Quantum algorithms (CO 781/CS 867/QIC 823, Winter 2013)
Andrew Childs, University of Waterloo

LECTURE 4: Period finding from $\mathbb{Z}$ to $\mathbb{R}$

In this lecture, we will explore quantum algorithms for determining the period of a function. Shor's factoring algorithm is based on a solution of the period-finding problem for a function over the integers. More recently, Hallgren considered the problem of solving a quadratic Diophantine equation known as *Pell's equation*, together with related problems involving number fields. Hallgren gave efficient quantum algorithms for these problems by generalizing period finding over the integers to period finding over the real numbers.

Factoring and order finding

Shor's factoring algorithm is based on a reduction of factoring to order finding (observed by Miller in the 1970s). This reduction is typically covered in a first course on quantum computing, so we will not discuss the details here.

In the order finding problem for a group G, we are given an element $g \in G$ and our goal is to find the order of g, the smallest $r \in \mathbb{N}$ such that $g^r = 1$. (Factoring L reduces to order finding in $G = \mathbb{Z}_L$.) One way to approach this problem is to consider the function $f\colon \mathbb{Z} \to G$ defined by $f(x) = g^x$. This function is periodic with period r, and there is an efficient quantum algorithm to find this period, which we review below.

Pell's equation

Given a squarefree integer d (i.e., an integer not divisible by any perfect square), the Diophantine equation

$$x^2 - dy^2 = 1 \tag{1}$$

is known as *Pell's equation*. This equation was already studied in ancient India and Greece, and is closely related to concepts in algebraic number theory.

The left-hand side of Pell's equation can be factored as

$$x^2 - dy^2 = (x + y\sqrt{d})(x - y\sqrt{d}). \tag{2}$$

Note that a solution of the equation $(x, y) \in \mathbb{Z}^2$ can be encoded uniquely as the real number $x+y\sqrt{d}$: since $\sqrt{d}$ is irrational, $x + y\sqrt{d} = w + z\sqrt{d}$ if and only if $(x, y) = (w, z)$. (Proof: $\frac{x-w}{z-y} = \sqrt{d}$.) Thus we can also refer to the number $x + y\sqrt{d}$ as a solution of Pell's equation.

There is clearly no loss of generality in restricting our attention to *positive* solutions of the equation, namely those for which $x > 0$ and $y > 0$. It is straightforward to show that if $x_1 + y_1\sqrt{d}$ is a positive solution, then $(x_1 + y_1\sqrt{d})^n$ is also a positive solution for any $n \in \mathbb{N}$. In fact, one can show that *all* positive solutions are obtained in this way, where $x_1 + y_1\sqrt{d}$ is the *fundamental solution*, the smallest positive solution of the equation. Thus, even though Pell's equation has an infinite number of solutions, we can in a sense find them all by finding the fundamental solution.

Unfortunately, it is not feasible to find the fundamental solution explicitly. The solutions can be very large—the size of $x_1 + y_1\sqrt{d}$ is only upper bounded by $2^{O(\sqrt{d}\log d)}$. Thus it is not even possible to *write down* the fundamental solution with $\mathrm{poly}(\log d)$ bits.

To get around this difficulty, we define the *regulator* of the fundamental solution,

$$R := \ln(x_1 + y_1\sqrt{d}). \tag{3}$$

Since $R = O(\sqrt{d}\log d)$, we can write down $\lceil R \rceil$ using $O(\log d)$ bits. Now R is an irrational number, so determining only its integer part may seem unsatisfactory. But in fact, given the integer part of R, there is a classical algorithm to compute n digits of R in time $\mathrm{poly}(\log d, n)$. Thus it suffices to give an algorithm that finds the integer part of R in time $\mathrm{poly}(\log d)$. The best known classical algorithm for this problem takes time $2^{O(\sqrt{\log d \log\log d})}$ assuming the generalized Riemann hypothesis, or time $O(d^{1/4}\,\mathrm{poly}(\log d))$ with no such assumptions.

Hallgren's algorithm for solving Pell's equation is based on defining an efficiently computable periodic function whose period is the regulator. Defining this function would require us to introduce a substantial amount of algebraic number theory, so we omit the details here (for a partial account, see the lecture notes from 2011; for a more thorough treatment, see the review article by Jozsa). Instead, we will focus on the quantum part of the algorithm, which solves the period-finding problem.

Period finding over the integers

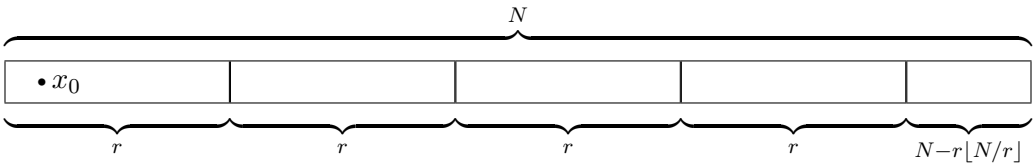
Recall that Shor's algorithm for factoring the number L works by finding the period of the function $f\colon \mathbb{Z} \to \mathbb{Z}_L$ defined by $f(x) = a^x \bmod L$ (where a is chosen at random). In other words, we are trying to find the smallest positive integer r such that $a^x \bmod L = a^{x+r} \bmod L$ for all $x \in \mathbb{Z}$. Note that since the period does not, in general, divide a known number N, we cannot simply reduce this task to period finding over $\mathbb{Z}_N$; rather, we should really think of it as period finding over $\mathbb{Z}$ (or, equivalently, the hidden subgroup problem over $\mathbb{Z}$).

Of course, we cannot hope to represent arbitrary integers on a computer with finitely many bits of memory. Instead, we will consider the function only on the inputs $\{0, 1, \ldots, N-1\}$ for some chosen N, and we will perform Fourier sampling over $\mathbb{Z}_N$. We will see that this procedure can work even when the function is not precisely periodic over $\mathbb{Z}_N$. Of course, this can only have a chance of working if the period is sufficiently small, since otherwise we could miss the period entirely. Later, we will see how to choose N if we are given an a priori upper bound of M on the period. If we don't initially have such a bound, we can simply start with $M = 2$ and repeatedly double M until it's large enough for period finding to work. The overhead incurred by this procedure is only $\mathrm{poly}(\log r)$.

Given a value of N, we prepare a uniform superposition over $\{0, 1, \ldots, N-1\}$ and compute the function in another register, giving

$$\frac{1}{\sqrt{N}} \sum_{x \in \{0,\ldots,N-1\}} |x\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{x \in \{0,\ldots,N-1\}} |x, f(x)\rangle. \tag{4}$$

Next we measure the second register, leaving the first register in a uniform superposition over those values consistent with the measurement outcome. When f is periodic with minimum period r, we obtain a superposition over points separated by the period r. The number of such points, n, depends on where the first point, $x_0 \in \{0, 1, \ldots, r-1\}$, appears. When restricted to $\{0, 1, \ldots, N-1\}$, the function has $\lfloor N/r \rfloor$ full periods and $N - r\lfloor N/r \rfloor$ remaining points, as depicted below. Thus $n = \lfloor N/r \rfloor + 1$ if $x_0 < N - r\lfloor N/r \rfloor$ and $n = \lfloor N/r \rfloor$ otherwise.



Discarding the measurement outcome, we are left with the quantum state

$$\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} |x_0 + jr\rangle \tag{5}$$

where x_0 occurs nearly uniformly random (it appears with probability n/N) and is unknown. To obtain information about the period, we apply the Fourier transform over $\mathbb{Z}_N$, giving

$$\frac{1}{\sqrt{nN}} \sum_{j=0}^{n-1} \sum_{k \in \mathbb{Z}_N} \omega_N^{k(x_0+jr)} |k\rangle = \frac{1}{\sqrt{nN}} \sum_{k \in \mathbb{Z}_N} \omega_N^{kx_0} \sum_{j=0}^{n-1} \omega_N^{jkr} |k\rangle. \tag{6}$$

Now if we were lucky enough to choose a value of N for which $r \mid N$, then in fact $n = N/r$ regardless of the value of x_0, and the sum over j above is

$$\sum_{j=0}^{n-1} \omega_N^{jkr} = \sum_{j=0}^{n-1} \omega_n^{jk} \tag{7}$$

$$= n\delta_{k \bmod n, 0}. \tag{8}$$

In this especially simple case, the quantum state is

$$\frac{n}{\sqrt{nN}} \sum_{k \in \mathbb{Z}_N} \omega_N^{kx_0} \delta_{k \bmod n, 0} = \frac{1}{\sqrt{r}} \sum_{k \in n\mathbb{Z}_r} \omega_N^{kx_0} |k\rangle, \tag{9}$$

and measurement of k is guaranteed to give an integer multiple of $n = N/r$, with each of the r multiples occurring with probability $1/r$. But more generally, the sum over j in (6) is the geometric series

$$\sum_{j=0}^{n-1} \omega_N^{jkr} = \frac{\omega_N^{krn} - 1}{\omega_N^{kr} - 1} \tag{10}$$

$$= \omega_N^{(n-1)kr/2} \frac{\sin \frac{\pi krn}{N}}{\sin \frac{\pi kr}{N}}. \tag{11}$$

The probability of seeing a particular value k is given by the normalization factor $1/nN$ times the magnitude squared of this sum, namely

$$\Pr(k) = \frac{\sin^2 \frac{\pi krn}{N}}{nN \sin^2 \frac{\pi kr}{N}}. \tag{12}$$

From the case where $n = N/r$, we expect this distribution to be strongly peaked around values of k that are close to integer multiples of N/r. The probability of seeing $k = \lfloor jN/r \rceil = jN/r + \epsilon$ for some $j \in \mathbb{Z}$, where $\lfloor x \rceil$ denotes the nearest integer to x, is

$$\Pr(k = \lfloor jN/r \rceil) = \frac{\sin^2(\pi jn + \frac{\pi \epsilon rn}{N})}{nN \sin^2(\pi j + \frac{\pi \epsilon r}{N})} \tag{13}$$

$$= \frac{\sin^2 \frac{\pi \epsilon rn}{N}}{nN \sin^2 \frac{\pi \epsilon r}{N}}. \tag{14}$$

Now using the inequalities $4x^2/\pi^2 \leq \sin^2 x \leq x^2$ (where the lower bound holds for $|x| \leq \pi/2$, and can be applied since $|\epsilon| \leq 1/2$), we have

$$\Pr(k = \lfloor jN/r \rceil) \geq \frac{4(\frac{\epsilon r n}{N})^2}{nN(\frac{\pi \epsilon r}{N})^2} \tag{15}$$

$$= \frac{4n}{\pi^2 N} \tag{16}$$

$$= \frac{4}{\pi^2 r}. \tag{17}$$

This bound shows that Fourier sampling produces a value of k that is the closest integer to one of the r integer multiples of N/r with probability lower bounded by a constant.

To discover r given one of the values $\lfloor jN/r \rceil$, we can divide by N to obtain a rational approximation to j/r that deviates by at most $1/2N$. Then consider the continued fraction expansion

$$\frac{\lfloor jN/r \rceil}{N} = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}. \tag{18}$$

Truncating this expansion after a finite number of terms gives a *convergent* of the expansion. The convergents provide a sequence of successively better approximations to $\lfloor jN/r \rceil/N$ by fractions that can be computed in polynomial time (see for example Knuth's *The Art of Computer Programming*, volume 2). Furthermore, it can be shown that any fraction p/q with $|p/q - \lfloor jN/r \rceil/N| < 1/2q^2$ will appear as one of the convergents (see for example Hardy and Wright, Theorem 184). Since j/r differs by at most $1/2N$ from $\lfloor jN/r \rceil/N$, the fraction j/r will appear as a convergent provided $r^2 < N$. By taking N is sufficiently large, this gives an efficient means of recovering the period.

Period finding over the reals

Now suppose we are given a function $f\colon \mathbb{R} \to S$ satisfying $f(x + r) = f(x)$ for some $r \in \mathbb{R}$, and as usual, assume that f is injective within each (minimal) period. Now we'll see how to adapt Shor's procedure to find an approximation to r, even if it happens to be irrational.

To perform period finding on a digital computer, we must of course discretize the function. We have to be careful about how we perform this discretization. For example, suppose that $S = \mathbb{R}$. If we simply evaluate f at equally spaced points and round the resulting values (perhaps rescaled) to get integers, there is no reason for the function values corresponding to inputs separated by an amount close to the period to be related in any way whatsoever. It could be that the discretized function is injective, carrying absolutely no information about the period.

Instead we will discretize in such a way that the resulting function is *pseudoperiodic*. We say that $f : \mathbb{Z} \to S$ is *pseudoperiodic at $k \in \mathbb{Z}$ with period $r \in \mathbb{R}$* if for each $\ell \in \mathbb{Z}$, either $f(k) = f(k + \lfloor \ell r \rfloor)$ or $f(k) = f(k - \lceil \ell r \rceil)$. We say that f is *ϵ-pseudoperiodic* if it is pseudoperiodic for at least an ϵ fraction of the values $k = 0, 1, \ldots, \lfloor r \rfloor$. We assume that the discretized function is ϵ-pseudoperiodic for some constant ϵ, and that it is injective on the subset of inputs where it is pseudoperiodic. Note that the periodic function encoding the regulator of Pell's equation can be constructed so that it satisfies these conditions.

Now let's consider what happens when we apply Fourier sampling to a pseudoperiodic function. As before, we will Fourier sample over $\mathbb{Z}_N$, with N to be determined later (again, depending on

some a priori upper bound M on the period r). We start by computing the pseudoperiodic function on a uniform superposition:

$$\sum_{x\in\{0,\ldots,N-1\}} |x\rangle \mapsto \sum_{x\in\{0,\ldots,N-1\}} |x, f(x)\rangle. \tag{19}$$

Now measuring the second register gives, with constant probability, a value for which f is pseudoperiodic. Say that this value is $f(x_0)$ where $0 \le x_0 \le r$. As before, we see $n = \lfloor N/r \rfloor + 1$ points if $x_0 < N - r\lfloor N/r \rfloor$ or $n = \lfloor N/r \rfloor$ points otherwise (possibly offset by 1 depending on how the rounding occurs for the largest value of x, but let's not be concerned with this detail). We will write $[\ell]$ to denote an integer that could be either $\lfloor \ell \rfloor$ or $\lceil \ell \rceil$. With this notation, we obtain

$$\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} |x_0 + [jr]\rangle. \tag{20}$$

Next, performing the Fourier transform over $\mathbb{Z}_N$ gives

$$\frac{1}{\sqrt{nN}} \sum_{j=0}^{n-1} \sum_{k\in\mathbb{Z}_N} \omega_N^{k(x_0+[jr])} |k\rangle = \frac{1}{\sqrt{nN}} \sum_{k\in\mathbb{Z}_N} \omega_N^{kx_0} \sum_{j=0}^{n-1} \omega_N^{k[jr]} |k\rangle. \tag{21}$$

Now we have $[jr] = jr + \delta_j$, where $-1 < \delta_j < 1$, so the sum over j is

$$\sum_{j=0}^{n-1} \omega_N^{k[jr]} = \sum_{j=0}^{n-1} \omega_N^{kjr} \omega_N^{k\delta_j}. \tag{22}$$

We would like this to be close to the corresponding sum in the case where the offsets δ_j are zero (which, when normalized, is $\Omega(1/\sqrt{r})$ by the same calculation as in the case of period finding over $\mathbb{Z}$). Consider the deviation in amplitude,

$$\left| \sum_{j=0}^{n-1} \omega_N^{kjr} \omega_N^{k\delta_j} - \sum_{j=0}^{n-1} \omega_N^{kjr} \right| \le \sum_{j=0}^{n-1} |\omega_N^{k\delta_j} - 1| \tag{23}$$

$$= \frac{1}{2} \sum_{j=0}^{n-1} \left| \sin \frac{\pi k \delta_j}{N} \right| \tag{24}$$

$$\le \frac{1}{2} \sum_{j=0}^{n-1} \left| \frac{\pi k \delta_j}{N} \right| \tag{25}$$

$$\le \frac{\pi k n}{2N}. \tag{26}$$

At least insofar as this bound is concerned, the amplitudes may not be close for all values of k. However, suppose we only consider values of k less than $N/\log r$. (We will obtain such a k with probability about $1/\log r$, so we can condition on this event with only polynomial overhead.) For such a k, we have

$$\left| \frac{1}{\sqrt{nN}} \sum_{j=0}^{n-1} \omega_N^{k[jr]} \right| = \Omega(1/\sqrt{r}) - O(\tfrac{1}{\sqrt{nN}} \cdot \tfrac{n}{\log r}) \tag{27}$$

$$= \Omega(1/\sqrt{r}) - O(\tfrac{1}{\sqrt{r}\log r}) \tag{28}$$

$$= \Omega(1/\sqrt{r}). \tag{29}$$

Thus, as in the case of period finding over $\mathbb{Z}$, Fourier sampling allows us to sample from a distribution for which some value $k = \lfloor jN/r \rceil$ (with $j \in \mathbb{Z}$) appears with reasonably large probability (now $\Omega(1/\operatorname{poly}(\log r))$ instead of $\Omega(1)$).

Finally, we must obtain an approximation to r using these samples. Since r is not an integer, the procedure used in Shor's period-finding algorithm does not suffice. However, we can perform Fourier sampling sufficiently many times that we obtain two values $\lfloor jN/r \rceil, \lfloor j'N/r \rceil$ such that j and j' are relatively prime, again with only polynomial overhead. We prove below that if $N \geq 3r^2$, then j/j' is guaranteed to be one of the convergents in the continued fraction expansion for $\lfloor jN/r \rceil / \lfloor j'N/r \rceil$. Thus we can learn j, and hence compute $jN/\lfloor jN/r \rceil$, which gives a good approximation to r: in particular, $|r - \lfloor jN/\lfloor jN/r \rceil \rceil| \leq 1$.

Lemma. *If $N \geq 3r^2$, then j/j' appears a convergent in the continued fraction expansion of $\lfloor jN/r \rceil / \lfloor j'N/r \rceil$. Furthermore, $|r - \lfloor jN/\lfloor jN/r \rceil \rceil| \leq 1$.*

Proof. A standard result on the theory of approximation by continued fractions says that if $a, b \in \mathbb{Z}$ with $|x - \frac{a}{b}| \leq \frac{1}{2b^2}$, then a/b appears as a convergent in the continued fraction expansion of x (see for example Hardy and Wright, *An Introduction to the Theory of Numbers*, Theorem 184.) Thus it is sufficient to show that

$$\left| \frac{\lfloor jN/r \rceil}{\lfloor j'N/r \rceil} - \frac{j}{j'} \right| < \frac{1}{2j'^2}. \tag{30}$$

Letting $\lfloor jN/r \rceil = jN/r + \mu$ and $\lfloor j'N/r \rceil = jN/r + \nu$ with $|\mu|, |\nu| \leq 1/2$, we have

$$\left| \frac{\lfloor jN/r \rceil}{\lfloor j'N/r \rceil} - \frac{j}{j'} \right| = \left| \frac{jN/r + \mu}{j'N/r + \nu} - \frac{j}{j'} \right| \tag{31}$$

$$= \left| \frac{jN + \mu r}{j'N + \nu r} - \frac{j}{j'} \right| \tag{32}$$

$$= \left| \frac{r(\mu j' - \nu j)}{j'(j'N + \nu r)} \right| \tag{33}$$

$$\leq \left| \frac{r(j + j')}{2j'^2 N - j'r} \right| \tag{34}$$

$$\leq \frac{r}{j'N - r/2} \tag{35}$$

where in the last step we have assumed $j < j'$ wlog. This is upper bounded by $1/2j'^2$ provided $j'N \geq r/2 + 2j'^2 r$, which certainly holds if $N \geq 3r^2$ (using the fact that $j' < r$).

Finally

$$r - \frac{jN}{\left\lfloor \frac{jN}{r} \right\rceil} = r - \frac{jN}{\frac{jN}{r} + \mu} \tag{36}$$

$$= r - \frac{jNr}{jN + \mu r} \tag{37}$$

$$= \frac{\mu r^2}{jN + \mu r} \tag{38}$$

which is at most 1 in absolute value since $N \geq 3r^2$, $|\mu| \leq 1/2$, and $j \geq 1$. □

Other algorithms for number fields

To conclude, we mention some further applications of quantum computing to computational algebraic number theory.

Hallgren's original paper on Pell's equation also solves another problem, the *principal ideal problem*, which is the problem of deciding whether an ideal is principal, and if so, finding a generator of the ideal. Factoring reduces to the problem of solving Pell's equation, and Pell's equation reduces to the principal ideal problem; but no reductions in the other direction are known. Motivated by the possibility that the principal ideal problem is indeed harder than factoring, Buchmann and Williams designed a key exchange protocol based on it. Hallgren's algorithm shows that quantum computers can break this cryptosystem.

Subsequently, further related algorithms for problems in algebraic number theory have been found by Hallgren and, independently, by Schmidt and Vollmer. Specifically, they found polynomial-time algorithms for computing the unit group and the class group of a number field of constant degree. These algorithms require generalizing period finding over $\mathbb{R}$ to a similar problem over $\mathbb{R}^d$.