SHOR'S ALGORITME FOR KVANTE FAKTORISERING

NIELS NYGAARD

Det er velkendt at mens det er meget nemt at få en computer til at gange to tal sammen er det meget sværere at gå den anden vej, at få en computer til at faktorisere et tal i primfaktorer. Det at det er så svært at faktorisere bruges i de fleste krypterings algoritmer til at beskytte kommunikationer over internettet. Den hidtil bedste algoritme, der går under navnet "kvadratisk si"bruger således omkring $\exp(\left(\frac{64}{9}\right)^{\frac{1}{3}} N^{\frac{1}{3}} (\log N)^{\frac{2}{3}})$ operationer på at finde ikke trivielle faktorer i et tal N. Som det ses vokser antallet af operationer eksponentielt med N. De største tal der bruges i kryptering er af størrelsesordenen 2^{2000} og det er derfor ikke realistisk at faktorisere dem selv ved hjælp af de hurtigste super-computere.

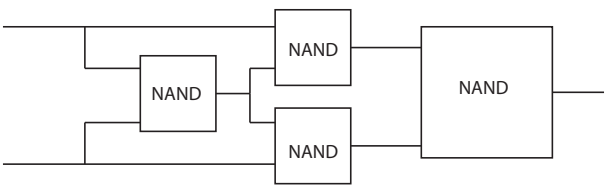
En digital computer virker som bekendt ved at repræsentere tal i binær form, som en følge af 0 og 1, kaldet "bits", f.eks. $2 = 10, 3 = 11, \ldots, 20 = 10100, \ldots$. Et tal af størrelsesordenen ovenfor skal derfor bruge omkring 2000 bits for at repræsenteres i binær form, men da en digital computer typisk har lagringskapacitet af flere milliarder bits er det ikke et problem at arbejde med sådanne tal. En moderne computer behandler data i bidder (som kaldes "ord") af længde 64 bits og processen foregår ved at sende data igennem en række logiske porte som kan implementere multiplikation, addition og andre aritmetiske operationer. Et eksempel på en logisk port er "Not And"eller "NAND"port som tager to bits ind og spytter en enkelt bit ud efter reglen

$$00 \mapsto 1, 01 \mapsto 1 \tag{1}$$

$$10 \mapsto 1, 11 \mapsto 0 \tag{2}$$

Hvis vi opfatte 0 som "falsk"og 1 som "sandt"er det præcis sandhedstavlen for "not (a and b)". Det viser sig at man ved at sammensætte NAND porte på forskellig vis kan konstruere alle logiske porte.

Når vi nu skal tale om kvante fænomener, er et af dem det meget karakteristiske at selv partikler kan interferere med hinanden. Et af de fænomener som førte til opdagelsen af kvantemekanikken var lige præcis et eksperiment hvor en stråle af elektroner som sendes igennem to spalter udviser et interferens mønster fuldstændig som en lysstråle. Det betyder at et system af flere partikler ikke bare opfører sig som en samling enkelt partikler, men at systemet har sin egen opførsel som ikke



FIGUR 1. XOR port sammensat at NAND porte

kan bestemmes ud fra viden om de enkelte partikler. Man taler om at systemet er en "overlejring"af de enkelte partiklers bølgenatur.

Tilsvarende i den matematiske model af en kvante computer, som jo er alt hvad vi har idag, er de enkelte celler ikke bare en bit, men en overlejring af de to bits 0 og 1. Når vi taler om bits og ord i kvante computer sammenhæng bruger vi en notation fra kvantemekanik og skriver dem som "ket"vektorer dvs. we sætter dem in i paranteser | og ⟩ som f.eks. $|0\rangle$, $|001101001\rangle$. Den fundamentale enhed i en kvante computer er således en "qbit"som ikke kun kan være i tilstand 0 eller 1, men kan være i en overlejrings tilstand

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

hvor α og β er complekse tal som opfylder

$$|\alpha|^2 + |\beta|^2 = 1$$

Et system af n qbits kan vœre i en overlejrings tilstand

$$\alpha_0|000\ldots0\rangle + \alpha_1|00\ldots01\rangle + \cdots + \alpha_{2^n-1}|111\ldots1\rangle$$

hvor ligeledes

$$|\alpha_0|^2 + |\alpha_1|^2 + \cdots + |\alpha_{2^n-1}|^2 = 1$$

En anden af mærkværdighederne ved kvante mekanikken er at når man måler på et system så influerer selve det at fortage målingen på systemets tilstand. Hvis vi måler en enkelt qbit i tilstand $\alpha|0\rangle + \beta|1\rangle$ så får vi resultatet $|0\rangle$ med sandsynlighed $|\alpha|^2$ og resultatet $|1\rangle$ med sandsynlighed $|\beta|^2$. Efter målingen er vores qbit i tilstand $|0\rangle$ hvis resultatet af målingen var $|0\rangle$ og i tilstand $|1\rangle$ hvis målingen gav $|1\rangle$. Hvis man så måler igen får man med 100% det samme resultat som den første måling.

For et n-qbit system kan man måle f.eks. de sidste m qbits. Hvis vi skriver en basis tilstand som $|x\rangle|y\rangle$ hvor $|x\rangle$ en basis tilstand af de første $n-m$ qbits og $|y\rangle$ en basis tilstand af de sidste m qbits. Så kan vi skrive den overlejrede tilstand på formen

$$\sum_{xy} \alpha_{xy}|x\rangle|y\rangle$$

with

$$\sum_{xy} |\alpha_{xy}|^2 = 1$$

Hvis vi nu måler de sidste m qbits får vi en basis tilstand $|z\rangle$ med sansynlighed

$$Prob(z) = \sum_x |\alpha_{xz}|^2$$

Efter målingen er de sidste m qbits i basis tilstanden $|z\rangle$ og de første $n-m$ qbits i den overlejrede tilstand

$$\frac{1}{\sqrt{\sum_x |\alpha_{xz}|^2}} \sum_x \alpha_{xz}|x\rangle$$

Hvis vi f.eks. betragter et 3-qbit system i tilstanden

$$\frac{1}{\sqrt{3}}|000\rangle + \frac{i}{\sqrt{3}}|010\rangle - \frac{1}{\sqrt{3}}|011\rangle$$

og måler den sidste qbit så får vi resultatet $|0\rangle$ med sandsynlighed $\frac{2}{3}$ og $|1\rangle$ med sandsynlighed $\frac{1}{3}$. Efter målingen er 2-qbit systemet af de første to qbits i første tilfœlde i tilstanden

$$\sqrt{\frac{3}{2}}\left(\frac{1}{\sqrt{3}}|00\rangle + \frac{i}{\sqrt{3}}|01\rangle\right) = \frac{1}{\sqrt{2}}|00\rangle + \frac{i}{\sqrt{2}}|01\rangle$$

og i det andet tilfœlde i basis tilstanden $|01\rangle$.

I den matematiske model af kvante computeren er en port en $2^n \times 2^n$ unitær matrix U dvs. en matrix der opfylder $\overline{U}^t U = Id$. Det betyder bl.a. at søjlerne er 2^n dimensionale komplekse vektorer med norm $= 1$.

Et simpelt eksempel er den såkaldte Hadamard port som er givet ved matricen $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$ som virker på en enkelt qbit ved at

$$\alpha|0\rangle + \beta|1\rangle \mapsto \alpha\left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\right) + \beta\left(\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle\right) = \frac{\alpha+\beta}{\sqrt{2}}|0\rangle + \frac{\alpha-\beta}{\sqrt{2}}|1\rangle$$

Vi kan lade Hadamard porten virke på en n-dimensional basistilstand $|a_n a_{n-1} \dots a_1\rangle$ ved $H^{(n)}|a_n a_{n-1} \dots a_1\rangle = |Ha_n\rangle|Ha_{n-1}\rangle \dots |Ha_1\rangle$. For eksempel

$$\begin{aligned} H^{(2)}|01\rangle &= |H|0\rangle\rangle|H|1\rangle\rangle = |\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\rangle|\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle\rangle \\ &= \frac{1}{2}|00\rangle - \frac{1}{2}|01\rangle + \frac{1}{2}|10\rangle + \frac{1}{2}|11\rangle \end{aligned}$$

Hvis vi anvender $H^{(n)}$ på basistilstanden $|000\dots0\rangle$ får vi

$$\begin{aligned} H^{(n)}|000\dots0\rangle &= |H|0\rangle\rangle|H|0\rangle\rangle|H|0\rangle\rangle \dots |H|0\rangle\rangle \\ &= |\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\rangle|\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\rangle \dots |\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\rangle \\ &= \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x\rangle \end{aligned}$$

dvs. vi har fået en uniform distribution over alle tallene fra 0 til $2^n - 1$. Hvis vi måler denne tilstand får vi et virkeligt tilfældigt tal (og ikke bare pseudo-tilfældigt) mellem 0 og $2^n - 1$.

Nu skal vi se lidt på den algebra der ligger til grund for Shor's algoritme.

Lad $M = pq$ hvor p og q er to forskellige primtal. Vores opgave er givet M at finde p og q.

Lad a være et vilkårligt tal som er indbyrdes primisk med M altså $(a, M) = 1$ og betragt undergruppen af gruppen af enheder $(\mathbb{Z}/M)^*$ frembragt af a, altså den cycliske undergruppe $< a > \leq (\mathbb{Z}/M)^*$. Lad r være ordenen a a så $a^r \equiv 1 \mod M$. Antag at r er et lige tal. Man kan vise at sandsynligheden for at et vilkårligt valgt tal primisk med M har lige orden i gruppen $(\mathbb{Z}/M)^*$ er $> \frac{1}{2}$. Vi betragter nu $b = a^{r/2}$ så er $b^2 \equiv 1 \mod M$ og derfor gælder at $M|(b-1)(b+1)$. Fra den kinesiske restsætning ved vi at $\mathbb{Z}/M \cong \mathbb{Z}/p \times \mathbb{Z}/q$. I $\mathbb{Z}/p \times \mathbb{Z}/q$ er der præcis 4 elementer der tilfredstiller $b^2 = 1 = (1,1)$. De fire elementer er $(1,1), (1,-1), (-1,1), (-1,-1)$, da $b \neq 1$ er det første element udelukket så der er en sandsynlighed på $2/3$ for at b er enten $(1,-1)$ eller $(-1,1)$. Sandsynligheden for at et vilkårligt tal tilfredsstiller begge betingelser, altså r er lige og b er enten $(1,-1)$ eller $(-1,1)$ er så mindst $1/3$. For et sådant b gælder at M hverken dividerer $b-1$ eller $b+1$, derfor er både $(M, b+1)$ og $(M, b-1)$ ikke trivielle faktorer i M.

Her er et eksempel:

$M = 21$, $a = 5$. $< a >= \{5, 4, 20, 16, 17, 1\}$ så a har orden 6. $a^3 = 20 \equiv -1$ så a tilfredsstillller den første betingelse , men ikke den anden.

Hvis $a = 10$ er $< a >= \{10, 16, 13, 4, 19, 1\}$, så igen er ordenen af a lig med 6. Her er $a^3 = b = 13$ og $(b-1, M) = (12, 21) = 3$ og $(b+1, M) = (14, 21) = 7$.

Bemærk at den største fælles divisor er meget hurtig at beregne ved hjælp af Euclid's algoritme.

Det der er svært at beregne, er ordenen af a i gruppen $(\mathbb{Z}/M)^*$. Vi kan omskrive problemet lidt ved at definere funktionen $f : \mathbb{Z} \to \mathbb{Z}$, $f(k) = a^k \mod M$. Dette er en periodisk funktion med periode $= r$, altså $f(k + rj) = f(k)$, så problemet med at beregne r er ækvivalent med at beregne perioden a funktionen f.

Vi betragter nu et et system af n qbits hvor $N = 2^n > M^2$, så hvis vi vil faktorisere er tal af størresordenen 2^{2000} skal vi bruge lidt over 4000 qbits, det er et relativt lille antal når man tænker på at en moderne digital computer har milliarder a bits.

Lad os nu antage at vi har en port som beregner funktionen f, dvs. en unitær matrix U_f så $U_f|x\rangle|000\ldots0\rangle = |x, f(x)\rangle$ som virker på ord af længde $2n$ så U_f er en $2N \times 2N$ matrix.

Vi starter med basistilstanden $|000\ldots00\rangle$ af længde $2n$ og anvender $H^{(n)}$ på de første n qbits. Det giver os overlejringen

$$\frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle|000\ldots00\rangle$$

så når vi anvender U_f får vi overlejringen

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

Vi har så lige pludselig et system af $2n$ qbits som lagrer alle f's værdier. Skulle vi lagre alle disse værdier i en digital computer skulle vi bruge $n \cdot 2^n$ bits, det er et tal som er mange gange større end de samlede antal partikler i hele universet. Det er denne massive parallelisme der gør at en kvante computer vil blive så meget hurtigere end en digital computer. Selv om vi har lagret alle værdierne kan vi ikke få fat i dem, hvis vi måler får vi kun en vilkårlig værdi, så derfor skal vi være ret kløgtige for at få noget nyttigt ud af denne overlejringstilstand.

Nu måler vi de sidste n qbits. Resultatet er af formen $|f(x_0)\rangle$. Da f er periodisk med periode r kan vi antage at $0 \leq x_0 < r$. Det efterlader de første n qbits i overlejrings tilstanden

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

hvor A er bestemt ved at $x_0 + r(A-1) < N$ og $x_0 + Ar \geq N$ så $\frac{N}{r} - 1 \leq \frac{x_0}{r} + (A-1) < \frac{N}{r}$. Da $0 \leq \frac{x_0}{r} < 1$ får vi

$$A - 1 < \frac{N}{r} < A + 1$$

Vi kan ikke aflæse perioden fra denne overlejringstilstand, måling giver os kun en enkelt værdi og vi kender hverken x_0 eller j. Vi bliver derfor nødt til at finde på noget andet. Det vi skal bruge er den diskete Fourier transformation DFT som er defineret ved at den tager en basistilstand af længde m, $|x\rangle$ og sender den i overlejrings tilstanden

$$DFT|x\rangle = \frac{1}{2^{m/2}} \sum_y e^{2\pi i \frac{x \cdot y}{2^m}} |y\rangle$$

Den diskrete Fourier transformation er også defineret ved en unitær matrix of definerer derfor en port i kvantecomputeren.

Vi anvender denne port på den målte tilstand

$$DFT\left(\frac{1}{\sqrt{A}}\sum_{j=0}^{A-1}|x_0+rj\rangle\right)=\frac{1}{\sqrt{N}}\frac{1}{\sqrt{A}}\sum_{j=0}^{A-1}\sum_{y=0}e^{2\pi i\frac{x_0\cdot y+rjy}{N}}|y\rangle$$
$$=\frac{1}{\sqrt{AN}}\sum_{y=0}^{N-1}e^{2\pi i\frac{x_0\cdot y}{N}}\sum_{j=0}^{A-1}e^{2\pi i\frac{rjy}{N}}|y\rangle$$

Nu kan vi måle og vi får resultatet $|y\rangle$ med sandsynlighed

$$Prob(y)=\frac{1}{AN}\left|\sum_{j=0}^{A-1}e^{2\pi i\frac{rjy}{N}}\right|^2$$

We have

$$\sum_{j=0}^{A-1}e^{2\pi i\frac{rjy}{N}}=\sum_{j=0}^{A-1}(e^{2\pi i\frac{ry}{N}})^j=\frac{e^{2\pi i\frac{ry}{N}A}-1}{e^{2\pi i\frac{ry}{N}}-1}=\frac{e^{i\theta_y A}-1}{e^{i\theta_y}-1}$$

hvor

$$\theta_y=2\pi\frac{ry}{N}$$

Vi betragter nu mængderne $\{0,N,2N,\ldots,rN\}$ og $\{0,r,2r,3r,\ldots,Nr\}$. For ethvert multiplum aN hvor $a\leq r$, findes der prœcis et y så $yr\leq aN\leq(y+1)r$ det betyder at for ethvert aN findes der et y så $|aM-yr|\leq\frac{r}{2}$. For ethvert multiplum aN vœlger vi et sådant y og vi har dermed r forskellige y'er med denne egenskab og for disse r vœrdier af y gœlder

$$-\frac{r}{2}\leq yr-kN\leq\frac{r}{2}$$

Det er klart at $\frac{e^{i\theta_y A}-1}{e^{i\theta_y}-1}$ kun afhœnger af $yr \mod N$ så hvad angår beregningen af denne størrelse kan vi antage at

$$-\frac{r}{2}\leq yr\leq\frac{r}{2}$$

og derfor

$$-\frac{r}{N}\pi\leq 2\pi\frac{yr}{N}\leq\frac{r}{N}\pi$$

Vi ved allerede at $A-1<\frac{N}{r}$ så $\frac{r}{N}(A-1)<1$ og dermed gœlder at $-\pi<\theta_y j<\pi$ for $j\leq A-1$. Nu har vi

$$\left|\frac{e^{i\theta_y A}-1}{e^{i\theta_y}-1}\right|=\left|\frac{e^{i\theta_y(A-1)}-1}{e^{i\theta_y}-1}+e^{i\theta_y(A-1)}\right|\geq\left|\frac{e^{i\theta_y(A-1)}-1}{e^{i\theta_y}-1}\right|-1$$

. I intervallet $(0,\frac{\pi}{A-1})$ er funktionen $\theta_y\mapsto|e^{i\theta_y(A-1)}-1|$ en konkav funktion og den ligger derfor altid over korden som er linien mellem $(0,0)$ og punktet $(\frac{\pi}{A-1},2)$ dvs.

$$|e^{i\theta_y(A-1)}-1|\geq\frac{2(A-1)}{\pi}\theta_y$$

På den anden side er $|e^{i\theta_y}-1|\leq\theta_y$ så alt i alt har vi

$$\left|\frac{e^{i\theta_y A}-1}{e^{i\theta_y}-1}\right|\geq\frac{2(A-1)}{\pi}-1=\frac{2A}{\pi}-(1+\frac{2}{\pi})$$

og derfor

$$Prob(y) \geq \frac{1}{AN}\left(\frac{2A}{\pi} - (1+\frac{2}{\pi})\right)^2 \geq \frac{A}{N}\frac{4}{\pi^2} = \frac{1}{r}\frac{4}{\pi^2} + \frac{1}{N}\frac{4}{\pi^2}$$

Da $\frac{4}{N\pi^2} \simeq 0$ er sandsynligheden for at udfaldet af målingen er en af de r tilstande vi har udvalgt af størrelsesordenen $r\frac{1}{r}\frac{4}{\pi^2} \simeq 0.40$

For enhver af disse værdier af y gælder

$$\frac{k}{r} - \frac{1}{2N} \leq \frac{y}{N} \leq \frac{k}{r} + \frac{1}{2N}$$

Nu er der højst et rationalt tal med nævner $< M$ der ligger inden for en afstand $\leq \frac{1}{2N}$ af $\frac{y}{N}$. Det er klart fordi afstanden mellem to forskellige rationale tal med nævner $< M$, $\frac{a}{b}$ og $\frac{c}{d}$ er $|\frac{ad-bc}{bd}| \geq \frac{1}{M^2}$ og $\frac{1}{M^2} > \frac{1}{2N}$

Dette rationale tal kan beregnes ved hjælp af kædebrøker, men det vil vi ikke komme ind på her.

Hvis nu $(k,r) = 1$ har vi fundet r, hvis ikke finder vi kun en faktor i r. Vi kan så køre beregningen igen og få et andet rationalt tal $\frac{k'}{r}$. Antag at $(k,k') = 1$. Lad $r = p_1^{n_1}p_2^{n_2}\dots p_m^{n_m}$ (nogen af eksponenterne kan være 0) være primtals faktoriseringen af r og $k = p_1^{\ell_1}p_2^{\ell_2}\dots p_u^{\ell_u}$, $k' = p_{u+1}^{\ell_{u+1}}p_{u+2}^{\ell_{u+2}}\dots p_m^{\ell_m}$. Hvis r_1 og r_2 er nævnerne i de rationale tal $\frac{k}{r}$ og $\frac{k'}{r}$ efter at vi har forkortet de fælles faktorer ud, så har vi $r_1 = p_1^{n_1-\min(n_1,\ell_1)}p_2^{n_2-\min(\ell_2,n_2)}\dots p_u^{n_u-\min(n_u,\ell_u)}p_{u+1}^{n_{u+1}}\dots p_m^{n_m}$ og $r_2 = p_1^{n_1}p_2^{n_2}\dots p_u^{n_u}p_{u+1}^{n_{u+1}-\min(\ell_{u+1},n_{u+1})}\dots p_m^{n_m-\min(\ell_m,n_m)}$. Vi får derfor at $r = [r_1,r_2]$, det mindste fælles multiplum.

Hvad er sandsynligheden for at to vilkårlige hele tal k og k' er indbyrdes primiske? Et primtal p dividerer et givet tal med sandsynligheden $\frac{1}{p}$. Sandsynligheden for at p dividerer to givne tal er derfor $\frac{1}{p^2}$. Sandsynligheden for at p ikke dividerer begge tal er så $(1-p^2)$. Vi får at sandsynligheden for at ingen primtal dividerer begge tallene er $\prod_p(1-p^2) = \zeta(2) \simeq 0.607$.

Ved at køre algoritmen et vist antal gange, uafhængigt af størrelsen af M, kan vi derfor med høj sandsynlighed finde r og der er meget hurtigt at beregne om det vilkårlige tal a og den beregnede orden r, virker. Sandsynligheden for at finde a og r der virker, er uafhængig af M og vi kan derfor ved at køre algoritmen et bestemt antal gange med meget høj sandsynlighed, finde en løsning