Factoring on a Quantum Computer

The Essence Shor's Algorithm

Wolfgang Polak

wp@pocs.com

Thanks to:
Eleanor Rieffel
Fuji Xerox Palo Alto Laboratory

Why is factoring interesting?

- Factoring and computing discrete logarithms is presumed hard on classical computers. There is no proof.
- Best algorithm for factoring an n-bit number: $O\left(\exp\left((\frac{64}{9}n)^{1/3}(\log n)^{2/3}\right)\right)$.
- All practical public key encryption systems are based on one of these problems.
  - RSA (Rivest, Shamir and Adleman) factoring
  - ElGamal (Taher ElGamal) discrete logarithms
  - DSA (Digital Signature Algorithm) discrete logarithms over elliptic curves.
- There are no known (practical) alternatives for public key encryption.
- A quantum computer can solve both problems in polynomial time.
- Dire consequences: digital signatures become forgeable, e-commerce seizes, etc.

Overview

- Quantum Mechanics
  - Quantum States
  - State Transformations
  - Measurement
- Quantum Computation
  - Use of quantum state transformation, measurement to compute
  - Time: number of primitive transformations required.
  - Space: size (dimension) of the quantum state required.
- Discrete Fourier Transform
  - Useful for finding the period of a function
  - Efficient implementation on a quantum computer.
- Factoring
  - Reduced to period finding
  - Use quantum Fourier transform

Quantum Mechanics

I think I can safely say that nobody understands quantum mechanics.

Richard Feynman

Work from a set of axioms (postulates in physics) - simplified for discrete quantum systems:

Ax 1: The state space of a binary quantum system is a 2-dimensional complex vector space.

Ax 2: The state space of multiple binary quantum systems is the tensor product of the individual state spaces.

Ax 3: Transformations of quantum states are unitary.

Ax 4: Measurement of a quantum system is a probabilistic projection on one of several orthogonal subspaces.

Binary Quantum Systems

Ax 1: The state space of a binary quantum system is a 2-dimensional complex vector space.

- More precisely: $\mathbf{CP}^1$, the complex projective space of dimension 1.
- Convention: (i) quantum states are unit vectors (ii) two states are the same if they differ by a constant factor (phase).
- Bra-Ket notation: vectors that represent quantum states [Paul Dirac]:

$\lvert x\rangle$	Vector labeled x	column vector, $\mathbf{x}$, $\vec{x}$
$\langle x\rvert$	Conjugate transpose	row vector, $\mathbf{x}^\dagger$
$\langle x\lvert\rvert y\rangle = \langle x\lvert y\rangle$	Inner product	$\mathbf{x}^\dagger \cdot \mathbf{y}$
$\langle x\lvert y\rangle = 0$	Orthogonal vectors	$\mathbf{x}^\dagger \cdot \mathbf{y} = 0$
$\langle x\lvert x\rangle = \lvert\lvert x\rangle\rvert^2$	Length	$\lvert\mathbf{x}\rvert^2$
$\langle x\lvert x\rangle = 1$	Unit vector	$\mathbf{x}^\dagger \cdot \mathbf{x} = 1$
$\lvert x\rangle\langle y\rvert$	Outer product	a matrix

Example: $(\lvert x\rangle\langle y\rvert)\lvert y\rangle = \lvert x\rangle(\langle y\lvert\rvert y\rangle) = \lvert x\rangle$.

Binary Quantum Systems - Example

Examples of binary quantum systems:

electron spin, ground/excited state, photon polarization.

$\|\uparrow\rangle$	vertical polarization
$\|\rightarrow\rangle$	horizontal polarization
$\{\|\uparrow\rangle, \|\rightarrow\rangle\}$	basis, i.e. $\langle\uparrow\|\rightarrow\rangle = 0$
$\frac{1}{\sqrt{2}}(\|\uparrow\rangle + \|\rightarrow\rangle)$	linear combination
$\frac{1}{\sqrt{2}}(\|\uparrow\rangle + \mathbf{i}\|\rightarrow\rangle)$	aka superposition
$\{\frac{1}{\sqrt{2}}(\|\uparrow\rangle + \|\rightarrow\rangle), \frac{1}{\sqrt{2}}(\|\uparrow\rangle - \|\rightarrow\rangle)\}$	a different basis

Example:

$$\begin{aligned}
\frac{1}{\sqrt{2}}(\langle\uparrow| + \langle\rightarrow|)\frac{1}{\sqrt{2}}(|\uparrow\rangle - |\rightarrow\rangle) &= \frac{1}{2}(\langle\uparrow|\uparrow\rangle - \langle\rightarrow|\rightarrow\rangle + \langle\uparrow|\rightarrow\rangle - \langle\rightarrow|\uparrow\rangle)\ . \\
&= \frac{1}{2}(1 - 1 + 0 - 0) \\
&= 0
\end{aligned}$$

Caution: $+|\rightarrow\rangle$ and $-|\rightarrow\rangle$ are the same quantum state, but $1/\sqrt{2}(|\uparrow\rangle + |\rightarrow\rangle)$ and $1/\sqrt{2}(|\uparrow\rangle - |\rightarrow\rangle)$ are different

Quantum Bits

- A quantum bit or qubit is a 2-dimensional quantum system.
- The state space $\mathbf{B}^{(1)}$ of a qubit has (computational) basis states $|0\rangle$ and $|1\rangle$ encoding 0 and 1.
- Unlike classical bits, qubits can be in superposition states, e.g. $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.
  - does NOT mean 0.5,
  - NEITHER randomly 0 or 1,
  - maybe 0 and 1 “at the same time.”
- A general qubit state is $a|0\rangle + b|1\rangle$ for complex a, b with $|a|^2 + |b|^2 = 1$.
- Qubit: unit of quantum information, different from classical information.

The essence of quantum computation
is not the use of quantum effects (every transistor does that)
it is the use of quantum, not classical information.

Multi-qubit State Spaces

Ax 2: The state space of multiple binary quantum systems is the tensor product of the individual state spaces.

- If V_1 has basis $\{a_1, \ldots, a_k\}$ and V_2 has basis $\{b_1, \ldots, b_n\}$, $V_1 \otimes V_2$ has basis $\{a_i \otimes b_j | 1 \leq i \leq k, 1 \leq j \leq n\}$.
- $\dim(V_1 \otimes V_2) = \dim(V_1) \cdot \dim(V_2)$.
- Notation: $|x\rangle \otimes |y\rangle = |x\rangle|y\rangle = |xy\rangle$.
- The 2-qubit state space $\mathbf{B}^{(2)}$ has computational basis
$$\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}.$$
- The state space $\mathbf{B}^{(n)}$ of an n-qubit system is a 2^n dimensional complex vector space with computational basis
$$\{|00\ldots00\rangle, |00\ldots01\rangle, \ldots, |11\ldots10\rangle, |11\ldots11\rangle\}.$$
- Notation: $|6\rangle = |110\rangle$ (when n is understood)
- A general n-qubit state: $\sum_{i=0}^{2^n-1} a_i|i\rangle$, s.t. $\sum_i |a_i|^2 = 1$.

Quantum State Transformations

Ax 3: Transformations of quantum states are unitary.

Possible quantum state transformations are subject to physical constraints

unitary $\equiv U^\dagger = U^{-1} \equiv$ length preserving, linear $\equiv$ basis change
$\equiv$ inner product preserving $\equiv$ rotation $\Longrightarrow$ reversible.

It suffices to specify transformation for some basis:

$$I: \begin{array}{ccc} |0\rangle & \rightarrow & |0\rangle \\ |1\rangle & \rightarrow & |1\rangle \end{array} \qquad X: \begin{array}{ccc} |0\rangle & \rightarrow & |1\rangle \\ |1\rangle & \rightarrow & |0\rangle \end{array} \qquad Z: \begin{array}{ccc} |0\rangle & \rightarrow & |0\rangle \\ |1\rangle & \rightarrow & -|1\rangle \end{array}$$

$$C_{not}: \begin{array}{ccc} |00\rangle & \rightarrow & |00\rangle \\ |01\rangle & \rightarrow & |01\rangle \\ |10\rangle & \rightarrow & |11\rangle \\ |11\rangle & \rightarrow & |10\rangle \end{array}$$

$$H: \begin{array}{ccc} |0\rangle & \rightarrow & \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ |1\rangle & \rightarrow & \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \end{array}$$

Realizing Transformations

Complex, high-dimensional transformation can be composed from primitive ones called quantum gates.

- Sequential composition (product): $HXH[=Z]$,
- Tensor product: $I \otimes X \otimes I$, a transformation of one of 3 qubits.
- Complete set of gates: can be composed to realize any unitary transformation.
- Computational complexity: number of primitive gates required to realize a unitary transformation.

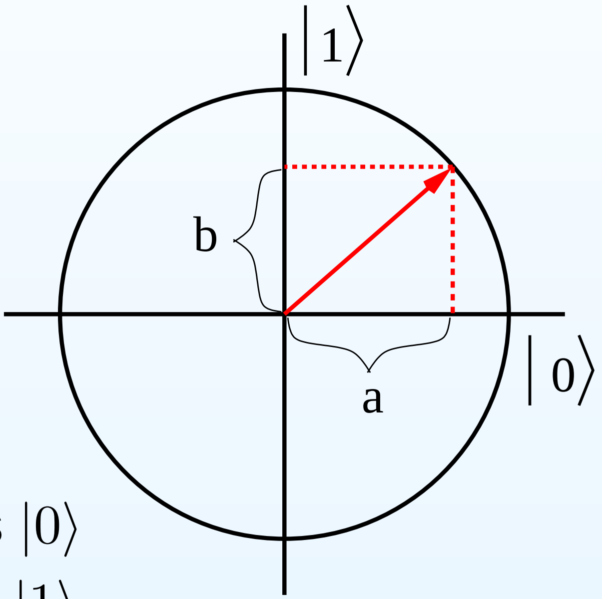
A complete (infinite) gate set: C_{not} together with $R(\beta)$ and $P(\alpha)$

$$R(\beta): \quad \begin{array}{rcl} |0\rangle & \rightarrow & \cos\beta|0\rangle + \sin\beta|1\rangle \\ |1\rangle & \rightarrow & -\sin\beta|0\rangle + \cos\beta|1\rangle \end{array} \qquad P(\alpha): \quad \begin{array}{rcl} |0\rangle & \rightarrow & e^{\mathbf{i}\alpha}|0\rangle \\ |1\rangle & \rightarrow & e^{-\mathbf{i}\alpha}|1\rangle \end{array}$$

There is no complete finite set of gates.

Quantum Measurement

Ax 4: Measurement of a quantum system is a probabilistic projection on one of several orthogonal subspaces.



Single qubit case:

- Basis $\{|0\rangle, |1\rangle\}$ defines two subspaces, $x|0\rangle$ and $y|1\rangle$.
- Measuring qubit $a|0\rangle + b|1\rangle$ in the computational basis $\{|0\rangle, |1\rangle\}$
  - returns 0 with probability $|a|^2$, state becomes $|0\rangle$
  - returns 1 with probability $|b|^2$, state becomes $|1\rangle$

Measurement changes the state unless it is one of the basis states of the measurement.

Measurement of a qubit provides only one classical bit of information.

Qubit Measurement in Context

Measuring qubit k of an n qubit system.

- Single qubit basis $\{|0\rangle, |1\rangle\}$ defines two subspaces:

$$\begin{aligned} S_0 &= \mathbf{B}^{(k-1)} \otimes \{|0\rangle\} \otimes \mathbf{B}^{(n-k-1)} \\ S_1 &= \mathbf{B}^{(k-1)} \otimes \{|1\rangle\} \otimes \mathbf{B}^{(n-k-1)} \end{aligned}$$

- Write $|\psi\rangle \in \mathbf{B}^{(n)}$ as $|\psi\rangle = c_0|\psi_0\rangle + c_1|\psi_1\rangle$ with $|\psi_0\rangle \in S_0$, $|\psi_1\rangle \in S_1$.
- Measuring qubit k of $|\psi\rangle$
  - results in 0 with probability $|c_0|^2$, changing $|\psi\rangle$ to $|\psi_0\rangle$
  - results in 1 with probability $|c_1|^2$, changing $|\psi\rangle$ to $|\psi_1\rangle$.

Example: measure 2nd qubit of $|\psi\rangle = a|001\rangle + b|100\rangle + c|110\rangle$.

Let $c_0 = \sqrt{1-|c|^2}$ then $|\psi\rangle = c_0\left(\frac{a}{c_0}|001\rangle + \frac{b}{c_0}|100\rangle\right) + c|110\rangle$

- 0 with probability $|c_0|^2 = 1 - |c|^2$, new state $\frac{a}{c_0}|001\rangle + \frac{a}{c_0}|100\rangle$
- 1 with probability $|c|^2$, new state $|110\rangle$.

Measure multiple qubits one qubit at a time (commutative).

Measurement – No Cloning

Theorem:

An unknown quantum state $|x\rangle$ cannot be copied, not by measurement, not by any other means.

Proof:

- Measurement yields only one classical bit of information.
- Assume U_c were a cloning transformation, such that for all $|x\rangle$:

$$U_c(|x\rangle|0\rangle) = |x\rangle|x\rangle.$$

- Consider orthogonal $|a\rangle$ and $|b\rangle$ and let $|c\rangle = \frac{1}{\sqrt{2}}(|a\rangle + |b\rangle)$.
  - By linearity:
    $U_c(|c\rangle|0\rangle) = \frac{1}{\sqrt{2}}(U_c(|a\rangle|0\rangle) + U_c(|b\rangle|0\rangle)) = \frac{1}{\sqrt{2}}(|a\rangle|a\rangle + |b\rangle|b\rangle)$.
  - By cloning: $U_c(|c\rangle|0\rangle) = |c\rangle|c\rangle = \frac{1}{2}(|a\rangle|a\rangle + |a\rangle|b\rangle + |b\rangle|a\rangle + |b\rangle|b\rangle$.

  Thus, there cannot be a cloning transformation.

Note: given a, b, $a|0\rangle + b|1\rangle$ can be constructed efficiently.

Entangled States

- Most n-qubit states cannot be written as the tensor product of 2 states. These states are called entangled.
- Examples: $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, $\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$

$$
\begin{aligned}
& (a_0|0\rangle + b_0|1\rangle) \otimes (a_1|0\rangle + b_1|1\rangle) \\
= \;& a_0a_1|00\rangle + a_0b_1|01\rangle + b_0a_1|10\rangle + b_0b_1|11\rangle \\
\neq \;& a_0a_1|00\rangle + 0|01\rangle + 0|10\rangle + b_0b_1|11\rangle \\
= \;& \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)
\end{aligned}
$$

- Entanglement depends on the tensor decomposition of the state.
- Measurement of $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$:
  - Measurement of first qubit yields either $|0\rangle$ or $|1\rangle$.
  - Measurement projects the state to either $|00\rangle$ or $|11\rangle$.
  - Measurement of the second qubit will give the same result as measurement of the first.

Quantum Computation

A quantum computation consists of

- initialization of n-qubit “register”,
- quantum state transformation of n-qubit state by a
  - sequence of primitive (1,2,3-qubit) transformations that collectively perform the transformation of the register,
- measurement of some of the qubits of the register,
- classical control to,
  - interpret results of quantum measurement
  - iterate quantum steps

For each classical algorithm with time/space complexity t/s there exist a classical reversible algorithm with time $O(t^{1+\epsilon})$ and space $O(s \log t)$ complexity.

For each classical reversible algorithm there is a unitary transformation with the same complexity.

Quantum Parallelism

For any classical function $f : \mathbf{Z}_{2^n} \to \mathbf{Z}_{2^m}$ there is a unitary transformation U_f on $n+m$ qubits such that for $x \in \mathbf{Z}_{2^n}$:

$$U_f|x\rangle|0\rangle = |x\rangle|f(x)\rangle.$$

By linearity, U_f works on superpositions:

$$U_f\left(\frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1}|x,0\rangle\right) = \frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1}U_f|x,0\rangle = \frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1}|x,f(x)\rangle.$$

Apparent exponential number of computations!?

Exponential size superposition can be created in linear time

Hadamard transformation $H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, thus

$$H^{(n)}|0\ldots0\rangle = \frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1}|x\rangle.$$

Is Quantum Parallelism Useful?

How can we exploit

$$\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x, f(x)\rangle?$$

- Measurement of all qubits gives $\langle x_0, f(x_0)\rangle$ for some random x_0.
- Measurement of the last m qubits gives some u and collapses the state to

$$c \sum_{x \in f^{-1}(u)} |x\rangle |u\rangle$$

i.e., the first n qubits are a superposition of the preimage of u.

Two strategies:

1. adjust amplitudes a_x so that values of interest are read with higher probability
2. compute properties of f or its preimage through Fourier Transform

Discrete Fourier Transform

N-th root of unity $\omega_N = \exp(\frac{2\pi\mathbf{i}}{N})$:

- $\omega_N^k = 1$ for $k = 0 \bmod N$.
- $\sum_{i=0}^{N-1} \omega_N^{ik} = 0$ for $k \neq 0 \bmod N$.

Discrete Fourier transform $F : \mathbf{C}^N \to \mathbf{C}^N$. As matrix, $F_{ij} = \frac{1}{\sqrt{N}}\omega_N^{ij}$.

F is unitary.

Proof: rows F_i of F are unit length and orthogonal, i.e., $F_i F_j^\dagger = \delta_{ij}$.

Let $pk = N$: If $\Big[$ $\mathbf{v}_i$ is non-zero iff $i + m$ is a multiple of p $\Big]$
then $\Big[$ $(F\mathbf{v})_i$ is non-zero iff i is a multiple of $\frac{N}{p}$ $\Big]$.

$$(F\mathbf{v})_i = \sum_{j=0}^{N-1} F_{ij}\mathbf{v}_j = c\sum_{r=0}^{k-1} \omega_N^{ipr} = c\sum_{r=0}^{k-1} \omega_k^{ir} = \begin{cases} kc & \text{if } i = 0 \bmod k \\ 0 & \text{otherwise} \end{cases}$$

Result is approximate if p does not divide N.

Fast Fourier Transform

FFT = efficient algorithm of DFT for $N = 2^n$ based on recursive decomposition of F:

$$
\begin{aligned}
F^{(0)} &= 1 \\
F^{(k)} &= \frac{1}{\sqrt{2}} \begin{pmatrix} I^{(k-1)} & D^{(k-1)} \\ I^{(k-1)} & -D^{(k-1)} \end{pmatrix} \begin{pmatrix} F^{(k-1)} & 0 \\ 0 & F^{(k-1)} \end{pmatrix} R^{(k)} \\
D^{(k-1)}_{ij} &= \begin{cases} 0 & \text{if } i \neq j \\ \omega^{i}_{2^k} & \text{otherwise} \end{cases} \\
R^{(k-1)}_{ij} &= \begin{cases} 1 & \text{if } 2i = j \\ 1 & \text{if } 2i - 2^k + 1 = j \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

Quantum Fourier Transforms

$$\begin{aligned} \mathrm{Q}: \quad \mathbf{B}^{(n)} &\rightarrow \mathbf{B}^{(n)} \\ |x\rangle &\rightarrow Q|x\rangle \end{aligned}$$

$$Q|i\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \omega_N^{ij} |j\rangle, \quad N = 2^n$$

$$Q \sum_{i=0}^{N-1} a_i |i\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \sum_{i=0}^{N-1} a_i \omega_N^{ij} |j\rangle = \sum_{j=0}^{N-1} (Fa)_j |j\rangle$$

FFT recursive decomposition applies. In bra/ket notation:

$$\begin{aligned} Q^{(1)} &= H \\ Q^{(k)} &= \frac{1}{\sqrt{2}} M^{(k)} (I \otimes Q^{(k-1)}) R^{(k)} \\ M^{(k)} &= \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle)\langle 0| \otimes I^{(k-1)} + \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle)\langle 0| \otimes D^{(k-1)} \\ D^{(k)} &= D^{(k-1)} \otimes (|0\rangle\langle 0| + \omega_{2^{k+1}} |1\rangle\langle 1|) \\ R^{(k)} &= \text{trivial swap of qubits} \end{aligned}$$

Fast vs Quantum Fourier Transform

- classical FFT requires explicit representation of exponentially (in n) many complex coefficients.
- in QFT coefficients are implicit in amplitudes of a single superposition of index values of an n-qubit state.

	FFT	QFT
Data structure	$\mathbf{C}^{2^n}$	$\mathbf{B}^{(n)}$
Space complexity	2^n complex numbers	n qubits
Time complexity	$O(n2^n)$	$O(n^2)$

Factoring by Period-Finding

- The order of $a \bmod M$ is the least $p > 0$ such that $a^p = 1 \bmod M$.
- p is finite when a and M are relative prime.
- $a^k = a^{k+p} \bmod M$ iff $a^p = 1 \bmod M$
- Let $g(k) = a^k \bmod M$ then $g(k) = g(k+p)$ and p is the period of g.
- If p, the order of $a \bmod M$, is even

$$(a^{p/2}+1)(a^{p/2}-1) = a^p - 1 = 0 \bmod M.$$

- If neither $a^{p/2}+1$ nor $a^{p/2}-1$ is a multiple of M,
  - gcd$(a^{p/2}+1, M)$ or
  - gcd$(a^{p/2}-1, M)$

  is a non-trivial factor of M.
- Shor's algorithm factors M by using QFT to compute the period of $g(k) = a^k \bmod M$.

Shor's Algorithm

Input: an n-bit number M, output: a non-trivial factor of M:

1. pick random a, $0 < a < M$. If gcd$(a, M) \neq 1$ we have a factor.
2. for $g(k) = a^k$ mod M compute the $2n$ qubit state $\frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} |k, g(k)\rangle = U_g(H^{(n)} \otimes I^{(n)})|0, 0\rangle$.
3. measure right-most n qubits, yielding some u.
4. the state collapses to $c \sum_{k=0}^{2^n-1} v_k |k, u\rangle$ with $v_k = \begin{cases} 1 & \text{if } g(k) = u \\ 0 & \text{otherwise} \end{cases}$

   v_k is non-zero iff $k + m$ is a multiple of p.
5. perform QFT on the first n qubits giving $c' \sum_{j=0}^{2^n-1} w_j |j\rangle$, $w = Fv$.
6. measure the result, some j_0. j_0 will be close to $\frac{2^n}{p}$.
7. conjecture a likely period p from j_0 (uses continued fraction expansion)
8. see if gcd$(a^{p/2} \pm 1, M)$ is a nontrivial factor
9. repeat steps 1 through 8 if necessary.

Conclusion

If the computers that you build are quantum,
Then spies everywhere will all want 'em.
Our codes will all fail,
And they'll read our email,
Till we get crypto that's quantum, and daunt 'em.

Jennifer and Peter Shor