Uniwersytet Jagielloński
Wydział Matematyki, Fizyki i Informatyki
Instytut Informatyki

Zastosowanie algorytmów kwantowych dla problemów grup rozwiązalnych

Iwona Cieślik

Praca magisterska
Promotor: dr Tomasz Gorazd

Kraków 2002

Pragnę wyrazić wdzięczność
mojemu promotorowi dr Tomaszowi Gorazdowi
za pomoc okazaną w trakcie przygotowania tej pracy,
cenne wskazówki i osobiste zaangażowanie.

Wstęp

Wiek XX jest z pewnością wiekiem fizyki kwantowej, informacji i komputerów. Niejako ze scalenia tych trzech dziedzin nauki wynurzyła się nowa, niezwykła i pełna zagadek teoria komputerów kwantowych, czyli urządzeń, które wykonują obliczenia korzystając z praw mechaniki kwantowej. Jest ona bardzo młodą dziedziną nauki liczącą sobie ostatnie kilkanaście lat.

Rzecz całą zapoczątkował słynny fizyk amerykański z Caltech (California Institute of Technology) Richard Feynman, który w 1982 roku wprowadził pojęcie komputera kwantowego - urządzenia kwantowego, za pomocą którego można byłoby efektywnie symulować dowolny inny układ kwantowy w sposób niedostępny dla klasycznych komputerów. Komputer kwantowy do obliczeń wykorzystywałby nie stany napięcia 0 (brak napięcia) i 1 (jest napięcie), ale stany mikrocząsteczek, w tym elektronów. Kilka lat później w 1985 roku fizyk z Uniwersytetu Oksfordzkiego David Deutsch opublikował swoją przełomową pracę, w której zauważył, że dzięki komputerom kwantowym można będzie rozwiązywać problemy trudne dla zwykłych maszynach.

Propozycja ta początkowo nie wzbudzała większego zaciekawienia. Prawdziwe zainteresowanie komputerami kwantowymi pojawiło się dopiero, gdy Peter Shor (1994) przedstawił kwantowy algorytm do faktoryzacji liczb całkowitych (tj. rozkładu liczb na czynniki pierwsze), działający ekspotencjalnie szybciej niż najlepsze algorytmy klasyczne. Niewielki nawet komputer kwantowy mógłby zatem w bardzo krótkim czasie złamać wszystkie kody i zabezpieczenia współczesnych systemów informatycznych (wykorzystujących właśnie duże liczby pierwsze), co dawałoby jego posiadaczom oczywistą i ogromną przewagę. Następnie Lov Grover (1996) przedstawił kwantowy algorytm przeszukiwania tablicy danych w czasie proporcjonalnym do pierwiastka czasu potrzebnego klasycznemu komputerowi do wykonania tego zadania.

Obliczenia kwantowe być może staną się naszą przyszłością. Miniaturyzacja podzespołów komputerowych postępuje bardzo szybko i zbliża się do poziomu mikroskopijnego, gdzie dominują prawa świata kwantowego. A więc, wcześniej czy później, pojawią się przeszkody w udoskonalaniu komputerów klasycznych, związane na przykład z tym, że tranzystory i połączenia elektryczne między nimi nie mogą być cieńsze niż średnica atomu. Keyes (1998) obliczył, że jeśli rozwój miniaturyzacji będzie postępował tak szybko jak dotychczas, około 2020 roku osiągnie poziom atomowy i każdemu bitowi będzie odpowiadał jeden elektron. A więc nie tylko ciekawość naukowców, ale też potrzeba rozwoju technologicznego szuka ratunku w sile obliczeń kwantowych.

Informatyka kwantowa jest więc wielkim wyzwaniem zarówno dla nauki, jak i dla technologii. Jej potencjalne możliwości wydają się imponujące i bardzo obiecujące. Wymienia się możliwość bezpiecznego kodowania informacji, pełne zabezpieczenie przed

hakerstwem (co związane jest z wrażliwością stanów kwantowych na jakąkolwiek ingerencję) czy nawet możliwość realizacji teleportacji kwantowej, czyli przesyłania na odległość stanów kwantowych jako przepisu odbudowywania układów z cząstek dostępnych u odbiorcy (wykonano już eksperymenty z teleportacją fotonów na całkiem duże odległości, rzędu kilometra). Niebywałe możliwości komputerów kwantowych związane są z zupełnie odmienną od klasycznych urządzeń zasadą ich działania. Można je obrazowo porównać do nieskończonej liczby klasycznych procesorów działających równolegle. Ich siła jest oparta na kilku fenomenach i prawach świata kwantowego oferujących radykalnie nowe i silne narzędzia, istotnie różne od tych, z którymi spotykamy się prowadząc obliczenia klasyczne: zespolone amplitudy, kwantowa interferencja, kwantowa równoległość, kwantowe splątanie, unitarność (a więc i odwracalność) wszystkich przekształceń. Aby zrozumieć wszystkie te cechy i umieć stosować je przy projektowaniu algorytmów kwantowych musimy zrozumieć podstawowe własności, na których opiera się mechanika kwantowa. W szczególności konieczne jest poznanie teorii przestrzeni Hilberta, która jest matematyczną strukturą stosowaną do opisu tego, co dzieje się w kwantowym świecie.

W niniejszej pracy chcielibyśmy przedstawić podstawy algorytmiki kwantowej i pokazać wykorzystanie jej do rozwiązywania problemów z zakresu teorii grup. W szczególności zajmiemy się skończonymi grupami rozwiązalnymi. Sercem pracy będzie kwantowy algorytm działający w czasie wielomianowym, obliczający rząd grupy rozwiązalnej. Został on opracowany na podstawie artykułu J.Watrousa ([15]). Kilka innych problemów, takich jak testowanie przynależności elementu do grupy, równość podgrup czy testowanie normalności podgrupy danej grupy, można zredukować do problemu wyznaczania rzędu grupy. Zatem można je rozwiązać kwantowo w czasie wielomianowym. Ubocznym efektem działania naszego algorytmu jest wyznaczenie superpozycji elementów podgrupy rozważanej grupy rozwiązalnej. Dzięki temu do grup ilorazowych grup rozwiązalnych możemy zastosować istniejące już algorytmy kwantowe (np. algorytmy dla grup abelowych).

Pracować będziemy z grupami "black-box", których elementy są jednoznacznie zakodowane przez ciągi binarne o długości pewnego ustalonego n. Dodatkowo podstawowe operacje są wykonywane na tych elementach w czasie jednostkowym. Nie potrafimy rozwiązać klasycznie problemu obliczania rzędu grupy "black-box" w czasie wielomianowym, nawet dla grup abelowych. Najlepszy znany klasyczny algorytm dla tego problemu przedstawił Luks ([10]). Działa on dla rozwiązalnych grup G odwracalnych macierzy nad skończonym ciałem w czasie wielomianowym plus największa liczba pierwsza dzieląca $|G|$. Nasz kwantowy algorytm rozwiązuje ten problem w czasie wielomianowym niezależnie od wielkości liczb pierwszych dzielących $|G|$.

Układ pracy jest następujący: Rozdział 1 stanowi wprowadzenie do obliczeń kwantowych, wyjaśniając podstawowe zagadnienia konstrukcji algorytmów kwantowych. Rozdział 2 zawiera niezbędne informacje dotyczące grup, w tym grup "black-box", a także twierdzenia i dowody wykorzystywane w pracy. Rozdział 3 opisuje główny algorytm obliczania rzędu rozwiązalnej grupy "black-box". Rozdział 4 przedstawia problemy, które mogą być rozwiązane przy wykorzystaniu naszego algorytmu.

Rozdział 1

Wprowadzenie do obliczeń kwantowych

1.1 Model obliczeń kwantowych

Model obliczeń kwantowych został sformułowany po raz pierwszy przez Davida Deutscha i nazywamy go kwantową maszyną Turinga (z ang. Quantum Turing machine - QTM). Jest on kwantowym odpowiednikiem jednego z popularnych matematycznych modeli obliczeń klasycznych, jakim jest probabilistyczna maszyna Turinga (z ang. Probabilistic Turing machine - PTM). Nasze wprowadzenie do obliczeń kwantowych rozpoczniemy od porównania tych dwóch modeli.

Definicja 1 *Jednotaśmowa probabilistyczna maszyna Turinga jest definiowana przez* (Σ, Q, δ)*, gdzie* Σ *jest skończonym alfabetem,* Q *skończonym zbiorem stanów (zawierającym pewien początkowy stan* q_0*), a* δ *funkcją przejścia określoną w następujący sposób*

$$\delta \colon \Sigma \times Q \times \Sigma \times Q \times \{L, R\} \to [0, 1]$$

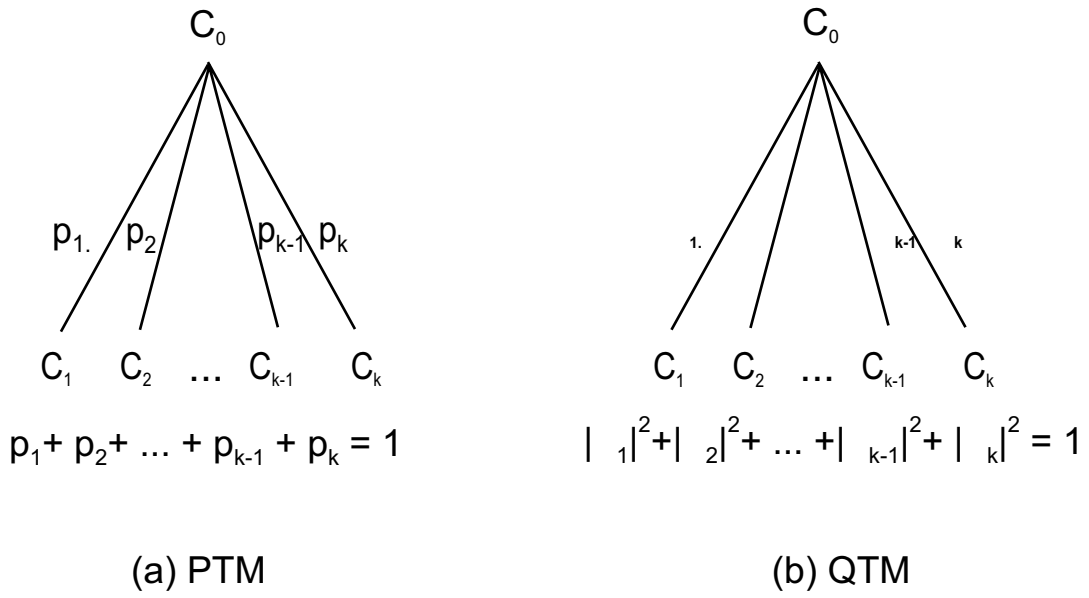
i spełniającą warunek (zwany lokalnym warunkiem prawdopodobieństwa)

(i) Jeśli $(\sigma_1, q_1) \in \Sigma \times Q$ *wówczas*

$$\sum_{(\sigma,q,d)\in\Sigma\times Q\times\{L,R\}} \delta(\sigma_1, q_1, \sigma, q, d) = 1.$$

Funkcja δ przypisuje prawdopodobieństwo każdemu możliwemu przejściu od pewnej konfiguracji c_0 do każdej z konfiguracji, będącej jej następnikiem $c_1, c_2, \ldots, c_k$. Jeśli przez p_i oznaczymy przypisane przez funkcję δ prawdopodobieństwa przejścia od c_0 do c_i wówczas warunek (i) oznacza, że (zob. rysunek 1.1a)

$$\sum_{i=1}^{k} p_i = 1.$$



Rysunek 1.1: *Lokalny warunek prawdopodobieństwa*

Prawdopodobieństwa przypisane przez funkcję przejścia maszyny Turinga możemy reprezentować za pomocą drzewa obliczeń. Każdy węzeł takiego drzewa odpowiada pewnemu stanowi (konfiguracji) maszyny Turinga, a każdy poziom reprezentuje jeden krok obliczeń. Korzeń drzewa odpowiada konfiguracji początkowej, a każdy inny węzeł konfiguracji osiągalnej z niezerowym prawdopodobieństwem w jednym kroku z konfiguracji reprezentowanej przez węzeł będący jego rodzicem. Wartości przypisane węzłom i łukom drzewa ściśle odpowiadają funkcji przejścia δ.

Dodatkowo drzewo obliczeń spełnia prawa rachunku prawdopodobieństwa, wynikające z własności funkcji δ określonej przez (i) w def. 1. Przede wszystkim prawdopodobieństwo każdej ścieżki rozpoczynającej się w korzeniu jest iloczynem prawdopodobieństw przypisanych poszczególnym łukom składającym się na tą ścieżkę. Stąd prawdopodobieństwo każdego węzła odpowiada prawdopodobieństwu osiągnięcia tego węzła w obliczeniu. Może się tak zdarzyć, że na jednym poziomie drzewa jest kilka wystąpień tej samej konfiguracji c: $c^{(1)}, c^{(2)}, \ldots, c^{(k)}$. Wówczas prawdopodobieństwo osiągnięcia tej konfiguracji jest równe sumie prawdopodobieństw odpowiadających konfiguracjom $c^{(1)}, c^{(2)}, \ldots, c^{(k)}$: $p = \sum_{i=1}^{m} p^{(i)}$. W końcu suma prawdopodobieństw wszystkich konfiguracji na poszczególnych poziomach drzewa musi być równa 1, niezależnie od konfiguracji początkowej.

Definicja 2 *Jednotaśmowa kwantowa maszyna Turinga jest definiowana przez* (Σ, Q, δ)*, gdzie* Σ *jest skończonym alfabetem,* Q *skończonym zbiorem stanów (zawierającym pewien początkowy stan* q_0*), a* δ *funkcją przejścia określoną w następujący sposób*

$$\delta \colon \Sigma \times Q \times \Sigma \times Q \times \{L, R\} \to \mathbb{C}_{[0,1]}$$

i spełniająca warunki

(i) Jeśli $(\sigma_1, q_1) \in \Sigma \times Q$ *wówczas*

$$\sum_{(\sigma,q,d)\in\Sigma\times Q\times\{L,R\}} |\delta(\sigma_1, q_1, \sigma, q, d)|^2 = 1$$

(tzw. lokalny warunek prawdopodobieństwa).

(ii) Globalny warunek prawdopodobieństwa (podany jako definicja 3).

Wartości funkcji δ, zwane **amplitudami** (amplitudami prawdopodobieństwa), są liczbami zespolonymi o wartościach bezwzględnych z przedziału $[0, 1]$. Są one przypisane każdemu przejściu maszyny od konfiguracji c_0 do każdej z możliwych konfiguracji $c_1, c_2, \ldots, c_k$ w taki sposób, że jeśli α_i to amplituda odpowiadająca przejściu od konfiguracji c_0 do c_i to $\sum_{i=1}^{k} |a_i|^2 = 1$, co wynika z własności (i) def. 2 oraz prawdopodobieństwo przejścia od c_0 do c_i jest równe $|a_i|^2$ (zob. rysunek 1.1b).

Podobnie jak w przypadku probabilistycznej maszyny Turinga działanie maszyny kwantowej możemy reprezentować przez drzewo obliczeń. Jest ono podobne do drzewa obliczeń klasycznych. Każdy łuk drzewa zamiast z prawdopodobieństwem jest skojarzony z amplitudą. Amplituda węzła jest iloczynem amplitud łuków leżących na ścieżce od korzenia do tego węzła, a łączna amplituda każdej z konfiguracji poziomu i jest sumą amplitud węzłów leżących na i-tym poziomie, odpowiadających tej konfiguracji.

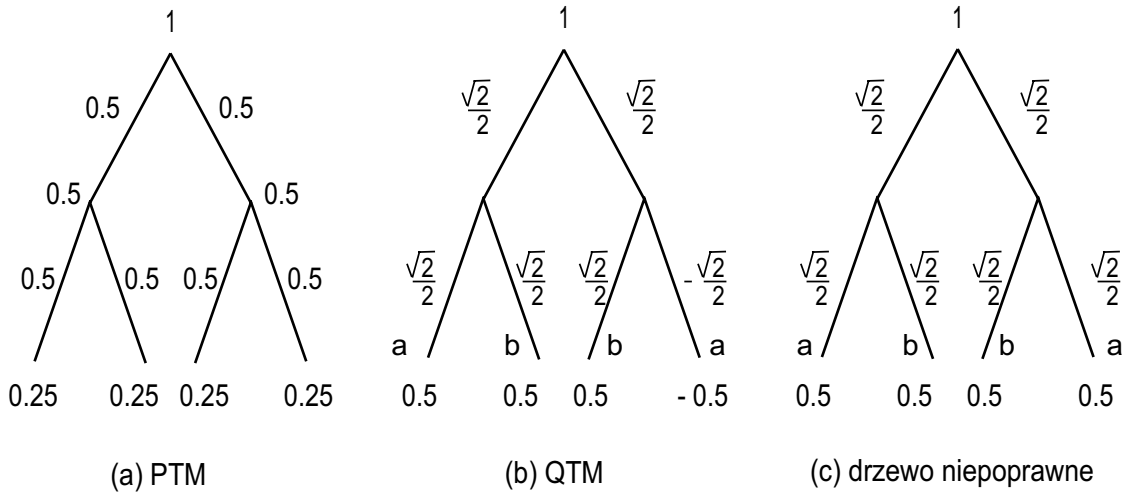
Definicja 3 (Globalny warunek prawdopodobieństwa) *Niech $c_1, c_2, \ldots, c_k$ będą wszystkimi wzajemnie różnymi konfiguracjami na pewnym poziomie j drzewa obliczeń kwantowych, a $\beta_1, \beta_2, \ldots, \beta_k$ będą łącznymi amplitudami przypisanymi wystąpieniu odpowiedniej konfiguracji c_i na j-tym poziomie drzewa. Zachodzi równość*

$$\sum_{i=1}^{k} |\beta_i|^2 = 1.$$

Dzięki temu, że QTM spełnia powyższy warunek suma prawdopodobieństw konfiguracji jednego poziomu, podobnie jak dla drzewa obliczeń klasycznych, jest zawsze równa 1, niezależnie od konfiguracji początkowej. Ze względu na niżej opisane zjawisko interferencji nie gwarantuje tego warunek o tym, że suma prawdopodobieństw łuków prowadzących od węzłów do ich dzieci jest równa 1 (def. 2,war.(i)), czego przykładem jest drzewo na rysunku 1.2c.

Z konstrukcji kwantowego drzewa obliczeń wynika, że prawdopodobieństwo osiągnięcia konfiguracji w pewnym kroku i jest równe kwadratowi modułu jej amplitudy. A co za tym idzie, prawdopodobieństwo osiągnięcia konkretnej konfiguracji końcowej to kwadrat sumy (a nie suma kwadratów!) modułów amplitud wszystkich liści odpowiadających tej konfiguracji. Takie wyliczanie prawdopodobieństwa ma niezwykle konsekwencje.

Na przykład niech pewna konfiguracja c koresponduje z dwoma liśćmi o amplitudach α i $-\alpha$. Wówczas prawdopodobieństwo osiągnięcia tej konfiguracji jest równe $|\alpha - \alpha|^2 = 0$!, pomimo że konfiguracje odpowiadające węzłom będącym rodzicami tych liści mogą mieć niezerowe prawdopodobieństwa. Prawdopodobieństwo wystąpienia pewnej konfiguracji końcowej jest prawdopodobieństwem, z jakim obserwator uzyska tą konfigurację jako wynik obliczeń. Czyli nie ma możliwości uzyskania na wyjściu konfiguracji o prawdopodobieństwie równym 0, mimo że podczas obliczenia pojawia się ścieżka prowadzą do niej. Jeśli natomiast oba liście mają amplitudę równą α, to prawdopodobieństwo uzyskania konfiguracji c jest równe nie $2\alpha^2$, ale $4\alpha^2$, czyli jest większe



Rysunek 1.2: *Przykładowe drzewa obliczeń*

niż suma prawdopodobieństw, jakie uzyskalibyśmy, gdyby te liście odpowiadały różnym konfiguracjom. Tą wzajemną zależność między różnymi gałęziami drzewa obliczeń nazywamy **interferencją**. Wynika ona z tego, że elektrony zachowują się czasem jak cząstki, a czasem jak fale, które interferują ze sobą. Eksperymenty fizyczne związane z tym zjawiskiem będą przedstawione w rozdziale 1.3.

Definicja 4 *Niech $\alpha_1, \alpha_2, \ldots, \alpha_k$ oznaczają amplitudy wszystkich wystąpień konfiguracji c na pewnym poziomie drzewa obliczeń. Mówimy o występowaniu interferencji pozytywnej, jeśli spełniona jest następująca nierówność*

$$|\sum_{i=1}^{k} \alpha_i|^2 > \sum_{i=1}^{k} |\alpha_i|^2.$$

Definicja 5 *Niech $\alpha_1, \alpha_2, \ldots, \alpha_k$ oznaczają amplitudy wszystkich wystąpień konfiguracji c na pewnym poziomie drzewa obliczeń. Mówimy o występowaniu interferencji destruktywnej, jeśli spełniona jest następująca nierówność*

$$|\sum_{i=1}^{k} \alpha_i|^2 < \sum_{i=1}^{k} |\alpha_i|^2.$$

Interferencja jest bardzo zaskakującą własnością świata kwantowego. Wymusza ona odmienny sposób patrzenia na obliczenia kwantowe, różny od tego, do którego jesteśmy przyzwyczajeni. Jeden z trików stosowanych w pisaniu algorytmów kwantowych opiera się na jej wykorzystaniu. Często tak projektujemy obliczenie, aby pożądany rezultat na skutek pozytywnej interferencji uzyskał wysokie prawdopodobieństwo, natomiast prawdopodobieństwo wyniku niepoprawnego na skutek interferencji destruktywnej stało się małe lub nawet spadło do 0.

Ze względu na zjawisko interferencji nie możemy traktować poszczególnych gałęzi drzewa obliczeń oddzielnie, ale musimy rozważać naraz całe drzewo. Dlatego też do rezultatu obliczenia kwantowego w każdym kroku odnosimy się równocześnie jako do **superpozycji** wszystkich gałęzi. Każdą konfigurację c oznaczamy $|\, c \,\rangle$, natomiast

superpozycję konfiguracji $c_1, c_2, \ldots, c_k$ o amplitudach $\alpha_1, \alpha_2, \ldots, \alpha_k$ oznaczamy przez $\sum_{i=1}^{k} \alpha_i |c_i\rangle$.

Funkcja przejścia δ kwantowej maszyny Turinga $\mathcal{M}$ wyznacza macierz przejścia $\mathcal{M}_M$, której rzędy i kolumny są etykietowane przez konfiguracje maszyny $\mathcal{M}$ w taki sposób, że $\mathcal{M}_M(i,j)$ jest amplitudą przejścia od stanu c_i do stanu c_j. Pola takiej macierzy są liczbami zespolonymi, takimi że norma Euklidesa każdej kolumny, które one tworzą jest równa 1, co wynika z tego, że funkcja przejścia δ spełnia lokalny warunek prawdopodobieństwa. Aby zachowane były wyżej przedstawione własności obliczeń kwantowych, macierz odpowiadająca maszynie Turinga musi być unitarna, tzn.

$$\mathcal{M}_M \mathcal{M}_M^* = \mathcal{M}_M \mathcal{M}_M^* = I,$$

gdzie $\mathcal{M}_M^*$ jest macierzą sprzężoną i transponowaną względem macierzy $\mathcal{M}_M$, a I jest macierzą jednostkową.

Niech v będzie wektorem utworzonym z amplitud wszystkich konfiguracji w danym kroku. Wówczas wykonanie kolejnego kroku maszyny odpowiada pomnożeniu macierzy przejścia $\mathcal{M}_M$ przez ten wektor: $\mathcal{M}_M v$.

Z unitarności macierzy przejścia wynika odwracalność maszyny Turinga. Oznacza to, że w każdym kroku na podstawie danej superpozycji wszystkich konfiguracji jesteśmy w stanie odtworzyć superpozycję konfiguracji poprzedniego kroku.

Kolejną zasadniczą różnicę między probabilistyczną maszyna Turinga, a kwantową maszyną Turinga możemy zobaczyć, gdy badamy wynik naszego obliczenia. W przypadku probabilistycznej maszyny Turinga w każdym konkretnym obliczeniu z drzewa obliczeń zostaje wybrana jedna ścieżka, która prowadzi do pewnej konfiguracji końcowej. Bez wpływu na obliczenie możemy zobaczyć, która ze ścieżek została wybrana i jaki uzyskaliśmy rezultat. Nie jest tak w przypadku obliczeń kwantowych. W kwantowej maszynie Turinga równocześnie obliczają się wszystkie ścieżki drzewa obliczeń. Ponieważ liczba węzłów danego poziomu drzewa rośnie wykładniczo ze względu na numer kroku obliczenia, QTM równocześnie przetwarza wykładniczą liczbę ścieżek i w konkretnym kroku obliczenia jest w superpozycji wykładniczej liczby konfiguracji! Warto zauważyć, że zachowanie QTM jest całkowicie wyznaczone przez unitarną macierz przejścia, a co za tym idzie jest deterministyczne.

Zasada nieoznaczoności Heisenberga (zasada mechaniki kwantowej) mówi, że nie możemy zmierzyć stanu kwantowego danego systemu. Dla teorii obliczeń kwantowych oznacza to, że dopóki kwantowa maszyna Turinga nie skończy liczyć, nie ma możliwości zobaczenia w jakim jest ona stanie i jakie daje rezultaty bez istotnego wpływu na obliczenia. Dodatkowo nie możemy zobaczyć superpozycji wszystkich konfiguracji, która jest wynikiem działania algorytmu. Po dokonaniu pomiaru otrzymujemy tylko jedną losowo wybraną konfigurację spośród wszystkich konfiguracji wyznaczonych podczas obliczenia (z prawdopodobieństwem będącym kwadratem modułu jej amplitudy), a pozostałe rezultaty obliczenia są nieodwracalnie tracone. Jest to jedyne miejsce, w którym zachowanie maszyny Turinga nie jest deterministyczne.

Powyższe uwagi możemy sformalizować w następujący sposób

Definicja 6 *Mówimy, że QTM będąca w superpozycji $\psi = \sum_{i=1}^{k} \alpha_i c_i$ jest obserwowana (lub jest mierzona), gdy próbujemy sprawdzić w jakim jest ona stanie. Otrzymujemy*

wówczas pewną konfigurację c_i z prawdopodobieństwem $|\alpha_i|^2$. Dodatkowo superpozycja, w jakiej była maszyna zostaje przekształcona w $\psi = c_i$.

1.2 Przestrzeń Hilberta

Przestrzeń Hilberta jest matematyczną strukturą wygodną do opisywania koncepcji, własności i praw mechaniki kwantowej. Można powiedzieć, że izolowany system kwantowy odpowiada tej przestrzeni, a stany systemu odpowiadają jej wektorom. Poniżej przytoczymy kilka podstawowych definicji teorii przestrzeni Hilberta, istotnych dla obliczeń kwantowych.

Definicja 7 *Przestrzeń Hilberta H jest przestrzenią wektorową nad ciałem $\mathbb{C}$, taką że*

a) jest ona wyposażona w funkcję zwaną iloczynem skalarnym $\langle\cdot|\,\cdot\rangle : H \times H \to \mathbb{C}$, spełniającą warunki

(i) $\langle\phi|\,\phi\rangle \geqslant 0 \qquad i \qquad \langle\phi|\,\phi\rangle = 0 \;\Leftrightarrow\; \phi = 0,$

(ii) $\langle\psi|\,\phi\rangle = \langle\phi|\,\psi\rangle^*,$

(iii) $\langle\psi|\,c_1\phi_1 + c_2\phi_2\rangle = c_1\langle\psi|\,\phi_1\rangle + c_2\langle\psi|\,\phi_2\rangle,$

b) jest przestrzenią zupełną z normą określoną przez iloczyn skalarny: $\|\phi\| = \sqrt{\langle\phi|\,\phi\rangle}$.

Definicja 8 *Dwa wektory z przestrzeni Hilberta ψ i ϕ są ortogonalne ($\psi \perp \phi$), jeśli $\langle\psi|\,\phi\rangle = 0$. Zbiór $S \subseteq H$ nazywamy ortogonalnym, jeśli każde jego dwa elementy są ortogonalne. Zbiór S jest ortonormalny, jeśli jest ortogonalny i wszystkie jego elementy mają normę równą 1.*

Definicja 9 *Zbiór $\mathcal{S}$ wektorów przestrzeni Hilberta H tworzy bazę (ortonormalną) przestrzeni H, jeśli jest ortonormalny i każdy z wektorów przestrzeni H można jednoznacznie wyrazić w postaci*

$$\phi = \sum_{\psi\in\mathcal{B}} \alpha_\psi \psi, \quad \text{gdzie} \quad \alpha_\psi \in \mathbb{C}.$$

Wiemy, że każde dwie bazy niezerowej przestrzeni wektorowej są równoliczne. Moc dowolnej bazy przestrzeni wektorowej nazywamy wymiarem przestrzeni. Dodatkowo dowolne dwie przestrzenie Hilberta o tym samym wymiarze są izomorficzne. Przestrzeń Hilberta d-wymiarową oznaczać będziemy przez H_d. Przykładem ortonormalnej bazy przestrzeni H_3 może być zbiór: $\{(0,0,1),(0,1,0),(1,0,0)\}$.

Twierdzenie 1 *Dla każdej podprzestrzeni W przestrzeni Hilberta H istnieje jednoznacznie wyznaczona podprzestrzeń $W^\perp$, taka że*

(i) $\forall\psi \in W \;\; i \;\; \forall\phi \in W^\perp$ *zachodzi równość* $\langle\psi|\,\phi\rangle = 0,$

(ii) $\forall\psi \in H \;\; \exists\phi_1 \in W \;\; i \;\; \exists\phi_2 \subseteq W^\perp$ *takie, że ψ może być jednoznacznie wyrażone w postaci $\psi = \phi_1 + \phi_2$.*

Piszemy wówczas, że $H = W \oplus W^{\perp}$, a mówimy W i $W^{\perp}$ tworzą ortogonalny rozkład przestrzeni H.

Dowód tego twierdzenia można znaleźć w [11].

Powyższą ideę w naturalny sposób można uogólnić mówiąc, że istnieje rozkład przestrzeni H na wzajemnie ortogonalne podprzestrzenie $W_1, W_2, \ldots, W_n$, takie że $\forall \psi \subseteq H$ istnieje reprezentacja $\psi = \phi_1 + \phi_2 + \ldots + \phi_n$, gdzie $\phi_i \in W_i$ dla $1 \leqslant i \leqslant n$, co zapisujemy w nastepujący sposób

$$H = W_1 \oplus W_2 \oplus \ldots \oplus W_n.$$

Notacja

Dla przestrzeni Hilberta zachodzą dwa następujące twierdzenia

Twierdzenie 2 *Dla każdego wektora ϕ z przestrzeni Hilberta H odwzorowanie $f_\phi : H \to \mathbb{C}$ zdefiniowane następująco*

$$f_\phi(\psi) = \langle \phi | \psi \rangle$$

jest odwzorowaniem liniowym (tzn. $f_\phi(c\psi) = c f_\phi(\psi)$ i $f_\phi(\psi_1 + \psi_2) = f_\phi(\psi_1) + f_\phi(\psi_2)$).

Dowód powyższego twierdzenia wynika z warunków, jakie spełnia iloczyn skalarny (def. 7, war. (iii)).

Twierdzenie 3 (Riesz-Fréchet) *Dla każdej ciągłej i liniowej funkcji $f : H \to \mathbb{C}$ istnieje takie $\phi_f \in H$, że $\forall \psi \in H \quad f(\psi) = \langle \phi_f | \psi \rangle$.*

Dowód twierdzenia można znaleźć w [11].

Przestrzeń wszystkich liniowych odwzorowań z przestrzeni Hilberta H (funkcjonałów) tworzy kolejną przestrzeń Hilberta, zwaną **dualną przestrzenią Hilberta** (lub **sprzężoną przestrzenią Hilberta**) H^* z iloczynem skalarnym $\langle f | g \rangle = \langle \phi_f | \psi_g \rangle$ $(\forall f, g \in H)$. Dla każdego ϕ odwzorowanie f_ϕ, określone wzorem $f_\phi(\psi) = \langle \phi | \psi \rangle$, jest funkcjonałem, stąd na mocy powyższych twierdzeń mamy bijekcję miedzy H a H^* i $H \cong H^*$.

Na bazie relacji miedzy tymi dwoma przestrzeniami została wprowadzona wygodna notacja "ket-bra", zaproponowana przez A.M. Diraca. Wektor przestrzeni Hilberta ψ oznaczamy $|\psi\rangle$ i nazywamy go **ket-wektorem**. Odpowiadający mu funkcjonał (wyznaczony przez twierdzenie 2) oznaczamy $\langle \psi |$ i nazywamy **bra-wektorem**. $\langle \cdot |$ może być też traktowany jak operator odwzorowujący każdy stan ϕ na funkcjonał $\langle \phi |$, tak że dla każdego stanu zachodzi $\langle \phi | (|\psi\rangle) = \langle \phi | \psi \rangle$.

W przypadku n-wymiarowej przestrzeni Hilberta, na ket-wektor $|\psi\rangle$ możemy patrzeć jako na n-wymiarową kolumnę, a na bra-wektor jak na n-wymiarowy rząd. Iloczyn skalarny $\langle \phi | \psi \rangle$ jest rezultatem mnożenia "rząd $\times$ kolumna". Przekształcenie $|\phi\rangle \leftrightarrow \langle \phi |$ odpowiada transpozycji i sprzężeniu, stąd iloczyn zewnętrzny $|\phi\rangle\langle\psi|$ jest macierzą $n \times n$, powstająca z mnożenia "kolumna $\times$ rząd".

Wektory przestrzeni Hilberta H o normie jednostkowej są nazywane **(czystymi) stanami** przestrzeni H.

Iloczyn tensorowy

Niech systemowi kwantowemu S odpowiada pewna przestrzeń Hilberta H. Jeśli jest on złożony z dwóch podsystemów S_1 i S_2, to H jest tzw. **produktem tensorowym** przestrzeni Hilberta H_1 i H_2, które odpowiadają odpowiednio podsystemom S_1 i S_2, co zapisujemy następująco

$$H = H_1 \otimes H_2.$$

Oznacza to, że wektory z H są produktami tensorowymi wektorów z H_1 i H_2.

Definicja 10 *Produkt tensorowy wektorów $x = (x_1, x_2, \ldots, x_m)$ i $y = (y_1, y_2, \ldots, y_n)$ jest mn-wymiarowym wektorem o elementach*

$$(x_1y_1, \ldots, x_1y_n, x_2y_1, \ldots, x_2y_n, \ldots, x_my_1, \ldots, x_my_n)$$

i oznaczamy go przez $x \otimes y$.

Produkt tensorowy można też zdefiniować dla macierzy. Niech

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & \ldots & b_{1m} \\ \vdots & & \vdots \\ b_{m1} & \ldots & b_{mm} \end{pmatrix}.$$

Wówczas $A \otimes B$ jest macierzą mn-wymiarową postaci

$$A \otimes B = \begin{pmatrix} a_{11}B & \ldots & a_{1n}B \\ \vdots & & \vdots \\ a_{n1}B & \ldots & a_{nn}B \end{pmatrix}.$$

Jeśli $\mathcal{B}_1, \ldots, \mathcal{B}_k$ są ortonormalnymi bazami przestrzeni Hilberta $H_1, \ldots, H_k$ to

$$\mathcal{B}_1 \otimes \ldots \otimes \mathcal{B}_k = \bigotimes_{i=1}^{k} \mathcal{B}_i = \{x_1 \otimes \ldots \otimes x_k \,:\, x_i \in \mathcal{B}_i\}$$

jest ortonormalną bazą przestrzeni Hiblerta

$$H = \bigotimes_{i=1}^{k} H_i.$$

Przykład 1 *Jeśli H_2 jest 2-wymiarową przestrzenią Hilberta z bazą $\mathcal{B}_2 = \{|0\rangle, |1\rangle\}$, wówczas*

$$\bigotimes_{i=1}^{n} \mathcal{B}_2 = \{|x_1\rangle \otimes |x_2\rangle \otimes \ldots \otimes |x_n\rangle \,:\, x_1, \ldots, x_n \in \{|0\rangle, |1\rangle\}\}$$

jest baza ortogonalną 2^n- wymiarowej przestrzeni

$$H_{2^n} = \bigotimes_{i=1}^{n} H_2.$$

Zamiast $|x_1\rangle \otimes |x_2\rangle \otimes \ldots |x_n\rangle$ używamy często prostszych notacji: $|x_1\rangle|x_2\rangle \ldots |x_n\rangle$ lub $|x_1, x_2, \ldots, x_n\rangle$, lub $|x_1\ x_2\ \ldots\ x_n\rangle$.

Obserwator

Definicja 11 *Niech H będzie przestrzenią Hilberta odpowiadającą pewnemu systemowi kwantowemu. Obserwator $\mathcal{O} = \{E_1, E_2, \ldots, E_k\}$ jest zbiorem rozłącznych, wzajemnie ortogonalnych podprzestrzeni, takich że*

$$H = E_1 \oplus E_2 \oplus \ldots \oplus E_k \qquad \textit{(suma ortogonalna)},$$

na którym jest określone odwzorowanie różnowartościowe $\mu\colon \{E_1, E_2, \ldots, E_k\} \to \mathbb{R}$.

Twierdzenie 4 *Niech $|\phi\rangle$ będzie stanem, a $\mathcal{O} = \{E_1, E_2, \ldots, E_k\}$ obserwatorem. $|\phi\rangle$ może być wyrażony jednoznacznie za pomocą liniowej superpozycji rzutowań $|\phi\rangle$ wzdłuż każdej z przestrzeni E_i*

$$|\phi\rangle = \sum_{i=1}^{k} \alpha_i |\phi_{E_i}\rangle,$$

gdzie $|\phi_{E_i}\rangle$ jest stanem w E_i oraz $\langle\phi_{E_i}|\,\phi_{E_i}\rangle = 1$ dla każdego i.

Obserwowanie stanu $|\phi\rangle$ przez $\mathcal{O}$ pociąga za sobą następujące konsekwencje:

1. Zostaje wybrana jedna z podprzestrzeni $E_1, E_2, \ldots, E_k$, powiedzmy że E_i i wyznaczona jest wówczas wartość $\mu(E_i)$. Prawdopodobieństwo wyboru E_i jest równe $|\alpha_i|^2$.

2. Po obserwacji, stan $|\phi\rangle$ przechodzi w stan $|\phi_{E_i}\rangle$.

3. Jedyną klasyczną informacją, jaką otrzymujemy z obserwacji systemu kwantowego przez $\mathcal{O}$, jest wartość funkcji μ. Cała informacja poza stanem $|\phi_{E_i}\rangle$ jest nieodwracalnie tracona.

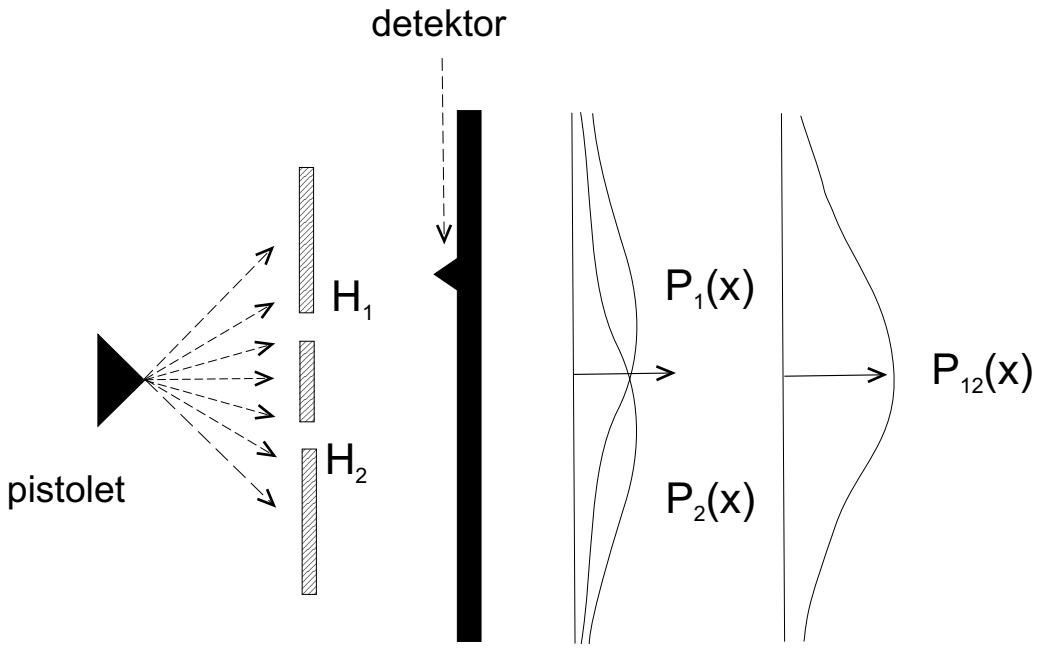
Definicja 12 *Obserwatorem standardowym nazywamy $\mathcal{B} = \{E_0, E_1\}$, gdzie E_i (dla $i = 0, 1$) jest liniową podprzestrzenią generowaną przez wektor $|i\rangle$. Obserwatorem dualnym nazywamy $\mathcal{O} = \{E_0', E_1'\}$, gdzie E_0' jest liniową podprzestrzenią generowaną przez wektor $|0'\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, a E_1' jest liniową podprzestrzenią generowaną przez wektor $|1'\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$.*

1.3 Eksperymenty

Prawa obliczeń kwantowych: przypisywanie amplitud do zdarzeń kwantowych, superpozycja, interferencja, specjalny sposób rozważania prawdopodobieństwa mogą wydawać się dziwne. Wszystko to ma swoje źródło w prawach mechaniki kwantowej. Aby lepiej zrozumieć zachowanie cząstek kwantowych spróbujemy je porównać z zachowaniem elementów świata klasycznego: pocisków i fal.

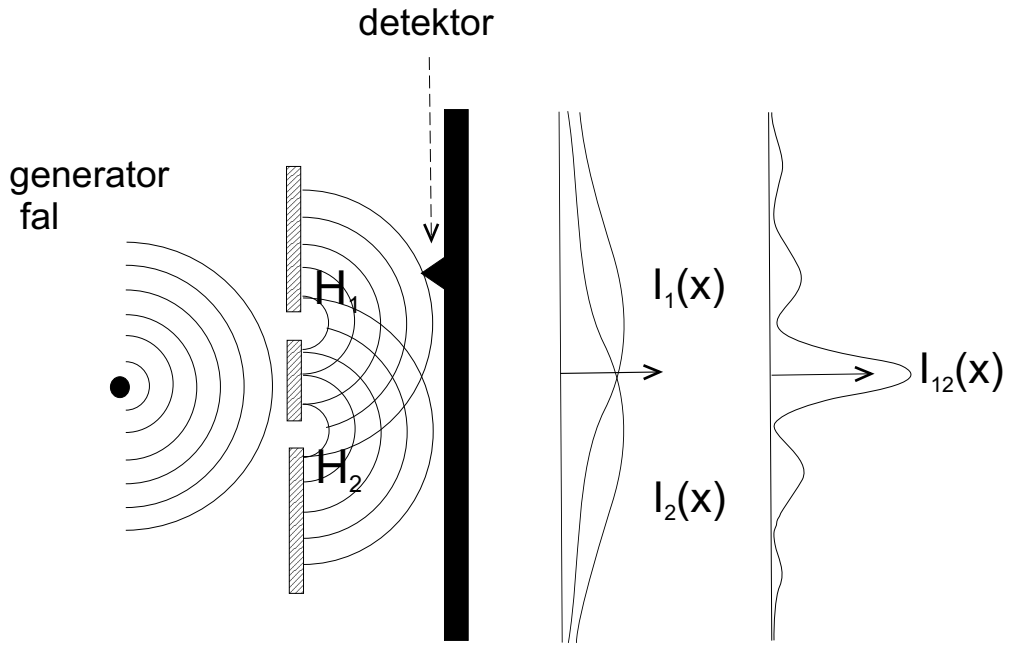
1.3.1 Eksperymenty klasyczne

Rozważmy pierwszy eksperyment opisany na rysunku 1.3. Mamy tu automatyczny pistolet, który wystrzeliwuje pociski w różnych kierunkach o losowym rozkładzie. Przed nim znajduje się ściana z dwiema niewielkimi szczelinami, każda wielkości jednego pocisku. Pociski przelatując przez te szczeliny mogą się odbijać i tym samym dowolnie zmieniać kierunki. Kolejna ścianka stojąca przed pistoletem zawiera detektor wykorzystywany do zliczania pocisków, które w nią uderzają. Pytanie w eksperymencie jest następujące: Jakie jest prawdopodobieństwo, że pocisk osiągnie wskazaną pozycje na drugiej ściance? W przypadku, gdy otwarta jest tylko jedna szczelina H_1 (H_2), otrzymujemy wynik opisany przez krzywą $P_1(x)$ ($P_2(x)$) a w przypadku dwóch otwartych szczelin, krzywą $P_{12}(x)$. Ogólny rezultat, zgodny z oczekiwaniami można opisać równaniem: $P_{12}(x) = \frac{1}{2}(P_1(x) + P_2(x))$ dla każdej pozycji x.



Rysunek 1.3: *Eksperyment z wystrzeliwanymi pociskami*

Drugi eksperyment, z falami na wodzie, przedstawia rysunek 1.4. Mamy tu wibrujący generator wytwarzający fale. Poruszają się one w kierunku ścianki z dwoma otworami i drugiej z detektorem, który wykrywa intensywność fal $I(x) = |h(x)|^2$, gdzie h jest amplitudą fali. Krzywa intensywności $I_1(x)$ ($I_2(x)$) odpowiada przypadkowi, gdy otwarta jest jedna szczelina H_1 (H_2). Rezultat dla przypadku dwóch otwartych szczelin przedstawia krzywa $I_{12}(x)$ i jest on dobrze znany z fizyki. Wynika ze zjawiska zwanego interferencją fal. W tym przypadku mamy $I_{12}(x) = |h_1(x) + h_2(x)|^2$, co oznacza że zachodząca interferencja w pewnych sytuacjach jest pozytywna, a w innych negatywna.



Rysunek 1.4: *Eksperyment z falami*

1.3.2 Eksperymenty kwantowe

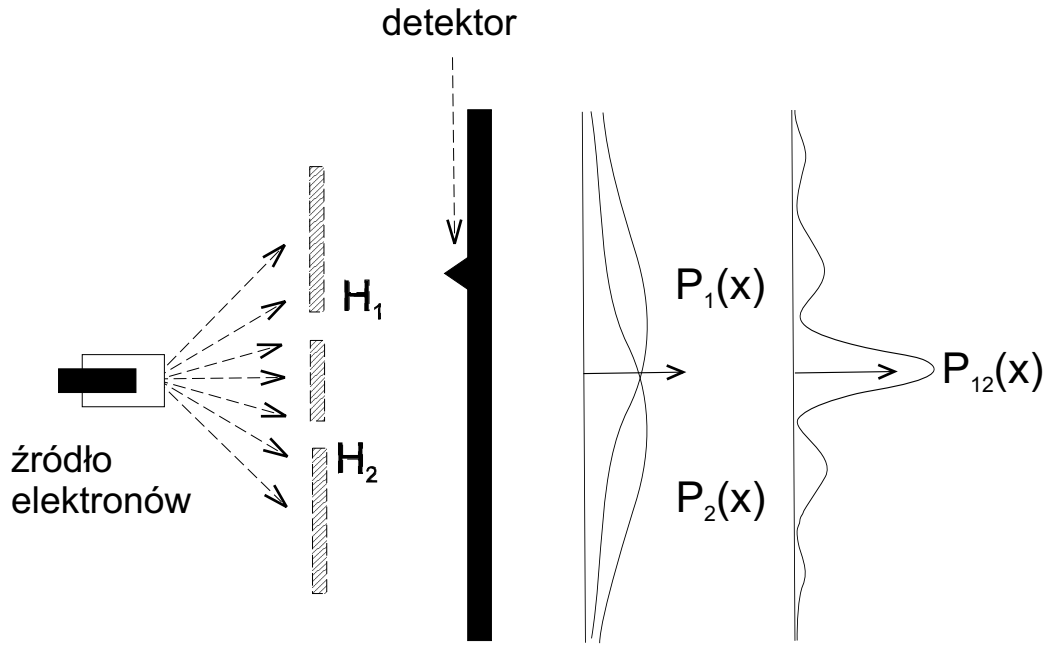
Trzeci eksperyment, opisany na rysunku 1.5, jest podobny do eksperymentu pierwszego. Mamy źródło produkujące elektrony. Pierwsza ścianka ma dwie, bardzo wąskie szczeliny tak małe, że w danej chwili przepuszczają one tylko jeden elektron, czasem zmieniając jego kierunek. Druga ścianka zawiera detektor używany do zliczania elektronów, które do niej dotarły.

P_1 i P_2 są eksperymentalnie wyznaczonymi krzywymi, oznaczającymi prawdopodobieństwa, że elektrony osiągnęły daną pozycję na drugiej ściance odpowiednio w przypadku otwartej szczeliny H_1 lub H_2. Nasz rezultat jest taki sam jak w eksperymencie 1. Oczekiwalibyśmy, że krzywa P_{12} wyznaczona dla sytuacji, gdy otwarte są dwie szczeliny, będzie również taka jak w eksperymencie 1. Niespodziewanie tak nie jest. Wygląda ona tak jak $I_{12}(x)$ z eksperymentu 2. Czyli elektrony, będące cząstkami czasem zachowują się jak fale!

Można powiedzieć, że każda cząstka zachowuje się tak, jakby przelatywała przez obie szczeliny na raz, po czym tworzyła fale, które interferują ze sobą tak jak w eksperymencie 2. Możemy z tego wyciągnąć dwa zaskakujące wnioski. Pierwszy dotyczy tego, że cząstka kwantowa może być jakby w dwu (a nawet wielu) stanach równocześnie. Mówimy, że może być w danej chwili w superpozycji wielu stanów. Stąd tak naprawdę nie jesteśmy w stanie dokładnie określić drogi, jaką przebył elektron. Drugą obserwacją jest to, że stany są od siebie zależne i interferują ze sobą.

Warto zauważyć, że rezultat eksperymentu nie jest zależny od częstotliwości generowania elektronów. Ten sam efekt interferencji powstanie, jeśli kolejny elektron zostanie wystrzelony dopiero gdy poprzedni uderzy w ściankę z detektorem.

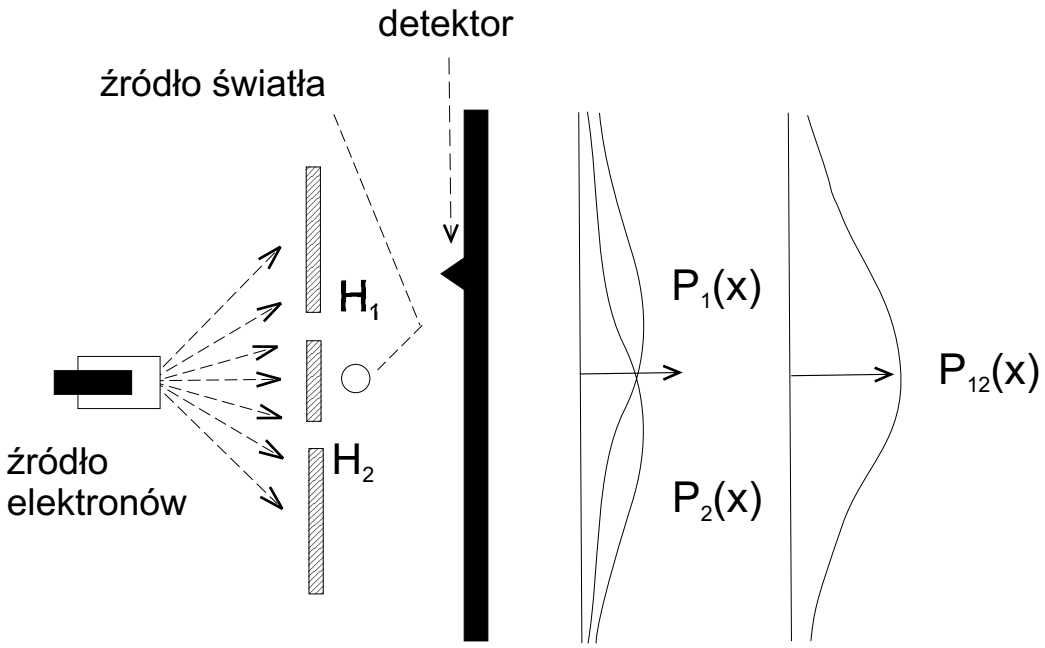
Aby zilustrować inną dziwną własność kwantową zmodyfikujemy powyższy ekspe-



Rysunek 1.5: *Eksperyment z dwoma szczelinami*

ryment. Jak widzimy na rysunku 1.6, w tym przypadku po prawej stronie pierwszej ścianki, między szczelinami mamy dodatkowe źródło światła. W czasie eksperymentu chcemy zaobserwować, przez którą szczelinę pierwszej ścianki przechodzi konkretny elektron. Jeśli będzie to szczelina H_1, w jej pobliżu na moment pojawi się światełko, jeśli H_2, pojawi się ono koło drugiej szczeliny. Podobnie jak w poprzednich eksperymentach, chcemy też wyznaczyć prawdopodobieństwa, z jakimi elektrony osiągają poszczególne pozycje na drugiej ściance. Krzywe dla przypadków, gdy otwarta jest tylko jedna szczelina są zgodnie z oczekiwaniami, takie jak we wcześniejsch ćwiczeniach. Natomiast krzywa dla przypadku dwóch otwartych szczelin jest ku naszemu zaskoczeniu inna, podobna do tej, którą widzieliśmy w eksperymencie 1. Rezultat ten możemy wyjaśnić tak, że zachowanie elektronów, a co za tym idzie zachowanie systemu kwantowego, jest zależne od tego czy jest on obserwowany, czy nie! Obserwacja niszczy efekt interferencji.

Między innymi ze względu na powyższe obserwacje, budowa komputera kwantowego, który dałoby się zastosować w praktyce, będzie bardzo trudna. Superpozycja stanów kwantowo-mechanicznych, która jest podstawą efektywności komputerów kwantowych jest bardzo niestabilna. Problem polega na tym, że każde oddziaływanie układu kwantowego z otoczeniem, na przykład zderzenia atomu z innym atomem lub zabłąkanym fotonem, jest pewnego rodzaju pomiarem. W jego wyniku superpozycja kolapsuje do pewnego konkretnego stanu, właśnie tego, który zarejestrowany został podczas obserwacji. Zjawisko to jest zwane dekoherencją i sprawia, że dalsze obliczenia kwantowe stają się niemożliwe. A więc wewnętrzna maszyneria komputera kwantowego musi być, w celu zachowania koherencji, w jakiś sposób odizolowana od otoczenia. Musi być ona jednak równocześnie dostępna tak, aby dane mogły zostać wprowadzone, procedury wykonane, a ich wynik odczytany.



Rysunek 1.6: *Eksperyment z dwoma szczelinami i obserwatorem*

1.4 Podstawowe elementy obliczeń kwantowych

1.4.1 Qubity

Najmniejszą, niepodzielną jednostką informacji klasycznej jest bit, przyjmujący jedną z dwóch możliwych wartości: 0 lub 1. W świecie kwantowym bitowi klasycznemu odpowiada bit kwantowy, inaczej **qubit**. Jest on wektorem przestrzeni Hilberta, opisującym stan najmniejszego możliwego systemu kwantowego.

Niech S będzie 2-wymiarową przestrzenią Hilberta (czyli najmniejszym systemem kwantowym) z bazą, na którą składają się dwa ortonormalne wektory $|0\rangle$ i $|1\rangle$.

Definicja 13 *Qubit (bit kwantowy) jest stanem kwantowym*

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{1.1}$$

gdzie $\alpha, \beta \in \mathbb{C}$ *i* $|\alpha|^2 + |\beta|^2 = 1$.

Notacja "qubit" jest też używana do oznaczenie zmiennej, która przyjmuje wartości postaci (1.1). W tym sensie możemy mówić o stanie qubitu. Takie spojrzenie na qubit jest zbliżone do tego, jak patrzymy na bit w obliczeniach klasycznych.

Niech qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. Możemy go mierzyć ze względu na nieskończoną ilość baz. Najpierw dokonamy jego pomiaru ze względu na bazę $\{|0\rangle, |1\rangle\}$, zwaną **bazą standardową**. Nastąpi wówczas rzutowanie $|\psi\rangle$ na tą bazę. Na wyjściu uzyskamy $|0\rangle$ ($|1\rangle$) z prawdopodobieństwem $|\alpha|^2$ ($|\beta|^2$). Pomiar nie jest więc deterministyczny. Dodatkowo po jego wykonaniu qubit przejdzie w $|0\rangle$ lub $|1\rangle$. Cała informacja o superpozycji jest nieodwracalnie tracona.

Teraz spróbujemy zmierzyć qubit $|\psi\rangle$ ze względu na bazę $\mathcal{D} = \{|0'\rangle, |1'\rangle\}$, zwaną **bazą dualną**, gdzie

$$|0'\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \qquad |1'\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle.$$

Ponieważ

$$|0\rangle = \frac{1}{\sqrt{2}}|0'\rangle + \frac{1}{\sqrt{2}}|1'\rangle \qquad |1\rangle = \frac{1}{\sqrt{2}}|0'\rangle - \frac{1}{\sqrt{2}}|1'\rangle$$

mamy

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big((\alpha+\beta)|0'\rangle + (\alpha-\beta)|1'\rangle\Big).$$

Pomiar $|\psi\rangle$ ze względu na bazę $\mathcal{D}$ daje $|0'\rangle$ ($|1'\rangle$) z prawdopodobieństwem $\frac{1}{2}|\alpha+\beta|^2$ ($\frac{1}{2}|\alpha-\beta|^2$).

1.4.2 Transformacje qubitów

Operacje na qubitach, jak wspominaliśmy wcześniej, są opisywane za pomocą macierzy unitarnych, nazywanych operatorami

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Macierz A przekształca stan qubitu z $\alpha|0\rangle + \beta|1\rangle$ na $(a\alpha + b\beta)|0\rangle + (c\alpha + b\beta)|1\rangle$. Na przykład przekształcenie dane przez macierz Hadamarda

$$H = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

zwane *rotacją Hadamarda*, przekształca stany $|0\rangle$ i $|1\rangle$ w poniższy sposób

$$H|0\rangle = |0'\rangle,$$
$$H|1\rangle = |1'\rangle,$$

a stany $|0'\rangle$, $|1'\rangle$ następująco

$$H|0'\rangle = |0\rangle,$$
$$H|1'\rangle = |1\rangle.$$

Transformację Hadamarda można też zapisać jako następujące przekształcenie stanów bazowych

$$H|i\rangle = \frac{1}{\sqrt{2}}(|0\rangle + (-1)^i|1\rangle).$$

Jak już wspomnieliśmy wcześniej, baza $\mathcal{B} = \{|0\rangle, |1\rangle\}$ jest zwana *bazą standardową*, a baza $\mathcal{D} = \{|0'\rangle, |1'\rangle\}$ jest zwana *bazą dualną* lub bazą Hadamarda, lub czasem bazą Fouriera. Dla transformaty Hadamarda zachodzi $H^2 = I$ i jak mogliśmy zobaczyć, zastosowanie jej przekształca bazę standardową w dualna i odwrotnie.

Trzy inne ważne operacje unitarne na qubitach to rotacja (o kąt θ) $R(\theta)$, przesunięcie fazowe $PS(\alpha)$ i skalowanie $Scal(\delta)$

$$R(\theta) = \begin{pmatrix} \cos\frac{\theta}{2} & \sin\frac{\theta}{2} \\ -\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}, \qquad PS(\alpha) = \begin{pmatrix} e^{i\frac{\alpha}{2}} & 0 \\ 0 & e^{-i\frac{\alpha}{2}} \end{pmatrix}, \qquad Scal(\delta) = \begin{pmatrix} e^{i\delta} & 0 \\ 0 & e^{i\delta} \end{pmatrix}.$$

1.4.3 Rejestr 2-qubitowy

Tensorowy produkt dwóch qubitów jest zwany *kwantowym rejestrem 2-qubitowym*. Odpowiada on przestrzeni Hilberta H_4. Trzy najważniejsze i najpopularniejsze bazy tej przestrzeni podane są w poniższej tabelce

Bazy	1 stan bazowy	2 stan bazowy	3 stan bazowy	4 stan bazowy
standardowa	$\|00\rangle$	$\|01\rangle$	$\|10\rangle$	$\|11\rangle$
dualna	$\|0'0'\rangle$	$\|0'1'\rangle$	$\|1'0'\rangle$	$\|1'1'\rangle$
Bella	$\Phi^+ = \frac{1}{\sqrt{2}}(\|00\rangle + \|11\rangle)$	$\Phi^- = \frac{1}{\sqrt{2}}(\|00\rangle - \|11\rangle)$	$\Psi^+ = \frac{1}{\sqrt{2}}(\|01\rangle + \|10\rangle)$	$\Psi^- = \frac{1}{\sqrt{2}}(\|01\rangle - \|10\rangle)$

Dla 2-qubitowego rejestru zapis $|01\rangle$ może być stosowany zamiennie z $|0,1\rangle$ lub $|0\rangle|1\rangle$, lub $|0\rangle \otimes |1\rangle$.

Zwykle stany bazy standardowej reprezentowane są w następujący sposób

$$|0\rangle = |00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},\ |1\rangle = |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},\ |2\rangle = |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix},\ |3\rangle = |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Ogólnie stan 2-qubitowego rejestru można przedstawić wyrażeniem

$$|\psi\rangle = a_{00}|00\rangle + a_{01}|01\rangle + a_{10}|10\rangle + a_{11}|11\rangle, \tag{1.2}$$

gdzie $|a_{00}|^2 + |a_{01}|^2 + |a_{10}|^2 + |a_{11}|^2 = 1$.

Pomiar 2-qubitowego rejestru

Gdy mierzymy stan rejestru postaci (1.2) ze względu na bazę standardową, to otrzymujemy na wyjściu 2 bity ij ($i, j \in \{0,1\}$) z prawdopodobieństwem $|a_{ij}|^2$ i stan $|\psi\rangle$ zostaje sprowadzony do $|ij\rangle$.

Czasem wygodne jest zmierzenie tylko jednego qubitu rejestru. Dzieje się to za pomocą obserwatora

$$\begin{array}{ll} \mathcal{B}_1 = \{E_1^0, E_1^1\} & \text{w przypadku pierwszego qubitu,} \\ \mathcal{B}_2 = \{E_2^0, E_2^1\} & \text{w przypadku drugiego qubitu,} \end{array}$$

gdzie E_1^i ($i = 0, 1$) jest podprzestrzenią rozpiętą na wektorach $\{|i0\rangle\, |i1\rangle\}$ a E_2^i ($i = 0, 1$) podprzestrzenią rozpiętą na $\{|0i\rangle\, |1i\rangle\}$.

Stąd jeśli mierzymy pierwszy qubit, możemy dostać na wyjściu bit 0 z prawdopodobieństwem $|a_{00}|^2 + |a_{01}|^2$ i rejestr przechodzi w

$$|\psi'\rangle = \frac{a_{00}|00\rangle + a_{01}|01\rangle}{\sqrt{|a_{10}|^2 + |a_{11}|^2}}.$$

Operacje kwantowe rejestru 2-qubitowego

Wśród unitarnych transformacji specjalną rolę odgrywa operacja

$$XOR\colon |x, y\rangle \to |x, x \oplus y\rangle$$

z macierzą

$$XOR = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Przy projektowaniu algorytmów kwantowych ważne jest pamiętanie o poniższej własności, której dowód można znaleźć w [7].

Twierdzenie 5 *Nie ma możliwości wykonania dokładnej kopii nieznanego stanu kwantowego. Tzn. nie istnieje unitarna transformacja U taka, że dla 1-qubitowego stanu $|\psi\rangle$ $U(|\psi, 0\rangle) = |\psi, \psi\rangle$. Twierdzenie to zachodzi dla każdej przestrzeni Hilberta.*

1.4.4 Rejestry n-qubitowe

Przestrzeń Hilberta dla rejestrów n-qubitowych

W łatwy sposób można uogólnić rejestr 2-qubitowy na n-qubitowy. Rozważając n-qubitowy rejestr pracujemy z 2^n-wymiarową przestrzenią Hilberta, z następującym zbiorem wektorów bazowych

$$\mathcal{B} = \{|i\rangle : i \in \{0, 1\}^n\}$$

lub w innym, bardziej popularnym zapisie

$$\mathcal{B} = \{|i\rangle : 0 \leqslant i < 2^n\},$$

które mówimy, że tworzą *bazę standardową*.

Ogólna postać stanu znajdującego się w n-qubitowym rejestrze jest następująca

$$|\psi\rangle = \sum_{i=0}^{2^n-1} a_i|i\rangle, \quad \text{gdzie } \sum_{i=0}^{2^n-1} |a_i|^2 = 1.$$

Rejestry n-qubitowe mają kilka godnych uwagi własności:

1. Ilość stanów bazowych i wielkość superpozycji rośnie wykładniczo ze względu na ilość qubitów.

2. W celu przechowania liczby całkowitej n potrzebujemy $\lceil \lg(n+1) \rceil$ qubitów, czyli musimy użyć przestrzeni Hilberta o wymiarze $2^{\lceil \lg(n+1) \rceil}$ z $2^{\lceil \lg(n+1) \rceil} > n$ ilością stanów bazowych.

3. Pomimo wykładniczego rozmiaru zbioru stanów bazowych istnieje szybki, liniowy sposób przejścia od jednego stanu do drugiego.

4. Wydaje się, że rejestr n-qubitowy jest w stanie zmagazynować wykładniczo więcej informacji niż rejestr n-bitowy. Tak jednak nie jest. W n-qubitowym rejestrze może znajdować się superpozycja aż 2^n qubitów, ale gdy dokonamy pomiaru tego rejestru uzyskamy n bitów, a pozostałe informacje zostaną utracone. Załóżmy, że chcemy zapamiętać w rejestrze n-qubitowym pewną ilość bitów, aby po chwili wszystkie bity dokładnie odczytać. Maksymalną ilością bitów, jakie mogą być w ten sposób przechowane w rejestrze n-qubitowym jest n, tyle samo co w klasycznym rejestrze n-bitowym.

Operacje n-qubitowe

Operacje na n-qubitowych rejestrach to unitarne transformacje, dane jako macierze $2^n \times 2^n$. Gdybyśmy chcieli wykonać klasycznie jeden taki krok obliczenia musielibyśmy wykonać $2^n(2 \cdot 2^n - 1)$ operacji arytmetycznych. (Mnożąc macierz $2^n \times 2^n$ przez wektor $2^n \times 1$ uzyskujemy wektor $2^n \times 1$. Wyliczenie każdej z jego 2^n wartości wymaga $2 \cdot 2^n - 1$ kroków.)

Jeśli unitarną macierz U zastosujemy tylko do i-tego qubitu, wówczas całe unitarne przekształcenie, jakie zostanie użyte na n-qubitowym rejestrze będzie postaci $(\bigoplus_{k=1}^{i-1} I) \oplus U \oplus (\bigoplus_{k=i+1}^{n} I)$. W celu zasymulowania takiego kroku klasycznie potrzebowalibyśmy $3 \cdot 2^n$ operacji arytmetycznych.

Pomiar rejestrów n-qubitowych

Jeśli mierzymy stan $|\phi\rangle$ n-qubitowego rejestru ze względu na bazę standardową, uzyskujemy na wyjściu n bitów z odpowiednio wyliczonymi prawdopodobieństwami i stan $|\phi\rangle$, superpozycja 2^n stanów bazowych, przechodzi tylko w jeden z tych stanów.

Rozważmy pomiar j-tego qubitu. Tej sytuacji odpowiada obserwator $\mathcal{B}_j = \{E_j^0, E_j^1\}$, gdzie E_j^0 (E_j^1) jest podprzestrzenią 2^n-wymiarowej przestrzeni Hilberta, rozpiętą na wszystkich wektorach bazowych, które na j-tym bicie mają 0 (1).
Pomiar j-tego qubitu daje

$$0\ (1) \text{ z prawdopodobieństwem } \sum_{i\,:\,i_j=0} |a_i|^2 \ (\sum_{i\,:\,i_j=1} |a_i|^2),$$

gdzie i_j oznacza j-ty bit w binarnej reprezentacji i.

Jeśli będziemy chcieć dokonać pomiaru pierwszych n qubitów $(n+m)$-wymiarowego rejestru kwantowego to wygodnie jest jego stan przedstawić w postaci

$$\sum_{i=0}^{2^n-1} \sum_{j=0}^{2^m-1} c_{ij}|i,j\rangle, \quad \text{gdzie } \sum_{i,j} |c_{ij}|^2 = 1.$$

Jeśli teraz zmierzymy pierwsze n qubitów, każdą z liczb $i \in [0, 2^n)$ otrzymamy z prawdopodobieństwem

$$p(i) = \sum_{j=0}^{2^m-1} |c_{ij}|^2$$

i stan $|\phi\rangle$ zostanie przekształcony w stan

$$|\phi_i\rangle = \frac{1}{\sqrt{p(i)}} \sum_{j=0}^{2^m-1} c_{ij}|i,j\rangle.$$

Zauważmy, że algorytm wykonywany na komputerze kwantowym jest algorytmem probabilistycznym, tzn. wykonując ten sam program kilka razy możemy otrzymać inne wyniki. Jest tak ze względu na losowość pomiaru.

Baza dualna

Baza dualna jest bazą składającą się ze stanów $\{|i'\rangle : 0 \leqslant i < 2^n\}$. Transformacja między bazą standardową, a bazą dualna jest zwana transformacją Hadamarda (lub transformacją Walsha)

$$H_n = \bigotimes_{i=1}^{n} H,$$

gdzie H jest 1-qubitową transformacją Hadamarda.

Łatwo możemy zobaczyć, że

$$H_n|0^{(n)}\rangle = \bigotimes_{i=1}^{n} H|0\rangle = \bigotimes_{i=1}^{n} |0'\rangle = |0'^{(n)}\rangle = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} |i\rangle$$

i ogólnie dla każdego $x \in \{0,1\}^n$

$$H_n|x\rangle = \frac{1}{\sqrt{2^n}} \sum_{y\in\{0,1\}^n} (-1)^{x\cdot y}|y\rangle,$$

gdzie iloczyn skalarny jest zdefiniowany nastepująco: $x \cdot y = \bigoplus_{i=1}^{n} x_i y_i$.

W algorytmach kwantowych często wykonujemy operacje inicjowania rejestru stanem, będącym równomiernym rozkładem wszystkich stanów bazowych

$$|\phi\rangle = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} |i\rangle. \tag{1.3}$$

Okazuje się, że jest łatwy sposób ustawienia "pustego" rejestru $|0^{(n)}\rangle$ na stan (1.3), przez zastosowanie transformacji Hadamarda H do każdego qubitu. Inaczej mówiąc, istnieje liniowa liczba operacji przekształcająca stan bazowy na superpozycję wykładniczej ilości równomiernie rozłożonych stanów bazowych.

Realizacja operacji kwantowych

Architektura komputera kwantowego, podobnie jak klasycznego, zbudowana jest z połączonych ze sobą bramek logicznych. Wejściem i wyjściem kwantowych bramek logicznych są qubity.

Definicja 14 *Bramka kwantowa o n wejściach i n wyjściach jest określona za pomocą unitarnego operatora* $U\colon H_{2^n} \to H_{2^n}$ *i jest reprezentowana za pomocą unitarnej macierzy o wymiarze* 2^n.

Zauważmy, że każde przekształcenie kwantowe jest unitarne, co w konsekwencji oznacza, że bramki kwantowe mają zawsze tą samą ilość wejść i wyjść oraz muszą być odwracalne.

Przykładem bramki 1-qubitowej może być bramka Hadamarda, realizująca unitarną transformację Hadamarda, daną macierzą

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Jej działanie przedstawia diagram

$$|x\rangle \to \boxed{H} \to |0\rangle + (-1)^x |1\rangle, \qquad \text{gdzie } x = 0, 1.$$

Inną bramką 1-qubitową jest fazowe przesunięcie ("phase shift"), które przekształca $|0\rangle \to |0\rangle$ i $|1\rangle \to e^{i\phi}|1\rangle$, co przedstawia macierz

$$\phi = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix}$$

i diagram

$$|x\rangle \to \boxed{\phi} \to e^{ix\phi}|x\rangle, \qquad \text{gdzie } x = 0, 1.$$

Dla obliczeń kwantowych ważną operacją jest XOR, która jest realizowana przez 2-qubitową bramkę CNOT (control-NOT)

$$XOR = CNOT = |a, b\rangle \to |a, a \oplus b\rangle.$$

Niech drugi rejestr zawiera qubit $|0\rangle$. Wówczas bramka ta może być wykorzystana do kopiowania $|0\rangle$ lub $|1\rangle$ pierwszego rejestru. Jeśli natomiast pierwszy rejestr to superpozycja $\alpha|0\rangle + \beta|1\rangle$ wejście zostanie przekształcone w splątany stan $\alpha|00\rangle + \beta|11\rangle$. Operacji tej odpowiada macierz

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Dokładnie bramki kwantowe zostały opisane w [1]. Autorzy pokazali m.in., że zbiór złożony z 2-qubitowej bramki CNOT i wszystkich kwantowych bramek 1-qubitowych tworzy zbiór uniwersalny, tzn. taki, że każda unitarna operacja na dowolnej ilości qubitów może być wykonana za pomocą pewnego układu bramek z tego zbioru.

1.4.5 Kwantowa Transformata Fouriera

Spośród unitarnych operacji wykonywanych na stanach kwantowych ważną rolę odgrywa kwantowa transformata Fouriera, dlatego poświęcamy jej oddzielny rozdział.

Transformata Fouriera jest jednym z silnych narzędzi stosowanych w matematyce. Szczególnie ważną rolę w obliczeniach odgrywa dyskretna transformata Fouriera (DFT), która jest unitarną transformacją q-wymiarowego wektora $\{f(0), \ldots, f(q-1)\}$ na $\{\bar{f}(0), \ldots, \bar{f}(q-1)\}$, gdzie

$$\bar{f}(c) = \frac{1}{\sqrt{q}} \sum_{a=0}^{q-1} e^{\frac{2\pi i a c}{q}} f(a), \tag{1.4}$$

dla $c \in \{0, \ldots, q-1\}$.

Definicja 15 *Kwantowa transformacja Fouriera z bazą q (możemy też powiedzieć w grupie $\mathbb{Z}_q$) jest unitarną transformacją, taką że*

$$QFT_q \colon |a\rangle \to \frac{1}{\sqrt{q}} \sum_{c=0}^{q-1} e^{\frac{2\pi i a c}{q}} |c\rangle \qquad dla\ 0 \leqslant a < q,$$

z macierzą unitarną

$$F_q = \frac{1}{\sqrt{q}} \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & \omega & \omega^2 & \ldots & \omega^{(q-1)} \\ 1 & \omega^2 & \omega^4 & \ldots & \omega^{2(q-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{(q-1)} & \omega^{2(q-1)} & \ldots & \omega^{(q-1)^2} \end{pmatrix},$$

gdzie $\omega = e^{\frac{2\pi i}{q}}$ jest pierwiastkiem $q-$stopnia z 1.

Jeśli zastosujemy QFT_q do kwantowej superpozycji $\frac{1}{\sqrt{q}} \sum_{a=0}^{q-1} f(a)|a\rangle$, transformacja ta przekształci ją według wzoru

$$QFT_q \colon \frac{1}{\sqrt{q}} \sum_{a=0}^{q-1} f(a)|a\rangle \to \frac{1}{\sqrt{q}} \sum_{c=0}^{q-1} \bar{f}(c)|c\rangle, \tag{1.5}$$

gdzie $\bar{f}(c)$ jest zdefiniowane przez (1.4).

Warto zauważyć, że

$$QFT_q \colon |0\rangle \to \frac{1}{\sqrt{q}} \sum_{c=0}^{q-1} |i\rangle, \tag{1.6}$$

czyli zachowuje się tak samo jak transformacja Hadamarda. Większość znanych i ważnych algorytmów używa QFT albo w jej właściwej postaci, albo jako transformację Hadamarda. Stąd pytanie o to, jak można szybko obliczyć QFT jest kluczowe w teorii obliczeń kwantowych. Kwantowa transformata Fouriera jest zwykle stosowana z bazą $q = 2^n$. Klasyczny algorytm wyliczający taką transformatę wymaga $O(n2^n)$ kroków (szybka transformata Fouriera). Zastosowanie obliczeń kwantowych redukuje ten czas do $O(n^2)$. Dzięki temu algorytmy kwantowe wykorzystujące tą transformatę mogą być wykonywane w czasie wielomianowym.

Prostą implementację QFT, dla przypadku gdy baza $q = 2^n$, przedstawili niezależnie od siebie CopperSmith (1994) i Deutsch. Jej opis można znaleźć w [7], [13] lub w [6].

1.5 Kwantowe splątanie (z ang. "entanglement")

Jedną z najbardziej specyficznych i ważnych cech systemów kwantowych jest kwantowe splątanie - zaskakująca własność fizyki kwantowej. Wyobraźmy sobie proces fizyczny polegający na emisji dwóch fotonów (elementarnych pakietów światła) o przeciwnych orientacjach (polaryzacjach) ich oscylującego pola elektrycznego - jednego w lewo, a drugiego w prawo. Pod koniec lat trzydziestych Albert Einstein wraz ze współpracownikami zauważył, że w chwili gdy dokonuje się pomiaru polaryzacji jednego z fotonów, polaryzacja drugiego natychmiast również się ustala, niezależnie od tego, jak daleko się on znajduje. Takie natychmiastowe działanie na odległość jest rzeczywiście zadziwiające. Zjawisko to pozwala, by układy kwantowe wchodziły ze sobą w niezwykłe związki zwane splątaniem, które w kwantowym komputerze odgrywa rolę kabla łączącego qubity. Własność ta jest często wykorzystywana przy projektowaniu algorytmów kwantowych. Na początek rozważymy 2 przykłady:

Przykład 1

Niech 2-qubitowy rejestr będzie w stanie

$$|\psi\rangle = |\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

Jeśli będziemy obserwować pierwszy qubit tego stanu używając standardowego obserwatora $\mathcal{B}_1$, wówczas otrzymamy bit 0 z prawdopodobieństwem $\frac{1}{2}$ i 1 również z prawdopodobieństwem $\frac{1}{2}$. Po takiej obserwacji stan $|\psi\rangle$ przejdzie w stan $|00\rangle$ w pierwszym przypadku albo w stan $|11\rangle$ w drugim. Jeśli teraz zmierzymy drugi qubit, okaże się, że jego wartość będzie wyznaczona jednoznacznie z prawdopodobieństwem 1. Czyli qubity rejestru kwantowego zainicjowanego stanem $|\psi\rangle$ są ze sobą związane.

Przykład 2

Niech teraz stan $|\psi\rangle$ 2-qubitowego rejestru będzie tensorowym produktem $|\psi_1\rangle \otimes |\psi_2\rangle$ dwóch stanów 1-qubitowych $|\psi_1\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ i $|\psi_2\rangle = \beta_0|0\rangle + \beta_1|1\rangle$:

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle = (\sum_{i=1}^{1} \alpha_i|i\rangle) \otimes (\sum_{j=1}^{1} \beta_j|j\rangle).$$

Jeśli będziemy obserwować pierwszy qubit takiego stanu $|\psi\rangle$ otrzymamy

$$\begin{array}{ll} 1 \text{ z prawdopodobieństwem} & |\alpha_0\beta_0|^2 + |\alpha_0\beta_1|^2 = |\alpha_0|^2, \\ 0 \text{ z prawdopodobieństwem} & |\alpha_1\beta_0|^2 + |\alpha_1\beta_1|^2 = |\alpha_1|^2. \end{array}$$

Dodatkowo po obserwacji pierwszego qubitu, stan $|\psi\rangle$ redukuje się do $|\psi_2\rangle$, a po obserwacji drugiego do $|\psi_1\rangle$. W tym przypadku qubity stanu $|\psi\rangle$ nie zależą od siebie i można na nich osobno wykonywać pewne operacje.

Jeśli kwantowy stan z przestrzeni Hilberta H nie może być przedstawiony za pomocą produktu tensorowego dwóch stanów z przestrzeni Hilberta o wymiarze mniejszym od wymiaru H, to nazywamy go **stanem splątanym**. Jeśli jest on stanem n-qubitowego rejestru, to mówimy również, że wszystkie jego qubity są splątane.

Stany splątane pojawiają się w naturalny sposób na skutek interakcji miedzy systemami kwantowymi. Dodatkowo niektóre kwantowe operacje tworzą je ze stanów niesplątanych. Na przykład operacja XOR zastosowana do stanu $\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)\otimes|0\rangle$ tworzy splatany stan $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$.

W skutek tego, że pomiar jednego qubitu qubitów splątanych natychmiast determinuje stan pozostałych, odległe części systemów kwantowych bardzo na siebie wpływają. Kwantowe przetwarzanie informacji oferuje więc silne narzędzie wykraczające poza realia świata obliczeń klasycznych. Jest to wykorzystywane do budowania bardzo efektywnych algorytmów. Zarazem zjawisko to jest jedną z poważniejszych trudności uniemożliwiających symulowanie systemu kwantowego z pomocą komputerów klasycznych.

1.6 Kwantowa równoległość

Wielki potencjał obliczeń kwantowych tkwi w tzw. "kwantowej równoległości". Każde przekształcenie kwantowe jest wykonywane za pomocą liniowych (unitarnych) operatorów. Po zastosowaniu operatora A do stanu

$$|\phi\rangle = \sum_{i=0}^{2^n-1} c_i|i\rangle$$

otrzymujemy

$$A|\phi\rangle = \sum_{i=0}^{2^n-1} c_i A|i\rangle,$$

czyli pojedyncza transformacja A powoduje wykonanie wykładniczej ilości operacji na stanach bazowych.

Kwantowy system może być równocześnie w wykładniczej ilości stanów bazowych. Liniowa ilość operacji może spowodować powstanie superpozycji wykładniczej ilości stanów, natomiast w jednym kroku może zostać wykonana wykładnicza ilość operacji. Dodatkowo wielkość równoległości wzrasta wykładniczo względem rozmiaru systemu. Ten wykładniczy wzrost wymaga jedynie liniowego zwiększenia wymiaru przestrzeni. Aby zrozumieć, czym tak naprawdę jest "kwantowa równoległość" rozważmy prosty przykład:

Przykład

Niech $f\colon \{0,1,\ldots,2^m-1\} \to \{0,1,\ldots,2^n-1\}$. Zakładamy, że mamy unitarny operator U_f, taki że jeśli zastosujemy go do $(m+n)$-qubitowego rejestru, złożonego z dwóch podrejestrów - m-qubitowego dla $x \in \{0,1,\ldots,2^m-1\}$ i n-qubitowego dla $b \in \{0,1,\ldots,2^n-1\}$, to otrzymamy

$$U_f|x,b\rangle = |x, b \oplus f(x)\rangle.$$

Za jego pomocą możemy więc wyliczać wartości funkcji f. Aby klasycznie obliczyć wszystkie wartości funkcji $f(x)$ musielibyśmy policzyć wartość funkcji 2^m razy (dla $m = 100 \sim 10^{30}$ razy). Na komputerze kwantowym możemy to zrobić w jednym kroku. Jeśli operator U_f zastosujemy dla równomiernej (jednostajnej) superpozycji wszystkich 2^m stanów bazowych, czyli do $|\phi\rangle = \frac{1}{\sqrt{2^m}} \sum_{x=0}^{2^m-1} |x\rangle$ otrzymamy

$$U_f|\phi, 0\rangle = U_f \frac{1}{\sqrt{2^m}} \sum_{x=0}^{2^m-1} |x, 0\rangle = \frac{1}{\sqrt{2^m}} \sum_{x=0}^{2^m-1} U_f|x, 0\rangle = \frac{1}{\sqrt{2^m}} \sum_{x=0}^{2^m-1} |x, f(x)\rangle = |\psi\rangle. \quad (1.7)$$

Czyli w jednym kroku, za pomocą transformacji U_f możemy obliczyć wszystkie wartości funkcji $f(x)$ dla $0 \leqslant x < 2^m$.

Z wyjątkiem trywialnego przypadku, wynikowy stan $|\psi\rangle$ z (1.7) jest splątany. Jest on superpozycją stanów postaci $|x, f(x)\rangle$ (dla $x = 1, \ldots, 2^n$). Drugi podrejestr (ostatnie n-qubitów) zawiera wartość funkcji f dla elementu znajdującego się w pierwszym podrejestrze (pierwsze m-qubitów). Gdy zmierzymy pierwszy podrejestr ze względu na bazę standardową, otrzymamy pewna wartość x_0 losowo wybraną ze zbioru $\{0, 1. \ldots, 2^m - 1\}$. Po tej operacji drugi podrejestr będzie odpowiadał wartości funkcji $f(x_0)$. Stan $|\phi\rangle$ z 2^m-elementowej superpozycji przekształci się w $|x_0, f(x_0)\rangle$. Pomiar drugiego rejestru da nam jednoznacznie $f(x_0)$. Powyższe rozważania pokazują, że qubity stanu (1.7) zależą od siebie.

Rozdział 2

Definicje i twierdzenia

2.1 Podstawowe informacje dotyczące grup

W rozdziale 3 przedstawimy kwantowy algorytm obliczania rzędu grupy rozwiązalnej, działający w czasie wielomianowym. Aby zrozumieć jego działanie, konieczne jest zapoznanie się z definicjami i własnościami dotyczącymi grup rozwiązalnych, czemu będzie poświęcony ten rozdział.

Definicja 16 *Niech G będzie grupą i niech $g, h \in G$. Iloczyn $ghg^{-1}h^{-1}$ nazywamy komutatorem elementów g i h oraz oznaczamy przez $[g, h]$.*

Definicja 17 *Niech G będzie grupą. Komutantem (grupa pochodną) grupy G nazywamy podgrupę $G' = [G, G]$ generowaną przez zbiór wszystkich komutatorów*

$$G' = \{[g, h] \ : \ g, h \in G\}.$$

Dla komutantów zachodzą dwie poniższe własności (zobacz [9])

1. $G' \lhd G$.
2. Grupa ilorazowa $G^{(k)}/G^{(k+1)}$ jest abelowa.

Możemy rozpatrzyć również komutant grupy G' : $(G')' = G''$, zwany drugą grupą pochodną (lub drugim komutantem) grupy G. Kontynuując to postępowanie otrzymamy k-tą grupę pochodną $G^{(k)} = (G^{(k-1)})'$. Zgodnie z wymienionymi wyżej własnościami komutantów otrzymujemy ciąg podgrup normalnych

$$G \rhd G^{(1)} \rhd G^{(2)} \rhd \ldots \rhd G^{(k-1)} \rhd G^{(k)} \rhd \ldots \tag{2.1}$$

z abelowymi grupami ilorazowymi $G^{(k)}/G^{(k+1)}$.

Definicja 18 *Skończoną grupę G nazywamy rozwiązalną, jeśli ciąg podgrup (2.1) urywa się na grupie jedynkowej, tj. istnieje takie m, że $G^{(m)} = \{1\}$.*

Równoważnie można zdefiniować grupę rozwiązalną w następujący sposób:

Definicja 19 *Skończona grupa G jest rozwiązalna, jeśli istnieją elementy $g_1, ..., g_m \in G$, takie że jeśli dla każdego $i \in \{0, \ldots, m\}$ $H_i = \langle g_1, \ldots, g_i \rangle$, to*

$$\{1\} = H_0 \lhd H_1 \lhd \ldots \lhd H_m = G.$$

Definicja 20 *Rzędem elementu $g \in G$ względem podgrupy H nazywamy najmniejszą liczbę p, taką że $g^p \in H$, tzn. $r_H = \min\{p > 0 : \ g^p \in H\}$.*

Twierdzenie 6 *Niech G będzie grupą rozwiązalną, a $H_0, \ldots, H_m$ ciągiem normalnych podgrup G określonych przez definicję 19. Rząd grupy G możemy przedstawić jako iloczyn rzędów elementów g_j względem odpowiednich podgrup H_j:*

$$|G| = \prod_{j=1}^{m} r_j.$$

Dowód

Z twierdzenia Lagrange'a wiemy, że rząd grupy jest równy rzędowi dowolnej podgrupy pomnożonemu przez rząd grupy ilorazowej, którą ona wyznacza. Stąd rząd naszej grupy rozwiązalnej G możemy wyznaczyć w następujący sposób:

$$\begin{aligned} |G| \ = \ |H_m| \ = \ |H_m/H_{m-1}| \cdot |H_{m-1}| \ = \ |H_m/H_{m-1}| \cdot |H_{m-1}/H_{m-2}| \cdot |H_{m-2}| \ = \ldots = \\ = \ |H_m/H_{m-1}| \cdot |H_{m-1}/H_{m-2}| \cdot \ldots \cdot |H_1/H_0| \cdot |H_0| \ = \ \prod_{j=1}^{m} |H_j/H_{j-1}|. \end{aligned}$$

Niech r_j oznacza rząd elementu g_j względem H_{j-1}, czyli $r_j = \min\{p > 0 : \ g_j^p \in H_{j-1}\}$. Ponieważ $H_{j-1} \lhd H_j$, to $g_j \notin H_{j-1}$ i w konsekwencji

$$H_j/H_{j-1} = \{H_{j-1}, gH_{j-1}, g^2H_{j-1}, \ldots, g^{r_j-1}H_{j-1}\} \Rightarrow |H_j/H_{j-1}| = r_j.$$

Otrzymujemy stąd, że

$$|G| = \prod_{j=1}^{m} |H_j/H_{j-1}| = \prod_{j=1}^{m} r_j.$$

c.n.d.

2.2 Grupy "black-box"

Definicja 21 *Skończoną grupę, w której każdy element jest jednoznacznie zakodowany za pomocą ciągu binarnego (kodu) o ustalonej długości n, zwanej długością kodowania, nazywamy grupą "black-box". Rząd takiej grupy jest ograniczony przez $|G| \leqslant 2^n$. Grupa "black-box" jest dana jako lista generatorów. Do naszej dyspozycji jest też "czarne-pudełko", które wykonuje podstawowe operacje na kodach, takie jak mnożenie czy wskazanie elementu odwrotnego. Koszt każdej z tych operacji jest jednostkowy.*

Warto zauważyć, że nie każdy ciąg binarny o długości n odpowiada pewnemu elementowi grupy. Można założyć, że nasza "czarna skrzynka" zachowuje się dowolnie dla takich niepoprawnych kodów.

Algorytmy pracujące z grupą "black-box" próbują uzyskać specyficzne własności danej grupy, ale ich działanie nie zależy od reprezentacji grupy, a w szczególności od tego, w jaki sposób są wykonywane operacje na jej elementach. Od tej pory, kiedy będziemy mówić, że dana jest jakaś grupa lub podgrupa będziemy mieć na myśli, że dany jest zbiór ciągów binarnych odpowiadających generatorom grupy czy podgrupy.

Przyjmijmy, że daną mamy rozwiązalną grupę "black-box" G, czyli mamy dane generatory $g_1, \ldots, g_m$ jako ciągi binarne o długości n. Istnieje klasyczny, wielomianowy algorytm ([4]), który w czasie mn umie przetestować czy dana grupa G jest rozwiązalna.

Ta uniwersalna reprezentacja grup została wprowadzona przez Babai oraz Szemerédi w 1984 roku ([5]). Obszerniejszą analizę grup "black-box" można znaleźć w [2].

W świecie kwantowym każdej grupie "black-box" o długości kodowania n odpowiada operator U_G, działający na $2n$ qubitach, za pomocą którego możemy wykonywać działania w grupie w następujący sposób:

$$U_G|g\rangle|h\rangle = |g\rangle|gh\rangle,$$

gdzie g i h odpowiadają elementom grupy.

Mamy również do dyspozycji operator U_G^{-1} realizujący operację odwrotną do U_G. Przy jego użyciu możemy znajdować elementy odwrotne w grupie

$$U_G^{-1}|g\rangle|h\rangle = |g\rangle|g^{-1}h\rangle.$$

Opis realizacji powyższych operatorów można znaleźć w [12].

Wykonując działania kwantowe, mające na celu badanie pewnych właściwości danej grupy, będziemy operować na stanie będącym superpozycją elementów danej grupy "black-box". Poniżej zamieszczamy dokładną definicję takiego stanu:

Definicja 22 *Superpozycją elementów grupy "black-box" G nazywamy stan, będący wynikiem superpozycji stanów odpowiadających binarnemu przedstawieniu elementów grupy G. Oznaczamy ją następująco:*

$$|G\rangle = \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g\rangle.$$

2.3 Ułamki łańcuchowe

Ułamki łańcuchowe są innym sposobem zapisywania ułamków postaci $\frac{P}{Q}$.

Definicja 23 *Wyrażenie postaci*

$$\frac{P}{Q} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ldots}}}, \tag{2.2}$$

gdzie $a_0, a_1, a_2, \ldots \in \mathbb{Z}$, a $P, Q \in \mathbb{Z}^+$ nazywamy ułamkiem łańcuchowym. Równoważnie można go zapisać w postaci

$$\frac{P}{Q} = [a_0; a_1, a_2, a_3, \ldots]. \tag{2.3}$$

Pierwsza z liczb łańcucha (2.3), a_0, jest liczbą nieujemną, będącą częścią całkowitą rozpatrywanego ułamka $\frac{P}{Q}$. Jeśli $\frac{P}{Q} < 1$, wówczas $a_0 = 0$. Nietrudno zauważyć, że gdy liczba $\frac{P}{Q}$ jest wymierna, wówczas ciąg $a_1, a_2, a_3, \ldots$ jest skończony.

Przykład 1

$$\frac{45}{16} = \frac{16+16+13}{16} = 2 + \frac{13}{16} = 2 + \frac{1}{\frac{16}{13}} = 2 + \frac{1}{1+\frac{3}{13}} = 2 + \frac{1}{1+\frac{1}{\frac{13}{3}}} = 2 + \frac{1}{1+\frac{1}{4+\frac{1}{3}}}.$$

Alternatywnie możemy zapisać ułamek $\frac{45}{16}$ w postaci

$$\frac{P}{Q} = [2; 1, 4, 3].$$

Ułamek łańcuchowy $\frac{P}{Q} = [a_0; a_1, a_2, a_3, \ldots]$ możemy uzyskać za pomocą prostej procedury rekurencyjnej:

1. $a_i = \lfloor \delta_i \rfloor$,
2. $\delta_{i+1} = \dfrac{1}{\delta_i - a_i}$,

gdzie $\delta_0 = \frac{P}{Q}$.

Przedstawianie ułamków w postaci łańcuchowej jest wykorzystywane w różny sposób. Między innymi, postać ta może służyć do przybliżania liczb niewymiernych liczbami wymiernymi. W algorytmie, który zaprezentujemy w kolejnym rozdziale, będziemy stosować tą metodę do przybliżania ułamka o bardzo dużym mianowniku innym ułamkiem o mianowniku mniejszym. Efekt ten możemy uzyskać obcinając ostatnią liczbę łańcucha (2.3) (lub kilka liczb), przez co uzyskamy pożądany ułamek.

2.4 Pomocnicze twierdzenia i dowody

Definicja 24 *Mówimy, że funkcja określona w przedziale $[b, c]$ jest wklęsła w tym przedziale, jeżeli dla każdej liczby x_a ($a \in [0, 1]$) postaci*

$$x_a = ax_1 + (1-a)x_2 \quad \textit{dla dowolnych } x_1, x_2 \in [b, c],$$

zachodzi nierówność

$$f(x_a) \geqslant af(x_1) + (1-a)f(x_2).$$

Warunek wklęsłości funkcji $f(x)$ oznacza geometrycznie, że łuk wykresu funkcji $y = f(x)$ o końcach $M_1(x_1, f(x_1))$ i $M_2(x_2, f(x_2))$ znajduje się powyżej odcinka M_1M_2.

Zachodzi następujące twierdzenie:

Twierdzenie 7 *Jeżeli funkcja $f(x)$ jest w przedziale $[b,c]$ dwukrotnie różniczkowalna, a jej druga pochodna przyjmuje w tym przedziale wartości ujemne, to funkcja $f(x)$ jest w przedziale $[b,c]$ funkcją wklęsłą.*

Twierdzenie 8 *Dla dowolnego kąta θ zachodzą następujące nierówności:*

(a) $|1-e^{i\theta}| \leqslant |\theta|$,

(b) $|1-e^{iA\theta}| \geqslant \frac{2A|\theta|}{\pi}$, *jeśli* $A|\theta| \leqslant \pi$.

Dowód nierówności (a)

Aby udowodnić nierówność $|1-e^{i\theta}| \leqslant |\theta|$ przekształcimy jej lewą stronę

$$L = |1-e^{i\theta}| = |1-\cos\theta - i\sin\theta| = \sqrt{(1-\cos\theta)^2+(\sin\theta)^2} =$$

$$= \sqrt{1-2\cos\theta+\cos^2\theta+\sin^2\theta} = \sqrt{2-2\cos\theta} = \sqrt{2}\sqrt{1-\cos\theta} =$$

$$= \sqrt{2}\sqrt{1-\cos^2\frac{\theta}{2}+\sin^2\frac{\theta}{2}} = \sqrt{2}\sqrt{\sin^2\frac{\theta}{2}+\cos^2\frac{\theta}{2}-\cos^2\frac{\theta}{2}+\sin^2\frac{\theta}{2}} =$$

$$= \sqrt{2}\sqrt{2\sin^2\frac{\theta}{2}} = 2|\sin\frac{\theta}{2}| \leqslant 2|\frac{\theta}{2}| = |\theta| = P.$$

W ostatnim przejściu skorzystaliśmy z faktu, że $\sin x \leqslant x$.

c.n.d.

Dowód nierówności (b)

Aby udowodnić powyższą własność pokażemy, że $f(\theta) = |1-e^{iA\theta}|$ jest funkcją wklęsłą w przedziałach $[-\frac{\pi}{A},0]$ i $[0,\frac{\pi}{A}]$ (zob. definicja wklęsłości 24). Oznacza to, że odcinek łączący punkty $(0,f(0)) = (0,0)$ i $(\frac{\pi}{A},f(\frac{\pi}{A})) = (\frac{\pi}{A},2)$ znajduje się poniżej wykresu funkcji $f(\theta)$. Pokażemy też, że odcinek ten jest wykresem funkcji $g(\theta) = \frac{2A|\theta|}{\pi}$ dla $\theta \in [0,\frac{\pi}{A}]$. Stąd uzyskamy, że $f(\theta) \geqslant g(\theta)$ dla $\theta \in [0,\frac{\pi}{A}]$.

Dla przedziału $[-\frac{\pi}{A},0]$ rozumowanie przebiega analogicznie.

Aby pokazać wklęsłość funcji $f(\theta)$, zapiszemy ją w innej postaci

$$f(\theta) = |1-e^{iA\theta}| = |1-\cos A\theta - i\sin A\theta| = \sqrt{(1-\cos A\theta)^2+(\sin A\theta)^2} =$$

$$= \sqrt{1-2\cos A\theta+\cos^2 A\theta+\sin^2 A\theta} = \sqrt{2-2\cos A\theta} = \sqrt{2}\sqrt{1-\cos A\theta},$$

wyznaczymy jej pochodną

$$f'(\theta) = \sqrt{2}\frac{A\sin A\theta}{2\sqrt{1-\cos A\theta}}$$

i druga pochodną

$$f''(\theta) = \frac{A\sqrt{2}}{2} \cdot \frac{A\cos A\theta \cdot \sqrt{1-\cos A\theta} - \sin A\theta \cdot \frac{A\sin A\theta}{\sqrt{1-\cos A\theta}}}{1-\cos A\theta} =$$

$$= \frac{A^2\sqrt{2}}{2} \cdot \frac{\frac{\cos A\theta(1-\cos A\theta)}{\sqrt{1-\cos A\theta}} - \frac{\sin^2 A\theta}{\sqrt{1-\cos A\theta}}}{1-\cos A\theta} = \frac{A^2\sqrt{2}}{2} \cdot \frac{\cos A\theta - \cos^2 A\theta - \sin^2 A\theta}{\sqrt{1-\cos A\theta}(1-\cos A\theta)} =$$

$$= -\frac{A^2\sqrt{2}}{2} \cdot \frac{1-\cos A\theta}{\sqrt{1-\cos A\theta}\cdot(1-\cos A\theta)} = -\frac{A^2\sqrt{2}}{2} \cdot \frac{1}{\sqrt{1-\cos A\theta}} < 0,$$

co na mocy twierdzenia 7 oznacza, że funkcja $f(\theta)$ jest wklęsła.

Obliczymy teraz wartości funkcji f na końcach przedziału $[0, \frac{\pi}{A}]$

$$f(0) = \sqrt{2}\sqrt{1-\cos 0} = 0,$$

$$f(\frac{\pi}{A}) = \sqrt{2}\sqrt{1-\cos A\frac{\pi}{A}} = 2.$$

Odcinek łączący punkty $(0,0)$ i $(\frac{\pi}{A}, 2)$ jest zawarty w wykresie liniowej funkcji $g(\theta) = a\theta + b$, przechodzącej przez te punkty. Możemy ją prosto wyznaczyć

$$g(0) = 0,$$

$$g(\frac{\pi}{A}) = 2,$$

$$g(\frac{\pi}{A}) = a\frac{\pi}{A} + b \implies b = 0 \text{ i } a = \frac{2A}{\pi},$$

czyli $g(\theta) = \frac{2A\theta}{\pi}$.

Ponieważ funkcja f jest wklęsła w $[0, \frac{\pi}{A}]$, wykres funkcji $g(\theta)$ znajduje się pod wykresem funkcji $f(\theta)$, czyli

$$|1 - e^{iA\theta}| \geqslant \frac{2A\theta}{\pi} \quad \text{dla } 0 \leqslant \theta \leqslant \frac{\pi}{A}.$$

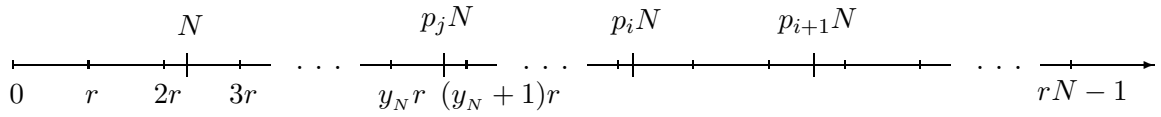
Dowód dla $-\frac{\pi}{A} \leqslant \theta \leqslant 0$ jest analogiczny.

c.n.d.

Twierdzenie 9 *Niech $r \in \{0, 1, \ldots, N-1\}$. Istnieje dokładnie r takich wartości $y \in \{0, 1, \ldots, N-1\}$, które spełniają poniższy warunek*

$$-\frac{r}{2} \leqslant yr \bmod N \leqslant \frac{r}{2}.$$

Bez straty ogólności przyjmujemy, że wyniki operacji *mod* N należą do przedziału $(-\frac{N}{2}, \frac{N}{2}]$.



Rysunek 2.1: Wielokrotności r i N

Dowód

Powyższą własność można zobaczyć analizując położenie względem siebie wielokrotności r i N na osi liczbowej od 0 do $rN-1$ (rysunek (2.1)). Każda z wielokrotności N leży między pewnym y_Nr a $(y_N+1)r$, gdzie $y_N \in \{0,1,\ldots,N-2\}$.

Wówczas albo $-\frac{r}{2} \leqslant y_Nr \ mod\ N \leqslant \frac{r}{2}$, albo $-\frac{r}{2} \leqslant y_{N+1} \ mod\ N \leqslant \frac{r}{2}$.
W przedziale $[0, rN-1]$ jest r wielokrotności N i tyle też będzie $y \in \{0,1,\ldots,N-1\}$, spełniających zadaną nierówność.

c.n.d.

Twierdzenie 10 *Jeśli y jest liczbą losowo wybraną ze zbioru $\{1,\ldots,r\}$, to prawdopodobieństwo, że $nwd(y,r)=1$ jest klasy $\Omega(\frac{1}{\log\log r})$.*

Prawdopodobieństwo, że dla y wylosowanego ze zbioru $\{1,\ldots,r\}$ zachodzi własność $nwd(y,r)=1$ jest równe $\frac{\phi(n)}{n}$, gdzie ϕ jest funkcją Eulera. Aby udowodnić powyższe twierdzenie wystarczy skorzystać z równości

$$\liminf_{n\to\infty} \frac{\phi(n)\log\log n}{n} = e^{-C},$$

gdzie ϕ jest funkcją Eulera, a C stałą Eulera. Jej dowód można znaleźć w [8].

Twierdzenie 11 *Niech $b_1, b_2, \ldots, b_p$ będą liczbami losowo wybranymi ze zbioru $\{1,\ldots,R\}$, a P_p będzie prawdopodobieństwem, że $nwd(b_i,p) \neq 1\ \forall i \in \{1,\ldots,p\}$. Dla ustalonego $\varepsilon \in (0,1)$ i pewnego $p = \Theta(\log\log R \cdot \log(\frac{1}{\varepsilon}))$ zachodzi $P_p \leqslant \varepsilon$.*

Dowód

Ustalamy $\varepsilon \in (0,1)$. Aby udowodnić powyższe twierdzenie, pokażemy, że $\exists d \geqslant 0$:

$$p = d\log\log R \cdot \log(\frac{1}{\varepsilon}) \quad \text{spełnia nierówność} \quad P_p \leqslant \varepsilon. \tag{2.4}$$

Z twierdzenia 10 wiemy, że dla dostatecznie dużych r

$$\exists c \geqslant 0 : P(nwd(b_i,r)=1) \geqslant \frac{c}{\log\log r}.$$

Zatem

$$\exists c \geqslant 0 \ : \ P_p \leqslant (1-\frac{c}{\log\log r})^p.$$

Chcemy, by szukane przez nas d spełniało nierówność

$$\Big(1-\frac{c}{\log\log r}\Big)^{d\log\log r\cdot\log(\frac{1}{\varepsilon})}\leqslant\varepsilon.$$

Logarytmując wyrażenie obustronnie otrzymujemy

$$d\log\log r\cdot\log(\frac{1}{\varepsilon})\cdot\log(1-\frac{c}{\log\log r})\leqslant\log\varepsilon$$

$$\Downarrow$$

$$-d\log\log r\cdot\log\varepsilon\cdot\log(1-\frac{c}{\log\log r})\leqslant\log\varepsilon.$$

Z założeń twierdzenia wiemy, że $\varepsilon\in(0,1)$, czyli $\log\varepsilon<0$. Dzieląc wyrażenie obustronnie przez $\log\varepsilon$ i dalej przekształcając uzyskujemy

$$-d\log\log r\cdot log(1-\frac{c}{\log\log r})\geqslant 1$$

$$\Downarrow$$

$$d\log\log r\cdot log(\frac{\log\log r}{\log\log r-c})\geqslant 1$$

$$\Downarrow$$

$$d\geqslant\frac{1}{\log\log r\cdot\log(\frac{\log\log r}{\log\log r-c})}$$

$$\Downarrow$$

$$d\geqslant\frac{1}{\log(\frac{\log\log r}{\log\log r-c})^{\log\log r}}.$$

Aby istniało d, takie że dla każdego r (poza $r=1,2$) spełniona jest powyższa nierówność, funkcja $f(r)=\dfrac{1}{\log(\frac{\log\log r}{\log\log r-c})^{\log\log r}}$ musi być ograniczona od góry. Zbadamy jak zachowuje się ona w granicy:

$$\Big(\frac{\log\log r}{\log\log r-c}\Big)^{\log\log r}=\Big(\frac{\log\log r-c+c}{\log\log r-c}\Big)^{\log\log r}=$$

$$\Big(1+\frac{c}{\log\log r-c}\Big)^{\log\log r}\xrightarrow{r\to\infty}e^c\quad\Rightarrow\quad\frac{1}{\log(\frac{loglogr}{\log\log r-c})^{\log\log r}}\xrightarrow{r\to\infty}\frac{1}{\log e^c},$$

czyli

$$f(r)\xrightarrow{r\to\infty}\frac{1}{\log e^c}.$$

Widać stąd, że $\exists d$, takie że dla prawie wszystkich r, jeśli

$$p=d\log\log r\cdot\log(\frac{1}{\varepsilon}),\qquad\text{to spełniona jest nierówność}\qquad P_p\leqslant\varepsilon.$$

Można więc dla naszego R wyznaczyć takie d_0, że

$$p = d_0 \log\log R \cdot \log(\frac{1}{\varepsilon}) \quad \text{spełnia nierówność} \quad P_p \leqslant \varepsilon.$$

c.n.d.

Twierdzenie 12 *Niech $u_1, v_1, u_2, v_2, k_1, k_2, r$ będą liczbami naturalnymi, takimi że*

(i) $\frac{u_1}{v_1} = \frac{k_1}{r}$ *i* $\frac{u_2}{v_2} = \frac{k_2}{r}$,

(ii) $nwd(k_1, r) = 1$ *i* $nwd(k_2, r) = 1$.

Jesli dodatkowo $nwd(k_1, k_2) = 1$, to $r = nww(v_1, v_2)$.

Dowód
Z warunków (i) i (ii) mamy, że $\exists\ b_1, b_2 \in \mathbb{N}$:

$$\begin{array}{ccc} u_1 \cdot b_1 = k_1 & & u_2 \cdot b_2 = k_2 \\ v_1 \cdot b_1 = r & \text{i} & v_2 \cdot b_2 = r. \end{array}$$

Ponieważ $v_1 \cdot b_1 = r$ oraz $v_2 \cdot b_2 = r$, r jest wielokrotnością liczb v_1 i v_2. Aby udowodnić tezę twierdzenia, wystarczy pokazać, że jest to wielokrotność najmniejsza.

Hipoteza: Istnieje liczba naturalna $A > 1$, taka że $\frac{r}{A} = nww(v_1, v_2)$.
Czyli istnieją $c, d \in \mathbb{N}$, takie że

$$\begin{array}{l} \frac{r}{A} = c \cdot v_1 \\ \frac{r}{A} = d \cdot v_2 \end{array} \implies \begin{array}{l} \frac{r}{A} = c \cdot \frac{r}{b_1} \\ \frac{r}{A} = d \cdot \frac{r}{b_2} \end{array} \implies \begin{array}{l} b_1 = Ac \\ b_2 = Ad \end{array} \implies nwd(b_1, b_2) \neq 1. \qquad (2.5)$$

Z faktu, że liczby k_1 i k_2 są względnie pierwsze otrzymujemy

$$nwd(k_1, k_2) = 1 \Rightarrow nwd(u_1 \cdot b_1, u_2 \cdot b_2) = 1 \Rightarrow nwd(b_1, b_2) = 1. \qquad (2.6)$$

Własności (2.5) i (2.6) dają razem sprzeczność.
Zatem otrzymujemy, że $nww(v_1, v_2) = r$.

c.n.d.

Rozdział 3

Obliczanie rzędu grupy rozwiązalnej

W tym rozdziale przedstawimy kwantowy algorytm znajdowania rzędu rozwiązalnej grupy "black-box" i dodatkowo wyznaczający stan będący superpozycją wszystkich jej elementów.

Twierdzenie 13 *Istnieje kwantowy algorytm działający w następujący sposób: Mając dane generatory $g_1, g_2, \ldots, g_m$, takie że grupa "black-box" $G = \langle g_1, g_2, \ldots, g_m \rangle$ o długości kodowania n jest rozwiązalna, algorytm wylicza rząd dla tak zdefiniowanej grupy z prawdopodobieństwem błędu ograniczonym przez ε, działając w czasie wielomianowym od $mn + \log(1/\varepsilon)$. Dodatkowo algorytm z tym samym prawdopodobieństwem generuje kwantowy stan ϕ przybliżający stan $|G\rangle = \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g\rangle$.*

Zakładamy, że mamy dane $d_1, \ldots, d_m \in G$, generatory grupy "black-box" G o długości kodowania n. (Własności grup rozwiązalnych oraz grup "black-box" znajdują się w rozdziale 2.) Za pomocą algorytmu klasycznego, wielomianowego od nm ([4]) można sprawdzić czy dana grupa $G = \langle d_1, d_2, \ldots, d_m \rangle$ jest grupą rozwiązalną.

Z definicji rozwiązalności grupy (def. 19) wiemy, że istnieją takie elementy $g_1, g_2, \ldots, g_m \in G$, że jeśli dla każdego $i \in \{0, \ldots, m\}$ $H_i = \langle g_1, \ldots, g_i \rangle$, to

$$\{1\} = H_0 \lhd H_1 \lhd \ldots \lhd H_m = G.$$

Elementy te możemy uzyskać za pomocą wspomnianego już algorytmu z [4], wielomianowego od nm.

Przedstawiany przez nas algorytm obliczania rzędu grupy G wykorzystuje fakt, że $|G|$ można przedstawić za pomocą iloczynu rzędów elementów g_j (dla $j \in \{1, \ldots, m\}$) względem odpowiednich podgrup H_j (zobacz def. 20 i tw. 6), czyli

$$|G| = \prod_{j=1}^{m} r_j,$$

gdzie r_j jest rzędem elementu g_j względem podgrupy H_{j-1}. Przypominamy, że rząd elementu g_j względem podgrupy H_{j-1} oznacza najmniejszą taką liczbę $p > 0$, że $g^p \in H$, czyli $r_j = \min\{p : \ g^p \in H_{j-1}\}$.

Zarysujemy teraz ideę omawianego algorytmu. Będzie on działał w m etapach. W każdym z nich obliczone zostaną kolejne r_j ($j \in \{1, \ldots, m\}$). Mnożenie uzyskanych w ten sposób liczb da nam pożądany wynik.

Rząd elementów g_j względem H_{j-1} będziemy wyliczać w czasie wielomianowym za pomocą algorytmu, będącego modyfikacją pomysłu P.W. Shora z 1994 roku znajdowania okresu funkcji $x^p \ mod \ n$, wykorzystywanego do rozkładu liczb na czynniki pierwsze (zob. [13] lub [14]). Metoda ta będzie wymagała kilku kopii stanu $|H_{j-1}\rangle$, który jest superpozycją wszystkich elementów podgrupy H_{j-1} (zob. def. 22). Musimy więc umieć efektywnie otrzymywać taki stan. Jak pokażemy dokładnie poniżej, stany $|H_0\rangle, |H_1\rangle, \ldots, |H_m\rangle$ będziemy uzyskiwać w połączeniu z obliczaniem $r_1, r_2, \ldots, r_m$. Rozpoczniemy z dużą (wielomianową względem m) ilością stanów $|H_0\rangle$. Część z nich użyjemy do wyliczenia r_1, a pozostałe przekształcimy na stany $|H_1\rangle$. W ten sposób w j-tym kroku będziemy wyznaczać r_j, przy wykorzystaniu części stanów $|H_{j-1}\rangle$ i konstruować $|H_j\rangle$ z pozostałych. Po m-tym kroku będziemy mieć wyznaczone $r_1, r_2, \ldots, r_m$ i stan $|H_m\rangle$. Mnożenie liczb $r_1, r_2, \ldots, r_m$ da nam poszukiwany rząd grupy G. Równocześnie jako skutek uboczny naszego algorytmu będziemy mieć stan $|H_m\rangle$, będący superpozycją wszystkich elementów grupy G.

Z powyższego szkicu wynika, że potrzebujemy dwóch procedur: jednej wyliczającej rząd elementu względem podgrupy i drugiej przekształcającej kilka kopii stanu $|H\rangle$ na kopie stanu $|\langle g \rangle H\rangle$. Tym dwóm problemom poświecone będą kolejne dwa podrozdziały. W 3.3 przedstawimy główny algorytm, który wykorzystując wspominane procedury w czasie wielomianowym wyznacza rząd rozwiązalnej grupy G.

3.1 Obliczanie rzędu elementu grupy względem podgrupy

Mamy daną grupę G, element $g \in G$ oraz grupę H, będącą podgrupą grupy G. Grupy G i H są grupami "black-box" o długości kodowania n. Oznacza to zgodnie z def. 21, że mamy dane jedynie ich generatory a nie wszystkie elementy grupy. Konsekwencją takiej reprezentacji jest to, że sprawdzenie czy dany element należy do grupy nie jest problemem trywialnym. W tym rozdziale pokażemy, jak przy użyciu algorytmu kwantowego obliczyć rząd elementu g względem podgrupy H w czasie wielomianowym od n (gdzie n oznacza długość kodowania grupy G).

Spróbujemy zarysować ideę opisywanego algorytmu. Pracować będziemy na dwóch rejestrach A i R i wykorzystywać przede wszystkim kwantową transformatę Fouriera opisaną w rozdziale 1.4.5. Algorytm opiera sie na pomyśle Shor'a znajdowania rzędu elementu w grupie $\mathbb{Z}_N$, który można znaleźć w [13] lub [14]. Główna różnica między tymi algorytmami polega na tym, że w algorytmie Shor'a oba rejestry A i R inicjuje się stanem $|0\rangle$, natomiast w naszym rozwiązaniu jeden z nich inicjujemy stanem $|0\rangle$ a drugi stanem $|H\rangle$, co umożliwia pracę z podgrupą H grupy G (stan $|H\rangle$ będący superpozycją wszystkich elementów podgrupy H został zdefiniowany w rozdziale 2.2 - definicja 22).

Na pierwszym rejestrze zastosujemy kwantową transformatę Fouriera z odpowiednio dużą bazą N, następnie za pomocą pewnego mnożenia splątamy stany obu reje-

strów ze sobą i dokonamy pomiaru drugiego rejestru. Na skutek tych operacji pierwszy rejestr przejdzie w stan będący superpozycją elementów pewnej warstwy $g^x H$ (dla $x \in \{0, \ldots, N-1\}$). Dla tego stanu jeszcze raz zastosujemy kwantową transformację Fouriera i dokonamy pomiaru. Uzyskany w ten sposób element $y \in \{0, \ldots, N-1\}$ posłuży nam do wyznaczenia szukanego rzędu elementu g względem podgrupy H.

Algorytm

Wejście :

1. Element $g \in G$, gdzie G jest grupą "black-box" z długością kodowania n.
2. Stan $|H\rangle$, będący superpozycją elementów grupy "black-box" H (która jest podgrupą grupy G) z długością kodowania n.

Wyjście : Rząd elementu g względem podgrupy H, czyli $r = \min\{p : g^p \in H\}$.

Algorytm korzysta z dwóch rejestrów: N-qubitowego A dla pewnego dużego N, które określimy później i 2^n-qubitowego R. Stany zawarte w tych rejestrach oznaczymy odpowiednio przez $|\phi\rangle$ i $|\psi\rangle$. Stan $|\phi\rangle$ inicjujemy przez $|0\rangle$, a $|\phi\rangle$ jak wspomnieliśmy wyżej przez stan $|H\rangle$. Na tak przygotowanych rejestrach wykonujemy następujące operacje:

1. Na rejestrze A stosujemy transformatę Fouriera z bazą N, według wzoru 1.6

$$|\phi, \psi\rangle = |0\rangle|H\rangle \xrightarrow{QFT_N} \Big(\frac{1}{\sqrt{N}} \sum_{a=0}^{N-1} |a\rangle\Big)|H\rangle.$$

2. Zawartość rejestru R mnożymy lewostronnie (odwracalnie) przez g^a, gdzie a to zawartość rejestru A (stosujemy tutaj podnoszenie do potęgi, a później operator U_G - zob. 2.2). Stany w rejestrach zostają w wyniku tej operacji splątane (zob. rozdz.1.5)

$$|\phi, \psi\rangle = \frac{1}{\sqrt{N}} \sum_{a=0}^{N-1} |a\rangle|g^a H\rangle.$$

3. Mierzymy rejestr R. W ten sposób otrzymujemy pewien element warstwy $g^{x_0} H$, gdzie $x_0 \in (0, \ldots, r-1)$. Ponieważ stany w naszych rejestrach są splątane, $|\phi\rangle$ przejdzie w superpozycję wszystkich takich $x \in \{0, \ldots, N-1\}$, że $g^x H = g^{x_0} H$:

$$|\phi\rangle = \frac{1}{\sqrt{|\{x : \ g^x H = g^{x_0} H\}|}} \sum_{x: \ g^x H = g^{x_0} H} |x\rangle.$$

Ponieważ r jest rzędem elementu g względem podgrupy H, liczby $x : \ g^x H = g^{x_0} H$ możemy zapisać w postaci

$$x = x_0 + jr, \quad \text{gdzie } j \in \mathbb{N}.$$

Moc zbioru wszystkich takich elementów jest równa E, które spełnia poniższe nierówności

$$x_0 + (E-1)r < N \qquad \text{i}$$
$$x_0 + Er \geqslant N.$$

Przekształcając warunki dla E otrzymujemy

$$\begin{aligned} x_0 + (E-1)r < N \;\Rightarrow\; & (E-1)r < N - x_0 \\ & E - 1 < N/r - x_0/r \\ & E - 1 < N/r \end{aligned}$$

oraz

$$\begin{aligned} x_0 + Er \geqslant N \;\Rightarrow\; & Er \geqslant N - x_0 \\ & E \geqslant N/r - x_0/r \\ & E \geqslant N/r - 1, \qquad \text{bo } \; x_0 < r. \end{aligned}$$

Uwzględniając powyższe obliczenia stan, który znajduje się w rejestrze A możemy zapisać w postaci

$$|\phi\rangle = \frac{1}{\sqrt{E}} \sum_{j=0}^{E-1} |x_0 + jr\rangle.$$

4. Na rejestrze A stosujemy transformatę Fouriera z bazą N, według wzoru 1.5

$$|\phi\rangle = \frac{1}{\sqrt{E}} \sum_{j=0}^{E-1} |x_0 + jr\rangle \xrightarrow{QFT_N} \frac{1}{\sqrt{N}} \cdot \frac{1}{\sqrt{E}} \sum_{y=0}^{N-1} \sum_{j=0}^{E-1} e^{\frac{2\pi i (x_0 + jr) y}{N}} |y\rangle.$$

5. Naszym kolejnym zadaniem jest wyznaczenie szukanego r z uzyskanego stanu $|\phi\rangle$. W tym celu mierzymy rejestr A i uzyskujemy pewne $y \in \{0, 1, \ldots, N-1\}$.

W dalszej części naszych rozważań pokażemy, w jaki sposób z wyznaczonego przez algorytm y, otrzymać szukany rząd elementu g względem podgrupy H. Udowodnimy najpierw, że uzyskane y z dużym prawdopodobieństwem spełnia warunek

$$-\frac{r}{2} \leqslant yr \; mod \; N \leqslant \frac{r}{2}. \tag{3.1}$$

Zakładamy, że wyniki operacji *mod* N należą do przedziału $(-\frac{N}{2}, \frac{N}{2}]$.

Warunek (3.1) oznacza, że $\exists k < r$, takie że

$$-\frac{r}{2} \leqslant yr - kN \leqslant \frac{r}{2}.$$

Przekształcając powyższą nierówność otrzymujemy

$$\exists k < r\colon \; -\frac{r}{2N} \leqslant \frac{yr}{N} - k \leqslant \frac{r}{2N}$$

$$\Downarrow$$

$$\exists k < r\colon \; -\frac{1}{2N} \leqslant \frac{y}{N} - \frac{k}{r} \leqslant \frac{1}{2N}. \tag{3.2}$$

Wykorzystując ostatnią nierówność będziemy przybliżać ułamek $\frac{y}{N}$ za pomocą metody ułamków łańcuchowych (zob. 2.3) i uzyskamy $\frac{k}{r}$. Okaże się, że aby z dużym

prawdopodobieństwem znaleźć r, szukany rząd elementu g wzgledem podgrupy H, wystarczy powtarzać cały proces kilka razy i wyznaczyć najmniejszą wspólną wielokrotność mianowników uzyskanych w powyższy sposób ułamków.

Wyznaczymy najpierw prawdopodobieństwo, że uzyskane y spełnia 3.1.

Prawdopodobieństwo otrzymania konkretnego y jest równe kwadratowi sumy amplitud wszystkich możliwych dróg osiągnięcia stanu $|y\rangle$

$$P(y) = \Big|\frac{1}{\sqrt{NE}} \sum_{j=0}^{E-1} e^{\frac{2\pi i(x_0+jr)y}{N}}\Big|^2 = \frac{1}{NE}|e^{\frac{2\pi i x_0 y}{N}}|\Big|\sum_{j=0}^{E-1} e^{\frac{2\pi i jry}{N}}\Big|^2 = \frac{1}{NE}\Big|\sum_{j=0}^{E-1} e^{\frac{2\pi i jry}{N}}\Big|^2. \quad (3.3)$$

Suma w powyższym wzorze jest sumą ciągu geometrycznego

$$\Big|\sum_{j=0}^{E-1} e^{\frac{2\pi i jry}{N}}\Big| = \Big|\sum_{j=0}^{E-1} e^{i\theta_y j}\Big| = \frac{|e^{iE\theta_y}-1|}{|e^{i\theta_y}-1|}, \qquad \text{gdzie } \theta_y = \frac{2\pi(ry \bmod N)}{N}. \quad (3.4)$$

Załóżmy teraz, że nasze y spełnia (3.1) i rozważmy odpowiadający mu kąt

$$\theta_y = \frac{2\pi(ry \bmod N)}{N}.$$

Zachodzi dla niego nierówność

$$-\frac{2\pi}{N}\cdot\frac{r}{2} \leqslant \theta_y \leqslant \frac{2\pi}{N}\cdot\frac{r}{2}$$

$$\Downarrow$$

$$-\frac{r\pi}{N} \leqslant \theta_y \leqslant \frac{r\pi}{N}. \quad (3.5)$$

Dodatkowo z warunków na E mamy

$$E-1 < \frac{N}{r} \quad \Rightarrow \quad \frac{r\pi}{N} < \frac{\pi}{E-1}.$$

Razem otrzymujemy, że

$$\theta_y < \frac{\pi}{E-1} \quad \Rightarrow \quad (E-1)\theta_y < \pi.$$

Z twierdzenia 8 wiemy, że dla θ_y spełniającego powyższy warunek, zachodzą nierówności

$$|1-e^{i(E-1)\theta_y}| \geqslant \frac{2(E-1)|\theta_y|}{\pi} \text{ oraz } |1-e^{i\theta_y}| \leqslant |\theta_y|. \quad (3.6)$$

Liczymy prawdopodobieństwo $P(y)$, że nasze y spełnia warunek (3.1). Otrzymamy je wykorzystując prawdopodobieństwo otrzymania dowolnego y (wzór 3.3), które oszacujemy wykorzystując warunki, jakie spełniają y oraz θ_y.

Przekształcimy najpierw wzór (3.4)

$$|\sum_{j=0}^{E-1} e^{i\theta_y j}| = \Big|\frac{e^{iE\theta_y}-1}{e^{i\theta_y}-1}\Big| = \Big|\frac{e^{i(E-1)\theta_y}\cdot e^{i\theta_y}-1-e^{i(E-1)\theta_y}+e^{i(E-1)\theta_y}}{e^{i\theta_y}-1}\Big| =$$

$$= \Big|\frac{e^{i(E-1)\theta_y}-1}{e^{i\theta_y}-1} + e^{i(E-1)\theta_y}\Big| \geqslant \Big|\frac{e^{i(E-1)\theta_y}-1}{e^{i\theta_y}-1}\Big| - 1.$$

Stosując nierówności (3.6) uzyskujemy

$$|\sum_{j=0}^{E-1} e^{i\theta_y j}| \geqslant \frac{2(E-1)|\theta_y|}{\pi|\theta_y|} - 1 = \frac{2E}{\pi} - (1+\frac{2}{\pi}).$$

Podstawiając uzyskaną zależność do wzoru na prawdopodobieństwo (3.3) otrzymujemy

$$P(y) \geqslant \frac{1}{NE}\Big|\frac{2E}{\pi} - (1+\frac{2}{\pi})\Big|^2.$$

Wykorzystując fakt, że $E > \frac{n}{r} - 1$ i dokonując prostych przekształceń uzyskujemy

$$P(y) \geqslant \frac{1}{NE}\cdot\frac{4E^2}{\pi^2} - \frac{4E}{NE\pi}(1+\frac{2}{\pi}) + \frac{(1+\frac{2}{\pi})^2}{NE} \geqslant \frac{4(\frac{N}{r}-1)}{N\pi^2} - \frac{4(1+\frac{2}{\pi})}{N\pi} + \frac{(1+\frac{2}{\pi})^2}{NE} =$$

$$= \frac{4\frac{N}{r}}{N\pi^2} - \frac{4}{N\pi^2} - \frac{4(1+\frac{2}{\pi})}{N\pi} + \frac{(1+\frac{2}{\pi})^2}{NE} = \frac{4}{r\pi^2} - \frac{4}{N\pi^2} - \frac{4(1+\frac{2}{\pi})}{N\pi} + \frac{(1+\frac{2}{\pi})^2}{NE}.$$

Liczbę N dobierzemy tak, aby była dostatecznie duża, dlatego ułamki o mianowniku zawierającym N będą bardzo małe i możemy je pominąć.

Łącznie prawdopodobieństwo $P(y)$, że po wykonaniu pięciu, opisanych wcześniej kroków algorytmu, y uzyskany poprzez zmierzenie rejestru A spełnia nierówność (3.1), jest ograniczone od dołu przez

$$P(y) \geqslant \frac{4}{r\pi^2}.$$

Z twierdzenia 9 wiemy, że jest dokładnie r takich $y \in \{0,1,\ldots,N-1\}$, dla których $-\frac{r}{2} \leqslant yr \ mod \ N \leqslant \frac{r}{2}$, czyli prawdopodobieństwo P_w, że w wyniku pięciu kroków algorytmu uzyskamy jakiekolwiek y spełniające rozważany warunek, jest większe niż $\frac{4}{\pi^2}$

$$P_w = rP(y) \geqslant \frac{4}{\pi^2}.$$

Mając y i N, o których zakładamy, że spełniają (3.1), czyli też (3.2) wskażemy teraz jak wyliczyć r będące szukanym rzędem elementu g względem podgrupy H. Zakładamy dodatkowo istnienie takiego $M \in \mathbb{N}$, że $r < M < M^2 < N$ (dobierzemy N na tyle duże, by tak było).

Pokażemy najpierw, że istnieje tylko jeden taki ułamek $\frac{k}{r}$, gdzie $k \in \{0,\ldots,r-1\}$ i $r < M$, że $-\frac{1}{2N} \leqslant \frac{y}{N} - \frac{k}{r} \leqslant \frac{1}{2N}$.

Hipoteza: Przypuśćmy, że istnieją dwa ułamki $\frac{k_1}{r_1}$ i $\frac{k_2}{r_2}$ spełniające

$$-\frac{1}{2N} \leqslant \frac{y}{N} - \frac{k_1}{r_1} \leqslant \frac{1}{2N} \tag{1}$$

$$-\frac{1}{2N} \leqslant \frac{y}{N} - \frac{k_2}{r_2} \leqslant \frac{1}{2N} \tag{2}$$

i niech $\frac{k_1}{r_1} > \frac{k_2}{r_2}$.

Dwa ułamki $\frac{a}{b}$ i $\frac{c}{d}$ o mianownikach mniejszych od pewnego $K \in \mathbb{N}$ nie mogą być bliżej siebie niż $1/K^2$, gdyż

$$\left|\frac{a}{b} - \frac{c}{d}\right| = \left|\frac{ad-bc}{bd}\right| > \left|\frac{ad-bc}{K^2}\right| \geqslant \frac{1}{K^2}.$$

Nasze ułamki mają mianowniki mniejsze od M, skąd otrzymujemy, że

$$\frac{k_1}{r_1} > \frac{k_2}{r_2} + \frac{1}{M^2}. \qquad (*)$$

Dodatkowo z założenia mamy

$$M^2 < N \Rightarrow -M^2 > -N \Rightarrow -\frac{1}{M^2} < -\frac{1}{N}. \qquad (**)$$

Przy użyciu powyższych zależności oszacujemy teraz wielkość $\frac{y}{N} - \frac{k_1}{r_1}$ w następujący sposób:

$$\frac{y}{N} - \frac{k_1}{r_1} <^{(*)} \frac{y}{N} - \frac{k_2}{r_2} - \frac{1}{M^2} <^{(**)} \frac{y}{N} - \frac{k_2}{r_2} - \frac{1}{N} \leqslant^{(2)} \frac{1}{2N} - \frac{1}{N} = -\frac{1}{2N}.$$

Równocześnie z (1) mamy, że $\frac{y}{N} - \frac{k_1}{r_1} \geqslant -\frac{1}{2N}$, co daje nam sprzeczność.

Unikalny ułamek $\frac{k}{r}$ spełniający (3.2) uzyskamy poprzez przybliżenie ułamka $\frac{y}{N}$ najbliższym ułamkiem o mianowniku mniejszym niż M. Może on być znaleziony w czasie wielomianowym przy użyciu metody ułamków łańcuchowych (zob. 2.3). Za jej pomocą znajdziemy pewne względnie pierwsze liczby u_1 i v_1, takie że $\frac{u_1}{v_1} = \frac{k}{r}$. Jeśli liczby k i r są względnie pierwsze, to wyznaczone v_1 jest dokładnie równe szukanemu r - rzędowi elementu g względem podgrupy H. W przeciwnym razie v_1 jest jedynie jednym z dzielników r.

Aby uzyskać r z dużym prawdopodobieństwem, będziemy całą procedurę wykonywać kilka razy. W pierwszym przebiegu (z prawdopodobieństwem co najmniej $\frac{4}{\pi^2}$) uzyskamy $y_1 \in \{1, \ldots, N-1\}$, dla którego istnieje $k_1 \in \{0, \ldots, r-1\}$, takie że

$$-\frac{1}{2N} \leqslant \frac{y_1}{N} - \frac{k_1}{r} \leqslant \frac{1}{2N}$$

i wyznaczymy nieskracalny ułamek $\frac{u_1}{v_1}$ o własności

$$\frac{u_1}{v_1} = \frac{k_1}{r}.$$

Po drugim przebiegu (z prawdopodobieństwem co najmniej $\frac{4}{\pi^2}$) będziemy mieć pewne $y_2 \in \{1, \ldots, N-1\}$, odpowiadający mu $k_2 \in \{0, \ldots, r-1\}$ i nieskracalny ulamek $\frac{u_2}{v_2}$, taki że

$$\frac{u_2}{v_2} = \frac{k_2}{r}.$$

Z twierdzenia 12 wiemy, że dla powyższych ułamków $\frac{u_1}{v_1}$ i $\frac{u_2}{v_2}$, jeśli $nwd(k_1, k_2) = 1$, to $r = nww(v_1, v_2)$. Spróbujemy teraz oszacować prawdopodobieństwo, że losowo wybrane k_1 i k_2 będą względnie pierwsze. Zauważmy najpierw, że liczba pierwsza p dzieli k_1 z prawdopodobieństwem równym $1/p$. Stąd mamy, że prawdopodobieństwo, iż liczba pierwsza p dzieli równocześnie k_1 i k_2 jest równe $1/p^2$. Liczby k_1 i k_2 są względnie pierwsze, jeśli nie ma takiej liczby pierwszej, która dzieliłaby jednocześnie k_1 i k_2, czyli

$$P(k_1, k_2 \text{ - względnie pierwsze}) = \prod_{p-l.pierwsza} (1 - \frac{1}{p^2}).$$

Korzystamy teraz z tego, że funkcja dzeta Riemanna $\zeta(s) = \prod_{p-l.pierwsza} (1 - \frac{1}{p^s})^{-1}$ i otrzymujemy

$$P(k_1, k_2 \text{ - względnie pierwsze}) = \frac{1}{\zeta(2)} = \frac{6}{\pi^2} \simeq 0.607.$$

Jeżeli uzyskane w powyższy sposób $r = nww(v_1, v_2)$, jest rzeczywiście rzędem elementu g względem podgrupy H, wówczas $|H\rangle = |g^r H\rangle$. Jeśli nie, to stany te bedą ortogonalne. Metoda rozróżniania stanów ortogonalnych została przedstawiona np. w [3]. Możemy więc sprawdzić, czy uzyskane r jest szukaną wartością.

Otrzymane prawdopodobieństwa P_w i $P(k_1, k_2$ - względnie pierwsze) nie zależą ani od n, ani od m, czyli nie zależą od wielkości danych wejściowych. Powtarzając więc przedstawiony algorytm dostateczną ilość razy (niezależną od rozmiaru danych wejściowych) i obliczając najmniejszą wspólną wielokrotność mianowników otrzymanych ułamków, z dużym prawdopodobieństwem uzyskamy szukane r - rząd elementu g względem podgrupy H. John Watrous w [15] podaje, że jeśli chcemy uzyskać wynik z prawdopodobieństwem $1 - \varepsilon$, wymagana liczba powtórzeń procedury jest rzędu $\Theta(\log \frac{1}{\varepsilon})$.

Pozostało nam dobrać liczbę N. Musi ona być taka, aby istniała liczba M spełniająca warunek $r < M < M^2 < N$. Rząd danego elementu g, należącego do grupy G, względem podgrupy tej grupy z pewnością nie jest większy od rzędu całej grupy G. Jest ona grupą "black-box" o długości kodowania n. Oznacza to, że jej rząd nie może przekraczać liczby 2^n (zobacz rozdział (2.2)), czyli $r < 2^n$. Wystarczy więc przyjąć $N = 2^{2n+1}$.

Przedstawiony powyżej algorytm opiera się głównie na wykorzystaniu kwantowej transformaty Fouriera, której złożoność wynosi $O((\log N)^2)$ Dla ustalonego przez nas N złożoność całej procedury jest równa $O(n^2)\cdot$ ilość powtórzeń.

3.2 Konstrukcja superpozycji elementów grupy

W tym rozdziale spróbujemy opisać, jak kilka kopii stanu $|H\rangle$ może być przekształconych w kilka kopii stanu $|\langle g\rangle H\rangle$. Zakładamy dodatkowo, że g normalizuje H (tzn. $gH = Hg$, co implikuje że $\langle g\rangle H$ jest grupą i $H \lhd \langle g\rangle H$). W odniesieniu do problemu obliczania rzędu grupy rozwiązalnej procedura przedstawiona poniżej posłuży nam do przekształcenia kopii stanu $|H_{j-1}\rangle$ na kopie stanu $|H_j\rangle$. Algorytm wykorzystuje dwukrotnie kwantową transformatę Fouriera i operację lewostronnego mnożenia rejestru

przez pewien element grupy.

Algorytm

Wejście :

1. Element $g \in G$, normalizujący podgrupę H grupy G.
2. p kopii stanu $|H\rangle$, gdzie p jest odpowiednio duże.
3. Wartość $r = |\langle g \rangle H/H|$.

Wyjście : $(p-1)$ kopii stanu $|\langle g \rangle H\rangle$.

Algorytm korzysta z rejestrów $R_1, R_2, \ldots, R_p$ oraz $A_1, A_2, \ldots, A_p$. Te pierwsze zawierają kopie stanu $|H\rangle$, a te drugie są rejestrami, których stany bazowe odpowiadają $\mathbb{Z}_r$ i są zainicjowane stanem $|0\rangle$. Stany będące wynikiem operacji wykonanych na rejestrach A_i, R_i oznaczać będziemy odpowiednio przez ϕ_i i ψ_i.

Dla każdego $i = 1, \ldots, p$ wykonujemy następujące kroki:

1. Na rejestrze A_i stosujemy transformatę Fouriera z bazą r (zobacz wzór 1.6)
$$|\phi_i, \psi_i\rangle = |0\rangle|H\rangle \xrightarrow{QFT_r} \frac{1}{\sqrt{r}} \sum_{a_i=0}^{r-1} |a_i\rangle|H\rangle.$$
2. Zawartość rejestru R_i mnożymy lewostronnie przez g_i^a, gdzie a_i to zawartość rejestru A_i (stosujemy tutaj podnoszenie do potęgi, a później operator U_G - zob. rozdz. 2.2). Stany w rejestrach zostają w wyniku tej operacji splątane (zob. rozdz.1.5)
$$|\phi_i, \psi_i\rangle = \frac{1}{\sqrt{r}} \sum_{a_i=0}^{r-1} |a_i\rangle|g^{a_i}H\rangle.$$
3. Na rejestrze A_i ponownie stosujemy transformatę Fouriera z bazą r (zobacz wzór 1.5)
$$|\phi_i, \psi_i\rangle = \frac{1}{\sqrt{r}} \sum_{a_i=0}^{r-1} |a_i\rangle|g^{a_i}H\rangle \xrightarrow{QFT_r} \frac{1}{r} \sum_{a_i=0}^{r-1} \sum_{b_i=0}^{r-1} e^{\frac{2\pi i a_i b_i}{r}} |b_i\rangle|g^{a_i}H\rangle.$$
4. Mierzymy rejestr A_i.

   Po wykonaniu powyższych czterech kroków otrzymujemy pewne $b_1, \ldots, b_p \in \mathbb{Z}_r$ a w rejestrach $R_1, \ldots, R_p$ stany postaci $\psi_i = \frac{1}{\sqrt{r}} \sum_{a_i=0}^{r-1} e^{\frac{2\pi i a_i b_i}{r}} |g^{a_i}H\rangle$.

   Do dalszych obliczeń będziemy potrzebować, aby przynajmniej jedna z wartości b_i była względnie pierwsza z liczbą r. Z twierdzenia 11 wynika, że dla pewnego $p \in \Theta\Big((\log\log r) \cdot (\log(\frac{1}{\varepsilon})\Big)$ tak nie będzie w najgorszym przypadku z prawdopodobieństwem ε. Liczba r jest rzędem grupy ilorazowej $|\langle g \rangle H/H|$ i jest ograniczona przez wielkość 2^n. Czyli dla powyższego p, wśród $b_1, \ldots, b_p$ będzie liczba względnie pierwsza z r z prawdopodobieństwem co najmniej $1 - \varepsilon$.

Niech k bedzie takie, że b_k jest względnie pierwsze z r. Stan $|\psi_k\rangle$ wykorzystamy, aby przekształcić stany z rejestrów R_i (z wyjątkiem $i \neq k$) do poszukiwanego stanu $|\langle g\rangle H\rangle$. Będziemy do tego potrzebować operatora M_q, który zastosowany do rejestru mnoży lewostronnie jego zawartość przez q. Rozważmy operator $M_{g^j h^m}$, dla $h \in H, j, m \in \mathbb{Z}$. Pokażemy, że stan $|\psi_k\rangle$ jest jego wektorem własnym:

$$M_{g^j h^m}|\psi_k\rangle = \frac{1}{\sqrt{r}} \sum_{a_k=0}^{r-1} e^{\frac{2\pi i a_k b_k}{r}} M_{g^j h^m}|g^{a_k} H\rangle = \frac{1}{\sqrt{r}} \sum_{a_k=0}^{r-1} e^{\frac{2\pi i a_k b_k}{r}} |g^j h^m g^{a_k} H\rangle.$$

Skorzystamy teraz z faktu, że g jest elementem normalizującym grupę H

$$\begin{aligned} M_{g^j h^m}|\psi_k\rangle &= \frac{1}{\sqrt{r}} \sum_{a_k=0}^{r-1} e^{\frac{2\pi i a_k b_k}{r}} |g^{j+a_k} H\rangle = \frac{1}{\sqrt{r}} \sum_{a_k=0}^{r-1} e^{\frac{2\pi i (a_k - j) b_k}{r}} |g^{a_k} H\rangle = \\ &= \frac{1}{\sqrt{r}} \sum_{a_k=0}^{r-1} e^{\frac{2\pi i a_k b_k}{r}} e^{\frac{-2\pi i j b_k}{r}} |g^{a_k} H\rangle = e^{\frac{-2\pi i j b_k}{r}} \frac{1}{\sqrt{r}} \sum_{a_k=0}^{r-1} e^{\frac{2\pi i a_k b_k}{r}} |g^{a_k} H\rangle = \\ &= e^{\frac{-2\pi i j b_k}{r}} |\psi_k\rangle, \end{aligned}$$

co oznacza, że stan $|\psi_k\rangle$ jest wektorem własnym operatora $M_{g^j h^m}$ odpowiadającym wartości własnej $e^{\frac{-2\pi i j b_k}{r}}$.

5. Wykonujemy teraz mnożenie stanu ψ_k przez f^c, gdzie f oznacza zawartość rejestru R_i, a c jest liczbą całkowitą, taką że $c = b_i b_k^{-1}$. Istnienie b_k^{-1} mamy stąd, że b_k oraz r są względnie pierwsze. Otrzymane w ten sposób stany będą tymi, których szukamy. (Przekształcenia tego nie realizujemy dla $i = k$.)

$$\begin{aligned} |\psi_i, \psi_k\rangle &= \frac{1}{\sqrt{r}} \sum_{a_i=0}^{r-1} e^{\frac{2\pi i a_i b_i}{r}} |g^{a_i} H\rangle|\psi_k\rangle = \frac{1}{\sqrt{r|H|}} \sum_{a_i=0}^{r-1} \sum_{h\in H} e^{\frac{2\pi i a_i b_i}{r}} |g^{a_i} h\rangle|\psi_k\rangle \xrightarrow{M_{f^c}} \\ &\to \frac{1}{\sqrt{r|H|}} \sum_{a_i=0}^{r-1} \sum_{h\in H} e^{\frac{2\pi i a_i b_i}{r}} |g^{a_i} h\rangle M_{g^{a_i c} h^c}|\psi_k\rangle = \\ &= \frac{1}{\sqrt{r|H|}} \sum_{a_i=0}^{r-1} \sum_{h\in H} e^{\frac{2\pi i a_i b_i}{r}} e^{\frac{-2\pi i a_i c b_k}{r}} |g^{a_i} h\rangle|\psi_k\rangle = \\ &= \frac{1}{\sqrt{r|H|}} \sum_{a_i=0}^{r-1} \sum_{h\in H} e^{\frac{2\pi i a_i b_i - 2\pi i a_i b_i b_k^{-1} b_k}{r}} |g^{a_i} h\rangle|\psi_k\rangle = \\ &= \frac{1}{\sqrt{r|H|}} \sum_{a_i=0}^{r-1} \sum_{h\in H} |g^{a_i} h\rangle|\psi_k\rangle = |\langle g\rangle H\rangle|\psi_k\rangle. \end{aligned}$$

Po wykonaniu powyższej procedury w każdym z rejestrów $R_1, R_2, \ldots, R_p$, z wyjątkiem rejestru R_k, znajduje się stan będący superpozycją elementów grupy $\langle g\rangle H$ (z prawdopodobieństwem $1 - \varepsilon$, dla $p \in \Theta\Big((\log n) \cdot (\log(\frac{1}{\varepsilon}))\Big)$). Kopii stanu $|\langle g\rangle H\rangle$ jest więc dokładnie $p - 1$. Podczas algorytmu wykorzystujemy kwantową transformatę Fouriera z bazą r, które jest ograniczone przez rząd grupy, czyli $|G| < 2^n$. Pesymistyczna złożoność obliczeniowa kroków algorytmu wykorzystujących transformatę wynosi więc $O(n^2)$. Łącznie złożoność procedury wyznaczającej $p - 1$ kopii $|H_j\rangle$ to $p \cdot O(n^2)$. Warto podkreślić, że w tym algorytmie istotne było założenie, że $H_{j-1} \lhd H_j$.

3.3 Algorytm główny

W tym podrozdziale przedstawimy główny algorytm obliczający rząd grupy rozwiązalnej. Wykorzystuje on procedury omówione w poprzedniej części tego rozdziału: wyliczanie rzędu elementu grupy względem podgrupy oraz konstrukcję superpozycji elementów grupy.

Niech $g_1, \ldots, g_m \in G$ będą takie, że dla $H_j = \langle g_1, \ldots, g_j \rangle$ zachodzi

$$\{1\} = H_0 \triangleleft H_1 \triangleleft \ldots \triangleleft H_m = G. \tag{3.7}$$

Przez r_j oznaczać będziemy rząd elementu g_j względem podgrupy H_{j-1}. Przypominamy, że rząd grupy można wyrazić jako iloczyn rzędów elementów g_j względem podgrup H_{j-1}. $|G| = \prod_{j=1}^{m} r_j$. Niech k będzie parametrem, który ustalimy później. Będzie on zależał od tego, jaką dokładności obliczeń chcemy uzyskać.

Algorytm

Wejście: Elementy $g_1, \ldots, g_m \in G$, takie że dla $H_j = \langle g_1, \ldots, g_j \rangle$ zachodzi (3.7).
Wyjście: Rząd grupy $|G|$.

1. Przygotowujemy $k(m+1)$ kopii stanu $|H_0\rangle = \{1\}$.
2. Dla każdego $j = 1, \ldots, m$:
   (i) używając $(k-1)$ kopii $|H_{j-1}\rangle$ liczymy $r_j = r_{H_{j-1}}(g_j)$,
   (ii) pozostałe kopie $|H_{j-1}\rangle$ przekształcamy na kopie $|H_j\rangle$ tracąc przy tym jeden stan $|H_{j-1}\rangle$.
3. Zwracamy $|G| = \prod_{j=1}^{m} r_j$.

Spróbujemy teraz oszacować złożoność czasową całego algorytmu. Krok (i) będzie wykonany w czasie $k \cdot O(n^2)$, natomiast (ii) w czasie $km \cdot O(n^2)$. Kroki te są powtarzane m razy. Złożoność czasowa całego algorytmu jest więc równa $k \cdot O(n^2 m^2)$.

Procedura zwróci poprawny wynik, jeśli każdy z kroków zostanie wykonany bezbłędnie. By mieć algorytm działający z wysokim prawdopodobieństwem musimy tak dobrać parametry, by błąd każdego z kroków był stosunkowo mały, Załóżmy, że chcemy obliczyć $|G|$ z prawdopodobieństwem błędu co najwyżej ε. Wówczas obliczenie każdego r_j, jak i konwersja kopii $|H_{j-1}\rangle$ na kopię $|H_j\rangle$ muszą być takie, by prawdopodobieństwo błędu wynosiło co najwyżej $\varepsilon/2m$.

Procedura wyznaczająca rząd elementu względem podgrupy, aby dała poprawny wynik z prawdopodobieństwem $1 - \varepsilon/2m$ musi być wykonana co najmniej $\Theta(\log 2m/\varepsilon)$ razy. Krok (i) w każdym obiegu używa $(k-1)$ kopii stanu $|H_j\rangle$, a więc k wybierzemy tak, by było rzędu $\Theta(\log m/\varepsilon)$.

Procedura konwertująca kopie stanu $|H_{j-1}\rangle$ na kopie stanu $|H_j\rangle$ da poprawne rezultaty z prawdopodobieństwem $1 - \varepsilon/2m$, jeśli ilość przekształcanych stanów będzie rzędu $\Theta(\log n \cdot \log 2m/\varepsilon)$.

Aby obliczyć rząd grupy G z prawdopodobieństwem $1 - \varepsilon$ wystarczy abyśmy ustalili $k = \Theta(\log n \cdot \log m/\varepsilon)$. Widzimy więc, że nasz algorytm wyznaczy rząd grupy G z prawdopodobieństwem błędu ograniczonym przez ε w czasie wielomianowym od $mn + \log(1/\varepsilon)$.

Rozdział 4

Wnioski końcowe

4.1 Problemy redukujące się do obliczania rzędu grupy

Przedstawiony w poprzednim rozdziale algorytm oblicza rząd grupy rozwiązalnej, działając kwantowo w czasie wielomianowym. Dodatkowo istnieje algorytm, za pomocą którego w czasie wielomianowym możemy sprawdzić czy dana grupa jest rozwiązalna ([4]). Przy pomocy tych dwóch narzędzi możemy wielomianowo rozwiązać kilka ważnych problemów, redukujących się do obliczenia rzędu grupy.

Pierwszym takim problemem jest testowanie przynależności. Przypuśćmy, że mamy dane elementy $g_1, \ldots, g_k$ i h należące do pewnej grupy "black-box". Pytamy czy $h \in \langle g_1, \ldots, g_k \rangle$? Jest tak wtedy i tylko wtedy, gdy $|\langle g_1, \ldots, g_k \rangle| = |\langle g_1, \ldots, g_k, h \rangle|$. Jeśli teraz $\langle g_1, \ldots, g_k, h \rangle$ jest grupą rozwiązalną, to odpowiedź na wyżej postawione pytanie, czy $h \in \langle g_1, \ldots, g_k \rangle$, uzyskujemy kwantowo w czasie wielomianowym w następujący sposób:

1. Sprawdzamy czy $\langle g_1, \ldots, g_k \rangle$ jest grupą rozwiązalną.
2. Jeśli nie, to $h \notin \langle g_1, \ldots, g_k \rangle$.
3. Jeśli tak, to:
   - obliczamy rząd $r = |\langle g_1, \ldots, g_k, h \rangle|$,
   - obliczamy rząd $p = |\langle g_1, \ldots, g_k \rangle|$,
   - jeśli $p = r$, to $h \in \langle g_1, \ldots, g_k \rangle$.

Kolejnym problemem redukującym się do obliczania rzędu grupy jest badanie czy dana grupa rozwiązalna jest podgrupą innej. Tzn. mamy dane dwie rozwiązalne grupy "black-box" (czyli ich generatory: $g_1, \ldots, g_k$ i $h_1, \ldots, h_l$) i chcemy odpowiedzieć na pytanie czy $\langle g_1, \ldots, g_k \rangle \leqslant \langle h_1, \ldots, h_l \rangle$? Rozwiązaniem jest wielomianowo wykonywany test $|\langle g_1, \ldots, g_k, h_1, \ldots, h_l \rangle| = |\langle g_1, \ldots, g_k \rangle|$.

Podobnym problemem jest testowanie równości dwóch grup rozwiązalnych o generatorach $g_1, \ldots, g_k$ i $h_1, \ldots, h_l$, co realizujemy sprawdzając czy $\langle g_1, \ldots, g_k \rangle \leqslant \langle h_1, \ldots, h_l \rangle$ i czy $\langle h_1, \ldots, h_l \rangle \leqslant \langle g_1, \ldots, g_k \rangle$.

Wielomianowo można też przetestować, czy dana podgrupa grupy rozwiązalnej jest jej podgrupą normalną. Mamy wówczas dane generatory grupy: $g_1, \ldots, g_k$ i podgrupy: $h_1, \ldots, h_l$. Aby zbadać normalność wystarczy dla każdego i oraz j sprawdzić czy $g_i^{-1} h_j g_i \in \langle h_1, \ldots, h_l \rangle$.

4.2 Problemy otwarte

W niniejszej pracy przedstawiliśmy wielomianowy, kwantowy algorytm obliczający rząd rozwiązalnej grupy "black-box" i wyznaczający superpozycję jej elementów. Pokazaliśmy też, jak w prosty sposób można wykorzystać powyższy algorytm do rozwiązywania w czasie wielomianowym innych problemów, związanych z grupami rozwiązalnymi.

Jest kilka innych pytań dotyczących rozwiązalnych grup "black-box", na które nie potrafimy dać odpowiedzi w czasie wielomianowym. Przykładami mogą być:

- Część wspólna grup - Mamy dane zbiory generatorów dla dwóch podgrup rozwiązalnej grupy "black-box". Pytamy czy dane podgrupy mają nietrywialne przecięcie?

- Część wspólna warstw - Mamy dane zbiory generatorów dla dwóch warstw rozwiązalnej grupy "black-box". Pytamy czy dane warstwy mają nietrywialne przecięcie?

Innym interesującym pytaniem, na które nie znamy jeszcze odpowiedzi jest czy istnieją wielomianowe, kwantowe algorytmy dla podobnych problemów zdefiniowanych dla dowolnych (nie rozwiązalnych) skończonych grup.

Bibliografia

[1] A. Barenco, C. Bennett, R. Cleve, D. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, H. Weinfurter, *Elementary gates for quantum computation*, manuscript, 1995.

[2] L. Babai, R. Beals, *A polynomial-time theory of black box groups*, Groups St.Andrews 1997 in Bath, vol.260 of London Math. Soc. Lecture Note Ser., Cambridge University Press, 1999

[3] H. Buhrman, R. Cleve, J. Watrous, R. de Wolf, *Quantum fingerprinting*, Physical Review Letters, vol. 87, nr 16, 2001

[4] L. Babai, G. Copperman, L. Finkelstein, E. Luks, A. Seress, *Fast Monte Carlo algorithms for permutation groups*, Journal of Computer and System Sciences, 50, 1995

[5] L. Babai, E. Szemerédi, *On the complexity of matrix group problems*, Proceeding of the 25th Annual Symposium on Foundations of Computer Science, 1984

[6] D. Coppersmith, *An approximate Fourier transform useful in quantum factoring*, IBM Research Report RC 19642, 1994

[7] J. Gruska, *Quantum Computing*, McGraw Hill, Cambridge University Press, 1999

[8] G.H. Hardy, E.M. Wright, *An introduction to the theory of numbers*, Oxford, 1960

[9] A.I. Kostrykin, *Wstęp do algebry*, PWN Warszawa, 1984

[10] E. Luks, *Computing in solvable matrix group*, Proceedings of the 33rd Symposium on the Foundations of Computer Science, 1992

[11] W. Mlak, *Wstęp do teorii przestrzeni Hilberta*, PWN Warszawa, 1982

[12] M. Mosca, *Quantum Computer Algorithms*, PhD thesis, University of Oxford, 1999

[13] P.W. Shor, *Polynomial time algorithm for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Computing 26, 1997

[14] J. Preskil, *Lecture notes for Physics 229: Quantum Information and Computation*, Cambridge University Press, 1998

[15] J. Watrous, *Quantum algorithms for solvable groups*, Proceedings of 33rd ACM Symposium on Theory of Computing, 2001

Spis treści