Quantenalgorithmus für die Faktorisierung ganzer Zahlen

Ausgehend von dem allgemeinen Algorithmus für das Hidden Subgroup Problem behandlen wir in diesem Abschnitt den Quantenalgorithmus für die Faktorisierung ganzer Zahlen aus $\mathbb{Z}$, welcher auf P. Shor zurückgeht.

Reduktion von FACT auf HSP

Wir führen das Faktorisierungsproblem auf das Hidden Subgroup Problem zurück. Zur Vereinfachung betrachten wir $N = pq$ mit p, q ungeraden Primzahlen. Der allgemeine Fall kann ähnlich behandelt werden. Wir setzen $S = (\mathbb{Z}/N\mathbb{Z})^\times$.

1 Lemma. *Für zufällig und gleichverteilt gewähltes $x \in S$ gilt mit Wahrscheinlichkeit $\geq 1/2$, daß* $\mathrm{ord}(x)$ *gerade und* $\gcd\{x^{\mathrm{ord}(x)/2} - 1, N\} \in \{p, q\}$ *ist.*

Beweis. Nach dem chinesischen Restsatz ist $\mathbb{Z}/N\mathbb{Z} \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$ und $S \cong (\mathbb{Z}/p\mathbb{Z})^\times \times (\mathbb{Z}/q\mathbb{Z})^\times$ unter dem üblichen Isomorphismus

$$\phi : \mathbb{Z}/N\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}, \quad \phi(x) = (x_p, x_q) = (x \bmod p, x \bmod q).$$

Die zufällige und gleichverteilte Wahl von $x \in G$ entspricht der zufälligen, gleichverteilten und unabhängigen Wahl von $x_p \in (\mathbb{Z}/p\mathbb{Z})^\times$ und $x_q \in (\mathbb{Z}/q\mathbb{Z})^\times$. Da $(\mathbb{Z}/p\mathbb{Z})^\times$ und $(\mathbb{Z}/q/Z)^\times$ nach Voraussetzung zyklische Gruppen mit geraden Ordnungen sind, sind $\mathrm{ord}(x_p)$ und $\mathrm{ord}(x_q)$ jeweils mit Wahrscheinlichkeit $1/2$ gerade. Ferner gilt $\mathrm{ord}(x) = \mathrm{lcm}\{\mathrm{ord}(x_p), \mathrm{ord}(x_q)\}$. Wir erhalten damit die folgende Aufstellung der möglichen Fälle:

$\mathrm{ord}(x_p)$	$\mathrm{ord}(x_q)$	$x_p^{\mathrm{ord}(x)/2}$	$x_q^{\mathrm{ord}(x)/2}$
ungerade	ungerade	–	–
gerade	gerade	± 1	± 1
ungerade	gerade	1	-1
gerade	ungerade	-1	1

Jede Zeile tritt dabei mit Wahrscheinlichkeit $1/4$ auf. In den letzten drei Fällen ist $\mathrm{ord}(x)$ gerade. In der zweiten Zeile steht mindestens eine -1, und in den beiden letzten Zeilen steht notwendigerweise unterschiedliches Vorzeichen. In den Fällen der letzten beiden Zeilen ergibt sich daher beispielsweise aus $\phi(x)^{\mathrm{ord}(x)/2} - 1 = (0, -2)$, daß $x^{\mathrm{ord}(x)/2} - 1 = 0 \bmod p$ und $x^{\mathrm{ord}(x)/2} - 1 \neq 0 \bmod q$, folglich $\gcd\{x^{\mathrm{ord}(x)/2} - 1, N\} = p \in \{p, q\}$. Daher ist insgesamt mit Wahrscheinlichkeit $\geq 1/2$ die Ordnung $\mathrm{ord}(x)$ gerade und $\gcd\{x^{\mathrm{ord}(x)/2} - 1, N\} \in \{p, q\}$. □

Reduktion von FACT auf ComputeOrder. Das Orakel O liefert unter Eingabe von $x \in S$ die Ordnung $\mathrm{ord}(x)$ zurück. Der Faktorisierungalgorithmus geht dann wie folgt vor: Rufe O für zufällige $x \in S$ auf und berechne $\gcd\{x^{\mathrm{ord}(x)/2} - 1, N\}$, falls $\mathrm{ord}(x)$ gerade ist. Dies liefert p oder q mit Wahrscheinlichkeit $\geq 1/2$.

Reduktion von ComputeOrder auf HSP. Das HSP Orakel P erhält als Eingabe $f : G \to S$ mit $f(x) = f(y) \Leftrightarrow xy^{-1} \in H$ für eine (unbekannte) Untergruppe H, und berechnet (Erzeuger von) H. Der Algorithmus zur Ordnungsberechnung von $x \in S$ geht dann wie folgt vor: Wende P auf $f : \mathbb{Z}/\phi(N)\mathbb{Z} \to S$, $f(z) = x^z$ an. Dies liefert $H = \langle r \rangle$ mit $r = \mathrm{ord}(x)$ zurück.

Wir könnten nun im Prinzip den allgemeinen Algorithmus zum HSP auf endlichen abelschen Gruppen anwenden. Es gibt jedoch zwei Probleme. Zum einen ist $\phi(N)$ und damit G unbekannt. Zum anderen können wir nur für spezielle G die Fouriertransformation mittels der Produktformel schnell berechnen. Zur Lösung des Problems ersetzen wir $\phi(N)$ durch eine Zahl $Q \gg N^2$. Zur Lösung des zweiten Problems wählen wir $Q = 2^n$. Ansonsten gehen wir wie in dem allgemeinen Algorithmus zum HSP bezüglich der Gruppe $\mathbb{Z}/Q\mathbb{Z}$ vor. Es bleibt zu zeigen, daß dann immer noch etwas sinnvolles berechnet wird (zum Beispiel, daß die Fouriertransformation auf $\mathbb{Z}/Q\mathbb{Z}$ ähnliches liefert wie die Fouriertransformation auf G).

Erweitertes Hidden Subgroup Problem

Wir betrachten $f : \mathbb{Z} \to S$ mit $f(x) = f(y) \Leftrightarrow x - y \in \mathbb{Z}r$. Die Aufgabe des erweiteren HSP ist, r zu berechnen. Der Definitionsbereich ist jetzt also nicht mehr eine endliche abelsche Gruppe.

Die Idee besteht nun darin, ein HSP bezüglich $h : G \to S$ mit einer zyklischen Gruppe G mittels des Restklassenhomomorphimusm $\mathbb{Z} \to G$ auf $\mathbb{Z}$ zu liften. Hier wird dann $Q = 2^n \gg r^2$ ausreichend groß gewählt, und r mittels eines „herkömmlichen“ HSP auf $\mathbb{Z}/Q\mathbb{Z} = \{0, \ldots, Q-1\}$ gelöst.

Im allgemeinen wird $r \nmid Q$ gelten, was wir im folgenden auch annehmen. Insofern wird durch r nun keine (versteckte) Untergruppe in $\mathbb{Z}/Q\mathbb{Z}$ mehr definiert. Man könnte sagen, daß nurmehr eine (versteckte) approximierte Untergruppe definiert wird.

Der Algorithmus benötigt wieder zwei Register. Register 1 enthält die Elemente von $\mathbb{Z}/Q\mathbb{Z} = \{0, \ldots, Q-1\}$, besteht also aus n Qubits. Register 2 nimmt die Funktionswerte von f auf. Die Schritte sind dann wie folgt.

1. Der Initialzustand ist
$$|0\rangle \otimes |0\rangle.$$

2. Anwendung der Fouriertransformation $F_{\mathbb{Z}/Q\mathbb{Z}}$ auf das erste Register liefert
$$Q^{-1/2} \sum_{x=0}^{Q-1} |x\rangle \otimes |0\rangle.$$

3. Berechnung der Funktionswerte f,
$$Q^{-1/2} \sum_{x=0}^{Q-1} |x\rangle \otimes |f(x)\rangle.$$
Hierbei ist f als Potenzierungsabbildung konkret gegeben, und kann bereits klassisch in Polynomzeit berechnet werden.

4. Messung im zweiten Register und ab jetzt nur noch erstes Register betrachten. Liefert
$$A^{-1/2} \sum_{j=0}^{A-1} |x_0 + jr\rangle$$
für ein (eindeutig bestimmtes) $x_0 \in \mathbb{Z}$ mit $0 \leq x_0 < r$ und $A \in \mathbb{Z}$ maximal mit $x_0 + (A-1)r \leq Q-1$. Dies heißt $A \leq (Q-1-x_0)/r + 1$, also ergibt sich $A = \lfloor Q/r \rfloor$ (für $x_0 = r-1$) oder $A = \lceil Q/r \rceil$ (für $x_0 = 0$).

5. Anwendung der Fouriertransformation $F_{\mathbb{Z}/Q\mathbb{Z}}$ liefert
$$(QA)^{-1/2} \sum_{y=0}^{Q-1} e^{2\pi i x_0 y} \sum_{j=0}^{A-1} e^{2\pi i j r y/Q} \, |y\rangle.$$

6. Schließlich Messung. Liefert
$$|y\rangle$$
mit $0 \leq y < Q-1$.

Die Frage ist, ob wir mit y wie im allgemeinen Verfahren etwas nützliches erfahren. Die Wahrscheinlichkeit, ein vorgegebenes y_0 zu erhalten, ist

$$\Pr(y = y_0) = \frac{A}{Q} \left| A^{-1} \sum_{j=0}^{A-1} e^{2\pi i j r y_0/Q} \right|^2 .$$

Zum Vergleich mit dem allgemeinen Verfahren betrachten wir den idealen Fall $r|Q$, wo r eine Untergruppe von $\mathbb{Z}/Q\mathbb{Z}$ definiert. Dann ist $\Pr(y = y_0) = 0$, außer wenn $y_0/Q = m/r$ beziehungsweise $y_0 = m(Q/r)$ für ein $m \in \mathbb{Z}$ mit $0 \leq m < r$ gilt. In diesem Fall ist $\Pr(y = y_0) = A/Q = 1/r$ und damit $\Pr(y \in Y_0) = r(1/r) = 1$ für $Y_0 = \{y_0 \,|\, \exists m : y_0/Q = m/r\}$. Dies ist wieder die Aussage der Orthogonalitätsrelationen. Die Summanden in der Summe sind übrigens von der Form $\chi_{y_0}(jr)$.

Zurück zu $r \nmid Q$ ergeben sich die Orthogonalitätsrelationen nur approximativ. Wenn $y_0/Q \approx m/r$, dann wird sich ergeben, daß auch $\Pr(y = y_0) \gtrapprox 1/r$ gilt. Ferner gibt es für $\approx r$ viele $m \in \mathbb{Z}$ mit $0 \leq m < r$ jeweils ein $y_0 \in \mathbb{Z}$ mit $0 \leq y_0 < Q$ mit $y_0/Q \approx m/r$. Mit $A/Q \approx 1/r$ und $Y_0 = \{y_0 \,|\, \exists m : y_0/Q \approx m/r\}$ folgt damit $\Pr(y \in Y_0) \approx r(1/r) = 1$. Im folgenden werden die $\approx$ genauer spezifiziert.

Im Intervall $[0, 1)$ liegen r Werte m/r, für $0 \leq m < r$. Um jedes m/r betrachten wir eine Umgebung $(m/r - 1/(2Q), m/r + 1/(2Q)]$. Wegen $Q \gg r$ sind diese Umgebungen disjunkt und in $[0, 1)$ enthalten. Außerdem gibt es in jeder Umgebung genau ein Element der Form y_0/Q mit $y_0 \in \mathbb{Z}$ und $0 \leq y_0 < Q$, da die Breite der Umgebung $1/Q$ beträgt. Zusammenfassend macht das r verschiedene Elemente y_0/Q mit

$$\left| \frac{y_0}{Q} - \frac{m}{r} \right| \leq \frac{1}{2Q},$$

korrespondierend zu $m \in \mathbb{Z}$ mit $0 \leq m < r$. Diese Approximation ist oben mit $y_0/Q \approx m/r$ gemeint. Die Menge Y_0 wird entsprechend definiert.

Für solche y_0 und für r, Q ausreichend groß ergibt sich mit $\alpha = y_0/Q - m/r$, $|\alpha| \leq 1/(2Q)$ folgende untere Abschätzung

$$\begin{aligned} \left| \sum_{j=0}^{A-1} e^{2\pi i j r y_0/Q} \right| &= \left| \sum_{j=0}^{A-1} e^{2\pi i j r \alpha} \right| \\ &\geq \left| \sum_{j=0}^{A-1} e^{2\pi i j r/(2Q)} \right| = \left| \frac{1 - e^{\pi i A r/Q}}{1 - e^{\pi i r/Q}} \right| \\ &\geq \frac{2 - \varepsilon}{\pi r/Q}. \end{aligned}$$

Der Ausdruck $2-\varepsilon$ bezeichnet eine beliebige, konstante Zahl, welche etwas kleiner 2 ist. aus Für die erste Ungleichung ist zu beachten, daß $|2\pi jr\alpha| \leq \pi$ gilt. Die Summe der Einheitsvektoren überstreicht daher einen Winkel $\leq \pi$, und eine Auffächerung auf einen Winkel immer noch $\leq \pi$ ergibt einen kürzeren Summenvektor. Die zweite Gleichung gilt nach der Formel für geometrische Reihen. Die zweite Ungleichung gilt, da zum einen $1-e^{\pi iAr/Q} \approx 2$ wegen $(1/r)(1-r/Q) \leq A/Q \leq (1/r)(1+r/Q)$ gilt, und zum anderen $e^{\pi ir/Q} \approx 1$ und somit die Bogenlänge $\pi r/Q$ länger als der Abstand von $e^{\pi ir/Q}$ und 1 ist. Dann folgt

$$\frac{4-\varepsilon}{\pi^2 r^2}\frac{Q}{A} \;\leq\; \Pr(y=y_0) \;\leq\; \frac{A}{Q}.$$

Für Q groß liegt A/Q nahe bei $1/r$. Für r groß sind die $\Pr(y=y_0)$ im wesentlichen gleichverteilt. Da die y_0 paarweise verschieden sind, summieren sich diese Wahrscheinlichkeiten auf, wenn man nur an dem Ereignis $y \in Y_0$ interessiert ist. Zusammenfassend erhalten wir folgenden Satz.

2 Satz. *Für r groß genug und $Q \gg r^2$ berechnen wir in einem Durchlauf des oben beschriebenen Quantenalgorithmus mit einer Wahrscheinlichkeit von mindestens*

$$\frac{1}{\pi}$$

ein $y_0 \in \mathbb{Z}$ mit $0 \leq y_0 < Q$ und

$$\left|\frac{y_0}{Q} - \frac{m}{r}\right| \leq \frac{1}{2Q},$$

für ein $m \in \mathbb{Z}$ mit $0 \leq m < r$. Die y_0 bzw. m werden dabei bis auf Faktoren der Größe π gleichverteilt aus den jeweiligen Mengen Y_0 bzw. $\{0,\ldots,r-1\}$ gewählt.

Beweis. Folgt aus der obigen Diskussion. Die Schranken für $\Pr(y=y_0)$ liegen um einen Faktor von höchstens π auseinander, und darüberhinaus gilt auch $\Pr(y=y_0) \geq 1/(\pi r)$. □

Soweit der Quantenteil des Algorithmus. Wir beschreiben nun, wie Information über r aus Q und einer Reihe von berechneten y_0 (mit klassischen Algorithmen effizient) erhalten werden kann.

Wir haben gesehen, daß y_0/Q unter den Elementen aus Y_0 durch m/r aufgrund der Intervallbreite und $Q \gg r$ eindeutig bestimmt wird. Umgekehrt wird aber auch m/r unter den Elementen $\{0, 1/r, \ldots, (r-1)/r\}$ durch y_0/Q eindeutig bestimmt. Die Forderung $Q \gg r^2$ bewirkt nun aber eine wesentlich stärkere Eindeutigkeitsaussage.

3 Lemma. *Für von m/r verschiedene rationale Zahlen m'/r' mit $0 < rr' < Q$ gilt*

$$\left|\frac{y_0}{Q} - \frac{m'}{r'}\right| > \frac{1}{2Q}.$$

Beweis. Angenommen, es gälte $\leq$. Dann folgt $|m/r - m'/r'| \leq |m/r - y_0/Q| + |m'/r' - y_0/Q| \leq 1/Q$. Auf der anderen Seite gilt aber auch $|m/r - m'/r'| = |(mr' - m'r)/(rr')| \geq 1/(rr') > 1/Q$, Widerspruch. □

Gilt nun $Q \gg r^2$, so ist m/r eindeutig unter allen Brüchen mit Nenner $\leq r$ durch y_0/Q bestimmt. Wie berechnet man m/r ausgehend von y_0/Q? Das Standardverfahren hierzu heißt Kettenbruchentwicklung und kann in den anderen Skripten nachgelesen werden. Hier soll ein gittertheoretischer Ansatz verfolgt werden (was im übrigen im Prinzip auf die gleiche Methode herausläuft). Wir betrachten das Gitter $\Lambda \subseteq \mathbb{R}^3$ der Dimension 2, welches von den Vektoren $(y_0, 1, 0)$ und $(Q, 0, 1)$ erzeugt wird.

4 Lemma. *Für $Q \gg r^2$ ist das erste Element v einer LLL-reduzierten Basis von Λ von der Form $\pm v = (r_1 y_0 - m_1 Q, r_1, -m_1)$, wobei $r_1 > 0$,* $\gcd\{r_1, m_1\} = 1$ *und $m_1/r_1 = m/r$ gilt.*

Beweis. Seien $m_1, r_1 > 0$ teilerfremd mit $m_1/r_1 = m/r$. Dann ist $u = (r_1 y_0 - m_1 Q, r_1, -m_1)$ ein erstes Basiselement der Länge $O(r)$.

Sei nun $w \in \Lambda$ mit $w \notin \mathbb{Z}u$. Dann gibt es $m_2, r_2 \in \mathbb{Z}$, $r_2 > 0$ mit $\pm w = (r_2 y_0 - m_2 Q, r_2, -m_2)$ und $m_2/r_2 \neq m_1/r_1$. Mit $a = |r_2 y_0 - m_2 Q|$ folgt $|y_0/Q - m_2/r_2| \leq a/(Qr_2)$ und wie eben $1/(r_1 r_2) \leq a/(Qr_2) + 1/(2Q)$ beziehungsweise $r_2 \geq 2Q/r_1 - 2a$. Unabhängig vom Wert von a gilt dann $||w|| \geq (a^2 + (2Q/r_1 - 2a)^2)^{-1/2} = \Omega(Q/r)$ (man kann den Ausdruck beispielsweise in a minimieren). Wegen $Q \gg r^2$ ist daher w mindestens 5 mal länger als u. Da v das erste sukzessive Minimum mit Faktor < 5 approximiert, muß $v = \pm u$ gelten. □

Schließlich wenden wir und der Frage zu, wie man aus m_1/r_1 den Wert von r bestimmen kann. Im allgemeinen können wir nur r_1 bestimmen und wissen dann $r_1 | r$. Wenn m und r jedoch teilerfremd sind, dann folgt $r = r_1$. Allgemeiner gilt $\mathbb{Z}(1/r) = \mathbb{Z}(m_1/r_1) + \cdots + \mathbb{Z}(m_s/r_s)$, also $r = \operatorname{lcm}\{r_1, \ldots r_s\}$, wenn die zu den m_i/r_i gehörigen m's teilerfremd sind.

5 Lemma. *Die Wahrscheinlichkeit, daß zwei zufällig, gleichverteilt und unabhängig gewählte ganze Zahlen ≥ 0 und $< r$ teilerfremd sind, ist größer als*

$$2 - \zeta(2) \approx 0.355.$$

Beweis. Es gibt r^2 viele mögliche Paare (m_1, m_2) mit $0 \leq m_i < r$. Die Anzahl der m_1, die durch ein j mit $1 < j < r$ teilbar sind, ist höchstens r/j. Die Anzahl der Paare, für die $j|m_1$ und $j|m_2$ mit irgendeinem $1 < j < r$ gilt, beträgt höchstens $\sum_{j=2}^{r-1}(r/j)^2$. Die Wahrscheinlichkeit, daß m_1 und m_2 nicht teilerfremd sind, beträgt damit höchstens

$$\sum_{j=2}^{r-1} j^{-2} \leq \zeta(2) - 1.$$

Die Wahrscheinlichkeit, daß m_1 und m_2 teilerfremd sind, beträgt damit mindestens $1 - (\zeta(2) - 1) = 2 - \zeta(2) = 2 - \pi^2/6 \approx 0.355$. □

Eine genauere Analyse der Aussage des Lemmas zeigt, daß die Wahrscheinlichkeit ungefähr $1/\zeta(2) \approx 0.607$ beträgt.

Wir fassen den Quantenalgorithmus zur Faktorisierung zusammen. Wir berechnen m_1/r_1 und m_2/r_2 mit $r_1|r$ und $r_2|r$ mit Wahrscheinlichkeit $\geq 1/\pi^2$. Wegen der (ungefähren) Gleichverteilung der m's gilt $r = \mathrm{lcm}\{r_1, r_2\}$ mit Wahrscheinlichkeit $\geq 1/3$. Insgesamt berechnen wir also r mit Wahrscheinlichkeit $1/(3\pi^2)$. Die Korrektheit von r kann dann leicht getestet werden. Bei inkorrektem r wird der Algorithmus wiederholt.