phi|}{\pi}$$

for $\Phi \in [-\pi, \pi]$ and

$$|1 - e^{i\Phi}| \leq |\Phi|$$

for all values of Φ. Let us first try to use this to derive a lower bound for $|1 - q^A|$. We can write this as

$$|1 - q^A| = |1 - e^{2\pi i \frac{rsA}{N}}|$$

and try to apply the above estimate with $\Phi = 2\pi \frac{\{rs\}_N A}{N}$. If $\{rs\}_N$ is in the range between $-\frac{r}{2}$ and $\frac{r}{2}$, then the absolute value of this angle is at most $\pi \frac{rA}{N}$. We know that $(A-1)r \leq N$, but not necessarily that $Ar \leq N$, so we cannot apply the above bound directly.

However, there is a little trick. Let $\Phi = 2\pi \frac{\{rs\}_N}{N}$. We can then apply the estimates for the numerator and the denominator to see that

$$\left|\frac{1-q^{A-1}}{1-q}\right| \geq \frac{2(A-1)|\Phi|}{\pi}\frac{1}{|\Phi|} = \frac{2(A-1)}{\pi}$$

Using this and the reversed triangle inequality $|x+y| \geq |x| - |y|$ we can now write

$$\begin{aligned}\left|\frac{1-q^{A}}{1-q}\right| &= \left|\frac{1-q^{A-1}}{1-q} + q^{A-1}\right| \\ &\geq \left|\frac{1-q^{A-1}}{1-q}\right| - \left|q^{A-1}\right| \\ &= \left|\frac{1-q^{A-1}}{1-q}\right| - 1 \\ &\geq \frac{2(A-1)}{\pi} - 1 = \frac{2A}{\pi} - (1+\frac{\pi}{2})\end{aligned}$$

Squaring this, we thus find that for the values of s in the specified range, we have

$$P(s) \geq \frac{4A}{\pi^2 N} - \frac{4(1+\frac{\pi}{2})}{\pi N}$$

Now we know that $rA \approx N$, so the first term is approximately $\frac{4}{\pi^2 r}$. Further, we know that $r \leq M \ll N$ by our choice of N, so the second term is negligible. We have therefore found an approximate lower bound

$$P(s) \geq (\frac{4}{\pi^2})\frac{1}{r} - \epsilon$$

for all values of s for which $\{rs\}_N \in [-\frac{r}{2}, \frac{r}{2}]$. To obtain an estimate for s to be in that range, we need to use the fact shown above that there are precisely r values of s so that the residual falls into this range. By summing up the probabilities, we now see that with probability bounded below by $\frac{4}{\pi^2} \approx 0.4$, the measured value s will be such that $\{rs\}_N$ is in that range. This condition simply means that with probability at least 0.4, the measured value of s will be within an interval of length $\frac{1}{2}$ around an integer times $\frac{N}{r}$.

Let us see why this helps. What we have shown is that with reasonable probability, our measurement will give us a value s such that there exists an integer d with

$$-\frac{r}{2} \leq rs - dN \leq \frac{r}{2}$$

Dividing this by rN gives

$$\left|\frac{s}{N} - \frac{d}{r}\right| \leq \frac{1}{2N}$$

We know the the fraction $\frac{s}{N}$, as we know s (from our measurement) and N (which we have chosen). We do not know d and r yet. However, this equation tells us that it is located in a strip of width $\frac{1}{N}$ around the known $\frac{s}{N}$.

Now two rational numbers $\frac{a}{b}$ and $\frac{c}{d}$ whose denominator is bounded by M need to have a distance of at most $\frac{1}{M^2}$ (which is an immediate consequence of the familiar formula)

$$\frac{a}{b} - cd = \frac{ad - bc}{bd}$$

As the order r we are looking for can be at least M, and $M^2 \leq N$, this shows that there is at most one fraction of the form $\frac{d}{r}$ in the region of size $\frac{1}{N}$ around $\frac{s}{N}$. As Shor points out in [2], we can therefore obtain the unknown fraction $\frac{d}{r}$ from the known fraction $\frac{s}{N}$ by looking for the nearest fraction that has denominator at most M, and there is an algorithm available for this purpose that uses continued fraction expansions. The situation is visualized in diagram 2.

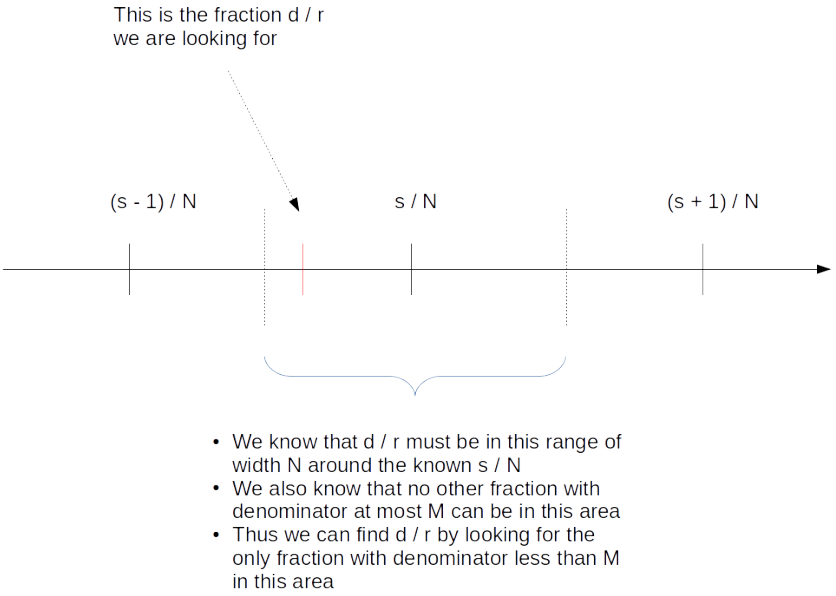

Figure 2

Once we have that fraction, we can be lucky and d and r are relatively prime. If this happens, we can extract r and are done. If not, we can run our quantum algorithm once more. Shor argues in [2] that the overall success probability for the algorithm is bounded below so that a reasonably small number of iterations will suceed with a very high probability.

A Estimates for points on the circle

Here we will take a closer look at the estimates for $1 - e^{i\Phi}$ used before. First, let us try to find a bound from above. The absolute value $|1 - e^{i\Phi}|$ is the length of the straight line connecting the point 1 and the point $e^{i\Phi}$, as shown in diagram 3.

Intuitively, it is clear that this path is shorter than going around, i.e. that

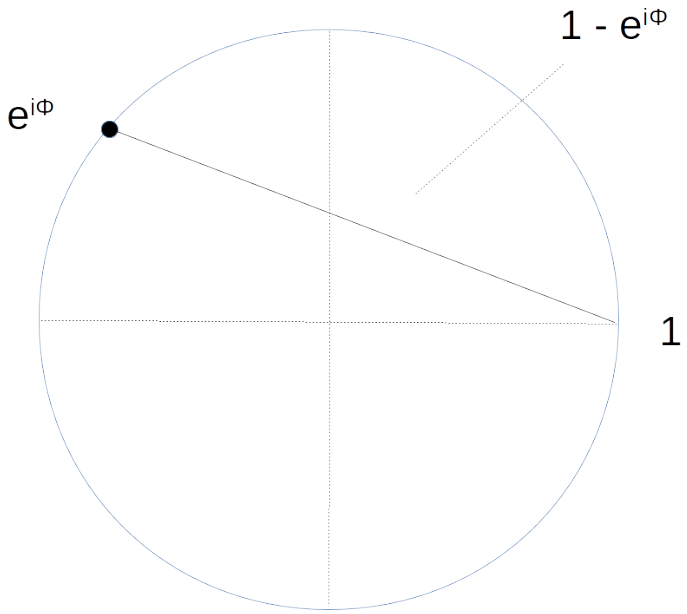

Figure 3

$$|1 - e^{\Phi}| \leq |\Phi|$$

for all angles Φ. Formally, this can be seen as follows. First, using the Euler identities, a short calculation shows that for all angles,

$$|1 - e^{i\Phi}|^2 = 2(1 - \cos\Phi)$$

Now consider the two functions f and g given by $f(\Phi) = 2(1 - \cos\Phi)$ and $g(\Phi) = \Phi^2$. At the origin, both functions are zero. Their derivatives are both zero at the origin as well, and for their second derivatives, we have $f'' \leq g''$ for all values. As furthermore both functions are symmetric around the origin, we can conclude that $f \leq g$ everyhwere. This proves our inequality.

Next, let us try to apply a similar argument to prove our second inequality. First, we rewrite our expression for $|1 - e^{i\Phi}|$ a bit. By the additional theorem for the cosine, we have

$$1 - \cos\Phi = 2\sin^2\frac{\Phi}{2}$$

so that we obtain the simple expression

$$|1 - e^{i\Phi}| = 2\sin\frac{\Phi}{2}$$

whenever the right hand side is non-negative. To use this to derive a lower bound, we therefore need a lower bound for the sine. So let us now consider the function F given by

$$F(x) = \sin x - \frac{2x}{\pi}$$

Clearly, this function has zeros at $x = 0$ and $x = \frac{\pi}{2}$. We claim that it does not have any zeros between these two points. There are several ways to see this. One possible argument is that the function represents the difference between the function cos and the straight line from the origin to the value of cos at $\frac{\pi}{2}$. As the cosine is concave in this region, this line is always below the graph of the cosine and does not intersect it in any other points within this interval. Thus we have shown that for $x \in [0, \frac{\pi}{2}]$, the inequality

$$\sin x \geq \frac{2x}{\pi}$$

holds. Consequently, we immediately find that for $\Phi \in [0, \pi]$, we have the inequality

$$|1 - e^{i\Phi}| = 2 \sin \frac{\Phi}{2} \geq 2\frac{\Phi}{\pi}$$

Using the symmetry of the right hand side with respect to reflexion at the origin, we therefore finally obtain that for all $\Phi \in [-\pi, \pi]$, we have the inequality

$$|1 - e^{i\Phi}| \geq 2\frac{|\Phi|}{\pi}$$

References

[1] M.A. Nielsen, I.L. Chaung, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2010

[2] P. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, SIAM J.Sci.Statist.Comput. Vol. 26 Issue 5 (1997), pp 1484–1509, available as arXiv:quant-ph/9508027v2

[3] R. Jozsa, *Quantum Algorithms and the Fourier Transform*, arXiv:quant-ph/9707033

[4] J. Preskill, *Quantum computation*, Lecture notes (PH229), chapter 6, available online at `http://www.theory.caltech.edu/people/preskill/ph229/`