Skrypt do Informacji Kwantowej

Tomasz Rybotycki
Rafał Demkowicz-Dobrzański
Janek Kołodyński
Konrad Banaszek

VI 2020

Spis treści

1 Powtórzenie mechaniki kwantowej

W mechanice kwantowej mamy pojęcie stanu (**stan czysty**). Jest to wektor w przestrzeni Hilberta, który można wyrazić w następujący sposób:

$$|\psi\rangle = \sum_{i=1}^{d} c_i \langle i|i\rangle \tag{1.1}$$

$$\sum_i |c_i|^2 = 1 \tag{1.2}$$

$$|\psi\rangle \equiv e^{i\xi} |\phi\rangle , \tag{1.3}$$

przy czym zakładamy, że przestrzeń jest d wymiarowa. Potrzebujemy $2d-2$ parametrów rzeczywistych, żeby opisać stan kwantowy w przestrzeni d wymiarowej.

W mechanice kwantowej ważne są też pomiary. Omówimy najpierw tak zwane **pomiary rzutowe** (lub **pomiary von Neumanna**). Pomiar rzutowy polega na tym, że podajemy zestaw operatorów rzutowych P_i, czyli takich, że

$$P_i^2 = P_i \tag{1.4}$$

$$\sum_i P_i = \mathbb{1} \tag{1.5}$$

Indeksy i to różne wyniki pomiaru. Prawdopodobieństwo p_i indeksu i liczy się tak, że

$$p_i = \langle\psi| P_i |\psi\rangle . \tag{1.6}$$

W wyniku pomiaru stan układu zmienia się w następujący sposób

$$|\psi\rangle \xrightarrow{i} \frac{P_i |\psi\rangle}{\sqrt{p_i}} := |\psi_i\rangle , \tag{1.7}$$

czyli $|\psi_i\rangle$ jest stanem otrzymanym po działaniu operatorem rzutowym.

Często jest tak, że $P_i = |i\rangle \langle i|$ i o takich operatorach mówi się, że są to **operatory rzutowe pierwszego rzędu** (ang. *rank 1*). Tutaj $|i\rangle$ jest pewną bazą w której dokonujemy pomiaru. Wtedy

$$p_i = \langle\psi|i\rangle \langle i|\psi\rangle = |\langle\psi|i\rangle|^2. \tag{1.8}$$

Na wykładach z mechaniki kwantowej pomiar często utożsamia się z **obserwablą**. Powyższy wywód jest bardziej fundamentalny. Mechanika kwantowa tym się różni od klasycznej tym, że pewna wielkość fizyczna, taka jak pęd, narzuca pewną bazę wektorów, w których ma ona dobrze określone wartości. Inna wielkość będzie miała inną bazę.

Obserwablę można określić jako matematyczną reprezetację pewnej wielkości fizycznej. Jeśli mamy bazę stanów o dobrze określonej wartości wielkości fizycznej $\mathcal{A}$. Definiujemy obserwablę A związaną z wielkością fizyczną $\mathcal{A}$ tak, że:

$$A = \sum a_i |i\rangle \langle i| . \tag{1.9}$$

Po co taka definicja? Jeszcze nie wiemy. Ale wrócimy do tego. Operator taki jest przydatny, bo możemy odtworzyć $a_i |i\rangle \langle i|$. Wynika to z tego, że jesteśmy w stanie wyznaczyć rozkład własny. A jest operatorem hermitowskim, a a_i to jego wartości własne. Jednym operatorem możemy na przykład określić hamiltonian.

Mając stan $|\psi\rangle$ dokonuję pomiaru wielkości $\mathcal{A}$ i pytam jaka jest średnia wartość tej wielkości?

$$\langle A\rangle = \sum_i p_i a_i = \sum_i |\langle\psi|i\rangle|^2 a_i = \sum_i \langle\psi|i\rangle \langle i|\psi\rangle a_i = \langle\psi| \sum_i a_i |i\rangle \langle i|\psi\rangle = \langle\psi| A |\psi\rangle \tag{1.10}$$

Widać, że nie trzeba rozkładu, aby wyliczyć wartość średnią.

Przejdziemy teraz do **układów mieszanych**. Rozważmy przykładowy układ mieszany. Nie mamy pełnej wiedzy o tym układzie. Załóżmy, że z prawdopodobieństwem q_k dostajemy stan $|\psi_k\rangle$, co opisuje naszą niepewną wiedzę.

Jeśli wykonujemy pomiar $\{P_i\}$, to wyniki i uzyskam z prawdopodobieństwem

$$p_i = \sum_k q_k \langle\psi_k| P_i |\psi_k\rangle , \tag{1.11}$$

jest to tak zwane **prawdopodobieństwo całkowite**. Licząc dalej otrzymamy

$$p_i = \sum_k q_k Tr(P_i |\psi_k\rangle \langle\psi_k|). \tag{1.12}$$

Możemy tak robić, bo

$$\langle\psi| A |\psi\rangle = Tr(A |\psi\rangle \langle\psi|). \tag{1.13}$$

A jak to udowodnić? Można to rozpisać w *dowolnej* bazie

$$|\psi\rangle = \sum_i \psi_i |i\rangle \tag{1.14}$$

$$\sum \psi_i^* A_j^i \psi_j = \sum A_j^i \psi^j \psi_i^* \tag{1.15}$$

Kontynuując możemy

$$p_i = Tr\left(P_i \sum_k q_k |\psi_k\rangle \langle\psi_k|\right) \tag{1.16}$$

$$\rho = \sum_k q_k |\psi_k\rangle \langle\psi_k| . \tag{1.17}$$

Wielkość ρ nazywamy **macierzą gęstości** i wykorzystujemy ją do określania stanów mieszanych. Matematycznie jest ona operatorem, którego ślad jest równy 1.

$$Tr(\rho) = 1, \tag{1.18}$$

a w dodatku $q_k \geqslant 0$, a to oznacza, że

$$\rho \geqslant 0, \tag{1.19}$$

czyli jest **nieujemnie** (w skrócie będzie mówione **dodatnio**) **określony**. To z kolei oznacza, że

$$\forall_{|\psi\rangle} \ \langle\psi| \rho |\psi\rangle \geqslant 0. \tag{1.20}$$

Dodatniość operatora jest mocniejsza niż jego hermitowskość, bo poniekąd ją wymusza i nadaje dodatkowo więcej warunków.

Macierz gęstości też można zapisać w bazie w następujący sposób

$$\rho = \sum_{ij} \rho_j^i |i\rangle \langle j| . \tag{1.21}$$

W szczególności może być tak, że

$$\rho = |\psi\rangle \langle\psi| \tag{1.22}$$

i wtedy ρ reprezentuje stan czysty.

Gdy mamy **wiele układów fizycznych** to w mechanice kwantowej do ich opisu wykorzystuje się **iloczyn tensorowy przestrzeni Hilberta**

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes ... \otimes \mathcal{H}_N. \tag{1.23}$$

Ogólny stan mogę zapisać tak:

$$\psi = \sum_{i_1,...,i_N} c_{i_1,...,i_N} |i_1\rangle \otimes ... \otimes |i_N\rangle \tag{1.24}$$

Zachodzi też

$$dim(\mathcal{H}_i) = d \implies dim(\mathcal{H}) = d^N. \tag{1.25}$$

Możemy zapisać też macierz gęstości dla dla takich stanów

$$\rho = \sum_{i_1,...,i_N,j_1,...,j_N} \rho^{i_1,...,i_N}_{j_1,...,j_N} |i_1\rangle \langle j_1| \otimes ... \otimes |i_N\rangle \langle j_N| . \tag{1.26}$$

Musimy sobie jeszcze przypomnieć jak **opisuje się ewolucje w mechanice kwantowej**. Robi się to ponownie operatorem unitarnym

$$|\psi(t)\rangle = U_t |\psi(0)\rangle \tag{1.27}$$

Gdy mamy równanie Schrödingera, to ewolucję aplikuję się w następujący sposób:

$$i\hbar \frac{d|\psi(t)\rangle}{dt} = H|\psi(t)\rangle \implies |\psi(t)\rangle = e^{\frac{-iHt}{\hbar}} |\psi(0)\rangle . \tag{1.28}$$

Jest to **ewolucja układu izolowanego**. W rzeczywistości praktycznie takich nie ma i będziemy się uczyć jak z tym sobie należy radzić. Na potrzeby naszego wykładu nie będziemy się specjalnie nad tym rozwodzić. Będziemy też często skrótowo pisać

$$|\psi'\rangle = U|\psi\rangle . \tag{1.29}$$

Jest to podsumowanie tego co potrzebujemy z mechaniki kwantowej.

1.1 Qubit

Aby oswoić się z przytoczonymi wcześniej informacjami możemy dokładniej omówić prosty układ kwantowy. Będziemy mówić o **qubicie**. Każdy dwuwymiarowy układ kwantowy będziemy określać **qubitem** (ew. **kubitem**). Qubit w ogólności można określić w następujący sposób:

$$dim(\mathcal{H}) = 2 \tag{1.30}$$

$$|\psi\rangle = a|0\rangle + b|1\rangle \tag{1.31}$$

W tym przypadku do opisu potrzebne nam są 2 rzeczywiste liczby, zgodnie z powyższym wywodem. Ponadto mamy warunek

$$|a|^2 + |b|^2 = 1 \tag{1.32}$$

Bez utraty ogólności mogę zapisać

$$a = \sin\left(\frac{\theta}{2}\right) e^{i\alpha} \tag{1.33}$$

$$b = \cos\left(\frac{\theta}{2}\right) e^{i\beta} \tag{1.34}$$

Pamiętamy jednak, że mamy pewną swobodność, jeśli chodzi o fazę

$$|\psi\rangle = e^{i\xi} |\psi\rangle \tag{1.35}$$

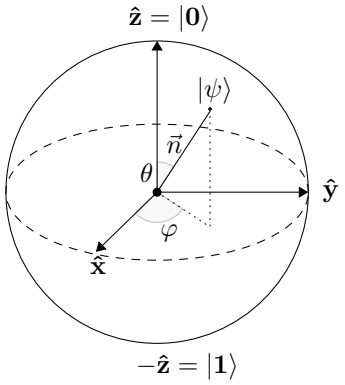
zatem jedną z faz można pominąć

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + \sin\left(\frac{\theta}{2}\right) e^{i\phi} |1\rangle , \tag{1.36}$$

gdzie ϕ zadaje względną fazę. Dodatkowo można określić zakres zmiennych

$$\theta \in [0, \pi], \ \phi \in [0, 2\pi[. \tag{1.37}$$

Dzięki temu możemy stan kwantowy qubitu *umieścić na sferze*.



Rysunek 1.1: Sfera Blocha z zaznaczonym omawianym stanem $|\psi\rangle$ oraz wektorem Blocha $\vec{n}$.

Można obliczyć macierz gęstości tego stanu.

$$|\psi\rangle\langle\psi| = \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) e^{i\phi} \end{bmatrix} \left[\cos\left(\frac{\theta}{2}\right) \ \sin\left(\frac{\theta}{2}\right) e^{-i\phi}\right] = \begin{bmatrix} \cos^2\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) e^{-i\phi} \\ \sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) e^{i\phi} & \sin^2\left(\frac{\theta}{2}\right) \end{bmatrix} =$$

$$= \frac{1}{2} \begin{bmatrix} 1+\cos(\theta) & \sin(\theta)e^{-i\phi} \\ \sin(\theta)e^{i\phi} & 1-\cos(\theta) \end{bmatrix} = \frac{1}{2}\left(\mathbb{1} + \cos(\theta)\sigma_z + \sin(\theta)\cos(\phi)\sigma_x + \sin(\theta)\sin(\phi)\sigma_z\right) = \frac{1}{2}\left(\mathbb{1} + \vec{n}\cdot\vec{\sigma}\right)$$

$$\vec{n} = \begin{bmatrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\theta) \end{bmatrix}$$

Można to poniekąd kojarzyć ze spinem 1/2 (jest to poniekąd fizyczny obrazek tego zagadnienia). Licząc

$$\langle\psi|\,\sigma_i\,|\psi\rangle = Tr(\sigma_i\,|\psi\rangle\langle\psi|) = \frac{1}{2} Tr\left(\sigma_i\left(\mathbb{1} + \vec{n}\cdot\vec{\sigma}\right)\right). \tag{1.38}$$

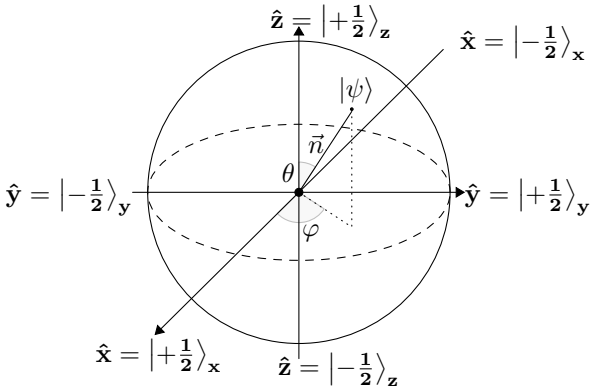
Pamiętamy, że w przypadku spinu mieliśmy

$$\vec{s} = \frac{\hbar}{2}\vec{\sigma} \tag{1.39}$$

Wektor Blocha, który reprezentuje nam stan, mówi o kierunku spinu, w którym stan jest najbardziej zorientowany. Mamy

$$\langle\psi|\,\vec{s}\,|\psi\rangle = \frac{\hbar}{2}\vec{n} \tag{1.40}$$

Jeśli teraz myślimy o spinie jako o qubicie, to możemy ponowić rysunek.



Rysunek 1.2: Sfera Blocha z zaznaczonym omawianym stanem $|\psi\rangle$ oraz wektorem Blocha $\vec{n}$. Oznaczenia osi mają teraz bardziej spinowe skojarzenia (mówią jakie wartości spinów możemy zmierzyć).

Widać, że sfera pojawia się tutaj naturalnie przez macierze Pauliego, powiązane ze spinem, który z kolei jest powiązany z grupą obrotów.

Co będzie, gdy będę miał **stan mieszany qubitu**? Pamiętamy, że

$$\rho = \sum_k q_k\,|\psi_k\rangle\,\langle\psi_k|\,. \tag{1.41}$$

Podstawiając to do powyższych rozważań, mamy

$$\rho = \sum_k q_k \frac{1}{2}(\mathbb{1} + \vec{n}_k \cdot \vec{\sigma}) = \frac{1}{2}\left(\sum_k q_k \vec{n}_k\right) \cdot \sigma_k \tag{1.42}$$

Widać, że struktura jest podobna, tylko tutaj wektorem Blocha $\vec{n}$ powinno się określić cały iloczyn $\sum_k q_k\vec{n}_k$. Zauważmy tutaj jednak, że nowy wektor Blocha spełnia

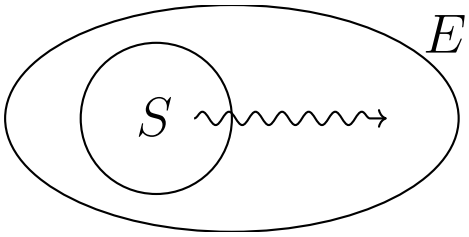
$$|\vec{n}| \geqslant 1, \tag{1.43}$$

przy czym równość zachodzi dla stanu czystego. Widać, że pełny opis stanu qubitu (biorący pod uwagę także stany mieszane) tworzy nie sferę, a **kulę Blocha**. Macierz gęstości qubitu w ogólnym stanie można zapisać jako

$$\rho = \frac{1}{2}\,(\mathbb{1} + \vec{n} \cdot \vec{\sigma})\,. \tag{1.44}$$

2 Ewolucja układów otwartych

W tym rozdziale omawiane będą kwantowe układy otwarte. O ile teoretycznie można rozważać odizolowane układy kwantowe, o tyle w rzeczywistości zawsze będziemy mieć do czynienie z jakimś otoczeniem. Schematycznie można to przedstawić na rysunku.



Rysunek 2.1: Układ (ang. *system*) S zawarty w środku otoczenia (ang. *environment*) E i oddziałujący z nim.

Wykonując jakieś działanie na układzie możemy założyć, że nie chcemy działać żadną operacją na otoczenie (czyli działamy na otoczenie identycznością). Można wtedy policzyć następujący ślad:

$$Tr_{SE}\left(\sum_{i_S,i_E,j_S,j_E}(\rho_{SE})^{i_S,i_E}_{j_S,j_E}\left|i_S\right\rangle\left\langle j_S\right|\otimes\left|i_E\right\rangle\left\langle j_E\right|\cdot A_S\otimes\mathbb{1}_E\right)=Tr_S\left(\sum_{i_S,j_S}\left(\sum_{i_E}(\rho_{SE})^{i_S,i_E}_{j_S,j_E}\right)\left|i_S\right\rangle\left\langle j_S\right|\cdot A_S\right)=Tr(\rho_S\cdot A_s) \tag{2.1}$$

Możemy tak zrobić, bo

$$Tr_{SE}(\cdot)=\sum_{i_S,i_E}\left\langle i_S\right|\otimes\left\langle i_E\right|\cdot\left|i_S\right\rangle\otimes\left|i_E\right\rangle=\sum_{i_S}\left\langle i_S\right|\left(\sum_{i_E}\left\langle i_E\right|\cdot\left|i_E\right\rangle\right)\left|i_S\right\rangle \tag{2.2}$$

W powyższym równaniu fragment

$$\left(\sum_{i_E}\left\langle i_E\right|\cdot\left|i_E\right\rangle\right) \tag{2.3}$$

nazywa się **częściowym śladem** Tr_E.

Mamy tutaj do czynienia z **zredukowaną macierzą gęstości**. Jak pokazać, że jest to macierz gęstości? Można zauważyć, że

$$(\rho_S)^{i_S}_{j_S}=\sum_{i_E}(\rho_{SE})^{i_S.i_E}_{j_S,i_E} \tag{2.4}$$

$$\rho_S=Tr_E(\rho_{SE})=\sum_{i_E}\left\langle i_E\right|\rho_{SE}\left|i_E\right\rangle \tag{2.5}$$

$$\left\langle\psi\right|\rho_S\left|\psi\right\rangle=\sum_{i_E}\left\langle\psi\right|\otimes\left|i_E\right\rangle\rho_{SE}\left|\psi\right\rangle\otimes\left|i_E\right\rangle\geqslant 0 \tag{2.6}$$

Zazwyczaj nie znamy otoczenia, dlatego opisujemy układ przez macierz gęstości (która w ogólności nie jest stanem czystym).

Uwaga

Jeżeli $\rho_{SE} = |\psi_{SE}\rangle \langle\psi_{SE}|$, czyli stan czysty, to w ogólności ρ_{SE} nie jest stanem czystym.

2.1 Puryfikacja (oczyszczenie) stanu

Wyobraźmy sobie, że mamy układ opisany stanem mieszanym opisanym przez ρ_S. Czy mając dowolną taką macierz gęstości, możemy napisać macierz gęstości tego układu wraz z otoczeniem? Czy zawsze mogę znaleźć reprezentacje tego obiektu jako zredukowaną macierz gęstości stanu większego? Formalnie można zapytać tak: czy zawsze istnieje $|\psi\rangle_{SE}$ takie, że $\rho_S = Tr_E(|\psi\rangle_{SE} \langle\psi|)$? Czy o każdej macierzy gęstości można *legalnie* myśleć w ten sposób?

Okazuje się, że tak. Pokazuje się to przez konstrukcję. Wiemy, że możemy legalnie napisać rozkład własny macierzy ρ_S, ponieważ macierze gęstości spełniają odpowiednie warunki.

$$\rho_S = \sum_k q_k |\psi_k\rangle \langle\psi_k| \tag{2.7}$$

$$q_k \geqslant 0 \tag{2.8}$$

Mając taką ogólną macierz gęstości, chcemy dla niej podać konstrukcję stanu będącego **jej puryfikacją**. Najlepiej popatrzeć na wzór 2.5, z którego liczy się częściowy ślad. Trzeba myśleć jak zdefiniować stan $|\Psi_{SE}\rangle$, żeby po obliczeniu częściowego śladu dawał macierz ρ_S.

$$|\Psi_{SE}\rangle = \sum_k \sqrt{q_k} |\psi_k\rangle \otimes |k\rangle_E , \tag{2.9}$$

gdzie $|k\rangle$ jest jakąś bazą ortonormalną na E.

Teraz sprawdźmy dlaczego to jest dobrze. Macierz gęstości odpowiadająca temu stanowi to

$$\rho_{SE} = |\Psi_{SE}\rangle \langle\Psi_{SE}| = \sum_{k,k'} \sqrt{q_k}\sqrt{q_{k'}} |\psi_k\rangle \langle\psi_{k'}| \otimes |k\rangle_E \langle k'| \tag{2.10}$$

Teraz ślad

$$Tr_E(\rho_{SE}) = \sum_k q_k |\psi_k\rangle \langle\psi_k| = \rho_S, \tag{2.11}$$

w którym z całej sumy zostają tylko diagonalne wyrazy.

Istnieje zatem konstrukcja na układzie rozszerzonym, która daje nam to co chcemy.

2.2 Ewolucja podukładu

Należy się teraz zastanowić jak opisać ewolucję podukładu. Mamy układ, który w chwili $t = 0$ jest opisany przez

$$\rho_S(0) \otimes \rho_E. \tag{2.12}$$

Staramy się przygotować układ w izolacji od otoczenia, a później dopiero *wrzucamy* go do tego otoczenia. Przy braku tego założenia trudno jest określić co będzie się dalej dziać. Mówimy, że **początkowo układ nie jest skorelowany z otoczeniem**. Później układ ewoluuje i w wyniku tego otrzymamy

$$U_{SE}(\rho_S(0) \otimes \rho_E)U_{SE}^\dagger =: \rho_{SE}(t), \tag{2.13}$$

zgodnie z tym jak ewoluuje układ izolowany (unitarnie). Jest tak, ponieważ

$$|\psi(t)\rangle = U\,|\psi(t=0)\rangle \tag{2.14}$$

$$\rho = \sum_k q_k\,|\psi_k\rangle\,\langle\psi_k| \to U\rho U^\dagger. \tag{2.15}$$

Niech $|i_E\rangle$ będzie bazą ortonormalną w E. My mamy dostęp tylko do podukładu, więc **chcemy się dowiedzieć jak wygląda zredukowana macierz gęstości i jak ona ewoluuje**. Chcemy znaleźć wyrażenie na

$$\rho_S(t) = Tr_E(\rho_{SE}(t)). \tag{2.16}$$

Można zapisać

$$\rho_S(t) = \sum_{i_E} \langle i_E|\,U_{SE}\rho_S(0)\otimes\rho_E\cdot U_{SE}^\dagger\,|i_E\rangle\,.$$

Następny zabieg nie jest konieczny, ale potrafi nieco uprościć obliczenia. Bez utraty ogólności możemy przyjąć, że ρ_E jest stanem czystym, tj.

$$\rho_E = |0\rangle\,\langle 0|\,. \tag{2.17}$$

Nie ogranicza nas to, ponieważ jeśli się uprzemy, że chcemy mieć ślad mieszany, to możemy go oczyścić i otrzymać stan czysty na otoczeniu rozszerzonym, bo i tak po tym będziemy śladować, więc różnicy nie ma. Otoczenie zawsze można powiększyć tak, żeby puryfikacja ρ_E była możliwa. Wybieramy stan $|0\rangle$, ponieważ baza jest kwestią umowną, a ślad i tak nie zależy od bazy, więc mam tutaj dowolność. Możemy zapisać stan układu po czasie t w następujący sposób

$$i_E \equiv i$$

$$\rho_S(t) = \sum_i {}_E\langle i|\,U_{SE}\rho_S(0)\otimes|0\rangle_E\,\langle S|\,U_{SE}^\dagger\,|i\rangle_E = \sum_i {}_E\langle i|\,U_{SE}\,|0\rangle_E\,\rho_S(0)\,{}_E\langle 0|\,U_{SE}^\dagger\,|i\rangle_E =$$

$$= \Big/\!\Big/\ K_i := {}_E\langle i|\,U_{SE}\,|0\rangle_E\ \Big/\!\Big/\,, = \sum_i K_i\rho_s K_i^\dagger$$

gdzie K_i nazywa się **operatorem Krausa**. Trzeba pamiętać, że $U = U(t)$, więc w $K_i = K_i(t)$, jednak często będziemy to pozostawiać w domyśle, bo nas obchodzi, ze coś wchodzi, a potem wychodzi.

W ogólności K_i nie musi być ani unitarny, ani hermitowski, ani jakiekolwiek inny. K_i są *jakimiś* macierzami zespolonymi. Operatory te są bardzo ogólne, ale zauważmy, że można na nie zapisać pewien warunek. Należy zwrócić uwagę, że

$$\sum_i K_i^\dagger K_i = \sum_i {}_E\langle 0|\,U_{SE}^\dagger\,|i\rangle_E\ {}_E\langle i|\,U_{SE}\,|0\rangle_E = {}_E\langle 0|\,U_{SE}^\dagger U_{SE}\,|0\rangle_E = {}_E\langle 0|\,\mathbb{1}_S\otimes\mathbb{1}_E\,|0\rangle_E = \mathbb{1}_S, \tag{2.18}$$

czyli dają jedynkę na układzie. Warunek ten jest poniekąd analogią warunku spełnianego przez macierze unitarne

$$U^\dagger U = \mathbb{1}. \tag{2.19}$$

Musi tak być, bo w układ może nie oddziaływać z otoczeniem (a tak by było w przeciwnym wypadku).

Wyprowadziliśmy zatem **ogólną postać ewolucji układu w kontakcie z otoczeniem**. Pomijając indeksy S (gdyż rozważania dotyczą już tylko układu S) ewolucję można podsumować następującymi równaniami:

$$\rho'' = \Lambda(\rho) = \sum_i K_i\rho K_i^\dagger, \tag{2.20}$$

$$\sum_i K_i^\dagger K_i = \mathbb{1}. \tag{2.21}$$

Uwaga

Można zauważyć, że operatorów K_i jest tyle, ile wynosi wymiar otoczenia.

Pojawia się teraz inne pytanie. **Czy mając operatory Krausa potrafimy powiedzieć jak wyglądała ewolucja? Czy można opisać puryfikację dynamiki?**

Ewolucja układu otwartego jest w postaci opisanej równaniem wyżej i każda taka ewolucja odpowiada jakiejś transformacji układu otwartego. Można powiedzieć jeszcze inaczej. Dla każdego zestawu operatorów Krausa istnieje ewolucja U_{SE} i ρ_E takie, że $\sum_i K_i \rho K_i^\dagger$ odpowiada ewolucji podukładu S. Można to trochę utożsamiać z **puryfikacją ewolucji**. Udowodnimy to ponownie przez konstrukcję. Zdefiniujmy operację U_{SE} w następujący sposób

$$\forall_{|\psi\rangle}\ U_{SE} |\psi\rangle \otimes |0\rangle_E := \sum_i (K_i |\psi\rangle_S) \otimes |i\rangle_E . \tag{2.22}$$

Przy takiej definicji możemy sprawdzić, jak U_{SE} zadziała na macierz gęstości

$$Tr_E \left(U_{SE} \rho_S \otimes |0\rangle_E \langle 0| U_{SE}^\dagger \right) = Tr_E \left(U_{SE} \sum_k q_k |\psi_k\rangle \langle\psi_k| \langle\psi_k| \otimes |0\rangle_E \langle 0| U_{SE}^\dagger \right) =$$

Możemy teraz sumę wyciągnąć, co pozwala zadziałać zgodnie z definicją

$$= \sum_k q_k Tr_E \left(\left(\sum_i K_i |k\rangle \otimes |i\rangle \right) \left(\sum_{i'} \langle k| K_{i'}^\dagger \otimes \langle i'| \right) \right) = \sum_k q_k Tr_E \left(\sum_{i,i'} K_i |\psi_k\rangle \langle\psi_k| K_{i'}^\dagger \otimes |i\rangle \langle i'| \right) =$$
$$= \sum_k q_k \sum_i K_i |\psi_k\rangle \langle\psi_k| K_i^\dagger = \sum_i K_i \rho_S K_i^\dagger .$$

Musimy jeszcze pokazać, że U jest unitarne, bo tylko w taki sposób, w mechanice kwantowej, ewoluuje układ. Aby to zrobić, wystarczy pokazać, że zachowuje iloczyny skalarne.

$$\begin{cases} |\psi\rangle \otimes |0\rangle \\ |\psi'\rangle \otimes |0\rangle \end{cases} \xrightarrow{U_{SE}} \begin{cases} \sum_i K_i |\psi\rangle \otimes |i\rangle \\ \sum_{i'} K_{i'} |\psi'\rangle \otimes |i'\rangle \end{cases}$$

Iloczyn skalarny:

$$\langle\psi|\psi'\rangle \stackrel{?}{=} \sum_i \langle\psi| K_i^\dagger K_i |\psi'\rangle = \langle\psi| \sum_i K_i^\dagger K_i |\psi'\rangle = \langle\psi|\psi'\rangle \tag{2.23}$$

Ale musimy to pokazać dla WSZYSTKICH wektorów na wejściu. To co aktualnie pokazaliśmy jest dla $|\psi\rangle \otimes |0\rangle$. Inaczej mówiąc, pokazaliśmy, że macierz U_{SE}, jeśli popatrzymy na jej kolumny powstałe z działania na nią podprzestrzenią $|\psi\rangle \otimes |0\rangle$ są *legalnymi kolumnami macierzy unitarnej*, a pamiętamy, że kolumny takiej macierzy tworzą bazę ortonormalną (tak samo zresztą jak jej wiersze). To co musimy zatem zrobić, to uzupełnić pozostałe, a to można zrobić zawsze!

Wniosek

$\Lambda(\rho)$ jest legalną transformacją macierzy gęstości układu otwartego *początkowo nieskorelowanego z otoczeniem* wtedy i tylko wtedy, gdy transformacja ta jest możliwa do zapisania w postaci

$$\Lambda(\rho) = \sum_i K_i \rho K_i^\dagger \tag{2.24}$$

$$\sum_i K_i^\dagger K_i = \mathbb{1} \tag{2.25}$$

3 Kanały kwantowe

Dzisiaj będziemy mówić o kanałach kwantowych. Dalej jesteśmy w temacie kwantowych układów otwartych. Mamy do czynienia ze stanami kwantowymi oraz macierzami gęstości takimi, że

$$|\psi\rangle \in \mathcal{H}, \tag{3.1}$$

$$\rho \in \mathcal{L}(\mathcal{H}). \tag{3.2}$$

Transformacja

$$\Lambda(\rho) = \sum_k K_k \rho K_k^\dagger \tag{3.3}$$

$$\sum_k K_k^\dagger K_k = \mathbb{1} \tag{3.4}$$

jest transformacją z macierzy gęstości w macierz gęstości

$$\Lambda : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2).$$

$\mathcal{L}(\mathcal{H})$ oznacza tutaj przestrzeń liniowych operatorów na przestrzeni Hilberta. Macierz Λ jest przede wszystkim operacją liniową, tj.

$$\Lambda(a\rho_1 + b\rho_2) = a\Lambda(\rho_1) + b\Lambda(\rho_2).$$

Można na to popatrzeć w następujący sposób

$$\rho' = \Lambda(\rho) \tag{3.5}$$

$$(\rho')^{i_2}_{j_2} = \sum_k \sum_{i_1,j_1} (K_k)^{i_2}_{i_1} \rho^{i_1}_{j_1} (K_k^\dagger)^{j_1}_{j_2} = \sum_{i_1,j_1} \sum_k (K_k)^{i_2}_{i_1} (K_k^\dagger)^{j_1}_{j_2} \rho^{i_1}_{j_1} \tag{3.6}$$

Można patrzeć na ρ' jak na macierz, ale można też rozpatrywać ρ' jako wektor w przestrzeni tensorowej:

$$(\rho')^{i_2}_{j_2} \to \rho'^{\,i_2,j_2} = |\rho'\rangle\rangle. \tag{3.7}$$

Ale równie dobrze można zastąpić

$$\rho^{i_1}_{j_1} = |\rho\rangle\rangle \tag{3.8}$$

$$\sum_k (K_k)^{i_2}_{j_2} (K_k^\dagger)^{j_1}_{j_2} = \Phi^{i_2,j_2}_{i_1,j_1} \tag{3.9}$$

i wtedy cały zapis jest

$$|\rho'\rangle\rangle = \Phi^{i_2,j_2}_{i_1,j_1} |\rho\rangle\rangle \tag{3.10}$$

i widać, że jest to liniowe przekształcenie.

Czasami o Φ niektórzy powiedzą, że jest to **superoperator**, ale nie figuruje to jako specjalny termin – uwypukla jedynie, że operator ten utworzony został w specjalny sposób.

Kolejną własnością Λ jest jej **dodatniość**. Oznacza to, że

$$\forall_{\rho \geqslant 0}\ \Lambda(\rho) \geqslant 0. \tag{3.11}$$

Znając wyprowadzenie Λ i fizyczną interpretację, jest to dosyć oczywiste, ale z samej struktury też powinno to wynikać. Jak to pokazać? Należy policzyć.

$$\langle\psi| \sum_k K_k \rho K_k^\dagger |\psi\rangle = \sum_k \langle\psi| K_k \rho K_k^\dagger |\psi\rangle = \sum_k \langle\psi_k| \rho |\psi_k\rangle \geqslant 0 \tag{3.12}$$

Okazuje się jednak, że Λ jest ponad to **całkowicie dodatnia** (ang. *completely positive – CP*). Oznacza to, że dowolne rozszerzenie Λ jest dodatnie. Formalnie

$$\Lambda_s \cdot \mathbb{1}_R \geqslant 0 \tag{3.13}$$

gdzie $\mathbb{1}_R$ jest identycznością nad rozszerzoną przestrzenią. Inaczej można zapisać

$$\sum_i K_i \otimes \mathbb{1}\rho_{SR}K_i^\dagger \otimes \mathbb{1} \geqslant 0. \tag{3.14}$$

Dla fizyka zapis taki jest w pewnym sensie oczywistym wymaganiem, bo operacja kwantowa nie zmienia przecież stanu rzeczy gdzieś dalej we wszechświecie (w ogólności). Nie psujemy dodatniości rozszerzając macierz operatora na resztę wszechświata.

Można zapytać, czy istnieją macierze gęstości, które nie są całkowicie dodatnie? Znajdziemy dużo takich...

Operacja transpozycji

$$\rho \to \rho^T \tag{3.15}$$

jest operacją dodatnią (transpozycja nie zmienia wartości własnych), ale nie **CP**. Ta cecha będzie pozwalała wykrywać splątanie między dwoma układami (ale to później).

Fakt (bez dowodu)

Każde odwzorowanie **CP** można zapisać w postaci Krausa.

Ostatnią cechą, którą posiada Λ jest **zachowanie śladu**. Zauważmy, że

$$Tr(\Lambda(\rho)) = Tr\left(\sum_k K_k\rho K_k^\dagger\right) = Tr\left(\sum_k K_k^\dagger K_k\rho\right) = Tr(\rho) = 1 \tag{3.16}$$

Możemy teraz powiedzieć czym jest **kanał kwantowy**.

Kanałem kwantowym nazywamy odwzorowanie całkowicie dodatnie, które zachowuje ślad (*ang. Compeletely Positive Trace Preserving – CPTP*).

W ogólności rozważać można

$$dim(\mathcal{H}_1) \neq dim(\mathcal{H}_2),$$

a wtedy macierze Krausa nie są macierzami kwadratowymi, a prostokątnymi.

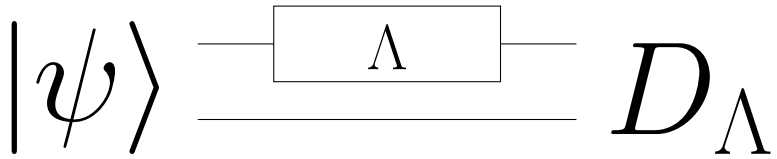
3.1 Izomorfizm Choia-Janiołkowskiego

Wprowadzimy teraz **izomorfizm Choia-Janiołkowskiego**. Mając odwzorowanie Λ rozważmy stan

$$|\psi\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^{d}|i\rangle \otimes |i\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_1,$$
$$dim(\mathcal{H}_1) = d,$$

czyli poniekąd na podwojonej przestrzeni $\mathcal{H}_1$. Układ jest bardzo silnie skorelowany, bo jak pierwszy stan jest i to drugi także. Rozważmy teraz obiekt

$$D_\Lambda = \Lambda \otimes \mathbb{1}\left(|\Psi\rangle\langle\Psi|\right),$$
$$|\Psi\rangle\langle\Psi| = \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2),$$
$$D_\Lambda \in \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_1).$$



Rysunek 3.1: Kanał kwantowy.

Macierz D_Λ jest **macierzą dynamiczną**. Skoro Λ jest **CP**, to D_Λ jest operatorem dodatnim. Z odwzorowaniem stowarzyszane jest zatem coś, co można interpretować jako stan kwantowy (myśl ta będzie jeszcze rozwinięta później).

$$|\Psi\rangle\langle\Psi| = \frac{1}{d}\sum_{i,j}|i\rangle\langle j| \otimes |i\rangle\langle j|$$

Podziałajmy tą transformacją

$$D_\Lambda = \frac{1}{d}\sum_k\sum_{i,j} K_k \otimes \mathbb{1} \otimes |i\rangle\langle j| K_k^\dagger \otimes \mathbb{1} = \frac{1}{d}\sum_k\sum_{i,j} K_k |i\rangle\langle j| K_k^\dagger \otimes |i\rangle\langle j| \tag{3.17}$$

$$(D_\Lambda)^{i_2,i_1}_{j_2,j_1} = \langle i_2|\langle i_1| D_\Lambda |j_2\rangle |j_1\rangle = ... \tag{3.18}$$

Zobaczmy co się dzieje. Na drugim układzie mamy $|i\rangle\langle j|$, zatem musimy otrzymać

$$... = \frac{1}{d}\sum_k \langle i_2| K_k |i_1\rangle\langle j_1| K_k^\dagger |j_2\rangle . \tag{3.19}$$

Można tutaj zauważyć, że

$$\langle i_2| K_k |i_1\rangle = (K_k)^{i_2}_{i_1}, \tag{3.20}$$
$$\langle j_1| K_k^\dagger |j_2\rangle = (K_k)^{j_2}_{j_1}, \tag{3.21}$$

o czym znowu można pomyśleć jak o wektorze

$$(K_k)^{i_2}_{i_1} = |K_k\rangle\rangle^{i_2,i_1}. \tag{3.22}$$

Skoro tak, to

$$(K_k^\dagger)^{j_1}_{j_2} = \langle\langle K_k|_{j_2,j_1} , \tag{3.23}$$

a stąd

$$D_\Lambda = \frac{1}{d}\sum_k |K_k\rangle\rangle\langle\langle K_k| , \tag{3.24}$$

więc jeśli z operatorów Krausa zrobić wektory to macierz dynamiczna jest sumą rzutów na te wektory. Żeby się upewnić dodajmy jeszcze, że

$$|K_k\rangle\rangle \in \mathcal{H}_2 \otimes \mathcal{H}_1. \tag{3.25}$$

Widać, że $D_\Lambda \geqslant 0$ oraz można zauważyć, że

$$Tr_{\mathcal{H}_2}(D_\Lambda) = \frac{1}{d}\sum_k \sum_{i,j} Tr_{\mathcal{H}_2}\left(K_k |i\rangle \langle j| K_k^\dagger \otimes |i\rangle \langle j|\right) = \tag{3.26}$$

$$= \frac{1}{d}\sum_{i,j} Tr\left(|i\rangle \langle j| \sum_k K_k^\dagger K_k\right) \cdot |i\rangle \langle j| = \frac{1}{d}\sum_{i,j} Tr(|i\rangle \langle j|)\, |i\rangle \langle j| = \frac{1}{d}\sum_{i,j} \delta_{i,j} |i\rangle \langle j| = \frac{1}{d}\sum_i |i\rangle \langle i| = \frac{1}{d}\mathbb{1}_{\mathcal{H}_1} \tag{3.27}$$

Widać, że zredukowany ślad daje nam $\mathbb{1}$. Jest to warunek / odpowiednik zachowania śladu. Gdy spojrzymy teraz jeszcze raz na elementy macierzowe, to zobaczymy, że

$$(D_\Lambda)^{i_2,i_1}_{j_2,j_1} = \frac{1}{d}\sum_l (K_k)^{i_2}_{i_1}(K_k^\dagger)^{j_1}_{j_2} = \Phi_\Lambda \,{}^{i_2,j_2}_{i_1,j_1}.$$

Można zauważyć, że jest to podobnie zdefiniowane, do wprowadzonej wcześniej wielkości Φ, jednak nieco inaczej mamy indeksy. Można je jednak poprawić.

$$\Phi_\Lambda \,{}^{i_2,j_2}_{i_1,j_1} \longrightarrow \Phi_\Lambda \,{}^{i_2,i_1}_{j_2,j_1}$$

Rysunek 3.2: Reshuffling indeksów.

Taką wymianę indeksów nazywa się **reshuffling**. Pamiętamy, że

$$D_\Lambda \in \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_1) \tag{3.28}$$

$$\Phi : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2) \tag{3.29}$$

Macierz Φ dla dwóch qubitów w ogólności można przedstawić w następujący sposób:

$$D_\Lambda = \begin{bmatrix} \Phi_{00}^{00} & \Phi_{01}^{00} & \Phi_{10}^{00} & \Phi_{11}^{00} \\ \Phi_{00}^{01} & \Phi_{01}^{01} & \Phi_{10}^{01} & \Phi_{11}^{01} \\ \Phi_{00}^{10} & \Phi_{01}^{10} & \Phi_{10}^{10} & \Phi_{11}^{10} \\ \Phi_{00}^{11} & \Phi_{01}^{11} & \Phi_{10}^{11} & \Phi_{11}^{11} \end{bmatrix} \tag{3.30}$$

Dla każdego $D_\Lambda \geqslant 0$ takiego, że

$$Tr_{\mathcal{H}_2}(D_\Lambda) = \frac{1}{d}\mathbb{1}_{\mathcal{H}_1} \tag{3.31}$$

istnieje odpowiadające mu **CPTP** Λ. Sprawdźmy, że definując

$$\Lambda(\rho) := dTr_{\mathcal{H}_1} D_\Lambda \mathbb{1}_{\mathcal{H}_2} \otimes \rho^T_{\mathcal{H}_1} = dTr_{\mathcal{H}_1}\left(\frac{1}{d}\sum_k \sum_{i,j} K_k |i\rangle \langle j| K_k^\dagger \otimes |i\rangle \langle j| \cdot \mathbb{1}_{\mathcal{H}_2} \otimes \sum_{i_1,j_2} (\rho^T)^{i_1}_{j_1} |i_1\rangle \langle j_1|\right) =$$

$$= \sum_k \sum_{i,j} K_k |i\rangle \langle j| K_k^\dagger (\rho^T)^j_i = \sum_k K_k \sum_{i,j} \rho^i_j |i\rangle \langle j| K_k^\dagger$$

Dzięki temu, że

$$Tr_{\mathcal{H}_2}(D_\Lambda) = \frac{1}{d}\mathbb{1}_{\mathcal{H}_1} \tag{3.32}$$

to

$$\sum_k K_k^\dagger K_k = \mathbb{1}_{\mathcal{H}_1}. \tag{3.33}$$

□

Uwaga! Zauważmy, że jak mając D_Λ, można poczynić następującą obserwację: **rozkład Krausa nie jest jednoznaczny**.

$$\Lambda(\rho) = \sum_k K_k \rho K_k^\dagger. \tag{3.34}$$

Skoro tak, to jeśli

$$\tilde{K}_k = \sum_l u_k^l K_l \tag{3.35}$$

to

$$\tilde{\Lambda}(\rho) = \sum_k \tilde{K}_k \rho \tilde{K}_k^\dagger = \sum_k \sum_{l,l'} u_k^l K_l \rho (u_k^{l'} K_i)^\dagger = \sum_{l,l'} \sum_k u_k^i (u^*)_k^{l'} K_l \rho K_{l'}^\dagger = \sum_l K_l \rho K_l^\dagger \tag{3.36}$$

Może być zatem tak, że taki sam kanał kwantowy może być rozpisany za pomocą różnej liczby macierzy Krausa. Jeżeli identyfikujemy D_Λ z Λ, to istnieje naturalna reprezentacja operatorów Krausa (**kanoniczne operatory Krausa**), które są zbudowane na bazie wektorów własnych D_Λ.

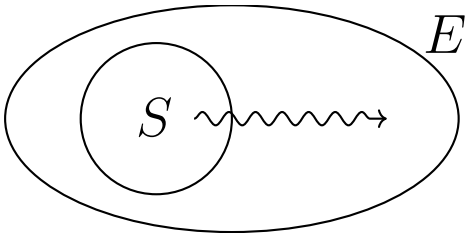
$$D_\Lambda = \frac{1}{d}\sum_k |K_k\rangle\rangle\langle\langle K_k| \,. \tag{3.37}$$

Pokazuje to jednocześnie, że każdą transformację **CP** można zapisać przy pomocy macierzy Krausa w liczbie wektorów własnych D_Λ.

4 Kwantowe Równanie Master

Jeżeli chodzi o ewolucję układów otwartych, to należy poruszyć jeszcze jeden temat. Jeśli w chwili początkowej układ nie jest skorelowany z otoczeniem $\rho_S \otimes |0\rangle$, to obserwując ten układ po jakimś czasie mamy

$$\rho_S(t) = \Lambda_t(\rho(0)) = \sum_k K_k(t)\rho(0)K_k^\dagger(t) \tag{4.1}$$



Rysunek 4.1: Układ i otoczenie.

To jest postać ogólna – nie zakłada jak oddziaływania przebiegają, ani jak otoczenie wpływa na układ. W związku z tym zależność czasowa może być *dosyć dzika*. Chcielibyśmy dołożyć pewne dodatkowe (fizyczne) założenia, aby dokładniej opisać zależność Λ_t od t. Naszym ostatecznym celem będzie dojście do równania różniczkowego na $\rho(t)$.

Tak na marginesie, przy ewolucji unitarnej

$$\frac{d\,|\psi(t)\rangle}{dt} = -\frac{i}{\hbar}H(t)\,|\psi(t)\rangle \tag{4.2}$$

$$\rho(t) = |\psi(t)\rangle\,\langle\psi(t)| \tag{4.3}$$

$$\frac{d\rho(t)}{dt} = -\frac{i}{\hbar}\left(H(t)\,|\psi(t)\rangle\,\langle\psi(t)| - |\psi(t)\rangle\right)\langle\psi(t)|\,H(t) = -\frac{i}{\hbar}[H(t),\rho(t)] \tag{4.4}$$

Jakie fizyczne założenia przyjmiemy?

1. Niezmienniczość względem przesunięcia czasu (S i E oddziałują poprzez H_{SE} niezależny od czasu).

2. Zakładamy, że otoczenie jest bardzo duże, przez co *szybko zapomina, że oddziaływało z S*. Będziemy myśleć w taki sposób, że S z E oddziałuje przez chwilkę (więc mogą się splątać, lub zadziałać na siebie w inny sposób), ale E jest duże i ma jakąś swoją dynamikę, więc wróci do swojego stanu początkowego ρ_E szybko.

$$\rho_S(0) \otimes \rho_E \to^{\delta t}_{S \text{ and } E \text{ interacts}} \ \ldots \simeq \rho_S(\delta t) \otimes \rho_E, \tag{4.5}$$

   bo w międzyczasie E szybko relaksuje do stanu początkowego ρ_E. Skala czasowa na jakiej obserwujemy zmianę ρ_S jest większa niż czas relaksacji E, zatem można przyjąć, ze cały czas ze strony otoczenia mamy do czynienia ze stanem ρ_E. Mówimy tutaj o **markowskości** czy **procesach Markowa** tj. procesach / układach, które nie mają pamięci. Bardzo często jest to dobre przybliżenie, jednak trzeba tutaj uważać na przyjętą skalę czasową (gdy schodzimy do zbyt mały czasów przybliżenie przestaje być zasadne).

Rozważmy teraz ewolucję od $t = 0$ do $t = \delta_t$. Wtedy

$$\rho_S(\delta_t) = \sum_k K_k(\delta_t)\rho_S(0)K_k^\dagger(\delta_t). \tag{4.6}$$

Jest to prawda, bo założyliśmy, że w chwili początkowej mamy $\rho_S(0) \otimes \rho_E$. Teraz zobaczmy co się dzieje od chwili $t = \delta_t$ do $t = 2\delta_t$. Tutaj korzystamy już z naszych założeń. Przez to, że zakładamy, że układ w dalszym ciągu jest nieskolerowany, to mogę znów wykorzystać cały wcześniej wprowadzony formalizm. Ponadto dynamika jest niezależna od czasu, zatem jest opisana tymi samymi Krausami.

$$\rho_S(2\delta_t) = K_k(\delta_t)\rho_S(\delta_t)K_k^\dagger(\delta_t). \tag{4.7}$$

Ostatecznie mogę zatem zapisać

$$\rho_S(t+\delta_t) = \sum_k K_k(\delta_t)\rho(t)K_k^\dagger(\delta_t).$$

Jest to już dosyć konkretne ograniczenie sposobu ewolucji macierzy gęstości. Będziemy teraz zmierzać, żeby wyciągnąć stąd postać równania różniczkowego na ρ. Zastanówmy się nad rozwinięciem macierzy Krausa w najniższych rzędach δ_t. Wyprowadzenie nie będzie może jakieś bardzo formalne, ale daje odpowiednie intuicje.

Co się musi dziać, gdy $\delta_t \to 0$? Musi być $\mathbb{1}$. Wystarczy nam jednak, że jeden z k operatorów zbiegał do jedynki – reszta może lecieć do 0. Formalnie musi być co najmniej jeden K_k taki, że

$$K_k(\delta_t) \to^{\delta_t \to 0} \mathbb{1}. \tag{4.8}$$

Niech to będzie K_0. Zatem mamy

$$K_0(\delta_t) = \mathbb{1} + Y\delta_t + O(\delta t^2), \tag{4.9}$$

pod warunkiem, że

$$K_k(\delta_t) \to^{\delta_t \to 0} 0 \tag{4.10}$$

$$k \geqslant 1. \tag{4.11}$$

Może być tak, że reszta Krausów dąży jednak do czegoś innego. Przykładowo

$$K_i(\delta_t) = \sqrt{p_i}\mathbb{1} + Y_i\delta_t + O(\delta_t^2)$$
$$\sum_i p_i = \mathbb{1}.$$

W takim wypadku, w zerowym rzędzie ρ też zostałoby odtworzone.

$$\sum_k K_o(\delta_t)\rho(t)K_i^\dagger(\delta_t) = \rho(t) + \left(\sum_i \sqrt{p_i}Y_i\rho(t) + \sum_i \sqrt{p_i}\rho(t)Y_i^\dagger\right)\delta_t + O(\delta_t^2) \tag{4.12}$$

Z dokładnością do $O(\delta_t)$ jest to ta sama dynamika jaką uzyskali byśmy biorąc

$$K_0 = \mathbb{1} + Y\delta_t \tag{4.13}$$

$$Y = \sum_i \sqrt{p_i}Y_i. \tag{4.14}$$

Innymi słowy, żałożenie jednej niezerowej macierzy nie sprawia, że tracimy ogólność. Zauważmy jednak, że gdyby to był jedyny nietrywialny operator Krausa, to

$$K_0^\dagger(\delta_t)K_0(\delta_t) = \mathbb{1} + (Y + Y^\dagger)\delta_t + O(\delta_t^2), \tag{4.15}$$

a to jednak musi być jedynką, na mocy warunku zachowania śladu

$$\sum_k K_k K_k^\dagger = \mathbb{1}, \tag{4.16}$$

co w liniowym rzędzie δ_t nie jest spełnione. Muszą być dodane jeszcze jakieś inne operatory Krausa takie, że $K_k(\delta_t) \to^{\delta_t \to 0} 0$. A jak takie operatory powinny się zachowywać, żebyśmy otrzymali to co chcemy? Widać, że trzeba wziąć

$$K_k(\delta_t) = R_k\sqrt{\delta_t} + O(\delta_t^{\frac{3}{2}}) + ... \tag{4.17}$$

Takich operatorów może być wiele – nie precyzujemy ile ich jest. Sprawdźmy, czy to nas faktycznie ustawia poprzez zbadanie warunku na zachowanie śladu.

$$\sum_k K_k^\dagger(\delta_t)K_k(\delta_t) = K_0^\dagger K_0 + \sum_{k\geqslant 1} K_k^\dagger K_k = \mathbb{1} + (Y + Y^\dagger) + \sum_{k\geqslant 1} R_k^\dagger R_k \delta_t + O(\delta_t^2) \tag{4.18}$$

Będzie zatem dobrze, jeżeli

$$(Y + Y^\dagger) = -\sum_{k\geqslant 1} R_k^\dagger R_k. \tag{4.19}$$

Pokazuje to, że musi istnieć związek między Krausem K_0 a pozostałymi. Wiemy, że Y jest jakimś dowolnym operatorem, niekoniecznie Hermitowskim. Zapiszmy go jako

$$Y = A - iH, \tag{4.20}$$

gdzie A i H są operatorami Hermitowskimi (można tak zrobić dla dowolnego operatora). Stąd mamy, że

$$2A = -\sum_k R_k^\dagger R_k$$
$$A = -\frac{1}{2}\sum_k R_k^\dagger R_k.$$

Skoro tak, to możemy napisać kanał

$$\rho(t+\delta_t) = \sum_k K_k(\delta_t)\rho(t)K_k^\dagger(\delta_t) = (\mathbb{1} + Y\delta_t)\rho(t)(\mathbb{1} + Y^\dagger\delta_t) + \sum_{k\geqslant 1} R_k\rho(t)R_k^\dagger\delta_t + O(\delta_t^2) \tag{4.21}$$

$$\rho(t-\delta t) = \rho(t) - i[H,\rho(t)]\delta_t + \delta_t\left[\sum_{k\geqslant 1} R_k\rho(t)R_k^\dagger - \frac{1}{2}R_k^\dagger R_k\rho(t) - \frac{1}{2}\rho(t)R_k^\dagger R_k\right] + O(\delta_t^2) \tag{4.22}$$

$$\frac{d\rho(t)}{dt} = -i[H,\rho(t)] + \sum_k R_k\rho(t)R_k^\dagger - \frac{1}{2}\left\{R_k^\dagger R_k, \rho(t)\right\} \tag{4.23}$$

I to nazywa się **kwantowym równaniem Master** lub częściej **równaniem Gorina-Kossakowskiego-Sudarshana-Lindblada**.R są tutaj dowolnymi operatorami. Na R_k mówi się **jump operators**, **noise operators** lub **Lindblad operators**.

Uwaga! To jest równanie liniowe na ρ

$$\frac{d\rho(t)}{dt} = \mathcal{L}(\rho(t)). \tag{4.24}$$

Widać, że po obu stronach jest $\rho(t)$, co oddaje fakt, że *nie czujemy* pamięci. Sam operator $\mathcal{L}$ nie czuje czasu sam z siebie. Można zatem pokazać, że $\mathcal{L}$ jest pewną macierzą, która działa na zwektoryzowane ρ.

$$\frac{d\rho(t)}{dt} \to e^{\mathcal{L}t}(\rho(0)) \tag{4.25}$$
$$\Lambda_t = \mathcal{L}t \tag{4.26}$$

Wcześniej, gdy mówiliśmy o kanałach kwantowych mówiliśmy, że coś wchodzi na wejściu, wychodzi na wyjściu i mamy Krausy, a teraz widzimy, że Λ można otrzymać przez odpowiednie wykorzystanie operatora liniowego $\mathcal{L}$. Możemy generować rodzinę Λ. Ponadto mają one własność składania

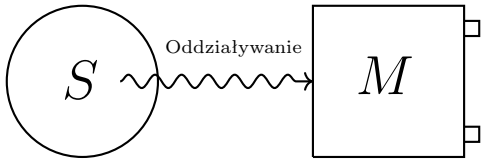
$$\Lambda_{t_1+t_2} = \Lambda_{t_1} \circ \Lambda_{t_2}.$$

Formalnie mamy tutaj do czynienia z **półgrupą dynamiczną**, co wynika z faktu, że nie mamy tutaj elementu odwrotnego (nie możemy rozkurczać sfery Blocha).

Podsumowując, otrzymaliśmy rodzinę odwzorowań CP Λ_t, które tworzą półgrupę dynamiczną.

4.1 Pomiary uogólnione i rozróżnialność stanów kwantowych

Mówimy o **pomiarach uogólnionych**, gdy układ oddziałuje z urządzeniem pomiarowym, a następnie pomiar rzutowy jest wykonywany na urządzeniu pomiarowym. Schemat jest bardzo podobny do tego jaki był z otoczeniem.



Rysunek 4.2: Układ S i urządzenie pomiarowe M.

Jak to opisać? Bardzo łatwo.

$$\rho_s \otimes |0\rangle_M \langle 0| \rightarrow^{U_{SM}} U_{SM}\rho_S |0\rangle_M \langle 0| U_{SM}^\dagger \tag{4.27}$$

Wykonujemy pomiar na M w bazie $|i\rangle_M$. Nasz operator rzutowy związany z pomiarem jest $\mathbb{1} \otimes |i\rangle_M \langle i|$. Jakie dostajemy prawdopodobieństwa?

$$p_i = Tr(U_{SM}\rho_s \otimes |0\rangle_M \langle 0| U_{SM}^\dagger \cdot \mathbb{1}_S \otimes |i\rangle_M \langle i|) = Tr(\rho_S \otimes |0\rangle_M \langle 0| U_{SM}^\dagger |i\rangle_M \langle i| U_{SM}) = Tr_S(\rho_S \langle 0| U_{SM}^\dagger |i\rangle \langle i| U_{SM} |0\rangle) =$$
$$= Tr_S(\rho_S K_i^\dagger K_i) = Tr_S(\rho_s M_i) = p_i, \tag{4.28}$$
$$K_i^\dagger K_i =: M_i, \tag{4.29}$$

przy czym operatory M_i nie koniecznie muszą być operatorami rzutowymi, ale wiemy, że

$$‘\sum_i M_i = \mathbb{1}, \tag{4.30}$$
$$M_i \geqslant 0. \tag{4.31}$$

Każdy operator o strukturze $A^\dagger A$ jest dodatni, wiec to naturalne. Prawdopodobieństwa będą zatem liczone zgodnie z regułą

$$\sum_i p_i = 1, \tag{4.32}$$
$$p_i \geqslant 0 \tag{4.33}$$

Zestaw operatorów dodatnich sumujących się do jedynki w matematyce nazywa się **pomiarem uogólnionym**.

5 Rozróżnialność stanów kwantowych

Ostatnio powiedzieliśmy, że warto rozważać pomiary uogólnione, a nie tylko pomiary rzutowe. Częstokroć jest to opis bardziej adekwatny. Wprowadziliśmy też operator $M_i = K_i^\dagger K_i$. Opisując uogólniony układ pomiarowy prawdopodobieństwa będą liczone w następujący sposób

$$p_i = Tr(\rho_s M_i), \tag{5.1}$$

przy czym należy pamiętać, że teraz nie ma gwarancji, że M_i jest pomiarem rzutowym. Pojawia się pytanie, czy jak dostaniemy dowolny zestaw operatorów M_i, to czy potrafimy odtworzyć jaki schemat pomiarowy został wykonany? Okazuje się, że **dla każdego zestawu operatorów pomiarowych**

$$M_i \geqslant 0 \tag{5.2}$$

$$\sum_i M_i = \mathbb{1} \tag{5.3}$$

istnieje U_{SM} i pomiar rzutowy na M taki, że

$$Tr(\rho_s M_i) = Tr(U_{SM} \rho_s \otimes |0\rangle_M \langle 0| U_{SM}^\dagger \mathbb{1} \otimes |i\rangle_M \langle i|). \tag{5.4}$$

Dowód

Jest praktycznie identyczny jak był dla Krausów wcześniej. Zdefiniujmy

$$U_{SM} |\psi\rangle_S \otimes |0\rangle_M := \sum_i \sqrt{M_i} |\psi\rangle \otimes |i\rangle \tag{5.5}$$

i sprawdźmy, że to działa. Weźmy stan czysty

$$\rho_S = |\psi\rangle \langle\psi| \tag{5.6}$$

i dostaniemy

$$Tr(U_{SM} |\psi\rangle \langle\psi| \otimes |0\rangle \langle 0| U_{SM}^\dagger \cdot \mathbb{1} \otimes |i\rangle \langle i|) = Tr\left(\left(\sum_{jk} \sqrt{M_j} |\psi\rangle \otimes |j\rangle\right)\left(\langle\psi| \sqrt{M_k} \otimes \langle k|\right) \cdot \mathbb{1} \otimes |i\rangle \langle i|\right) = \tag{5.7}$$

$$= Tr\left(\sqrt{M_i} |\psi\rangle \langle\psi| \sqrt{M_i}\right) = Tr\left(|\psi\rangle \langle\psi| M_i\right) \tag{5.8}$$

□

Uwaga! Nic się nie zmieni, gdy weźmiemy

$$U_{SM} |\psi\rangle \otimes |0\rangle = \sum_i V_i \sqrt{M_i} |\psi\rangle \otimes |i\rangle, \tag{5.9}$$

gdzie V_i jest unitarne. Prawdopodobieństwa są nieczułe na macierze unitarne, więc możliwe jest *kręcenie* stanem przy otrzymanym z pomiaru wyniku i. Jest to stanowcza różnica między pomiarami rzutowymi a uogólnionymi.

5.1 Rozróżnianie stanów kwantowych – przykład

Wyobraźmy sobie sytuację, że z prawdopodobieństwem q_i mamy stany ρ_i. Mamy zatem zespół stanów $\{q_i, \rho_i\}$. Naszym zadaniem jest wykonać pomiar na takim przychodzącym stanie (zakładamy, że mamy tylko jedną kopię) i naszym zadaniem jest powiedzieć który z tych stanów mamy. Chcemy sformułować problem tak, żeby podać najlepszy pomiar (żeby coś zoptymalizować). Jak zapisać pomiar? Powiemy, że to jest zestaw operatorów $\{M_j\}$ i będziemy szukać zestawu, który da nam najlepszą rozróżnialność. Nie musimy tutaj myśleć o oddziaływaniu

z otoczeniem czy innych zagadnieniach, jeżeli matematycznie pokażemy, że ten zestaw operatorów jest najlepszy, jaki może być. Często okazuje się, że optymalnym jest pomiar rzutowy. Wprowadźmy funkcję kosztu. Określa ona karę za zgadnięcie j jeśli mieliśmy odgadnąć i. Oznaczmy ją $C(j|i)$. Średni koszt dany jest wtedy przez

$$\bar{c} = \sum_i q_i p(j|i) C(j|i), p(j|i) = Tr(\rho_i M_j) \tag{5.10}$$

Szukamy

$$min \underset{\substack{M_j \geqslant 0 \\ \sum_i M_j = \mathbb{1}}}{\{M_j\}} = ? \tag{5.11}$$

Przykład

Dostajemy dwa stany czyste z równymi prawdopodobieństwami.

$$\rho_0 = |\psi_0\rangle \langle\psi_0| \tag{5.12}$$

$$\rho_1 = |\psi_1\rangle \langle\psi_1| \tag{5.13}$$

$$q_0 = q_1 = \frac{1}{2} \tag{5.14}$$

Przyjmiemy prostą funkcję kosztu

$$c(j|i) = 1 - \delta_{ij}. \tag{5.15}$$

W związku z tym, że są tylko dwa możliwe stany, to mamy tylko dwa operatory $\{M_0, M_1\}$ i musimy znaleźć jakie dwa operatory będą tutaj najlepsze. Chcemy zminimalizować

$$\bar{c} = \frac{1}{2}\left(\langle\psi_0| M_0 |\psi_0\rangle \cdot 0 + \langle\psi_0| M_1 |\psi_0\rangle \cdot 1 + \langle\psi_1| M_1 |\psi_1\rangle \cdot 0 + \langle\psi_1| M_0 |\psi_1\rangle \cdot 1\right) = \frac{1}{2}\left(\langle\psi_0| M_1 |\psi_0\rangle \cdot 1 + \langle\psi_1| M_0 |\psi_1\rangle \cdot 1\right) \tag{5.16}$$

$$M_1 = \mathbb{1} - M_0 \tag{5.17}$$

$$\bar{c} = \frac{1}{2}\left(1 + Tr(M_0(|\psi_1\rangle \langle\psi_1| - |\psi_0\rangle \langle\psi_0|))\right) \tag{5.18}$$

i chcę to zminimalizować po operatorach. Wiemy, że skoro M_0 jest dodatnie i M_1 jest dodatnie, to M_0 nie mieć wartości własnych większych niż 1. Możemy sparametryzować

$$|\psi_0\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)|1\rangle \tag{5.19}$$

$$|\psi_1\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle - \sin\left(\frac{\theta}{2}\right)|1\rangle. \tag{5.20}$$

Wtedy

$$\langle\psi_1|\psi_0\rangle = \cos(\theta). \tag{5.21}$$

Podstawmy sobie te stany i zbadajmy otrzymaną macierz.

$$|\psi_1\rangle \langle\psi_1| - |\psi_0\rangle \langle\psi_0| = \begin{bmatrix} 0 & -\sin(\theta) \\ -\sin(\theta) & 0 \end{bmatrix} = -\sin(\theta) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{5.22}$$

Szukamy takiego M_0, który maksymalizuje wyrażenie

$$Tr\left(M_0 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right). \tag{5.23}$$

Widać, że mamy tutaj do czynienia z macierzą σ_x, zatem możemy w bazie jej wektorów własnych rozpisać

$$Tr\left(M_0 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) = \langle +| M_0 |+\rangle - \langle -| M_0 |-\rangle = 1, \tag{5.24}$$

więc szukamy rzutu na $|+\rangle$.

$$M_0 = |+\rangle \langle +|, \tag{5.25}$$

stąd

$$\bar{c} = \frac{1}{2}(1 - \sin(\theta)) = \frac{1}{2}\left(1 - \sqrt{1 - \cos^2(\theta)}\right) = \frac{1}{2}\left(1 - \sqrt{1 - |\langle \psi_0|\psi_1\rangle|^2}\right). \tag{5.26}$$

Jeśli stany są ortogonalne, to

$$\langle \psi_0|\psi_1\rangle = 0 \tag{5.27}$$

i wtedy $\bar{c} = 1$ – stany ortogonalne są idealnie rozróżnialne. W przypadku $|\psi_0\rangle = |\psi_1\rangle$ mamy $\bar{c} = \frac{1}{2}$, więc znowu zgodnie z oczekiwaniem. Optymalny pomiar, który otrzymaliśmy, to

$$M_0 = |+\rangle \langle +| \tag{5.28}$$

$$M_1 = \mathbb{1} - |+\rangle \langle +| = |-\rangle \langle -|, \tag{5.29}$$

czyli mamy pomiar w bazie $\{|+\rangle, |-\rangle\}$.

Uwaga! Gdybyśmy dostali n-razy ten sam stan

$$|\psi_0\rangle^{\otimes N} \vee |\psi_1\rangle^{\otimes N} \tag{5.30}$$

to używamy w zasadzie tej samej procedury. Wtedy mamy znowu

$$\bar{c}_{opt,N} = \frac{1}{2}\left(1 - \sqrt{1 - |\langle \psi_0|\psi_1\rangle|^{2n}}\right) \to^{N\to\infty}_{|\langle \psi_0|\psi_1\rangle|<1} 0 \tag{5.31}$$

□

Pokazaliśmy, że minimalny średni koszty może być interpretowany jako prawdopodobieństwo błędu przy rozróżnianiu stanów. Mamy dwa stany i próbujemy je rozróżnić. Powiedzmy, że będziemy zgadywać tylko wtedy, gdy jesteśmy pewni, że zgadniemy dobrze. Problem taki nosi nazwę **rozróżniania jednoznacznego**. W tym protokole (rozróżniając znowu dwa stany), mamy 3 operatory pomiarowe.

1. M_0 – zgadujemy $|\psi_0\rangle$,
2. M_1 – zgadujemy $|\psi_1\rangle$,
3. $M_?$ – nie zgaduję.

$$M_0 + M_1 + M_? = \mathbb{1} \tag{5.32}$$

i tutaj wyraźnie widać, że to nie są operatory rzutowe, więc gdy mowa o operatorach uogólnionych, otwierają się pewne nowe horyzonty. Jaki powinien być operator M_0?

$$M_0 = \lambda \left|\psi_1^{\perp}\right\rangle \left\langle \psi_1^{\perp}\right| \tag{5.33}$$

Mamy tutaj symetrię problemów (uwaga ad. λ).

$$M_1 = \lambda \left|\psi_0^\perp\right\rangle \left\langle\psi_0^\perp\right| \tag{5.34}$$

$$M_? = \mathbb{1} - M_0 - M_1 \tag{5.35}$$

Naszym celem jest teraz to, żeby zgadywać jak najczęściej. Chodzi o to, żeby wynik $M_?$ wychodził najrzadziej, jak to możliwe.

$$p_? = \frac{1}{2}\left(\langle\psi_0| M_? |\psi_0\rangle + \langle\psi_1| M_? |\psi_1\rangle\right) = \frac{1}{2}\left(1 - \lambda|\left\langle\psi_0\middle|\psi_1^\perp\right\rangle + 1 - \lambda\left|\left\langle\psi_1\middle|\psi_0^\perp\right\rangle\right|^2\right) \tag{5.36}$$

Wracając do wygodniejszej postaci (trygonometrycznej):

$$p_? = 1 - \left|2\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)\right|^2 = 1 - \lambda\sin^2(\theta).$$

Chcemy, by λ była jak największa, pamiętając, że

$$M_? = \mathbb{1} - M_0 - M_1 \geqslant 0. \tag{5.37}$$

Otrzymamy

$$\mathbb{1} - \lambda\left(\left|\psi_1^\perp\right\rangle\left\langle\psi_1^\perp\right| + \left|\psi_0^\perp\right\rangle\left\langle\psi_0^\perp\right|\right) \geqslant 0 \tag{5.38}$$

$$\lambda \geqslant \frac{1}{2\cos^2\left(\frac{\theta}{2}\right)} = \frac{1}{1+\cos(\theta)} \tag{5.39}$$

i stąd

$$p_? = 1 - 2\sin^2\left(\frac{\theta}{2}\right). \tag{5.40}$$

6 Splątanie

Będziemy teraz mówić o splątaniu kwantowym. Zaczniemy od wprowadzenia notacji. Stan kwantowy

$$|\psi\rangle \in \mathcal{H}_d \tag{6.1}$$

i macierz gęstości

$$\rho \in B(\mathcal{H}_d) \tag{6.2}$$

lub

$$\rho \in \mathcal{L}(\mathcal{H}_d). \tag{6.3}$$

Pamiętamy, że

$$\rho \geqslant 0 \tag{6.4}$$
$$Tr(\rho) = 1 \tag{6.5}$$
$$\rho^\dagger = \rho. \tag{6.6}$$

Szczególnym stanem jest qubit

$$|\psi\rangle \in \mathbb{C}_2 \tag{6.7}$$

i ogólnie można napisać

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \tag{6.8}$$
$$|\alpha|^2 + |\beta|^2 = 1. \tag{6.9}$$

Każdy qubit mogę też reprezentować jako sferę Blocha

$$\rho = \frac{1}{2}(\mathbb{I} + \vec{n}\vec{\sigma}). \tag{6.10}$$

Macierze gęstości

Obserwablami nazywamy takie operatory, że

$$\hat{A} = \hat{A}^\dagger \tag{6.11}$$

Mamy układ kwantowy $|\psi_i\rangle$, ale możemy uzyskać taki stan z prawdopodobieństwem p_i. Dysponując zatem kilkoma stanami mamy zestaw $\{p_i, |\psi_i\rangle\}$. I tutaj, jeżeli chcemy na nim zmierzyć wartość wielkości $\hat{A}$, to

$$A_\psi = \langle\psi|\,\hat{A}\,|\psi\rangle \implies A = \sum_i p_i\,\langle\psi_i|\,\hat{A}\,|\psi_i\rangle = Tr\left(\sum_i p_i\,|\psi_i\rangle\,\langle\psi_i|\,\hat{A}\right) \tag{6.12}$$

Mówimy, że stan $|\psi\rangle\,\langle\psi| = \rho$ jest stanem czystym. Natomiast, jeżeli

$$\rho = \sum_i p_i\,|\psi_i\rangle\,\langle\psi_i|\,, \tag{6.13}$$

to stan jest mieszany. W ogólności $\langle\psi_i|\psi_j\rangle \neq \delta_{ij}$. Zauważmy, że

$$\rho = \sum_i p_i\,|\psi_i\rangle\,\langle\psi_i| = \sum_{k=1}^{r<d} \lambda_k\,|e_k\rangle\,\langle e_k| \tag{6.14}$$
$$\langle e_k|e_l\rangle = \delta_{kl}. \tag{6.15}$$

Czystością stanu nazywamy

$$`Tr(\rho^2) = \sum_k \lambda_k^2 \leqslant \sum_k \lambda_k = 1, \tag{6.16}$$

czyli widać, że czystość jest ograniczona z góry przez 1 (dla stanu czystego). Z drugiej strony zachodzi

$$Tr(\rho^2) \geqslant \frac{1}{d}, \tag{6.17}$$

co z kolei jest równe dla stanu najbardziej zmieszanego:

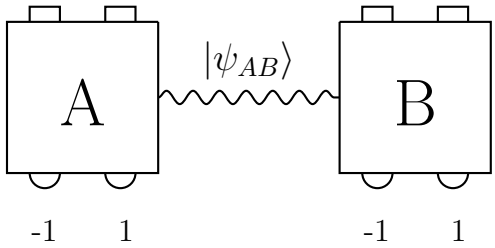
$$\rho = \frac{1}{d} \sum_{k=1}^{d} |e_k\rangle \langle e_k| . \tag{6.18}$$

Ale chcemy mówić o stanach splątanych, więc potrzebujemy przejść do większej przestrzeni.

Stany dwuczęściowe (*Bipartiate states*)

Będziemy rozpatrywać stany w postaci.

$$|\psi_{AB}\rangle \in B(\mathcal{H}_{d_A}^A \otimes \mathcal{H}_{d_B}^B) \tag{6.19}$$



Rysunek 6.1: Alicja i Bob współdzielą stan.

Mamy Alicję i Boba i oni przyciskiem wybierają pomiar i żarówka wskazuje jaki mają wynik. Stan, który dzielą, można zapisać w następujący sposób:

$$|\psi_{AB}\rangle = \sum_{i=1}^{d_A} \sum_{j=1}^{d_B} \alpha_{ij} |i\rangle_A \otimes |j\rangle_B \tag{6.20}$$

$$|ij\rangle \equiv |i,j\rangle \equiv |i\rangle \otimes |j\rangle , \tag{6.21}$$

przy czym póki co ograniczamy się do stanu czystego. Możemy wyznaczyć bazę obliczeniową (obsadzeniową, *computational base*). Standardowo:

$$\{|00\rangle , |01\rangle , |10\rangle , |11\rangle\} \tag{6.22}$$

Istnieje też baza Bella

$$\left\{ |\psi_\pm\rangle = \frac{1}{\sqrt{2}} (|01\rangle \pm |10\rangle) , |\phi_\pm\rangle = \frac{1}{\sqrt{2}} (|00\rangle \pm |11\rangle) \right\} \tag{6.23}$$

i można pokazać, że są te wektory ortonormalne. Chcąc zdefiniować stan zmieszany u Alicji i Boba chcemy określić

$$\rho_{AB} \in B(\mathcal{H}_{d_A}^A \otimes \mathcal{H}_{d_B}^B) \tag{6.24}$$

i pomijając oznaczenia AB mamy

$$\rho = \sum_i p_i \left|\psi_i\right\rangle_{AB} \left\langle\psi_i\right| \tag{6.25}$$

$$\left|\psi_i\right\rangle = \sum_{kl} \alpha_{kl}^{(i)} \left|kl\right\rangle \tag{6.26}$$

Licząc dalej mamy

$$\rho = \sum_i p_i \sum_{kl} a_{kl}^{(i)} (a_{kl}^{(i)})^* \left|kl\right\rangle \left\langle kl\right| = \sum_{kl} \sum_i p_i a_{kl}^{(i)} (a_{kl}^{(i)})^* \left|kl\right\rangle \left\langle kl\right| \tag{6.27}$$

W przypadku dwóch qubitów, w bazie standardowej, mamy przykładowo

$$\rho_{11}^{01} \left|01\right\rangle \left\langle 11\right| = \rho_{11}^{01} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} (0,0,0,1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho_{11}^{01} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.28}$$

Stan dwukubitowy mogę zatem zapisać w postaci macierzy stanu

$$\rho = \begin{bmatrix} \rho_{00}^{00} & \rho_{01}^{00} & \rho_{10}^{00} & \rho_{11}^{00} \\ \rho_{01}^{00} & \rho_{01}^{01} & \rho_{10}^{01} & \rho_{11}^{01} \\ \rho_{10}^{00} & \rho_{01}^{10} & \rho_{10}^{10} & \rho_{11}^{10} \\ \rho_{11}^{00} & \rho_{01}^{11} & \rho_{10}^{11} & \rho_{11}^{11} \end{bmatrix}. \tag{6.29}$$

Jesteśmy teraz gotowi do rozpoczęcia rozmowy o splątaniu.

Definicja

Stan produktowy to stan, który możemy zapisać w postaci

$$\left|\psi\right\rangle_{AB} = \left|\chi\right\rangle_A \otimes \left|\phi\right\rangle_B \tag{6.30}$$

Przykład

$$\left|01\right\rangle = \left|0\right\rangle \otimes \left|1\right\rangle. \tag{6.31}$$

Można to uogólnić na stany mieszane.

Definicja

Stany separowalne to takie, które są mieszaniną stanów produktowych.

Formalnie:

$$\rho_{AB} = B(\mathcal{H}_A \otimes \mathcal{H}_B) \tag{6.32}$$

$$\rho \in SEP \iff \rho = \sum_i p_i \left|\chi_i\right\rangle_A \left\langle\chi_i\right| \otimes \left|\phi_i\right\rangle_B \left\langle\phi_i\right| \tag{6.33}$$

Uwaga

Można powiedzieć, że stan produktowy to czysty stan separowalny.

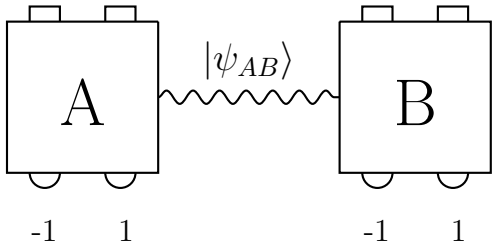
Definicja

Stan splątany to stan, który nie jest stanem separowalnym.

$$\rho \notin SEP. \tag{6.34}$$

Co się dzieje na stanie zmieszanym, gdy Alicja i Bob mierzą różne obserwable?

Motywacja



Rysunek 6.2: Rysunek pomocniczy do przyszłych rozważań.

Będziemy teraz mówić o klasycznych korelacjach. Rozważmy obserwable $\hat{p}$ braną przez Alicję i $\hat{x}$ braną przez Boba. Mierzymy zatem

$$\hat{A} \otimes \hat{B} = \hat{p} \otimes \hat{x} \tag{6.35}$$

Ta obserwacja daje pewną korelację (iloczyn) [można tutaj odnieść się do hasła korelacja EPR]

$$\langle AB \rangle = Tr\left(\rho_{AB} \hat{A} \otimes \hat{B}\right) = \tag{6.36}$$

i teraz jeśli $\rho \in SEP$ to mamy

$$=_{\rho \in SEP} \sum_{p_i} \langle \chi_i | A \chi_i \rangle \langle \phi_i | B \phi_i \rangle = \sum_i p_i \langle A \rangle_{\chi_i} \langle B \rangle_{\phi_i}$$

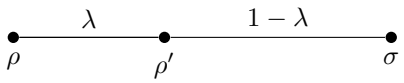
i dostajemy tutaj korelację klasyczną. Pokazuje to, że dla stanu separowalnego można by wszystko zasymulować klasycznie, nie znając mechaniki kwantowej. Jeżeli

$$\rho, \sigma \in SEP \implies \rho' = \lambda\rho + (1-\lambda)\sigma \tag{6.37}$$
$$\rho' \in SEP \tag{6.38}$$
$$0 \leqslant \lambda \leqslant 1 \tag{6.39}$$

Jeżeli wezmę dwa dowolne stany, to ich mieszanka też jest stanem, wiec powyższe jest dosyć oczywiste.

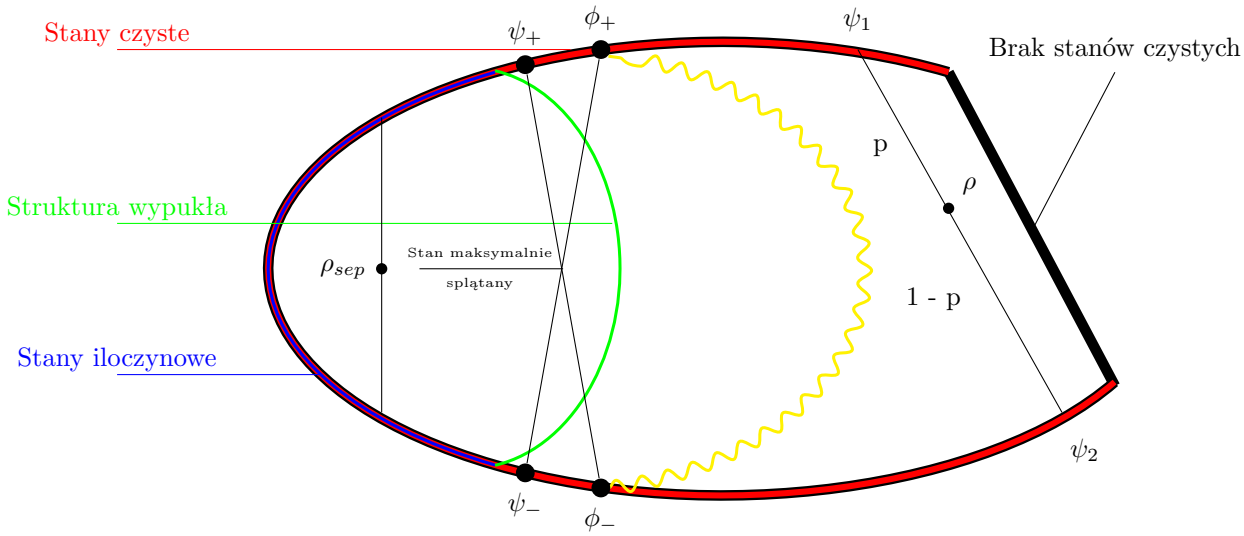
$$\rho, \sigma \in \mathcal{H}_A \otimes \mathcal{H}_b \implies \lambda\rho + (1-\lambda)\sigma \in B(\mathcal{H}_A \otimes \mathcal{H}_B) \tag{6.40}$$
$$\rho' = \lambda\rho + (1-\lambda)\sigma \tag{6.41}$$
$$\rho \geqslant 0 \tag{6.42}$$
$$Tr(\rho') = 1 \tag{6.43}$$
$$\rho' = (\rho')^\dagger \tag{6.44}$$

Możemy to zwizualizować. Co oznacza mieszanie stanów? Jeżeli będziemy mieli stany ρ i σ to możemy połączyć je linią, w następujący sposób



Rysunek 6.3: Wizualizacja stanu mieszanego.

Jest to poniekąd geometryczny sposób mieszania czegoś. Wszystkie stany leżą w przestrzeni wypukłej. Warunkiem dodatniości jest hiperpłaszczyzna w przestrzeni. Stany czyste leżą na brzegu. Reszta została zaznaczona na rysunku.



Rysunek 6.4: Graficzna reprezentacja omawianego zagadnienia.

Będziemy teraz mówić o dekompozycji (rozkładu) Schmidta stanów czystych, lub inaczej SVD macierzy gęstości.

$$|\psi\rangle_{AB} = \sum_{i=1}^{d_B} \sum_{j=1}^{d_A} \alpha_{ij} |ij\rangle \,, \tag{6.45}$$

gdzie α_{ij} jest macierzą współczynników stanu czystego.

$$\alpha_{ij} = \sum_{k=1}^{r} u_{ik}^* \lambda_k v_{kj}, \tag{6.46}$$

r oznacza tutaj Rank (rząd) Schmidta (liczba wartości osobliwych).

$$d_A \left\{ \left[\overbrace{\quad \alpha \quad}^{d_B} \right] \right. = d_A \left\{ \left[\overbrace{\quad U^* \quad}^{r} \right] \right. r \left\{ \left[\overbrace{\quad \lambda \quad}^{r} \right] \right. r \left\{ \left[\overbrace{\quad V \quad}^{d_B} \right] \right.$$

Rysunek 6.5: Wymiary omawianych macierzy.

Zachodzi

$$\sum_i u_{ki}^* u_{il} = \delta_{kl} \tag{6.47}$$

Wszystko to oznacza, że

$$|\psi\rangle_{AB} = \sum_{ij}\sum_{k} u^*_{ik}\lambda_k v_{kj}\,|ij\rangle = \sum_{k=1}^{r}\left(\sum_i u^*_{ik}\,|i\rangle\right)\otimes\left(\sum_j v_{kj}\,|j\rangle\right) \tag{6.48}$$

Obserwacja

Z SVD wynika, że $1 \leqslant r \leqslant min(d_A, d_B)$.

1. $r = 1 \implies |\psi_{AB}\rangle = |e\rangle_A \otimes |f\rangle_B$
2. $r > 1 \implies |\psi_{AB}\rangle$ jest splątane.
3. $r = min(d_A, d_B) \wedge \forall_k : |\lambda_k| = \frac{1}{\sqrt{r}}$, to stan taki mogę w ogólności napisać

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{r}}\sum_{k=1}^{r} e^{i\phi k}\,|e_k\rangle_A\,|f_k\rangle_B$$

mamy tutaj równe rozłożenie współczynników Schmidta i najwyższy rząd. Taki stan nazywamy **maksymalnie splątanym**.

W przypadku dwóch qubitów mówimy o częściowo splątanym stanie

$$|\psi_+(\lambda)\rangle = -\sqrt{\lambda}\,|01\rangle + \sqrt{1-\lambda}\,|10\rangle \tag{6.49}$$

$$0 \leqslant \lambda \leqslant 1, \tag{6.50}$$

gdzie λ można utożsamiać poniekąd z wagami (przy $\lambda = 0, 1$ mamy produktowe).

Czasami dla $\lambda = \frac{1}{2}$ mówi się o **ebitach** (*entangled bits*).

Stany mieszane

W stanach mieszanych nie ma metody (uniwersalnej) na weryfikację splątania. Są różne metody jak sprawdzać, czy dany stan jest splątany, czy nie. Warunkiem wystarczającym jest **ujemność częściowej transpozycji**. Mówimy wtedy, że

$$\rho \in NPT \implies \rho \notin SEP, \tag{6.51}$$

gdzie NPT to skrót od *negative partial transposition*.

$$\rho = \sum_{ijkl}\rho^{ij}_{kl}\,|ij\rangle\langle kl|\,(\rho)^T = \sum_{ijkl}\rho^{ij}_{kl}(|ij\rangle\langle kl|)^T = \sum_{ijkl}\rho^{ij}_{kl}\,|kl\rangle\langle ij| \tag{6.52}$$

Możemy zrobić transpozycję częściową, np. tylko po części Alicji

$$\rho^{T_A} = \sum_{ijkl}\rho^{ij}_{kl}(|i\rangle\langle k|)^T \otimes |j\rangle\langle l| = \sum_{ijkl}\rho^{kj}_{il}\,|ij\rangle\langle kl| \tag{6.53}$$

$$\rho^{T_B} = \sum_{ijkl}\rho^{il}_{kj}\,|ij\rangle\langle kl| \tag{6.54}$$

Można sprawdzić, że

$$(\rho^{T_A})^T = \rho^{T_B}. \tag{6.55}$$

Jeżeli mamy stan separowalny $\rho \in SEP$, to jest to równoważne, z

$$\rho \in SEP \iff \rho^{T_A} = \sum_i p_i(|\psi_i\rangle \langle\psi_i|^*) \otimes |\phi_i\rangle \langle\phi_i| \geqslant 0, \tag{6.56}$$

bo to też jest poprawny stan separowalny. Zatem, jeśli

$$\rho^{T_A} < 0 \implies \rho \text{ jest splątany.} \tag{6.57}$$

Jeżeli $d_A = d_B = 2$ oraz $d_A = 2, d_B = 3$ to mamy w dwie strony. Zajmowali się tym Horoteccy. Jest to **kryterium splątania**. Jest zawsze poprawne dla dwóch qubitów oraz qubita i qutrita.

7 Klonowanie i teleportacja

W tym rozdziale poruszone zostaną dwie kwestie – **zakaz klonowania** (ang. *no-cloning theorem*) oraz **teleportacja kwantowa**. Zakaz klonowania jest jednym z powodów, przez które niemożliwy jest przesył informacji z prędkością szybszą od światła. Drugi podrozdział zawierać będzie opis protokołu kwantowego umożliwiającą teleportację stanu kwantowego (pod warunkiem że strony mogą komunikować się klasycznie i współdzielą stan splątany).

7.1 Zakaz klonowania

Weźmy dwa stany, $|\psi\rangle, |\phi\rangle$. Mamy

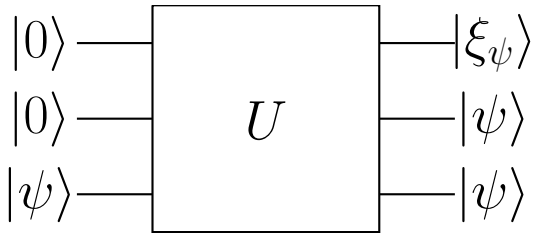
$$P_{err} = \frac{1}{2}\left(1 - \sqrt{1 - |\langle\psi|\phi\rangle|^2}\right). \tag{7.1}$$

A teraz weźmy większa ich liczbę $|\phi\rangle^{\otimes N}, |\phi\rangle^{\otimes N}$. Wtedy uniwersalna maszyna klonująca:

$$p_{err}^N = \frac{1}{2}\left(1 - \sqrt{1 - |\langle\psi|\phi\rangle|^{2N}}\right) \to^{N\to\infty} 0 \tag{7.2}$$

$$0 < |\langle\psi|\phi\rangle| < 1. \tag{7.3}$$

Można to przedstawić graficznie, w postaci krótkiego obwodu kwantowego.



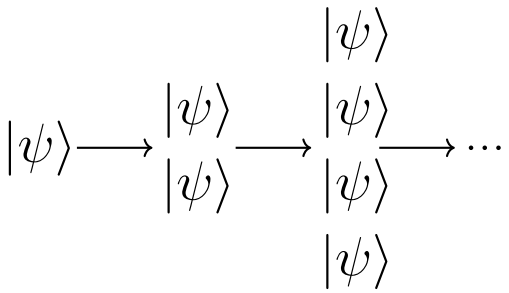
Rysunek 7.1: Uniwersalna maszyna (bramka) klonująca.

Mając taką maszynę można by klonować stan $|\psi\rangle$ w nieskończoność

$$|\psi\rangle \to^{U} |\psi\rangle \cdot |\psi\rangle, \tag{7.4}$$

$$|\psi\rangle \to^{U^N} |\psi\rangle^{\otimes N}, \tag{7.5}$$

uzyskując w efekcie dowolną ilość wejściowych stanów.



Rysunek 7.2: Gdyby istniała uniwersalna maszyna klonująca możliwym byłoby uzyskanie dowolnej liczby kopii wejściowego stanu.

A teraz omówimy protokół *Remote-State-Preparation* razem z omówioną wyżej maszyną klonującą. Weźmy stan Bella.

$$|\psi_-\rangle = U_A \otimes U_B |\psi_-\rangle = \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle) = \frac{1}{\sqrt{2}} (|+-\rangle - |-+\rangle) \tag{7.6}$$

Bob ma maszynę klonującą. Alicja mierzy

$$0 \to \{|0\rangle, |1\rangle\} \to \text{Bob ma } |1\rangle \text{ lub } |0\rangle. \tag{7.7}$$
$$1 \to \{|+\rangle, |-\rangle\} \to \text{Bob ma } |+\rangle \text{ lub } |-\rangle. \tag{7.8}$$

Ale Bob ma maszynę, więc klonuje. Skoro tak, to może wygenerować statystykę, a przez to może wykonkludować co Ala chciała przesłać, bo może wygenerować $N \to \infty$ kopii i wtedy błąd dąży do 0. Mamy zatem do czynienia z **komunikacją FTL** (ang. *faster than light*)!

Nauczyliśmy się wcześniej, że nie można rozróżniać nieortogonalnych stanów – tutaj widać, że jednak się da, więc gdzieś jest sprzeczność. Pokażemy, że uniwersalna maszyna klonująca nie istnieje, przez dowód *ad absurdum*.

Dowód

Nasze założenie jest następujące

$$|\psi\rangle \to |\psi\rangle^{\otimes N}. \tag{7.9}$$

Zakładamy, że istnieje uniwersalna maszyna klonująca. Skoro tak, to możemy

$$|0\rangle |0\rangle |\phi\rangle \to^U |\xi_\phi\rangle |\phi\rangle |\phi\rangle.$$
$$|0\rangle |0\rangle |\psi\rangle \to^U |\xi_\psi\rangle |\psi\rangle |\psi\rangle,$$

przy czym załóżmy, że stany $|\psi\rangle$ i $|\phi\rangle$ nie są ze sobą ortogonalne. Skoro jest to ewolucja unitarna, to musi być zachowana norma.

$$\left|\tilde{\psi}\right\rangle = U |\psi\rangle \tag{7.10}$$
$$\left\langle \tilde{\psi}_i \middle| \tilde{\psi}_j \right\rangle = \langle \psi_i | U^\dagger U |\psi_j\rangle \tag{7.11}$$
$$\langle \psi|\phi\rangle = \langle \xi_\psi|\xi_\phi\rangle \langle \psi|\phi\rangle \langle \psi|\phi\rangle \tag{7.12}$$

Ale myśmy założyli, że $|\psi\rangle$ nie jest ortogonalne z $|\phi\rangle$. Skoro tak, to

$$0 < |\langle \phi|\psi\rangle| < 1 \tag{7.13}$$

No i żądamy

$$\langle \psi|\phi\rangle (1 - \langle \psi|\phi\rangle \langle \xi_\phi|\xi_\psi\rangle) = 0, \tag{7.14}$$

ale

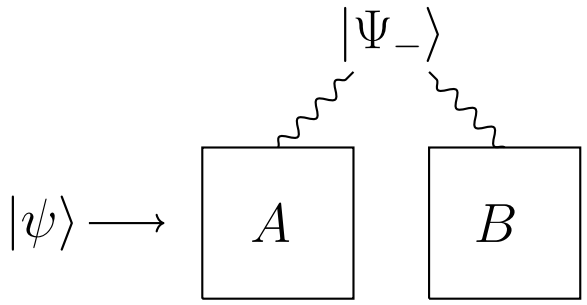
$$|1 - \langle \phi|\psi\rangle \langle \xi_\phi|\xi_\psi\rangle| < 1 \tag{7.15}$$

więc otrzymaliśmy sprzeczność.

□

7.2 Teleportacja kwantowa

Alicja i Bob dzielą stan $|\Psi_-\rangle$. Mamy dodatkowy qubit $|\psi\rangle$, niech ma go Ala, i pragniemy go przesłać do drugiego użytkownika. Jak to zrobić?



Rysunek 7.3: Ala i Bob współdzielą stan $|\Psi_-\rangle$. Ale chce przesłać bob

Można zacząć od takich operacji:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \tag{7.16}$$

$$|\psi\rangle_1\,|\Psi_-\rangle_{23} = ((a\,|0\rangle + \beta\,|1\rangle)) \otimes \left(\frac{1}{\sqrt{2}}\,|01\rangle - |10\rangle\right) = \frac{1}{\sqrt{2}}\,(\alpha(|001\rangle - |010\rangle) + \beta(|101\rangle - |110\rangle)) =$$

$$= \frac{1}{2}\,(-|\psi_-\rangle \otimes (\alpha\,|0\rangle + \beta\,|1\rangle) - |\psi_+\rangle \otimes (\alpha\,|0\rangle - \beta\,|1\rangle) + |\phi_-\rangle \otimes (\alpha\,|1\rangle + \beta\,|0\rangle) + |\phi_+\rangle \otimes (\alpha\,|1\rangle - \beta\,|0\rangle)) \tag{7.17}$$

Teraz, aby dokonać teleportacji stanu, Alicja wykonuje pomiar w bazie stanów Bella i otrzymuje wynik

$$|\Psi_+\rangle \to 1, \tag{7.18}$$
$$|\Psi_-\rangle \to 2, \tag{7.19}$$
$$|\Phi_+\rangle \to 3, \tag{7.20}$$
$$|\Phi_-\rangle \to 4. \tag{7.21}$$

Wynik komunikuje Bobowi drogą klasyczną. W zależności od wyniku Bob wykonuje na swoim qubicie operację

$$1 \to \mathbb{1}, \tag{7.22}$$
$$2 \to Z, \tag{7.23}$$
$$3 \to X, \tag{7.24}$$
$$4 \to iY. \tag{7.25}$$

i w ten sposób Bob będzie miał qubit w stanie $|\psi\rangle$. Stan warunkowy Boba po pomiarze Alicji można zapisać wprost:

$$\left|\tilde{\Psi}(i)\right\rangle_3 =_{12} \langle\Psi_i|\psi\rangle_1\,|\Psi_-\rangle_{23}\,. \tag{7.26}$$

Jest nieznormalizowany. Można sprawdzić, że

$$\left\langle\tilde{\psi}(i)\middle|\tilde{\psi}(i)\right\rangle = p(i). \tag{7.27}$$

Zatem

$$|\psi\rangle_3 = \sum_{i=1}^{n} U_i\left|\tilde{\psi}(i)\right\rangle_3 \tag{7.28}$$

Przy braku informacji o wyniku i, mamy

$$\rho_3 = \sum_{i=1}^{n} \left|\tilde{\psi}(i)\right\rangle \left\langle\tilde{\psi}(i)\right| =$$

$$= \frac{1}{4}(\alpha|0\rangle + \beta|1\rangle)(\alpha^*\langle 0| + \beta^*\langle 1|) + \frac{1}{4}(\alpha|0\rangle - \beta|1\rangle) + \frac{1}{4}(\alpha|1\rangle + \beta|0\rangle)(a^*\langle 1| + \beta^*\langle 0|) + \frac{1}{4}(\alpha|1\rangle - \beta|0\rangle)(\alpha^*\langle 1| - \beta^*\langle 0|) =$$

$$= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|), \tag{7.29}$$

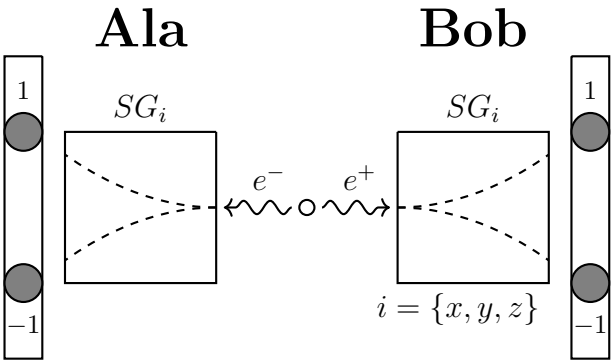
czyli stan maksymalnie zmieszany. Wniosek jest taki, że bez informacji od Alicji Bob nie wie w zasadzie nic, czyli nie ma tutaj mowy o przesyle informacji **FTL**.

8 Bell, GHZ, Hardy

Jedną z najbardziej charakterystycznych cech mechaniki kwantowej, jest to, że jest ona *nielokalna*. W rozdziale tym opisane zostanie co to w zasadzie oznacza i w jaki sposób owa nielokalność została dowiedziona.

8.1 Nierówności Bella

Motywacją ich powstania był paradoks EPR.



Rysunek 8.1: Schemat eksperymentu z elektronami i aparatami Sterna-Gerlacha.

Elektrony są tutaj w stanie

$$\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = |\psi_-\rangle \tag{8.1}$$

To sa rozważania jak przy Remote-State-Preparation, tylko z obserwablami. Mowa tutaj, cytując Einsteina, o *elements of physical reality.*

Alicja mierzy

$$\hat{\sigma}_Z^A\{|0\rangle, |1\rangle\} \implies |0\rangle_A \to |1\rangle_B, |1\rangle_A \to |0\rangle_B \tag{8.2}$$

Bob mierzy tak samo

$$\hat{\sigma}_Z^B \to -\frac{1}{2} \vee +\frac{1}{2} \tag{8.3}$$

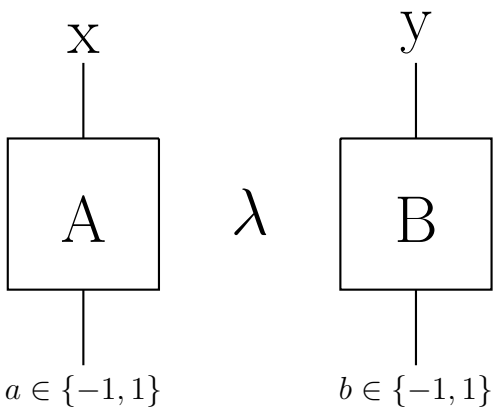
z pewnością. Natomiast w przypadku, gdyby Ala mierzyła w innej bazie

$$\hat{\sigma}_x^A \quad \{|+\rangle, |-\rangle\} \implies |+\rangle_A \to |-\rangle_B, |-\rangle_A \to |+\rangle_B \tag{8.4}$$

to Bob

$$\hat{\sigma}_Z^B \implies -\frac{1}{2} \vee \frac{1}{2}, \tag{8.5}$$

ale z prawdopodobieństwem 0.5 dla każdego. Tutaj Einstein miał pewien problem, bo przecież Bóg nie gra w kości (ang. *God doesn't play dice*). Próbowano to tłumaczyć przez lokalne zmienne ukryte (ang. *local hidden variables*).



Rysunek 8.2: Eksperyment opisany przy pomocy zmiennych ukrytych.

No i tutaj być może jest jeszcze jakaś dodatkowa informacja, której nie widzimy. Może elektrony są otoczone przez jakiś *eter* w stanie λ. Oznacza to, że rozkład prawdopodobieństwa dla EPR należy zapisać w postaci

$$p(ab|xy) = \sum_{\lambda} p(\lambda) p(ab|xy\lambda). \tag{8.6}$$

Załóżmy tutaj, że $x = y$, tj. że mierzyli w tej samej bazie. Wtedy prawdopodobieństwa są dane tabelą

a \ b	-1	1
- 1	0	1
1	1	0

a w drugim przypadku

a \ b	-1	1
- 1	0.5	0.5
1	0.5	0.5

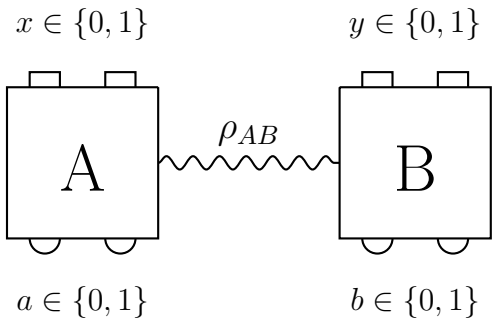
Jeśli tak by faktycznie było, to nie ma żadnego paradoksu EPR. Bell pokazał jednak warunek

$$\sum_{a,b,x,y} (-1)^{a+b+xy} p(ab|xy) \leqslant 2, \tag{8.7}$$

który obowiązuje w przypadku teorii zmiennych lokalnych. Jest to właśnie **nierówność Bella**. Dla kwantowej mechaniki jest ona łamana.

8.2 Nierówność CHSH

Będziemy teraz mówić o nierówności CHSH. Można to sobie wyobrażać ponownie jako sytuację z dwoma pudełkami. Mamy dwa różne pomiary i każdy z nich może mieć dwa wyniki.



Rysunek 8.3: Schemat eksperymentu CHSH.

Wyniki są w zbiorach

$$A = \{-1, 1\} \tag{8.8}$$
$$B = \{-1, 1\} \tag{8.9}$$

My zastanawiamy się nad prawdopodobieństwem

$$p(ab|xy) \equiv p(AB|xy) = Tr(\rho_{AB} \hat{M}_A^{(x)} \otimes \hat{M}_B^{(y)}) \tag{8.10}$$

Możemy sobie zdefiniować statystykę zmiennych A_x, B_y i ich korelatory. Oznaczmy $\langle A_x \rangle$ wynik A jeśli wybrałem bazę x.

$$\langle A_x \rangle = \sum_{A=\pm 1} A p(A|x) \tag{8.11}$$

i podobnie dla reszty. Mam

$$\langle A_x B_y \rangle = \sum_{A,B=\pm 1} AB p(AB|xy) \tag{8.12}$$

Teraz przejdźmy do $CHSH$. Będziemy rozpatrywać wartość oczekiwaną zmiennej losowej β zdefiniowaną tak

$$\langle \beta \rangle = \langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle \tag{8.13}$$

Próbując to tłumaczyć zmiennymi ukrytymi dochodzimy do wniosku, że

$$p(AB|xy) = \sum_{\lambda} p(\lambda) p(A|x\lambda) p(B|y\lambda). \tag{8.14}$$

Można pokazać, że

$$|\langle \beta \rangle| \leqslant 2 \tag{8.15}$$

W przypadku kwantowo-mechanicznych rozważań dojdziemy jednak do wniosku, że

$$|\langle \beta \rangle| = 2\sqrt{2}. \tag{8.16}$$

Mechanika kwantowa nie jest teorią spełniającą lokalny realizm.

Uwaga! Nie prowadzi to do przesyłu informacji **FTL**.

W okolicy podobnych rozważań można spotkać się z terminem *non-signaling condition*. Formalnie

$$\sum_{A=\pm 1} p(AB|xy) \equiv_{NS} \; p(B|y) \tag{8.17}$$

W przypadku mechaniki kwantowej jest to zawsze spełnione.

$$\sum_{A=\pm 1} p(AB|xy) = \sum_{A=\pm 1} Tr\left(\rho_{AB}\hat{M}_A^{(x)} \otimes \hat{M}_B^{(y)}\right) = Tr\left(\rho_{AB}(\mathbb{1} \otimes \hat{M}_B^{(y)})\right) \tag{8.18}$$

Spójrzmy jeszcze raz na korelator

$$\langle A_x B_y \rangle = \sum_{A,B=\pm 1} AB\; Tr\left(\rho_{AB}\hat{M}_A^{(x)} \otimes \hat{M}_B^{(y)}\right) = Tr\left(\rho_{AB}\left(\sum_A A\hat{M}_A^{(x)}\right) \otimes \left(\sum_B B\hat{M}_B^{(y)}\right)\right) \tag{8.19}$$

Alicja rzutowała na stan czysty. Zdefiniujmy

$$\sum_A A\hat{M}_A^{(x)} = \hat{\sigma}_{\vec{a}}^{(x)} \tag{8.20}$$

$$\sum_B B\hat{M}_B^{(y)} = \hat{\sigma}_{\vec{b}}^{(y)} \tag{8.21}$$

$$\hat{\sigma}_{\vec{a}} = (+1)\,|\vec{a}\rangle\langle\vec{a}| + (-1)\,|-\vec{a}\rangle\langle-\vec{a}| = \frac{1}{2}(\mathbb{1} + \vec{a}\hat{\vec{\sigma}}) - \frac{1}{2}\left(\mathbb{1} - \vec{a}\hat{\vec{\sigma}}\right) = \vec{a}\hat{\vec{\sigma}} \tag{8.22}$$

$$\hat{\sigma}_{\vec{a}'} = \vec{a}'\hat{\vec{\sigma}} \tag{8.23}$$

No i teraz CHSH

$$\langle\hat{\beta}\rangle = Tr\left(\rho_{AB}\hat{\beta}(\vec{a},\vec{a}',\vec{b},\vec{b}')\right) \tag{8.24}$$

$$\hat{B} = \hat{\sigma}_{\vec{a}} \otimes \hat{\sigma}_{\vec{b}} + \hat{\sigma}_{\vec{a}} \otimes \hat{\sigma}_{\vec{b}'} + \hat{\sigma}_{\vec{a}'} + \otimes\hat{\sigma}_{\vec{b}} - \hat{\sigma}_{\vec{a}'} \otimes \hat{\sigma}_{\vec{b}'} \tag{8.25}$$

A co jeśli dostaniemy stan mieszany?

$$\rho' = \lambda\rho + (1-\lambda)\sigma \tag{8.26}$$

Wtedy operator Bella

$$\langle\beta\rangle_{\rho'} = Tr(\rho', \hat{\beta}) = \lambda Tr(\rho\hat{\beta}) + (1-\lambda)Tr(\sigma\hat{\beta}) = \lambda\langle\hat{\beta}\rangle_\rho + (1-\lambda)\langle\hat{\beta}\rangle_\sigma \tag{8.27}$$

Żeby móc legalnie mówić o łamaniu lokalnego realizmu w tym przypadku, należy jeszcze pokazać, że istnieją *dobre pomiary* (takie, które pozwalają to pokazać) u Ali i Boba.

8.3 Wykres *CHSH* w 2D

Zdefiniujmy

$$\langle\hat{\beta}\rangle = \langle A_{11}\rangle + \langle A_{01}\rangle + \langle A_{10}\rangle - \langle A_{00}\rangle. \tag{8.28}$$

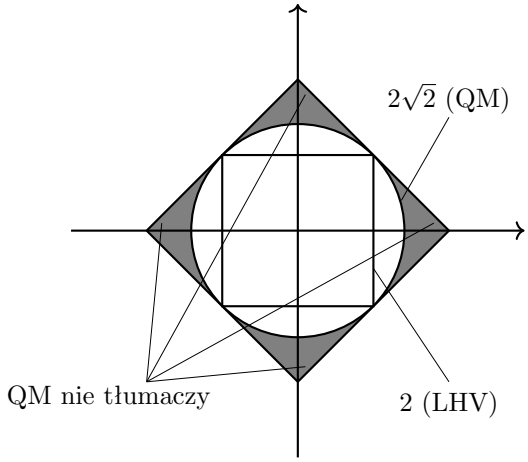
LHV (teoria lokalnych zmiennych ukrytych, ang. *local hidden variable*) mówi, że

$$\langle\beta\rangle \leqslant 2 \implies \langle\beta'\rangle \leqslant 2 \tag{8.29}$$

Kwantowo-mechanicznie można zapisać

$$\langle\beta\rangle \leqslant 2\sqrt{2} \implies \langle\beta'\rangle \leqslant 2\sqrt{2}. \tag{8.30}$$

Prawdopodobieństwa układają się wtedy jak na poniższym rysunku.



Rysunek 8.4: Graficzna reprezentacja CHSH.

Warunek *no-signalingu* (NS) jest następujący

$$\sum_{\lambda} p(AB|xy) = p(B|y) \implies |\langle\beta\rangle| \leqslant 4, \tag{8.31}$$

co pokazwali Popeson i Rohrliol ($PR\text{-}Box$). Pokazuje to, że QM nie jest jedyną teorią, która spełnia fizyczne założenia NS.

Rozpatrzmy teraz **stan Wernera**

$$\rho_p = p\,|\psi_-\rangle\,\langle\psi_-| + (1-p)\frac{\mathbb{1}}{4} \tag{8.32}$$

$$p > \frac{1}{3} \implies \text{splątanie} \tag{8.33}$$

Pytanie – czy wszystkie stany splątane łamią nierówności Bella? Czy stan separowalny może ją złamać? Łatwo można pokazać, że dla separowalnych nie jest to możliwe, bo

$$\rho_{AB} = p_i \rho_i^A \otimes \rho_i^B \tag{8.34}$$

$$p(ab|xy) = \mathrm{Tr}\left(\rho_{AB}\hat{M}_a^{(x)} \otimes \hat{M}_b^{(y)}\right) = \sum_i p_i \,\mathrm{Tr}\left(\rho_i^A \hat{M}_a^{(x)}\right) \mathrm{Tr}\left(\sigma_i^B \hat{M}_b^{(y)}\right) \tag{8.35}$$

czyli dostajemy to samo, co w LHV. Chcemy teraz policzyć $CHSH$ dla ρ_p.

$$\langle\hat{\beta}\rangle_{\rho_p} = p\langle\hat{\beta}\rangle_{\psi_+} + (1-p)\langle\hat{\beta}\rangle_{\frac{1}{4}} \tag{8.36}$$

Zauważmy jednak, że

$$\langle\hat{\beta}\rangle = \frac{1}{4}\,\mathrm{Tr}(\sigma_a \otimes \sigma_b + ...) = 0, \tag{8.37}$$

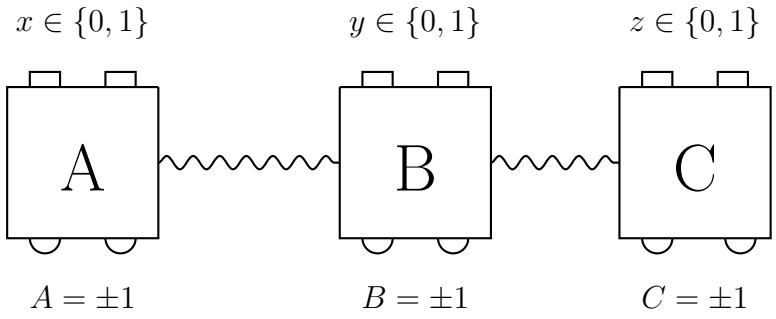
bo macierze Pauliego są bezśladowe. Zatem

$$\langle\hat{\beta}\rangle_{\rho_p} = p\langle\hat{\beta}\rangle_{\psi_-} = 2\sqrt{2}p. \tag{8.38}$$

Oznacza to, że dla $p = \frac{1}{\sqrt{2}}$ łamany jest lokalny realizm. Dodatkowo, zauważmy, że $\frac{1}{\sqrt{2}} > \frac{1}{3}$, zatem istnieją stany splątane, które nie łamią nierówności Bella.

□

Można pokazać, że QM jest nielokalna bez posługiwania się nierównościami.



Rysunek 8.5: Schemat omawianego eksperymentu.

Rozpatrzmy stan GHZ

$$|\psi_{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \tag{8.39}$$

Jest to stan maksymalnie splątany ze względu na dowolne przecięcie. Można go uogólnić na

$$|\psi^n_{GHZ}\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes n} + |1\rangle^{\otimes n}). \tag{8.40}$$

Wróćmy do trzech qubitów. Weźmy

$$(x, y, z) \in \{0, 1\} \implies \{\sigma_x \to 0, \sigma_y \to 1\} \tag{8.41}$$

i chcemy się dowiedzieć jaki będzie wynik równania

$$A_0B_0C_0 = \begin{cases} +1, +1, +1, p(111|000) = \mathrm{Tr}(\psi_{GHZ} |+\rangle \langle +| \otimes |+\rangle \langle +| \otimes |+\rangle \langle +|) \\ (-1), +1, +1... \\ \vdots \end{cases} = 1, \tag{8.42}$$

dla pojedynczego eksperymentu. Widać, że

$$\langle A_0B_0C_0\rangle = \mathrm{Tr}(|\psi_{GHZ}\rangle \langle\psi_{GHZ}| \sigma_x \otimes \sigma_x \otimes \sigma_x) = 1 \tag{8.43}$$

Można tak samo pokazać, że

$$A_0B_1C_1 = -1 \tag{8.44}$$
$$A_1B_1C_0 = -1 \tag{8.45}$$
$$A_1B_0C_1 = -1 \tag{8.46}$$

Czyli jeżeli jeden mierzy σ_z to wyniki są -1. Wróćmy teraz do zmiennych ukrytych

$$p(ABC|xyz) = \sum_\lambda p(\lambda)p(A|x\lambda)p(B|y\lambda)p(C|z\lambda) \tag{8.47}$$

λ jest dane, co znaczy, że $A_{0/1}$, $B_{0/1}$ i $C_{0/1}$ mają dane wartości. Wtedy

$$(A_0B_1C_1)(A_1B_1C_0)(A_1B_0C_1)(A_0B_0C_0) = A_0^2A_1^2B_0^2B_1^2C_0^2C_1^2 = 1, \tag{8.48}$$

ale wyżej pokazaliśmy, że mamy -1 trzy razy i jedną jedynkę. Oznacza to, że LHV nie jest poprawna.

9 Teoria informacji Shannona

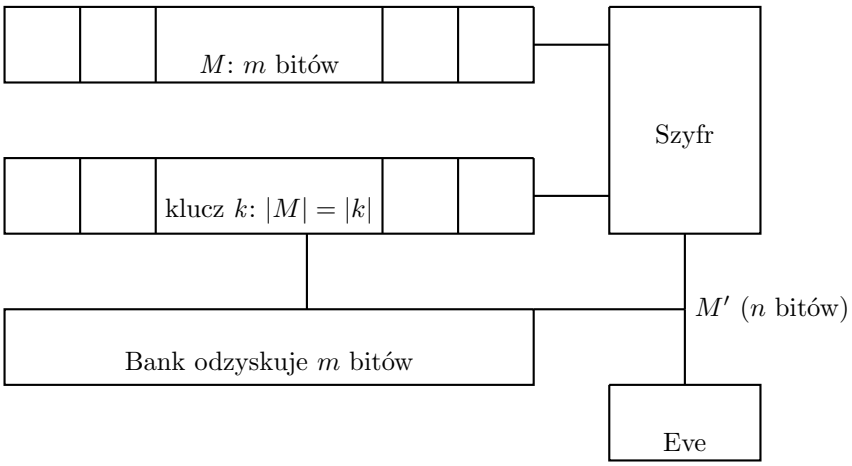
Klasyczna teoria informacji (Shannona) dobrze jest przedstawiona w pozycji *Elements of Information Theory* Thomasa Covera.

(Klasyczna) Teoria Informacji Shannona

Za ojca klasycznej teorii informacji uznaje się Clauda Shannona. Kiedy pracował w laboratoriach Bella współpracował z Alanem Turingiem. W 1948 stworzył dzieło *A mathematical theory of computation*. Stworzył podwaliny matematyczne teorii informacji. Najważniejsze rzeczy, którymi Shannon się zajmował:

1. Kompresja danych ([Shannon] Coding Theorem)
2. Komunikacja ([Shannon] Channel Capacity)
3. Kryptografia ([Shannon] Secret Channel Capacity)

To Shannon wprowadził koncepcję klucza jednorazowego użytku. Pokazał też, że aby wiadomość była zaszyfrowana w sposób nierozszyfrowywalny, to klucz musi być takiej samej długości co wiadomość.



Rysunek 9.1: Schemat przesyłania danych z podsłuchem.

Omówimy teraz kompresję danych. Weźmy alfabet X, który przyjmuje proste wiadomości

$$X = \{A, B, C, ...\}. \tag{9.1}$$

Wiadomością nazywamy sekwencję

$$M = X^m = X_0 X_1 ... \tag{9.2}$$

My tę wiadomość chcemy skompresować.

$$X \rightarrow^{Kodowanie}_{Encoding} E(X^m) = X^n \rightarrow^{Dekodowanie}_{Decoding} X^m \tag{9.3}$$

W teorii Shannona będziemy opisywać wiadomość jakimś rozkładem prawdopodobieństwa. Załóżmy, że każdy z symboli X dany jest rozkładem prawdopodobieństwa $p(X = x)$ i symbole w ciągu są **niezależne** (*IID* w literaturze – *independent and identically distributed*).

Przykład

Mamy alfabet $X = \{a, b, c, d\}$. Wysyłam takie wiadomości, że prawdopodobieństwa są dane liczbami

$$p(a) = \frac{1}{2}, \quad p(b) = \frac{1}{4}, \quad p(c) = p(d) = \frac{1}{8}.$$

Ile bitów muszę wykorzystać, żeby tę wiadomość wysłać? Jest kilka podejść

1. **Podejście proste**. Mamy 4 symbole, to przypiszmy im $a = 0_b, b = 1_b, c = 2_b, d = 3_b$. Pomijamy tutaj prawdopodobieństwo. Interesuje nas tylko liczba symboli. Mamy

$$\frac{n}{m} = 2, \tag{9.4}$$

   gdzie n to liczba bitów użytych do kodowania wiadomości złożonej z m symboli.Na każdy symbol potrzebuję zatem *średnio* 2 bity.

2. **Kodowanie sprytniejsze.**

$$a = 0 = 0_b \tag{9.5}$$
$$b = 10 = 2_b \tag{9.6}$$
$$c = 110 \tag{9.7}$$
$$d = 111 = 7_b \tag{9.8}$$

   Są tak dobrane, żeby kodowanie było jednoznaczne. Jaka jest średnia $\frac{n}{m}$ teraz? Można policzyć, że mniejsza, niż w przypadku wcześniejszym.

$$\left\langle \frac{n}{m} \right\rangle = \frac{1}{2} + 2\frac{1}{4} + 3\frac{2}{8} = \frac{1}{2} + \frac{1}{2} + \frac{6}{8} = \frac{14}{8} < 2 \tag{9.9}$$

   , czyli kodowanie jest lepsze.

Należy zauważyć, że teoria Shannona jest teorią asymptotyczną. Będziemy rozważać tylko granice $lim_{m\to\infty}$.... Shanon zadał pytanie – jaki jest optymalny, asymptotyczny stosunek ilości bitów do ilości symboli (to jest właśnie *Coding Theorem*). Okazuje się, że

$$lim_{m\to\infty}\left(\frac{n(m)}{m}\right) = H(x), \tag{9.10}$$

gdzie $H(x)$ jest tak zwaną entropią Shannona i określa na ile losowy jest rozkład prawdopodobieństwa X.

$$H(x) = -\sum_x p(x) log_2(p(x)). \tag{9.11}$$

Uwaga notacyjna

$$log \equiv log_1 0 \tag{9.12}$$
$$lg \equiv log_2 \tag{9.13}$$
$$ln \equiv log_e \tag{9.14}$$

Powiedzmy, że X ma d wyników. Gdy $d = 1$, to $H(x) = 0$, co można pokazać pamiętając, że $0log(0) = 0$. Maksymalny jest dla rozkładu płaskiego, wtedy $H(X) = lg(d)$. Zatem mamy granice

$$0 \leqslant H(x) \leqslant lg(d). \tag{9.15}$$

Przykład

Policzmy entropię dla wcześniej omawianego przykładu.

$$H(x) = \left(\frac{1}{2}lg\left(\frac{1}{2}\right) + \frac{1}{4}lg\left(\frac{1}{4}\right) + \frac{1}{8}lg\left(\frac{1}{8}\right) + \frac{1}{8}lg\left(\frac{1}{8}\right)\right) = \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \frac{3}{8} = 1.75 = \frac{14}{8} \tag{9.16}$$

□

Intuicja dowodu

Zastanówmy się nad ciągami X^m, kiedy $m \to \infty$. Mogę powiedzieć, że dla typowych ciągów $x \in X$ występuje $mp(x)$ razy średnio. Jest to konsekwencja prawa wielkich liczb.

$$\sum_{i=1}^{m} \frac{x_i}{m} \to E(x) \tag{9.17}$$

Prawdopodobieństwo ciągu typowego

$$p_{typ}^m = p(x_0x_1...x_m \in T) = \left(\prod_{x=1}^{d} p(x)\right)^{mp(x)} = 2^{lg\left(\prod_{x=1}^{d}\right)p(x)^{mp(x)}} = 2^{m\sum_{x=1}^{d} p(x)lg(p(x))} = 2^{-mH(x)}, \tag{9.18}$$

gdzie T jest zbiorem typowych ciągów, a d jest mocą alfabetu.

Indeksowanie typowych ciągów wymaga zatem $mH(x)$ bitów. Do tego jeszcze wrócimy na ćwiczeniach, bo się trochę poplątało.

□

Jeżeli chodzi o kompresję, to można tutaj znaleźć przepis na kodowanie przy pomocy tzw. **kodowania Huffmana**. Wiemy, że entropia Shannona jest dana wzorem

$$H(x) = -\sum_{x=1}^{1} p(x)lg(p(x)). \tag{9.19}$$

Powiedzmy jednak, że mogę wysyłać dwie skorelowane wiadomości, na dwóch oddzielnych alfabetach. Możemy zdefiniować entropię łączną.

$$X, Y \; p(x, y) \tag{9.20}$$

$$H(X, Y) = -\sum_{x,y=1}^{d_x, d_y} p(x, y)lg(p(x, y)) \tag{9.21}$$

Gdy są niezależne, to

$$p(x, y) = p(x)p(y) \implies H(X, Y) = -\sum_{x,y} p(x)p(y)\left(lg(p(x)) + lg(p(y))\right) = H(x) + H(y), \tag{9.22}$$

czyli ich entropie Shannona się dodają. W przypadku prawdopodobieństw warunkowych

$$p(X|Y = y), H(X|Y = y) = -\sum_{x=1}^{d_x} p(x|y)lg(p(x|y)) \tag{9.23}$$

$$H(X|Y) = \sum_{y=1}^{d_y} p(y)H(X|Y = y) = -\sum_{x,y} p(x, y)lg\left(\frac{p(x, y)}{p(y)}\right) = \tag{9.24}$$

$$= -\sum_{x,y} p(x, y)lg(p(x, y)) + \sum_{x,y} lg(p(x, y))lg(p(y)) = -H(X, Y) - H(Y) \tag{9.25}$$

Można pokazać, że korelacje między X i Y zmniejszają nieporządek, co znaczy, że

$$H(X,Y) \leqslant H(X) + H(Y) \tag{9.26}$$

Co więcej zachodzą także

$$H(X,Y) = H(X|Y) + H(Y) \tag{9.27}$$
$$H(X|Y) \leqslant H(X), \tag{9.28}$$

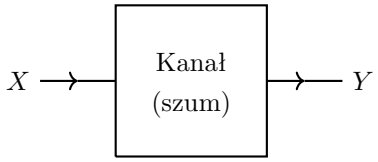
przy czym ostatnie jest oczywiste, bo warunek coś nam przecież mówi. Oczywistym jest też, że

$$H(X,Y|Z) = H(X|Y,Z) + H(Y|Z). \tag{9.29}$$

□

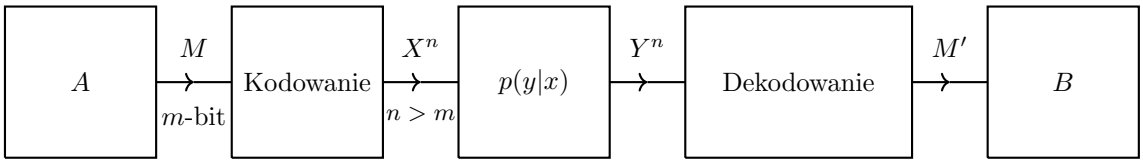
Komunikacja

Będziemy tutaj rozważać sytuację, w której mamy następujący kanał.



Rysunek 9.2: Rozważany kanał kwantowy.

Rozważmy następujący scenariusz. Niech M będzie ciągiem o m-bitach. Cały proces przebiega następująco.



Rysunek 9.3: Rozważany proces.

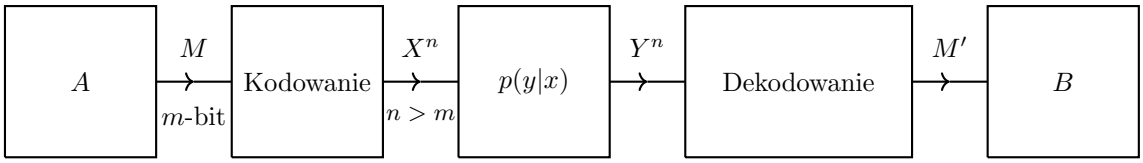
Przy czym Bob chce, aby $M = M'$.

Przykład

Korekcja błędów jest jednym z takich przykładów. Mamy dysk CD, który ma 700 MB, co określa nam liczbę bitów logicznych. Ratio tutaj jest jednak 192 do 588, więc bitów fizycznych mamy liczbę określoną przez ok. 2.1 GB.

10 Bezpieczna pojemność kanałów kwantowych i korekcja błędów

Kontynuujemy klasyczną teorię informacji, ale bardziej pod kątem kryptografii i bezpieczeństwa. Załóżmy, że mamy wiadomość M. Celem kryptografii jest takie zakodowanie, przesłanie i zdekodowanie wiadomości M, aby zdekodowana na końcu wiadomość M' była równa tej wejściowej.



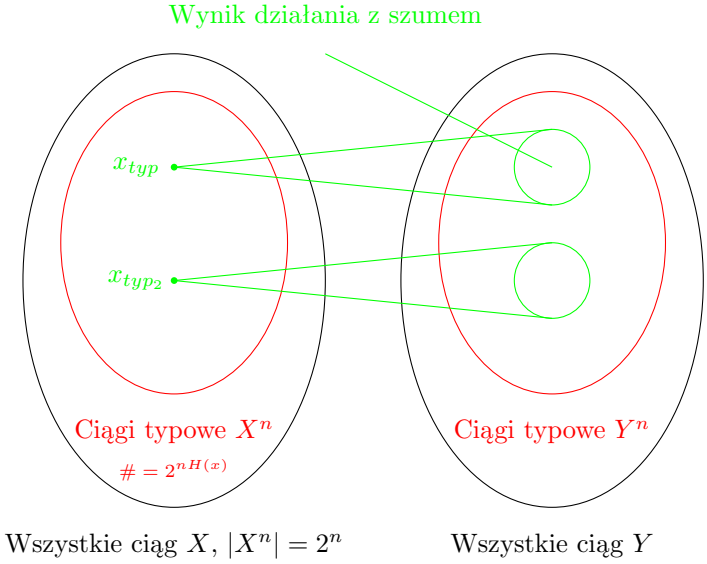
Rysunek 10.1: Schemat kryptografii.

Rozważamy teraz bitrate, czyli stosunek $\frac{m(n)}{n}$ ale w granicy $m \to \infty$. Twierdzenie Shannona mówi, że

$$\lim_{m\to\infty} \frac{m(n)}{n} = R \leqslant C_{p(y|x)} = max_{p(x)}\{I(X:Y)\}, \tag{10.1}$$

$$I(X:Y) = I(Y:X) = H(Y) - H(X|Y) \tag{10.2}$$

gdzie C jest przepustowością kanału, a $I(X:Y)$ to informacja wzajemna. Jak tego dowieść? Intuicja dowodu jest następująca. Załóżmy, że mamy zbiór wszystkich ciągów.



Rysunek 10.2: Zbiory ciągów i ciągi typowe.

Nas interesować będą tylko ciągi typowe. Ciąg typowy, zostanie zmapowany na wiele ciągów (przez szum) z Y^n. Weźmy dany $X = x$. Wtedy, zgodnie z twierdzeniem, mapujemy to na $2^{H(Y|X=x)}$, przy czym zakładamy tutaj, że $n = 1$ dla uproszczenia. Umiemy zatem określić na co może się zmienić jeden bit. Zastanówmy się teraz ile wynosi liczebność $X = x$ w naszym m-elementowym ciągu. Skoro ciąg jest typowy, to powinno być $\simeq np(X = x)$. Możemy zatem zapisać, że

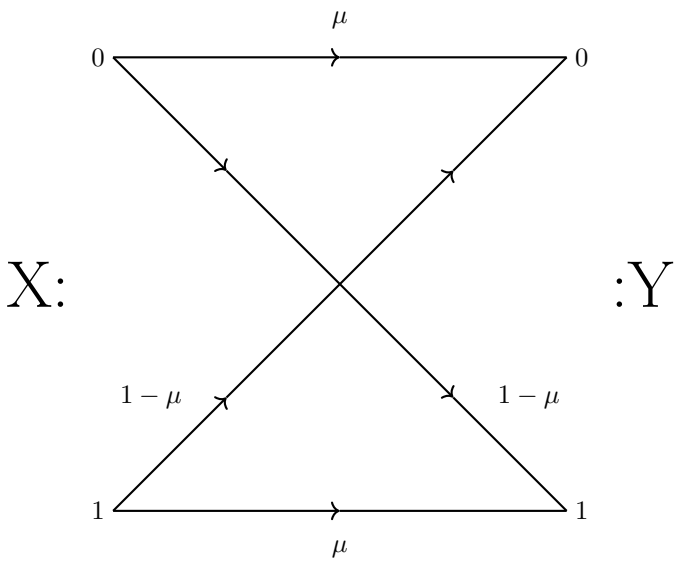
$$\prod_{X\in\{0,1\}} \left(2^{H(Y|X=x)}\right)^{np(x)} = 2^{nH(Y|X)}. \tag{10.3}$$

Muszę zatem tak dobierać typowe x, aby móc je rozróżniać na Y. Liczba sekwencji bitów na wejściu, które mogę jednoznacznie rozpoznać jest zatem równa

$$\frac{card(\text{ciągów typowych na wyjściu})}{card(\text{ciągów na wyjściu dla danego } x_{typ})} = \frac{2^{nH(y)}}{2^{nH(Y|X)}} = 2^{n((H(Y)-H(Y|X)))} = 2^{nI(X:Y)} \tag{10.4}$$

To kończy intuicję. W formalnych dowodach pokazuje się jeszcze, że błąd jest zaniedbywalny.

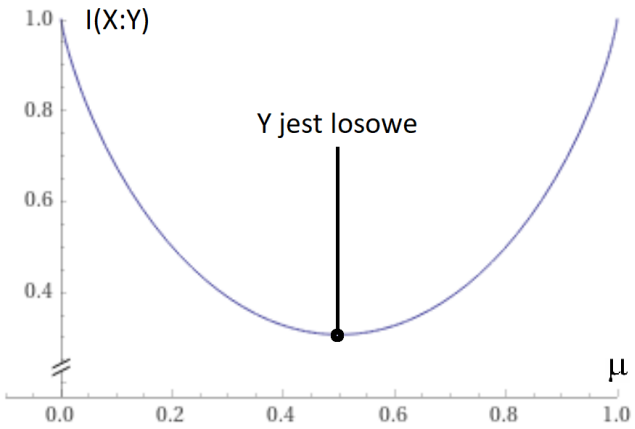
□



Rysunek 10.3: Kodowanie bitów.

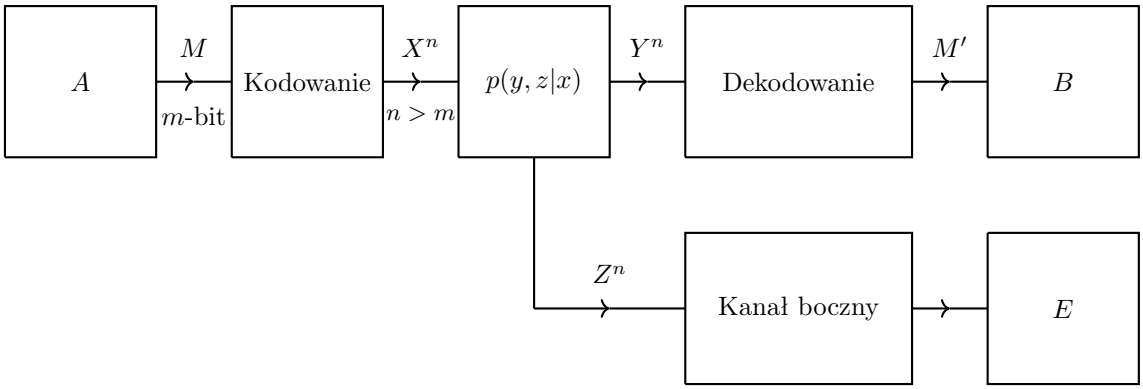
Optymalnie $p(X = 0) = p(X = 1) = \frac{1}{2}$.

$$I(X : Y) = 1 - h(\mu)$$
$$h(\mu) = -\mu lg(\mu) - (1 - \mu)lg(1 - \mu)$$



Rysunek 10.4: Wykres zależności informacji wzajemnej $I(X : Y)$ od parametru μ.

Rozważmy teraz bezpieczną pojemność kanału kryptograficznego. Uogólnijmy schemat na taki, który zawiera podsłuchiwacza.



Rysunek 10.5: Schemat kanału z podsłuchem. Cel to $M = M'$.

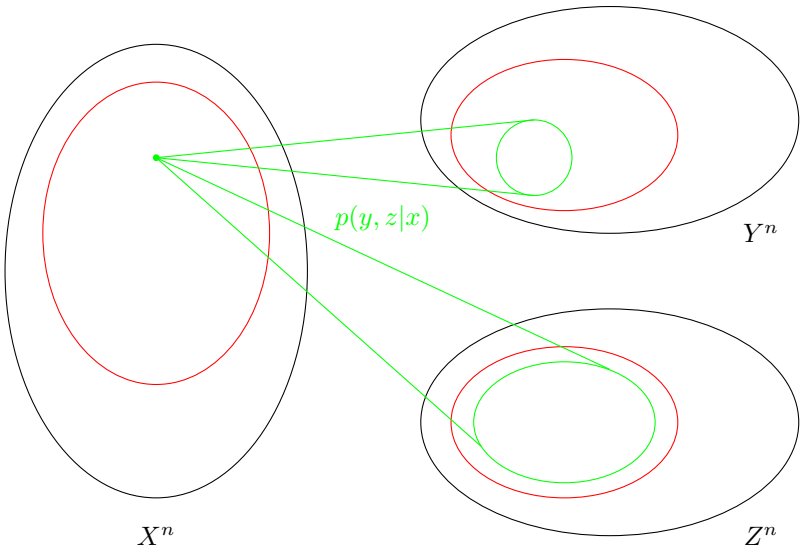
Z definicji

$$p(y, z|x) = p(y|x)p(z) \tag{10.5}$$

Jaki tutaj mamy bitrate?

$$R_S = lim_{n\to\infty}\frac{n_s}{n} \leqslant C_S = max_{p(x)}\,(I(X:Y) - I(X:Z)). \tag{10.6}$$

Jest to możliwe tylko jeśli istnieje takie kodowanie, dla kanału $p(y, z|x)$ takie, że $I(X:Y) > I(X:Z)$. Ponownie przestawiona będzie intuicja dowodu.

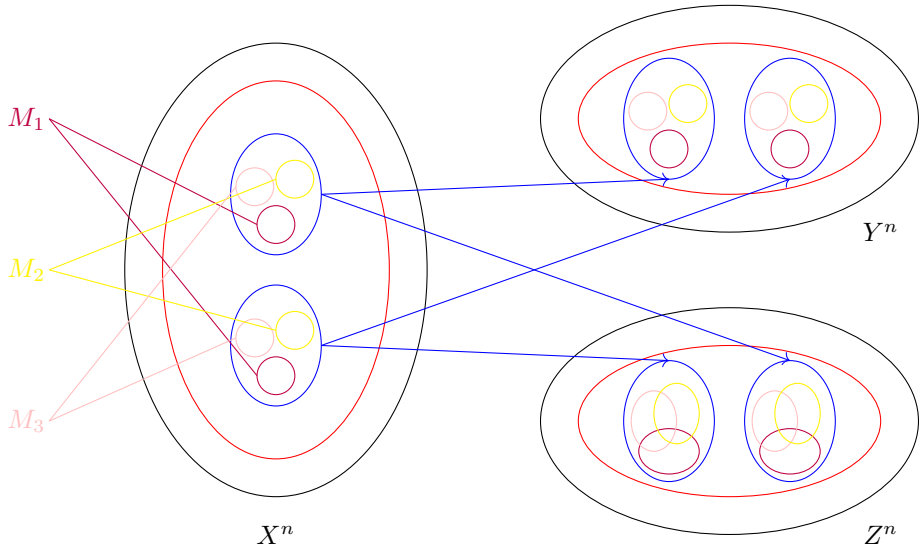


Rysunek 10.6: Transformacja ciągów przy dla schematu z podsłuchem. Niepewność w Z jest wyraźnie większa niż w Y.

Widać, że wynik u Ewy jest znacznie większy niż, u Boba, co oznacza, że

$$I(X:Y) > I(X:Z). \tag{10.7}$$

Szum jest większy u Ewy. Załóżmy teraz, że wchodzą wiadomości.



Rysunek 10.7: Ciągi M_i w podzbiorze Z przecinają się, co nie pozwala odszyfrować informacji.

W ramach gruboziarnistej struktury nadal jestem w stanie rozróżniać struktury drobnoziarniste. U Ewy mogę rozróżnić grubą, ale drobnoziarnistych wewnątrz już nie. Gruboziarnistością tutaj jest $I(X : Z)$. Gruboziarnista struktura może być postrzegana jako klucz publiczny. $I(X : Y)$ jest strukturą drobnoziarnistą. Jeżeli da się zbudować takie struktury, to mamy bezpieczną wymianę informacji między A i B.

Spójrzmy na to z kryptograficznego punktu widzenia. Kluczowe będą tutaj dwa pojęcia – **korekcja błędów** (EC – *error correction*) i **wzmocnienia prywatności** (PA). Co się będzie działo? Mamy Alicję, Boba i Ewę.

$lp.$	1	2	3	4	5	6	7
X	1	0	0	1	0	0	1
Y	1	0	1	1	1	0	1
Z	0	0	1	0	0	1	0

W pierwszej fazie (komunikując się w jedną stronę) Alicja i Bob chcą poprawić swoje błędy. Kompresują swoje ciągi w jakiś sposób, używając do tego funkcji haszujących, w taki sposób, żeby E już nic nie wiedziała. Podchodząc do tego kryptograficznie wiemy, że ujawnianie struktur gruboziarnistych jest przesyłem informacji $A \to B$.

Przykład

Weźmy strategię taką, że $X \to M$. Jeżeli $X = 0$, to zapisuję $X = 000$. Czyli jeden bit zapisuję trzema bitami (pod warunkiem, że $p_{error}(B) << \frac{1}{3}$ i $p_{error}(E) >> \frac{1}{3}$. Wtedy

$$X = 1001 \to 111\ 000\ 000\ 111$$
$$Y \to 101\ 001\ 000\ 011$$
$$Z \to 001\ 101\ 100\ 110$$

Trywialna korekcja błędu polega na wybraniu wartości o największej liczebności z każdej z trójek. Wtedy mamy

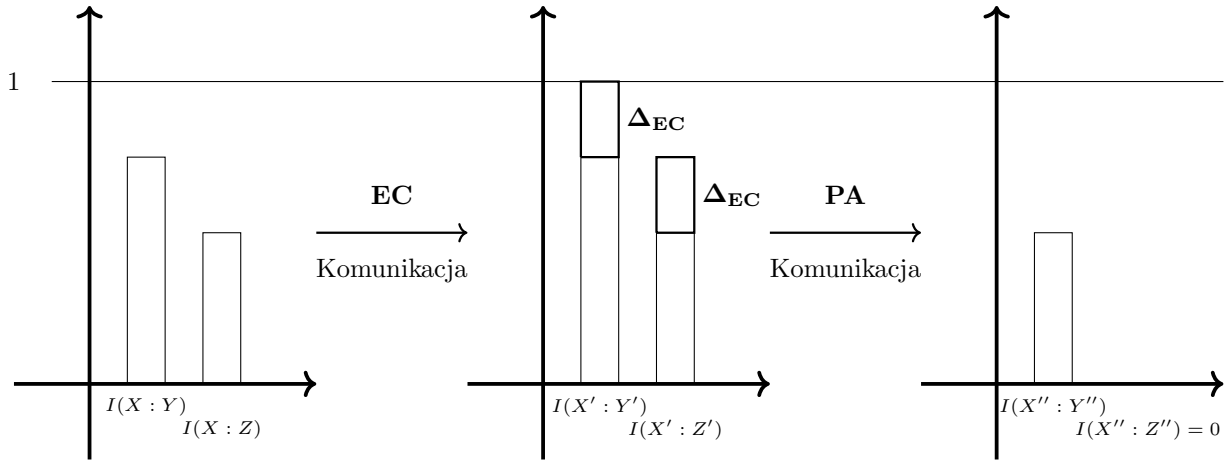
$$Y \to 1001$$
$$Z = 0101 \neq 1001$$

więc działa. □

Przykład 2

Teraz przykład na PA. Możemy dokonać losowej (znanej A, B i E permutacji), przez co zabijają korelacje. □

Możemy się zastanowić nad informacją per bit (bitrate).



Rysunek 10.8: Zmiany informacji wzajemnej w każdym kroku.

Czyli ostatecznie

$$I(X'':Y'') = I(X':Y') - I(X':Z') = I(X:Y) + \Delta_{EC} - (I(X:Z) + \Delta_{EC}) = I(X:Y) - I(X:Z), \qquad (10.8)$$

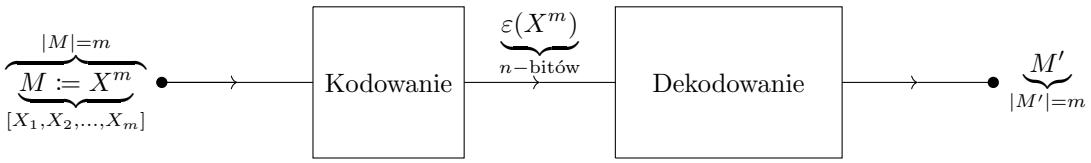
to samo z czego wyszliśmy.

11 Kwantowa teoria informacji

W tym rozdziale omówione zostaną podstawy kwantowej teorii informacji. Rozpoczniemy od omówienia zagadnienia kompresji. Później omówione zostaną różne rodzaje entropii. Na koniec omówione zostanie twierdzenie Holevo oraz kody dostępu swobodngo (*random access codes*, *RACs*).

11.1 Kwantowa kompresja

Pamiętając o twierdzeniu o kodowaniu Shannona możemy rozpocząć od schematu.



Rysunek 11.1: Klasyczny schemat kodowania.

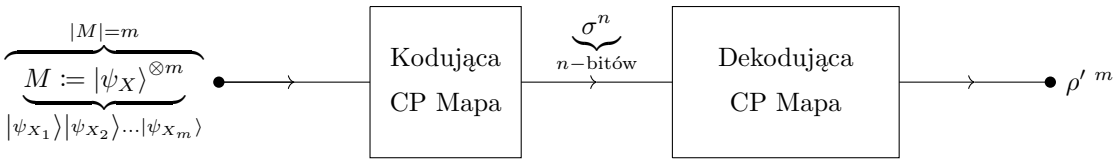
Należy zauważyć, że powyższy schemat dotyczy każdego $X = \{1, ..., d\}$. Oczekujemy, że $M = M'$. Wierna kompresja wymaga, żeby

$$\lim_{m\to\infty} \frac{n(m)}{m} \leqslant H(x), \tag{11.1}$$

przy czym równość zachodzi dla kodów Huffmana (asymptotyczny stopień kompresji: R_∞).

Pytanie

Co jeśli pozwolimy na kodowanie (klasycznej) informacji w kwantowych stanach (czystych)? X określa wtedy krotkę $\{p(X), |\Psi_X\rangle\}$. Mamy sytuację zaprezentowaną na rysunku 11.2.



Rysunek 11.2: Kodowanie kwantowe.

Chcemy, aby $\rho'^{\,m}$ reprezentowało każdą przetransmitowaną wiadomość M:

$$|\psi_{x_1}\rangle, ..., |\psi_{x_m}\rangle = \left|\vec{\psi}_x\right\rangle \rightarrow^{\text{zachodzące z}}_{\text{prawdopodobieństwem}} p(\vec{x}) = p(x_1)p(x_2)...p(x_m) \tag{11.2}$$

Wciąż jednak pozostajemy w reżimie IID, więc średnio każdy symbol ma postać

$$\rho = \sum_{x=1}^{d} p(x) |\psi_X\rangle \langle\psi_x| \tag{11.3}$$

Cała wiadomość ma zatem postać

$$\rho^{\otimes m} = \sum_{x_1,...,x_m=1}^{d} p(\vec{x}) |\psi_{\vec{x}}\rangle \langle\psi_{\vec{x}}| . \tag{11.4}$$

Oczekujemy

$$F(\rho^{\otimes m}, \rho^{lm}) = 1. \tag{11.5}$$

Rozpatrzmy to bardziej ogólnie

$$F(\rho,\sigma) = Tr^2\left(\sqrt{\rho}\sigma\sqrt{\rho}\right) = \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1 \tag{11.6}$$

$$F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2 \tag{11.7}$$

$$F(|\phi\rangle, \rho) = \langle\phi|\rho|\phi\rangle \tag{11.8}$$

Wierna kompresja

$$|\psi_{\vec{x}}\rangle\langle\psi_{\vec{x}}| \to \rho_{\vec{x}} = D\left(E\left(|\psi_{\vec{x}}\rangle\langle\psi_{\vec{x}}|\right)\right) \tag{11.9}$$

Zależy nam, żeby $\rho_{\vec{x}}$ było blisko $|\psi_{\vec{x}}\rangle$ w przypadku średnim.

$$\vec{F} = \sum_{\vec{x}} p(\vec{x})\langle\psi_{\vec{x}}|\rho'_{\vec{x}}|\psi_{\vec{x}}\rangle = \sum_{\vec{x}} p(\vec{x}) Tr\left(|\psi_{\vec{x}}\rangle\langle\psi_{\vec{x}}|\rho'_{\vec{x}}\right) \tag{11.10}$$

Wierna kompresja stopnia R:

$$\forall_{\epsilon>0}\ \exists_{m_0}\ \forall_{m\geqslant m_0} \quad \vec{F} > 1-\epsilon, \tag{11.11}$$

gdzie liczba qubitów n wykorzystana w kodowaniu dana jest wzorem

$$n = Rm. \tag{11.12}$$

Definicja

Entropią von-Neumanna nazywamy

$$S(\rho) = -\operatorname{Tr}\left(\rho lg(\rho)\right). \tag{11.13}$$

Rozpatrzmy rozkład spektralny operatora ρ

$$\rho = \sum_{\lambda} p_\lambda |e_\lambda\rangle\langle e_\lambda|. \tag{11.14}$$

Wtedy

$$S(\rho) = -\sum_{\lambda} p_\lambda lg(p_\lambda) = H(\lambda). \tag{11.15}$$

Można teraz zapisać **uśrednioną** wiadomość w bazie wektorów własnych w postaci

$$\sum_{\lambda_1,\lambda_2,\ldots,\lambda_m} p(\lambda_1)...p(\lambda_m)|e_{\lambda_1}\rangle\langle e_{\lambda_1}| \otimes ... \otimes |e_{\lambda_m}\rangle\langle e_{\lambda_m}| = \sum_{\lambda} p(\vec{\lambda})\left|e_{\vec{\lambda}}\right\rangle\left\langle e_{\vec{\lambda}}\right| \tag{11.16}$$

Przy czym mamy tutaj

$$\left\langle e_{\vec{\lambda}}\middle| e_{\vec{\lambda'}}\right\rangle. \tag{11.17}$$

Można to traktować klasycznie!

$$\rho^{\otimes m} \simeq \sum_{\vec{\lambda}\in\text{typ. sequences}} p(\vec{\lambda})\left|e_{\vec{\lambda}}\right\rangle\left\langle e_{\vec{\lambda}}\right|, \tag{11.18}$$

dla $m \to \infty$.

Tutaj

$$\mathbb{P}_T = \sum_{\vec{\lambda}\in\text{typ. sequences}} \left|e_{\vec{\lambda}}\right\rangle \left\langle e_{\vec{\lambda}}\right| \tag{11.19}$$

jest projektorem na typową podprzestrzeń, której wymiar wynosi

$$2^{mH(\lambda)} = 2^{mS(\rho)}. \tag{11.20}$$

Z teorii Shannona wynika, że użycie etykiet $\vec{\lambda}$ i przy $m \to \infty$ jest warunkiem wystarczającym, aby zapisać

$$mH(\lambda) = mS(\rho). \tag{11.21}$$

W jaki sposób? Mamy **kompresję Schumachera**.

$$R_\infty = S(\rho) \tag{11.22}$$

Kodowanie E:

W pierwszej kolejności należy dokonać projekcji na typową podprzestrzeń T i przesłać $n = mS(\rho)$-qubitową reprezentację wiadomości M:

$$|\psi_{\vec{x}}\rangle \simeq \left|\psi_{\vec{\lambda}}\right\rangle, \quad \vec{\lambda} \in T, \tag{11.23}$$

gdzie

$$\rho' = D(E(\rho)) = \mathbb{P}_T\rho\mathbb{P}_T + E(\rho) = \mathbb{P}_T\rho\mathbb{P}_T + |0\rangle\langle 0|^{\otimes m} \operatorname{Tr}\left(\rho(\mathbb{1} - \mathbb{P}_T)\right). \tag{11.24}$$

$\mathbb{P}_T\rho\mathbb{P}_T$ jest λ-reprezentacją n qubitów, które zostały wysłane. $E(\rho)$ jest wszystkim spoza T. Wtedy wysyłamy zafiksowany (atypowy) stan.

Dowód

Niech

$$\operatorname{Tr}\left(\rho^{\otimes m}\mathbb{P}_T\right) = 1 - \delta, \tag{11.25}$$

gdzie

$$\vec{F} = \sum_x p(\vec{x}) \operatorname{Tr}\left(|\psi_{\vec{x}}\rangle\langle\psi_{\vec{x}}| DE\left(|\psi_{\vec{x}}\rangle\langle\psi_{\vec{x}}|\right)\right) = \sum_x p(\vec{x}) \operatorname{Tr}\left(\psi_{\vec{x}}\left(\mathbb{P}_T\psi_{\vec{x}}\mathbb{P}_T + |0\rangle\langle 0| \operatorname{Tr}\left(\rho\mathbb{P}_T^\dagger\right)\right)\right) = \tag{11.26}$$

$$= \sum_x p(\vec{x}) \left(\operatorname{Tr}(\psi_{\vec{x}}\mathbb{P}_T\psi_{\vec{x}}\mathbb{P}_T) + \langle 0|\psi_{\vec{x}}|0\rangle \operatorname{Tr}\left(\rho\mathbb{P}_T^\dagger\right)\right) \geqslant \sum_x p(\vec{x}) \operatorname{Tr}(\psi_{\vec{x}}\mathbb{P}_T\psi_{\vec{x}}\mathbb{P}_T),$$

bo

$$|\psi_{\vec{x}}\rangle = \alpha_{\vec{x}}|\phi_{\vec{x}}\rangle + \beta_{\vec{x}}\left|\phi_{\vec{x}}^\dagger\right\rangle, \quad |\phi\rangle \in T \tag{11.27}$$

$$... = \sum_x p(\vec{x})|\alpha_{\vec{x}}|^4 = \sum_{\vec{x}} p(\vec{x})(1 - \beta_{\vec{x}}^2)^2, \tag{11.28}$$

ale

$$\operatorname{Tr}\left(\rho^{\otimes m}\mathbb{P}_T\right) = \sum_{\vec{x}} p(\vec{x})\langle\psi_{\vec{x}}|\mathbb{P}_T|\psi_{\vec{x}}\rangle = 1 - \sum_x p(\vec{x})|\beta_{\vec{x}}|^2 \geqslant 1 - \delta. \tag{11.29}$$

Stąd

$$\sum_{\vec{x}} p(\vec{x})|\beta_{\vec{x}}|^2 \leqslant \delta \tag{11.30}$$

$$... = 1 - \sum_{\vec{x}} p(\vec{x})2|\beta_{\vec{x}}|^2 + \sum_{\vec{x}} p(\vec{x})|\beta_{\vec{x}}|^4 \geqslant 1 - 2\delta \tag{11.31}$$

i wtedy

$$\vec{F} \geqslant 1 - 2\delta. \tag{11.32}$$

Należy tutaj zauważyć, że 2δ to właśnie ϵ z 11.11. Z typowości mamy

$$\delta \overset{m\to\infty}{\to} 0 \implies \vec{F} \to 1 \tag{11.33}$$

stąd

$$R_\infty = S(\rho) \tag{11.34}$$

jest wiernym stopniem kompresją (*faithful compression rate*). Dowód ten został pokazany przez Schumachera w 1995.

□

11.2 Kwantowe entropie

Rozpoczniemy od kilku definicji i własności.

Kwantowa informacja wzajemna

$$S(A:B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \tag{11.35}$$

Kwantowa entropia warunkowa

$$S(A|B) = S(\rho_{AB}) - S(\rho_B) \tag{11.36}$$

Należy zauważyć, że ta może być ujemna!

Kwantowa entropia relatywna (*quantum K-L divergence*)

$$S(\rho_A||\rho_B) = \mathrm{Tr}(\rho lg(\rho)) - \mathrm{Tr}(\rho lg(\sigma)) \tag{11.37}$$

Własności *S*

1. $S(\rho||\sigma) \geqslant 0$
2. podaddytywność (*subadditivity*): $S(A,B) \leqslant S(A) + S(B)$
3. $S(A,B,C) + S(B) \leqslant S(A,B) + S(B,C) \iff S(A:B) \leqslant S(A:BC)$
4. Niech ϵ będzie $CPTP$-mapą na B, taką, że

   $$\rho'_B = \epsilon(\rho_B), \tag{11.38}$$

   a wtedy

   $$S(A:B') \leqslant S(A:B). \tag{11.39}$$

5. $S(\sum_x p_x |x\rangle\langle x| \otimes \rho_x) = H(x) + \sum_x p_x \delta(\rho_x)$, dla $\langle x|x'\rangle = \delta_{xx'}$

Dowody

1.

$$\rho = \sum_i p_i |i\rangle\langle i|, \quad \sigma = \sum_k \tilde{p}_k |e_k\rangle\langle e_k|$$

$$S(\rho||\sigma) = \sum_i p_i lg(p_i) - \sum_i p_i \langle i| lg(\sigma) |i\rangle = \sum_i p_i \left(lg(p_i) - \sum_k |\langle i|e_k\rangle|^2 lg(\tilde{p}_k) \right) \geqslant$$

$$\geqslant \sum_i p_i \left(lg(p_i) - lg \sum_k \tilde{p}_k |\langle i|e_k\rangle|^2 \right) \geqslant 0,$$

przy czym ostatnia nierówność jest związana z klasyczną entropią relatywną (i $K-L$ *divergence*).

2.

$$\rho = \rho_{AB}, \quad \sigma = \rho_A \otimes \rho_B$$
$$S(\rho||\sigma) = \mathrm{Tr}(\rho_{AB} lg(\rho_{AB})) - \mathrm{Tr}(\rho_{AB} lg(\rho_A \otimes \rho_B)) = -S(A,B) - \mathrm{Tr}(\rho_{AB} lg((\rho_A \otimes \mathbb{1})(\rho_B \otimes \mathbb{1}))) =$$
$$= -S(A,B) - S(A) - S(B) \geqslant 0$$

3. Dowód nie jest tak prost jak w klasycznym przypadku, ale wciąż z podobną intuicją.

4.

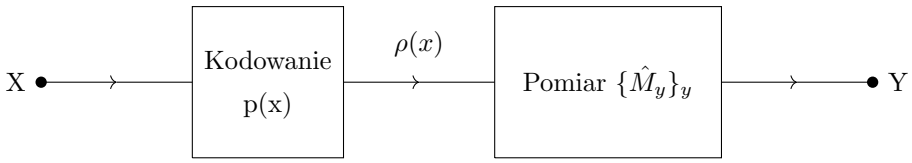
$$S(A:B') \leqslant S(A:B'C')$$
$$\rho_{AB'} = \sum_i \mathbb{1} \otimes K_i(\rho_{AB}) \mathbb{1} \otimes K_i^\dagger$$
$$\rho_{AB'C'} = \mathbb{1} \otimes U(\rho_{AB} \otimes |0\rangle \langle 0|) \mathbb{1} \otimes U^\dagger$$
$$S(\rho_{AB'C'}) = S(\rho_{AB} \otimes |0\rangle \langle 0|) = S(\rho_{AB}) + S(|0\rangle \langle 0|)$$
$$S(A:B') \leqslant S(A) + S(B',C') - S(A,B',C') \leqslant S(A:B)$$

5.

$$S\left(\sum_x p_x |x\rangle \langle x| \otimes \rho_x\right) = S\left(\sum_{x,r} p_x \lambda_r^{(x)} |x\rangle \langle x| \otimes \left|e_r^{(x)}\right\rangle \left\langle e_r^{(x)}\right|\right) = -\sum_{x,r} p_x \lambda_r^{(x)} lg\left(p_x \lambda_r^{(x)}\right) =$$
$$= -\sum p_x lg(p_x) - \sum_x p_x \sum_r \lambda_r^{(x)} lg(\lambda_r^{(x)}) = H(x) + \sum_x p_x S(\rho_x)$$

11.3 Twierdzenie Holevo

Twierdzenie to dotyczy skutecznego (wiernej) kodowania i dekodowania (klasycznej) informacji ze stanów kwantowych.



Rysunek 11.3: Schemat kodowania i dekodowania wiadomości.

$$p(y|x) = \mathrm{Tr}\Big(\rho(x)\hat{M}_y\Big) \tag{11.40}$$

Stąd mamy

$$p(x,y) = p(x)p(y|x) = p(x)\,\mathrm{Tr}\Big(\rho(x)\hat{M}_y\Big) \tag{11.41}$$

Twierdzenie

$$I(X:Y) \leqslant S(\vec{\rho}) - \sum_x p(x) S(\rho(x)) \tag{11.42}$$

$$\vec{\rho} = \sum_x p(x)\rho(x) \tag{11.43}$$

Prawą stronę tej nierówności nazywa się **wielkością Holevo** X (*Holevo quantity*).

Granica Holevo (*Holevo bound*) określa maksymalną dostępną informację (najlepszą znaną).

Dowód

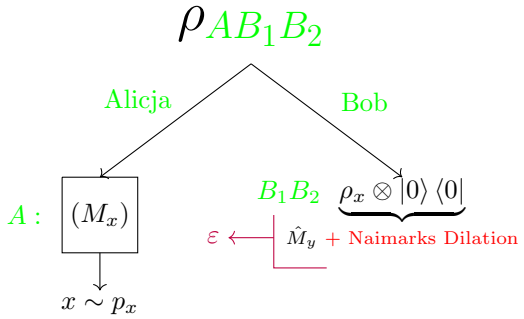
Rozważmy

$$\rho_{AB_1B_2} = \sum_x p_x \left|x\right\rangle_A \left\langle x\right| \otimes \rho_x \otimes \left|0\right\rangle_{B_2} \left\langle 0\right|, \tag{11.44}$$

z warunkiem, że

$$\langle x|x'\rangle = \delta_{xx'} \tag{11.45}$$

Można przedstawić to graficznie w sposób przedstawiony na rysunku 11.4.



Rysunek 11.4: Granica Holevo. W tym schemacie (część Boba) należy myśleć o protokole *RSP* (*remote state preparation*). Stan przygotowywany jest zgodnie z zespołem (ensemble) $\{p_x, \rho_X\}$.

$$\rho_{AB_1'B_2'} = E(\rho_A B_1 B_2) = \sum_x p_x \left|x\right\rangle \left\langle x\right| \otimes \sum_y \sqrt{\hat{M}_y}\rho_x\sqrt{\hat{M}_y} \otimes U_y \left|0\right\rangle \left\langle 0\right| U_y^\dagger \tag{11.46}$$

$$U_y \left|0\right\rangle \left\langle 0\right| U_y^\dagger = \left|y\right\rangle \left\langle y\right| \tag{11.47}$$

Mamy

$$S(A:BC) \geqslant S(A:B'C') \geqslant S(A:C') \tag{11.48}$$

$$S(A:BC) = S(A) + S(B) - S(A,BC) = H(X) + S(\vec{\rho}) - \sum_x p_x S(\rho_x) - H(X) = S(\vec{\rho}) - \sum_x p_x S(\rho_x) \tag{11.49}$$

$$S(A:C') = S(\rho_A) + S(\rho_{C'}) - S(\rho_{AC'}) = H(X) + H(Y) - H(X,Y) \tag{11.50}$$

$$\rho_{AC'} = \sum_x p_x \operatorname{Tr}\left(\rho_x \hat{M}_y\right) \left|x\right\rangle \left\langle x\right| \otimes \left|y\right\rangle \left\langle y\right| = \ldots = I(X:Y) \tag{11.51}$$

Stąd

$$I(X:Y) \leqslant S(\vec{\rho}) \leqslant S(\vec{\rho}) - \sum_{x=1}^{d} p_x S(\rho_x), \tag{11.52}$$

gdzie

$$\vec{\rho} = \sum_{x=1}^{d} p_x \rho_x \tag{11.53}$$

Ważna konsekwencja

Mając n-qubitów, w taki sposób, że $\vec{\rho} \in B(\mathbb{C}_2^{\otimes n})$ mamy

$$I(X:Y) \leqslant S(\vec{\rho}) = H(\lambda_{\vec{\rho}}) \leqslant lg(2^n) = n. \tag{11.54}$$

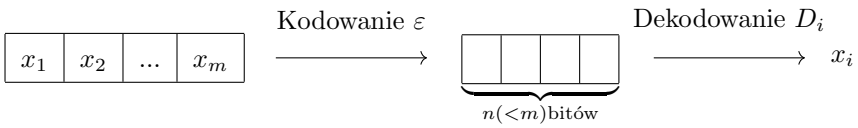
Oznacza to, że możemy w wiernie zakodować i zdekodować n bitów informacji w n-qubitach!

□

Czy jest zatem jakaś przewaga w używaniu qubitów do przechowywania informacji? Przykładem są (kwantowe) kody dostępu swobodnego!

11.4 Kwantowe kody dostępu swobodnego (*quantum random access codes*, QRAC)

Klasyczne kody dostępu swobodnego można przedstawić przy pomocy schematu.



Rysunek 11.5: Schemat klasycznych kodów dostępu swobodnego. $x_i \in [0,1]$ i początkowo mamy m-bitowy stan. Dla każdego x_i mamy inną funkcję dekodującą D_i. Dla jasności $i = 1, ..., m$.

Cel jest następujący. Dla każdego i należy skonstruować D_i w taki sposób, żeby bit x_i został odzyskany z prawdopodobieństwem (stąd *random* w angielskiej nazwie) co najmniej $> p$. RAC można przedstawić za pomocą krotki (m, n, p), gdzie m to bity na wejściu, n to tzw. bity złożone. Przykład RAC został przedstawiony na rysunku 11.6.

Spróbujmy

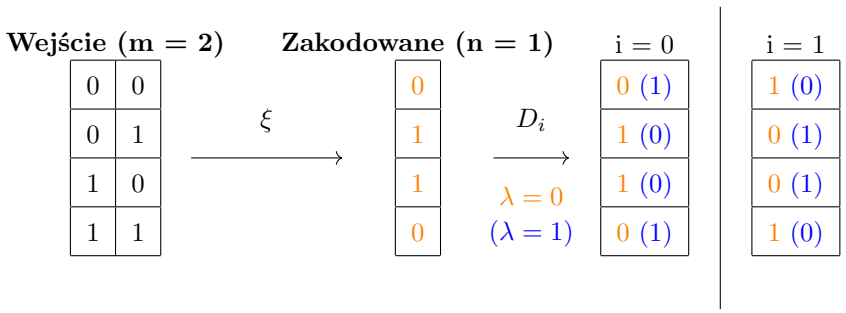
$$\xi(x, y) = x \otimes y. \tag{11.55}$$

Rzucając monetą $\lambda \in \{0, 1\}$ z prawdopodobieństwem $p(\lambda = 0) = p(\lambda = 1) = \frac{1}{2}$. Dla danej λ mamy

$$D_i(x, \lambda) = x \otimes (\lambda \otimes i) \tag{11.56}$$

Jednak nie jest to dobra strategia. Mamy $p = \frac{1}{2}$ co daje prawdopodobieństwo odgadnięcia i-tego bitu $\tilde{p} = \frac{1}{2}$, czyli zbyt mało, żeby było to jakkolwiek użyteczne.

Rozważmy teraz kwantową wersję. Ją także można przedstawić przy pomocy krotki (m, n, p), gdzie jedyną różnicą jest definicja n, które teraz oznacza liczbę qubitów.



Rysunek 11.6: Przykład RAC ($RAC(2, 1, p)$).

W ogólności (w odniesieniu do rysunku 11.6) mamy

$$\xi : \{0,1\}^2 \times R \to \{0,1\} \tag{11.57}$$
$$\forall_i \; D_i : \{0,1\} \times R \to \{0,1\} \tag{11.58}$$

gdzie R jest źródłem losowości.

Kwantowe kody dostępu swobodnego

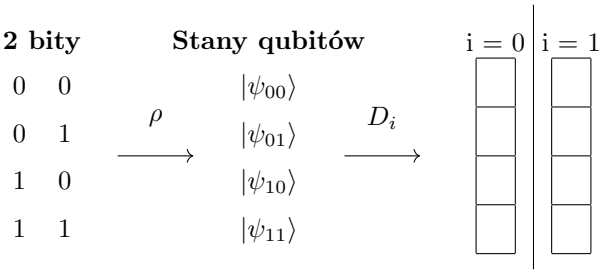
Rozpoczniemy od przykładu (rys. 11.7). Mamy tutaj kodowanie

$$\frac{\theta}{2} = \chi \tag{11.59}$$
$$|\psi_{00}\rangle = sin(\chi)\,|0\rangle + cos(\chi)\,|1\rangle \tag{11.60}$$
$$|\psi_{01}\rangle = -sin(\chi)\,|0\rangle + cos(\chi)\,|1\rangle \tag{11.61}$$
$$|\psi_{10}\rangle = cos(\chi)\,|0\rangle + sin(\chi)\,|1\rangle \tag{11.62}$$
$$|\psi_{11}\rangle = cos(\chi) - sin(\chi)\,|1\rangle \tag{11.63}$$

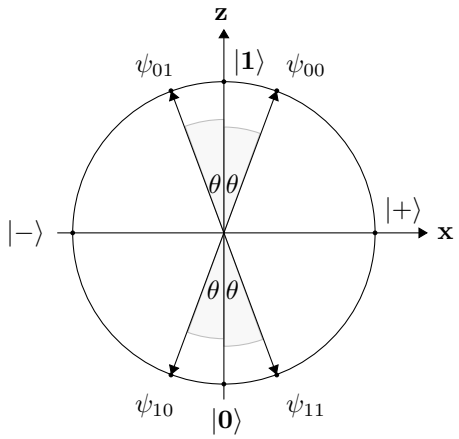
Stany zostały przedstawione w obrazie Blocha na rysunku 11.8.

Dekodowanie polegać będzie na

- $D_{i=0}$ pomiarze w bazie $\{|0\rangle, |1\rangle\}$ otrzymując wynik 1 lub 0, gdzie $|0\rangle \to 1, |1\rangle \to 0$,
- $D_{i=1}$ pomiarze w bazie $\{|+\rangle, |-\rangle\}$ otrzymując wynik 0 lub 1, gdzie $|-\rangle \to 1, |+\rangle \to 0$,



Rysunek 11.7: Przykład QRAC $\left(2, 1, \cos^2(\frac{\pi}{8}) \approx 0.85\right)$.

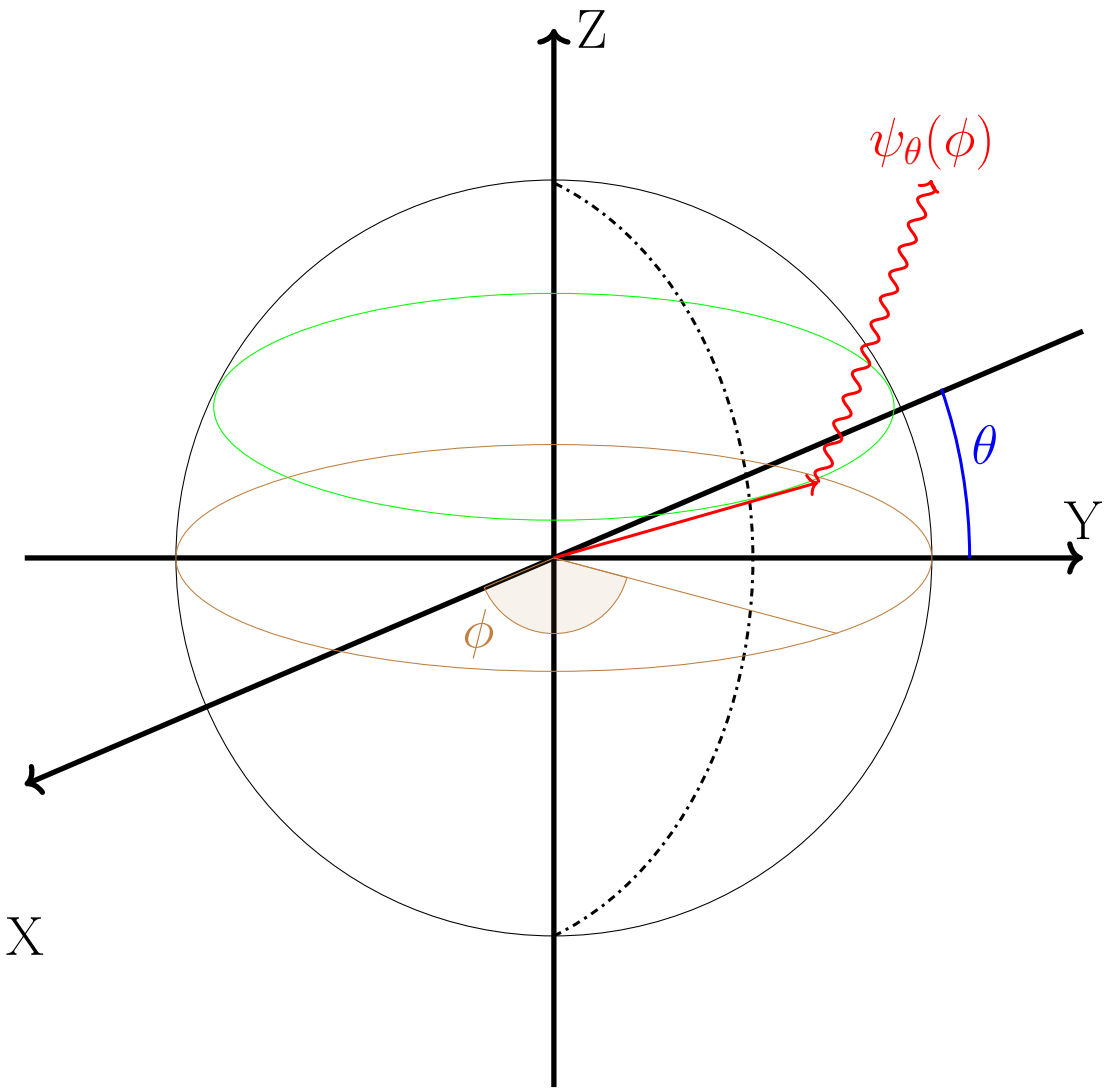


Rysunek 11.8: Kodowanie w obrazie Blocha.

Zauważmy, że dla $\theta = 0$ mierzymy w bazie $|0\rangle, |1\rangle$ i z $p = 1$ odzyskujemy pierwszy bit. Natomiast dla $\theta = \frac{\pi}{2}$ mamy pomiar w bazie $\{|+\rangle, |-\rangle\}$ i z prawdopodobieństwem $p = 1$ odzyskujemy bit drugi.

Przykłady

1. Kompresja ze stanami rozdystrybuowanymi równomiernie i równoległością.
2. Kodowanie i dekodowanie (Holevo)



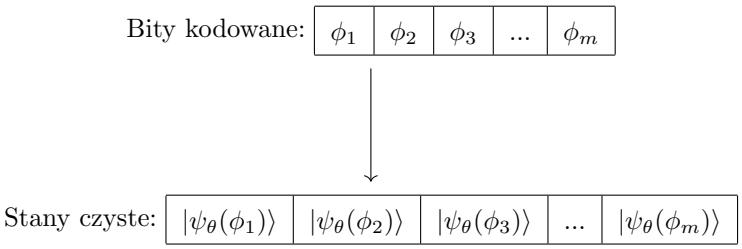
Rysunek 11.9: Kodowanie wiadomości.

Wiadomości $\phi \sim p(\phi) = \frac{1}{2\pi}$. Parametr θ jest tutaj ustalony.

Kompresja Schumachera

$$R_\infty = \frac{n(m)}{m} = S(\vec{\rho}) \tag{11.64}$$

$$\vec{\rho} = \int_0^{2\pi} d\phi p(\phi) \, |\phi_\theta(\phi)\rangle \, \langle\psi_\theta(\phi)| \tag{11.65}$$



Rysunek 11.10: Kodowanie ϕ.

Ponadto mamy tutaj wpływ granicy Holevo. Kodujemy do postaci $\{p(\phi), \psi_\theta(\phi)\}$. Mamy

$$I(X:Z) \leqslant S(\vec{\rho}) - \int d\phi p(\phi) S(|\psi_\theta(\phi)\rangle) \tag{11.66}$$

Ponadto, dla stanów czystych mamy

$$S(\psi) = 0, \tag{11.67}$$

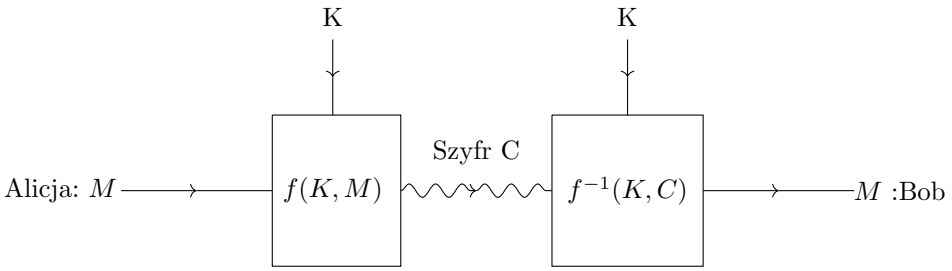
więc w efekcie

$$I(X:Z) = S(\vec{\rho}) = h\left(\frac{1+sin(\theta)}{2}\right) \leqslant lf(2) = 1, \tag{11.68}$$

przy czym równość zachodzi dla $\theta = 0$.

12 Kwantowa kryptografia

Rozpoczniemy od przedstawienia generalnego schematu w kryptografii. Alicja i Bob mają ustalony wcześniej współdzielony klucz K.



Rysunek 12.1: Graficzne przedstawienie kryptografii. Dwie strony, Alicja i Bob, współdzielą klucz. To co chcą zrobić, to przesłać sekretną wiadomość M. f jest tutaj funkcją szyfrującą, potencjalnie ogólnie znaną wszystkim zainteresowanym stronom.

Przy czym

1. $f(\cdot, \cdot)$ jest *praktycznie* nieodwracalną funkcją,
2. $\begin{cases} g(\cdot) = f(K, \cdot) \\ g^{-1}(\cdot) = f^{-1}(K, \cdot) \end{cases}$, jest odwracalną funkcją.

Należy tutaj myśleć o jednokierunkowych funkcjach / algorytmach typu SHA czy AES.

Przykład f

Mnożenie dużych liczb pierwszych.

$$f(prime_K, prime_M) = prime_K \cdot prime_M \tag{12.1}$$

Rozkład na czynniki pierwsze jest trudny (wg. teorii złożoności obliczeniowej należy do grupy BQP), chociaż nie dla komputerów kwantowych (algorytm Shora). Za to znając jedną z liczb mamy

$$g(\cdot) = f(prime_K, \cdot), \tag{12.2}$$

co jest łatwe do rozłożenia znając $prime_K$ (wystarczy podzielić).

$$g^{-1}(C) = \frac{C}{prime_K} = prime_M \implies M \tag{12.3}$$

Należy tutaj jednak zauważyć, że istnieją jednostronne-funkcje (funkcje hashujące), których złożoność obliczeniowa przekracza nawet komputery kwantowe (należy tutaj szukać pod hasłem *post-quantum cryptography*).

Jeśli chodzi o kryptografię, to zwykle rzecz polega na znalezieniu trudnego problemu i szyfrowaniu wiadomości w taki sposób, żeby odszyfrowanie wymagało rozwiązania tego problemu. W innym popularnym algorytmie kryptograficznym, RSA, problemem jest problem dyskretnego logarytmu. Skupmy się teraz na działaniu RSA.

Alicja ma dwa klucze, prywatny (K_{priv}) i publiczny (K_{public}). Jej klucz prywatny będzie jej później służył do odszyfrowania wiadomości zakodowanych kluczem publicznym, do którego każdy ma dostęp.

Problem logarytmu dyskretnego jest taki, że

$$b^P = a (mod\ n), \tag{12.4}$$

gdzie $n = prime_1 \cdot prime_2$, a nas interesuje znalezienie spełniającego to p. Ze strony Alicji protokół przebiega w następujący sposób.

1. $n = p_1 \cdot p_2$ (przy czym p_1 oraz p_2 to duże liczby pierwsze)
2. $\lambda(n) = NWW(p_1 - 1, p_2 - 1)$
3. Wybierz e względnie pierwsze z $\lambda(n)$, tj. $NWD(\lambda(n), e) = 1$
4. Znajdź d, takie, żeby $e \cdot d = 1 (mod\ \lambda(n))$

Wtedy mamy $K_{priv} = d$ i $K_{public} = (n, e)$. Zakodowana wiadomość ma wtedy postać

$$C = M^e (mod\ n) \tag{12.5}$$

a dekodowanie to po prostu

$$M = C^d (mod\ n). \tag{12.6}$$

Tutaj pojawia się pewien problem. To co my byśmy chcieli robić, to kodować **długie** wiadomości przy pomocy **krótkich** kluczy (AES - 256b, RSA $\geqslant$ 1024b). Shannon (w 1949) pokazał, że w celu zapewnienia wiadomości pełnej prywatności, klucz musiałby być tej samej długości co wiadomość. Ponadto, taki klucz musiałby być jednorazowy (*one-time pad*). Próbowano sobie z tym radzić na wiele sposobów. Omówimy przykładowy – **szyfr Vernama** (1919, patent) – który polega na zastosowaniu operacji XOR. Wtedy

$$K \oplus M = C \implies M = C \oplus K, \tag{12.7}$$

przy czym należy tutaj poczynić kilka uwag.

1. Bardzo istotnym jest posiadanie dobrego generatora liczb losowych, w calu generacji dobrego klucza K o długości $m = |M|$.
2. Potrzebny jest sposób na bezpieczne przekazanie klucza drugiej stronie...

Kwantowa mechanika może pomóc w tym drugim przypadku. Zobaczmy w jaki sposób.

12.1 Kwantowa dystrybucja klucza

Naszym celem będzie określenie sposobu w jaki Alice i Bob mogą określić sekretny (i losowy) ciąg bitów, który posłuży im jako klucz. Jeden ze sposobów na dokonanie tego, został zaproponowany przez Benneta i Brassarda w 1984. W tym protokole Alice i Bob wysyłają do siebie wielokrotnie qubity w odpowiednio przygotowanych stanach. Dzieje się to w następujący sposób.

1. Alicja losowo wybiera bazę, przykładowo $\{|0\rangle, |1\rangle\}$ lub $\{|+\rangle, |-\rangle\}$. Mają z Bobem ustalone, który stan z obu baz odpowiada klasycznemu 0 i 1. Koduje pożądany bit w tej bazie i przesyła do Boba.

   Przykładowa implementacja w systemie fotonowym polegałaby na przesyłaniu fotonów z bazy $\{|\leftrightarrow\rangle, |\updownarrow\rangle\}$ lub $\{|\nwarrow\rangle, |\nearrow\rangle\}$.
2. Bob wybiera bazę (z tego samego zestawu co Alicja), w której będzie dokonywał pomiaru.
3. Alicja i Bob komunikują się poprzez poświadczony publiczny kanał klasyczny. Bob ogłasza której bazy użył do pomiaru w każdej rundzie (nie ogłasza wyników!). Alicja następnie mówi, które rundy powinien zostawić.

Runda	1	2	3	4	5	6	...
Ala	0	-	+	1	-	0	...
Bob	0 / 1	0 / 1	+ / -	0 / 1	0 / 1	+ / -	...
Kompatybilność	T	N	T	T	N	N	...

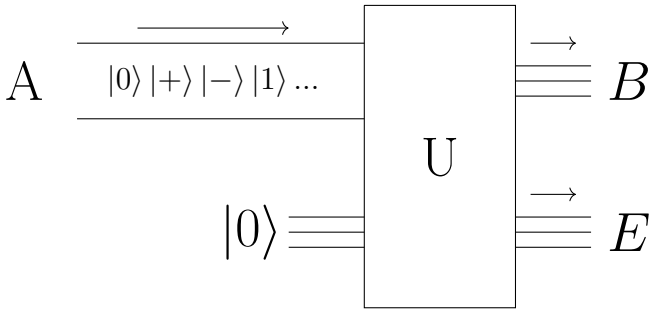
Rysunek 12.2: Przykładowa kompatybilność baz Alicji i Boba w poszczególnych rundach.

4. Faza przesiewania. Alicja i Bob zostawiają tylko te wyniki, w których ich bazy były kompatybilne.

5. Estymator błędu (podsłuchu)

   Alicja i Bob poświęcają kilka ($log(\#rund)$) aby zweryfikować aby zweryfikować procentową wielkość błędów (komunikacja dwu-kierunkowa). W ten sposób określają **QBER** (*quantum bit error-rate*). W najgorszym przypadku należy przyjąć, że wszystkie błędy są związane z posłuchem!

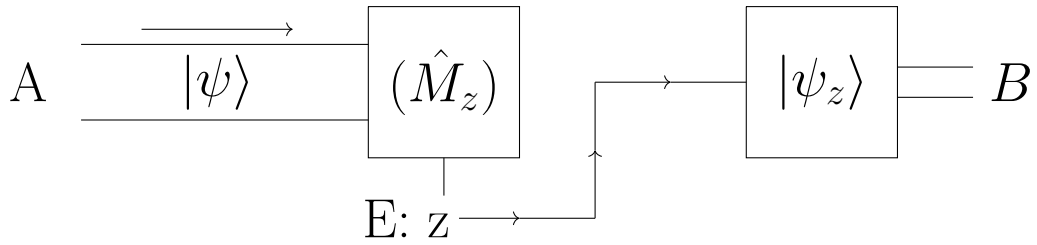
   Jak zbadać bezpieczeństwo tego protokołu?



Rysunek 12.3: Badanie bezpieczeństwa protokołu kwantowej dystrybucji klucza.

Ewa (podsłuchująca strona, od *eavesdropper*) czeka aż Alicja i Bob ustalą, które rundy miały kompatybilne bazy. Jej celem jest jak najcelniejsze odtworzenie ciągu znaków, który otrzymają Alicja i Bob. Należy wziąć tutaj pod uwagę najbardziej generalny atak, jaki może przeprowadzić Ewa, na który pozwala mechanika kwantowa. Rozważymy kilka typów ataków.

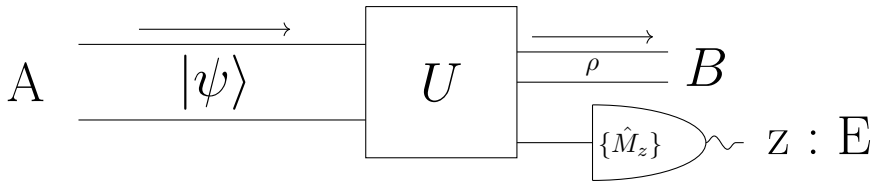
(a) Przechwycenie i przesyłanie (*intercept and resend*)



Rysunek 12.4: Schemat ataku z przechwyceniem i przesłaniem.

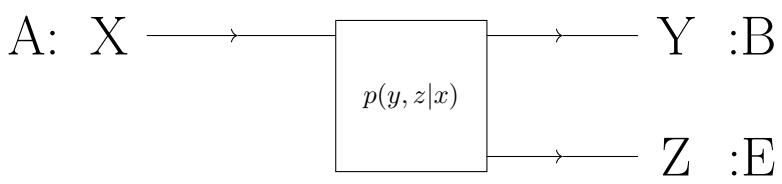
Ewa dokonuje pomiaru stanu przesłanego przez Alicję, a następnie, w zależności od wyniku pomiaru Z przygotowuje stan $|\psi_Z\rangle$ i przesyła do Boba. W każdej rundzie strategia pomiaru musi być inna (*warunek IID*).

(b) Ataki indywidualne (silniejsze)



Rysunek 12.5: Schemat ataku indywidualnego.

Ponownie, strategia musi być inna w każdej rundzie (warunek IID). Po fazie przesiewania mamy:



Rysunek 12.6: Atak indywidualny, po fazie przesiewania.

Aplikując twierdzenie Csiszara-Körnera mamy

Alicja	X	0	1	0	1	0	0	1	...
Bob	Y	0	1	1	1	0	1	1	...
Ewa	Z	1	1	1	0	1	0	1	...

$\Longrightarrow$ QBER: 25 %

Rysunek 12.7: Końcowa sytuacja.

6. Alicja i Bob dokonują korekcji błędu (*EC, error correction*) i amplifikacji prywatności (*PA, privacy amplification*) aby uzyskać klucz. Mamy

$$C_s = I(A:B) - I(A:E), \tag{12.8}$$

gdzie C_s to *asymptotic key-rate*. Ponadto powyższe jest prawdziwe tylko dla $I(A:B) > I(A:E)$. Wiemy ponadto, że

$$I(A:B) = 1 - h(QBER), \tag{12.9}$$

ale

$$I(A:E) = ? \tag{12.10}$$

Musimy wziąć pod uwagę dwa optymalne ataki, o których mówiliśmy wcześniej. Dzięki temu będziemy mogli określić informację wzajemną Alicji i Ewy w funkcji $QBER$.

(a) Proste przechwycenie i przesłanie.

Ewa naśladuje Boba:

- Wybiera jedną z ustalonych baz pomiarowych z równym prawdopodobieństwem.
- Przesyła zmierzony stan do Boba.

W takim przypadku mamy

$$QBER = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} = 25\%, \tag{12.11}$$

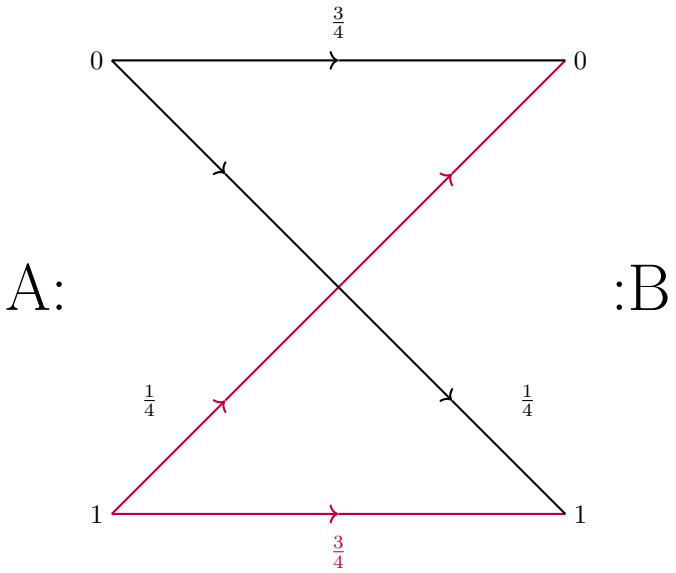
gdzie pierwszy czynnik odpowiada prawdopodobieństwu, że Ewa wybierze złą bazę, a drugi, że w przypadku gdy Ewa wybierze złą bazę, Bob dostanie zły wynik. Mamy zatem

$$I(A:B) = 1 - h(QBET) \simeq 0.189. \tag{12.12}$$

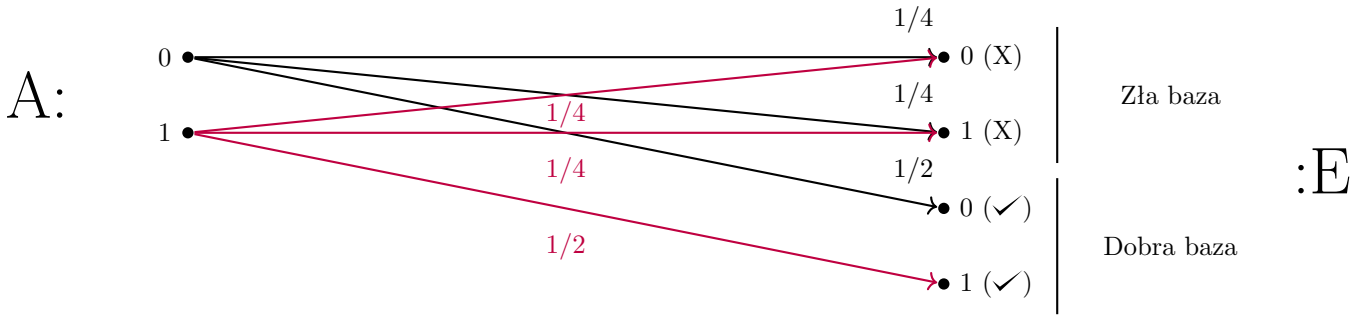
Jeśli chodzi o Ewę, to mamy za to

$$e_E = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \tag{12.13}$$

gdzie pierwszy czynnik oznacza, że Ewa wysłała złą bazę, a drugi, że dostała zły wynik po wybraniu złej bazy.



Rysunek 12.8: Bit-flip.



Rysunek 12.9: Faktyczna wzajemna informacja Alicji i Ewy. Zauważmy, że $I(A:E) = \frac{1}{2}$. Ewa wie kiedy mierzy w dobrej bazie, a kiedy nie.

Takie rozumowanie nie jest jednak poprawne! Ewa ma przecież dostęp (bo jest to ogłaszane na kanale publicznym), które rundy zostały zaakceptowane! Okazuje się (jak pokazano na rysunku 12.9), że jest jeszcze gorzej, a mianowicie $I(A:E) > I(A:B)$. Ostatecznie można powiedzieć, że jeśli

$$QBER \geqslant QBER_H = 25\% \tag{12.14}$$

to Alicja i Bob powinni zakończyć protokół bez przesyłania informacji.

Dodatkowo można pokazać, że $I(B:E) = I(A:E) = \frac{1}{2}$, jednak zostaje to pozostawione jako ćwiczenie dla czytelnika.

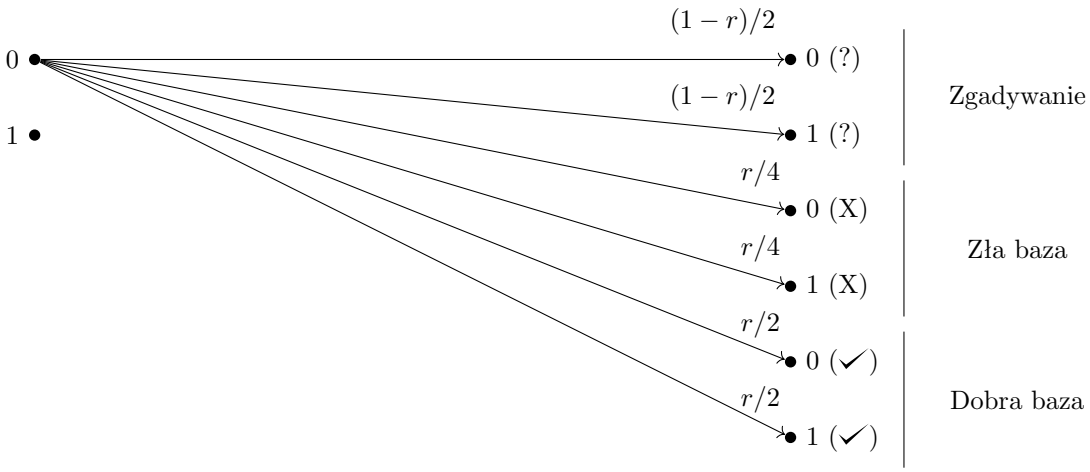
(b) Uogólnione przechwytywanie i przesyłanie

W uogólnionym przypadku tylko ułamek r jest przechwycony przez Ewę. Pozostałych $1-r$ Ewa musi

zgadnąć. W takim przypadku

$$QBER = \frac{r}{4} \implies I(A:B) = 1 - h\left(\frac{r}{4}\right) \tag{12.15}$$

Jeśli chodzi o Ewę, to mamy



Rysunek 12.10: Informacja wzajemna Alicji i Ewy w przypadku uogólnionym.

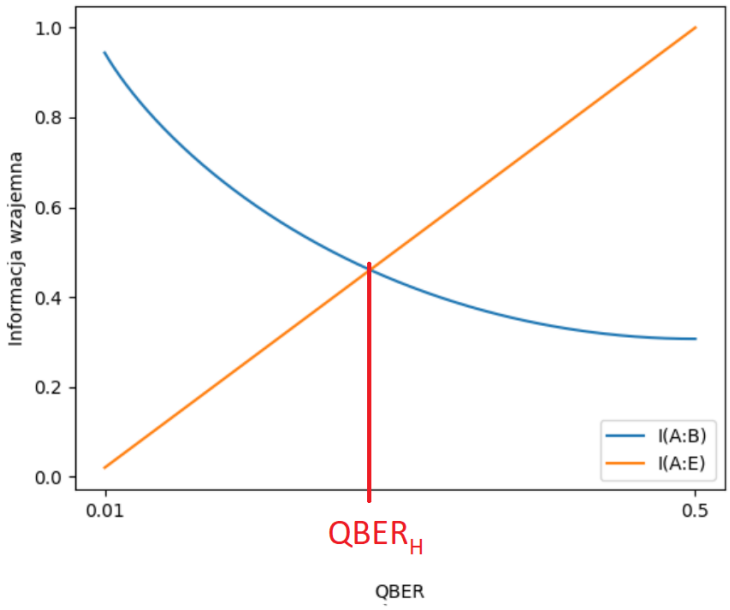
Stąd mamy

$$I(A:E) = \frac{r}{2}. \tag{12.16}$$

Po aplikacji korekcji błędu i wzmocnienia prywatności mamy

$$C_s = I(A:B) - I(A:E) = 1 - h\left(\frac{r}{4}\right) - \frac{r}{2} = 1 - h(QBER) - 2QBER \tag{12.17}$$

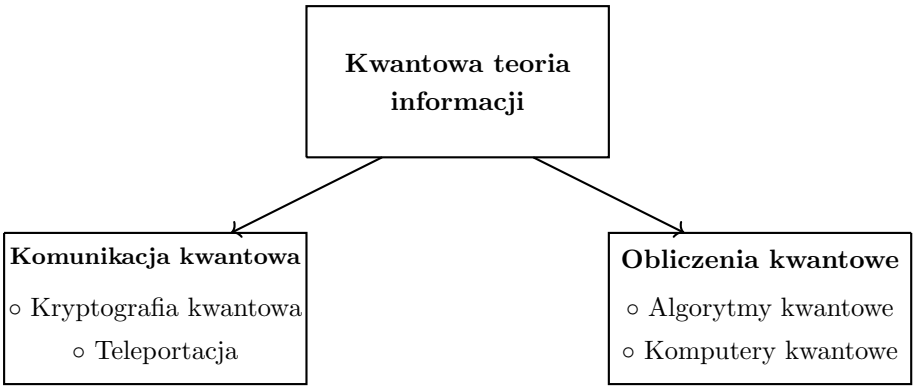
Można to przedstawić na wykresie. Dla $QBER \geqslant QBER_H = 17.1\%$ protokół powinien zostać przerwany.



Rysunek 12.11: Zależności informacji wzajemnych w funkcji $QBER$. $C_s > 0$.

13 Obliczenia kwantowe

Kwantową teorię informacji można zdefiniować jako przetwarzanie informacji korzystając z praw fizyki kwantowej operując na pojedynczych układach kwantowych (atomach, fotonach...). Można ją podzielić na węższe specjalizacje.

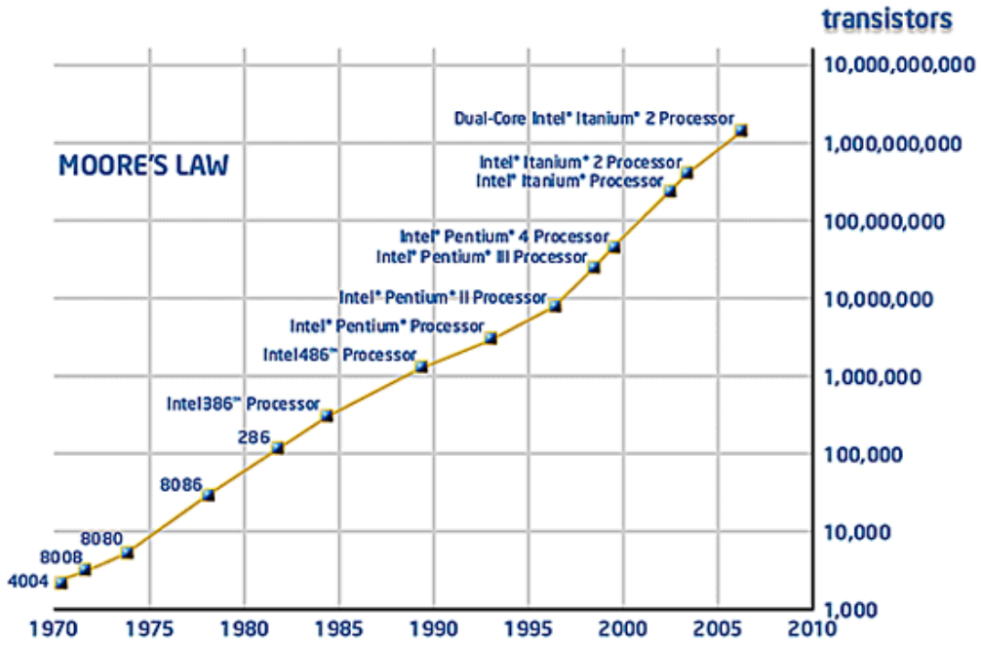


Rysunek 13.1: Podział kwantowej teorii informacji.

Do tej pory mówiliśmy jedynie o komunikacji kwantowej. Pora przejść do obliczeń. Należy przed tym jednak zaznaczyć, że aktualne komputery także korzystają z praw fizyki kwantowej (struktury półprzewodników, tranzystory, momenty magnetyczne atomów / spiny ...), jednakże

1. 1 bit na dysku twardym ma wymiary $250\ nm \times 250\ nm \times 25\ nm$ co przykłada się blisko na 12.5 mln. atomów ($d \simeq 0.5\ nm$).
2. 1 tranzystor w CPU ma wymiary $50\ nm \times 50\ nm \times 25\ nm$ co daje ok. 500 tys. atomów.

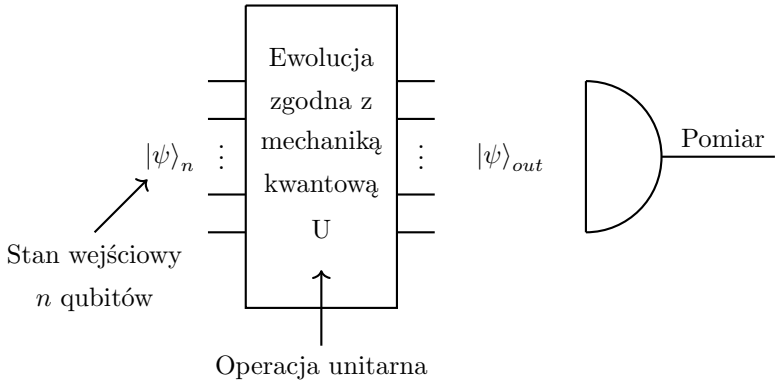
Nie jesteśmy jeszcze na etapie używania pojedynczych atomów do obliczeń i wykorzystywać pełne możliwości fizyki kwantowej. Można zadać sobie pytanie – kiedy będziemy? Odpowiedzi należy szukać w prawie Moore'a.



Rysunek 13.2: Graficzna reprezentacja prawa Moore'a. Wykres przedstawia jak zwiększa się liczba tranzystorów w procesorach w poszczególnych latach.

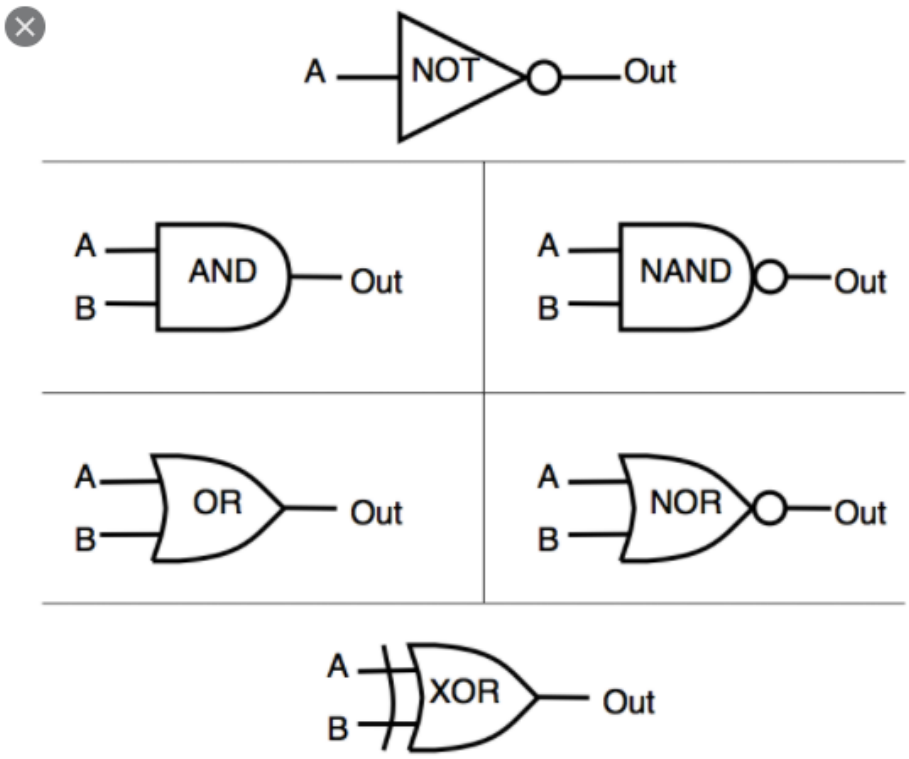
Rozmiary tranzystorów zmniejszają się ok. 2 razy co 2 lata. Oznacza to, że rozmiar atomu powinniśmy osiągnąć w latach 2030 - 2050! Zauważmy jednak, że nawet gdy to nastąpi, nie będzie to oznaczało, że mamy komputer kwantowy. Będziemy musieli dodatkowo umieć utrzymać kwantową superpozycję i dopiero to pozwoli nam w pełni wykorzystać potencjał fizyki kwantowej.

Algorytmy kwantowe można przedstawić przy pomocy ogólnego schematu



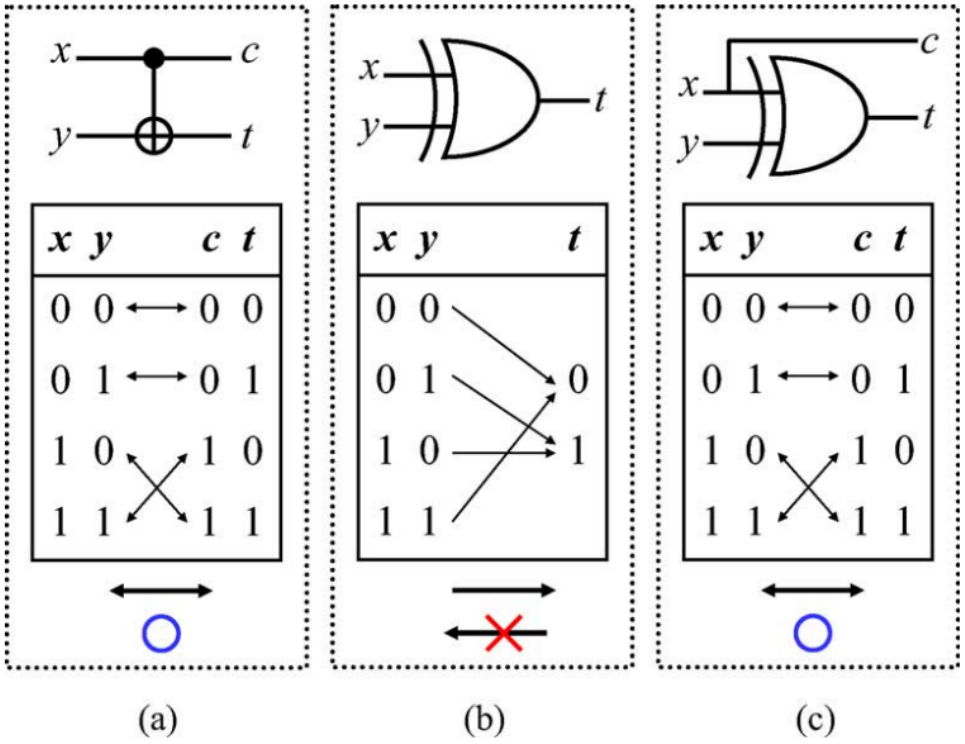
Rysunek 13.3: Ogólny algorytm kwantowy.

Klasyczne komputery buduje się zazwyczaj z pewnych elementarnych bramek. Te najbardziej popularne zostały przedstawione poniżej.



Rysunek 13.4: Typowe bramki logiczne. Źródło:
`https://anton-computing.weebly.com/classwork/logic-gates`.

Wiadomo także, że klasyczne komputery zwykle korzystają z bramek nieodwracalnych (tj. takich, na podstawie wyjścia których nie da się odtworzyć wejścia). Możliwe byłoby jednak wykorzystanie bramek odwracalnych. Zostało to zaprezentowane na rysunku poniżej (na przykładzie bramki XOR).



Rysunek 13.5: (a) Kwantowa bramka CNOT, (b) nieodwracalna bramka XOR (c) odwracalna bramka XOR. Źródło: `https://ieeexplore.ieee.org/document/4531956?arnumber=4531956`.

Należy jednak zauważyć, że bramki odwracalne komplikują obwody, dlatego nie stosuje się ich za często. Należy jednak pamiętać, że fizyka generalnie jest odwracalna. Nieodwracalność bierze się z tego, że zaczynamy ignorować jakieś stopnie swobody.

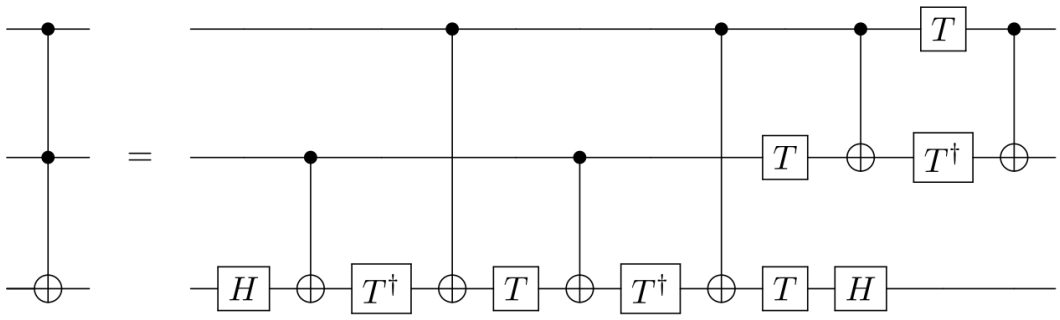
Komputery kwantowe działają tylko na bramkach odwracalnych. Wynika to z faktu, że operacje unitarne są odwracalne. Możliwe jest wykonanie nieodwracalnej operacji (np. śladując po niektórych qubitach), ale zazwyczaj niszczy to kwantowe superpozycje, więc nie bardzo ma sens. Ograniczymy się zatem do operacji unitarnych. Zastanówmy się zatem z zestawu jakich bramek da się złożyć dowolną operację unitarną i które są odwracalne. Pierwszą z nich jest zaprezentowana na rysunku 13.5 bramka CNOT. Jej działanie jest pokazane poniżej obwodu. Jej macierz można zapisać jako

$$U_{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{13.1}$$

Widać zatem, że zestawienie kwantowej bramki $CNOT$ i klasycznych XORów nie jest bezzasadne.

Fakt

Każda wieloqubitowa bramka unitarna może być rozłożona na jednoqubitowe bramki unitarne oraz operacje CNOT. Na rysunku 13.6 zaprezentowano jak realizuje się bramkę CCNOT (podwójnego zaprzeczenia) przy pomocy bramek jednoqubitowych i CNOT.



Rysunek 13.6: Realizacja bramki CCNOT przy pomocy bramek jednoqubitowych i CNOT. Źródło: Wikipedia.

W ogólności potrzebujemy wielu bramek jednoqubitowych (bo mamy nieskończenie wiele obrotów na sferze Blocha). Są jednak pewne szczególne bramki. Pierwszą jest bramka Hadamarda.

Rysunek 13.7: Bramka Hadamarda.

Jej postać macierzowa to

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad H^\dagger H = \mathbb{1}. \tag{13.2}$$

Mamy też bramkę fazową, przedstawioną na rysunku poniżej.

Rysunek 13.8: Bramka fazowa.

Który przedstawia się macierzą

$$U_\phi = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix} \tag{13.3}$$

Przejdźmy teraz do powodu dla którego kwantowe komputery mogą liczyć szybciej niż klasyczne tj. do **kwantowego paralelizmu**. Idea jest następująca. Weźmy sobie jednobitową funkcję f

$$f : \{0,1\} \to \{0,1\}. \tag{13.4}$$

Wyobraźmy sobie teraz, że kodujemy tę funkcję na bramce kwantowej U_f.

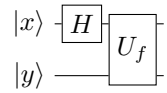
$$|x\rangle\,|y\rangle \xrightarrow{U_f} |x\rangle\,|y \oplus f(x)\rangle\,, \tag{13.5}$$

gdzie $|x\rangle$ to wejście, $|y\rangle$ wyjście, a $\oplus$ oznacza dodawanie mod 2.

Taka funkcję można wyrazić następującą tabelą

$$|00\rangle \to |0, 0 \oplus f(0)\rangle$$
$$|01\rangle \to |0, 1 \oplus f(0)\rangle$$
$$|10\rangle \to |1, 0 \oplus f(1)\rangle$$
$$|11\rangle \to |1, 1 \oplus f(1)\rangle$$

Zastanówmy się teraz nad działaniem następującego obwodu.



Rysunek 13.9: Rozpatrywany obwód dla jednobitowej f.

Załóżmy, że zaczynamy w stanie $|0\rangle\,|0\rangle$. Mamy wtedy:

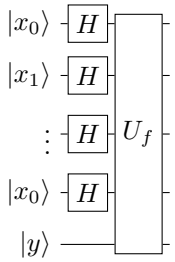
$$|0\rangle\,|0\rangle \overset{H\otimes\mathbb{1}}{\to} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\,|1\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) \overset{U_f}{\to} \frac{1}{\sqrt{2}}(|0\rangle\,|f(0)\rangle + |1\rangle\,|f(1)\rangle) = |\psi\rangle$$

Zauważmy, że obliczyliśmy f tylko raz (jedno wykorzystanie bramki U_f) a w wyniku otrzymaliśmy stan $|\psi\rangle$ w którym f jest policzona zarówno dla argumentu 1 jak i 0. Liczenie jednocześnie $f(0)$ i $f(1)$ wynika z faktu, że w pierwszej kolejności wprowadziliśmy nasz stan w superpozycję (przy pomocy bramki Hadamarda). Można rozważyć przypadek ogólniejszy. Weźmy

$$f : \{0, 1\}^N \to \{0, 1\}, \tag{13.6}$$

czyli mamy funkcję na N bitach.

$$|x_1, ..., x_N, y\rangle \overset{U_f}{\to} |x_1, ..., x_N, y \oplus f(x_1, ..., x_N)\rangle \tag{13.7}$$



Rysunek 13.10: Rozpatrywany obwód dla wielobitowej f.

Jak to działa na stan początkowy $|0, 0, ..., 0, 0\rangle$?

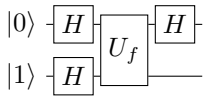
$$|0\rangle^{\otimes N}\,|0\rangle \overset{H^{\otimes N}\otimes\mathbb{1}}{\to} \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right)^{\otimes N} \otimes |0\rangle = \frac{1}{\sqrt{2^N}}\left(|0\rangle ... |0\rangle + |0\rangle ... |1\rangle + |1\rangle ... |1\rangle\right) \overset{U_f}{\to} \tag{13.8}$$

$$\overset{U_f}{\to} \frac{1}{\sqrt{2^N}}(|0\rangle ... |0\rangle \otimes |f(0, ..., 0)\rangle + |0\rangle ... |1\rangle \otimes |f(0, ..., 1)\rangle + |1, ..., 1\rangle\,|f(1, ...1)\rangle) \tag{13.9}$$

Należy zauważyć, że w nawiasie po pierwszym znaku równości mamy wszystkie liczby z zakresu $[0, 2^N - 1]$. Oznacza to, że jedną operacją f policzyliśmy wartość f dla 2^N wartości. Nie należy jednak cieszyć się przedwcześnie, gdyż nie istnieje pomiar, który pozwoli odczytać wszystkie te wartości na raz. Należy się zatem zastanowić, czy nie istnieją przypadkiem problemy, dla których obliczanie wszystkich wartości f byłoby elementem pośrednim, a istotny wynik jest faktycznie do zwrócenia w jednym pomiarze? Takie algorytmy to (m. in) algorytmy z wyrocznią.

13.1 Algorytmy z wyrocznią

Pierwszym algorytmem z wyrocznią, o którym mówi się na kursach informacji kwantowej, jest **algorytm Deutscha**. Algorytm ten jest relatywnie prosty, ale całkiem nieprzydatny. Rozważmy funkcję $f : \{0,1\} \to \{0,1\}$. Pytamy się, czy f jest różnowartościowa. Zauważmy, że klasyczną funkcję trzeba policzyć dwa razy (mamy dwie możliwe wartości wejścia). Jeśli jednak użyjemy metod informatyki kwantowej, wystarczy tylko jedna ewaluacja. Rozważmy obwód



Rysunek 13.11: Algorytm Deutscha w formie obwodu.

Działanie U_f jest następujące

$$U_f \left|x, y\right\rangle = \left|x, y \oplus f(x)\right\rangle \tag{13.10}$$

W takim razie możemy obliczyć wyjście

$$\begin{aligned} &\frac{1}{\sqrt{2}}(\left|0\right\rangle + \left|1\right\rangle)\frac{1}{\sqrt{2}}(\left|0\right\rangle - \left|1\right\rangle) \overset{U_f}{\to} \frac{1}{2}(\left|0\right\rangle(\left|f(0)\right\rangle - \left|1 \oplus f(0)\right\rangle + \left|1\right\rangle(\left|f(1)\right\rangle - \left|1 \oplus f(1)\right\rangle)) = \\ &= \frac{1}{2}\left((-1)^{f(0)}\left|0\right\rangle(\left|0\right\rangle - \left|1\right\rangle) + (-1)^{f(1)}\left|1\right\rangle(\left|0\right\rangle - \left|1\right\rangle)\right) = \frac{1}{2}\left((-1)^{f(0)}\left|0\right\rangle + (-1)^{f(1)}\left|1\right\rangle(\left|0\right\rangle - \left|1\right\rangle)\right) \overset{H\otimes\mathbb{1}}{\to} \\ &\overset{H\otimes\mathbb{1}}{\to} \frac{1}{2\sqrt{2}}\left((-1)^{f(0)}(\left|0\right\rangle + \left|1\right\rangle) + (-1)^{f(1)}(\left|0\right\rangle - \left|1\right\rangle)\right) \otimes (\left|0\right\rangle - \left|1\right\rangle) = \\ &= \left(\frac{1}{2}\left((-1)^{f(0)} + (-1)^{f(1)}\right)\left|0\right\rangle + \frac{1}{2}\left((-1)^{f(0)} - (-1)^{f(1)}\right)\left|1\right\rangle\right) \otimes \frac{1}{\sqrt{2}}(\left|0\right\rangle - \left|1\right\rangle) \end{aligned} \tag{13.11}$$

I teraz, czerwony nawias może mieć dwie wartości – 0 lub 1. Jeżeli jest równa 1, to znaczy, że f jest stała, natomiast dla wyniku 0 mamy f różnowartościową. Zielony nawias podobnie, może przyjmować wartości 1 lub 0. Dla 1 f jest różnowartościowa, a dla 0 f jest stała. Mierzymy pierwszy qubit. Wynik $\left|0\right\rangle$ oznacza funkcję stałą, a $\left|1\right\rangle$ różnowartościową. Dzięki temu, że U_f liczyliśmy jednocześnie na wszystkich wartościach wejściowych, wystarczyło użyć jej tylko raz!

Drugim algorytmem z wyrocznią jaki omówimy jest **algorytm Grovera**. Jest to nic innego jak kwantowe przeszukiwanie bazy danych. Załóżmy, że baza składa się z $N(= 2^n)$ elementów. Niech f będzie funkcją identyfikującą poszukiwany element (swojego rodzaju *wyrocznią*). Jej działanie można przedstawić w następujący sposób

$$f(x) = \begin{cases} 1 \iff x \text{ jest poszukiwanym elementem} \\ 0 \iff x \text{ nie jest poszukiwanym elementem} \end{cases} \tag{13.12}$$

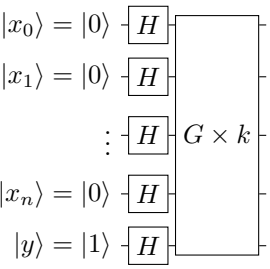
Jeżeli szukamy jednego elementu klasycznie, musimy średnio obliczyć funkcję f $\frac{N}{2}$ razy. Czy kwantowo da się to zrobić lepiej? Załóżmy, że mamy kwantową wersję f

$$U_f \left|x\right\rangle \left|y\right\rangle = \left|x\right\rangle \left|y \oplus f(x)\right\rangle, \tag{13.13}$$

przy czym $\left|x\right\rangle$ jest tutaj n-qubitowy, a $\left|y\right\rangle$ jedno-qubitowy. Zauważmy, że

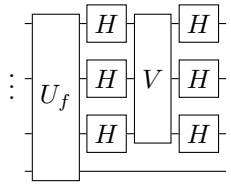
$$U_f \left|x\right\rangle \frac{1}{\sqrt{2}}(\left|0\right\rangle - \left|1\right\rangle) = (-1)^{f(x)}\frac{1}{\sqrt{2}}(\left|0\right\rangle - \left|1\right\rangle) \tag{13.14}$$

Od teraz będziemy ignorować ostatni qubit ($\left|y\right\rangle$). Będzie on cały czas pozostawał w stanie $\frac{1}{\sqrt{2}}(\left|0\right\rangle - \left|1\right\rangle)$. Schemat algorytmu Grovera można przedstawić na rysunku.



Rysunek 13.12: Schemat algorytmu Grovera. $G \times k$ oznacza k krotne wykonanie bramki G. Bramka G oznacza jedną iterację algorytmu Grovera.

Pojedyncza iteracja algorytmu Grovera (bramka G) może być rozbita na mniejsze operacje przedstawione przez obwód.



Rysunek 13.13: Bramka G.

Bramka V, to taka operacja unitarna, że

$$\begin{cases} V|0\rangle = |0\rangle \\ V|x\rangle = -|x\rangle \iff x > 0 \end{cases} \tag{13.15}$$

Można ją zapisać jako

$$V = 2|0\rangle\langle 0| - \mathbb{1}. \tag{13.16}$$

Można zauważyć, że

$$H^{\otimes N} V H^{\otimes N} = 2|\psi\rangle\langle\psi| - \mathbb{1}, \tag{13.17}$$

gdzie $|\psi\rangle = \frac{1}{\sqrt{N}} \sum_x |x\rangle$. Oznacza to, że

$$G = (2|\psi\rangle\langle\psi| - \mathbb{1})U_f. \tag{13.18}$$

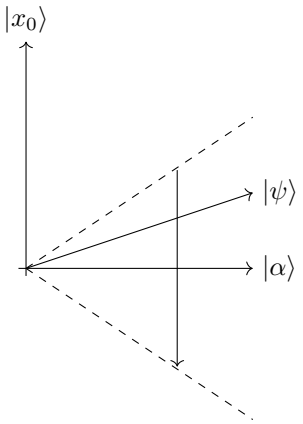
Niech teraz $|x_0\rangle$ będzie odpowiadać szukanemu stanowi. Niech $|\alpha\rangle = \frac{1}{\sqrt{N-1}} \sum_{x \neq x_0} |x\rangle$ będzie superpozycją pozostałych stanów. Rozważmy jak działa G na superpozycję $|x_0\rangle$ i $|\alpha\rangle$. Zauważmy, że

$$|\psi\rangle = \sqrt{\frac{N-1}{N}}|\alpha\rangle + \frac{1}{\sqrt{N}}|x_0\rangle. \tag{13.19}$$

Działanie G jest zatem

$$G(a|\alpha\rangle + b|x_0\rangle) = (2|\psi\rangle\langle\psi| - \mathbb{1})(a|\alpha\rangle - b|x_0\rangle) = ..., \tag{13.20}$$

czyli można je utożsamić z odbiciem względem kierunku $|\alpha\rangle$.



Rysunek 13.14: Obrazowe działanie algorytmu Grovera – odbicia stanów.

Kontynuując obliczenia ze wzoru [13.20] rozpoczynając od od obrotu względem $|\psi\rangle$

$$
\begin{aligned}
... &= 2\,|\psi\rangle \left(\sqrt{\frac{N-1}{N}} a - \frac{1}{\sqrt{N}} b \right) - (a\,|\alpha\rangle - b\,|x_0\rangle) = \\
&= 2\frac{N-1}{N} a\,|\alpha\rangle - 2\frac{1}{N} b\,|x_0\rangle + \frac{2\sqrt{N-1}}{N} a\,|x_0\rangle - \frac{2\sqrt{N-1}}{N} b\,|a\rangle - (a\,|\alpha\rangle - b\,|x_0\rangle) = \\
&= \left(2a - \frac{2a}{N} - \frac{2\sqrt{N-1}}{N} b - a \right) |\alpha\rangle + \left(b - \frac{2}{N} b + \frac{2\sqrt{N-1}}{N} a \right) |x_0\rangle = \\
&= \left(a \left(1 - \frac{2}{N} \right) - b \frac{2\sqrt{N-1}}{N} \right) |\alpha\rangle + \left(b \left(1 - \frac{2}{N} \right) + a \frac{2\sqrt{N-1}}{N} \right) |x_0\rangle
\end{aligned}
$$

Analiza wyniku pozwala stwierdzić, że mamy tutaj do czynienia z obrotem o kąt θ:

$$\cos(\theta) = 1 - \frac{2}{N}, \quad \theta = \arccos\left(1 - \frac{2}{N}\right). \tag{13.21}$$

W algorytmie Grovera startuje się ze stanu

$$|\psi\rangle = \sqrt{\frac{N-1}{N}}\,|\alpha\rangle + \frac{1}{\sqrt{N}}\,|x_0\rangle = \cos\left(\frac{\theta}{2}\right) + sin\left(\frac{\theta}{2}\right)|x_0\rangle\,. \tag{13.22}$$

W każdym kroku obracamy się o kąt θ. Po k iteracjach mamy

$$G^k\,|\psi\rangle = \cos\left(\frac{2k+1}{2}\theta\right) + sin\left(\frac{2k+1}{2}\theta\right)|x_0\rangle\,. \tag{13.23}$$

Jeżeli N jest bardzo duże, to mamy

$$\theta \simeq \frac{2\sqrt{N-1}}{N} \simeq \frac{2}{\sqrt{N}}. \tag{13.24}$$

Chcemy, aby $\frac{2k+1}{2}\theta \simeq \frac{\pi}{2}$, co daje równanie

$$(2k+1)\frac{2}{\sqrt{N}} = \pi \implies k \simeq \sqrt{N}. \tag{13.25}$$

Czyli kwantowe przyspieszenie jest ewidentnie widoczne (kwadratowe) w porównaniu z algorytmem klasycznym!

13.2 Algorytm Shora

W tej części omówione zostanę kwantowa transformacja Fouriera oraz algorytm Shora. Algorytm Shora jest kwantowym algorytmem rozkładu n-bitowej liczby na czynniki pierwsze. Działa w czasie n^3 – wielomianowym (sic!). Obecnie klasyczne algorytmy robią to w czasie $2^{\sqrt[3]{n}}$, czyli mniejszym niż wykładniczy, ale wyraźnie większym niż wielomianowy.

Esencją algorytmu Shora jest kwantowa transformata Fouriera (QFT). Zacznijmy od klasycznej transformaty Fouriera. Załóżmy, że ktoś ma jakieś dane w ilości N, weźmy $x_0, ..., x_{N-1}$. Zakładam, że $N \sim 2^n$, czyli N to z grubsza 2 do liczby qubitów n. Załóżmy, że ktoś sampluje muzykę, wtedy x_i to amplitudy. Chcemy poznać częstotliwości sygnałów, więc robimy transformatę Fouriera:

$$x_0, ..., x_{N-1} \rightarrow^{\mathcal{F}} y_0, ..., y_{N-1}$$

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} e^{2\pi i j k} N x_j$$

Jest to dyskretna, klasyczna transformacja Fouriera. Jest na skończonej próbce, więc zdarzają się tam pewne *artefakty*. Transformata mówi coś niecoś o częstotliwościach w sygnale. Dla nas ważnym jest teraz to, aby powiedzieć ile operacji trzeba wykonać do przeprowadzenia klasycznej transformaty Fouriera.

Ile klasycznych operacji trzeba wykonać? Liczba operacji (naiwnie) to $N^2(2^{2n})$. Istnieje jednak sposób na szybsze wykonanie tych operacji (*Fast Fouerier Transform, FFT*), której złożoność jest $Nlg(n)$ (czyli $(n2^n)$). Jest to jednak limit dla klasycznego podejścia.

Przejdźmy teraz do kwantów. QFT definiuje się jako operację unitarną U_F taką, że

$$U_F : |j\rangle \rightarrow \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{2\pi i j k} N |k\rangle ,$$

czyli mamy n qubitów, dających możliwe stany $|0\rangle , ..., |N-1\rangle$ – analogicznie jak było w przypadku klasycznym. W pierwszym przypadku musimy sprawdzić, czy jest to transformacja unitarna. Musimy sprawdzić, czy unormowanie jest zachowane. Szybko można zauważyć, że tak. Następnie trzeba pokazać unitarność. Zauważmy, że

$$\langle j'| U_F^\dagger U_f |j\rangle = \frac{1}{N} \sum_{k=0}^{N-1} e^{\frac{2\pi i (j-j')k}{N}} = \begin{cases} 1 & j = j' \\ \frac{1}{N} \frac{(1-e^{\frac{2\pi i (j-j')}{N}})^N}{1-e^{\frac{2\pi i (j-j')}{N}}} = 0 & j \neq j' \end{cases} , \tag{13.26}$$

co pokazuje, że operacja jest unitarna.

Mając U_F możemy się zastanowić co będzie, gdy podziałamy tym U_F stan

$$|\psi\rangle = \sum_j x_j |j\rangle$$

czyli dane klasyczne wpisujemy jako amplitudy. Dostaniemy

$$|\psi\rangle = \sum_j x_j |j\rangle \rightarrow^{U_F} \sum_j x_j \frac{1}{\sqrt{N}} \sum_k e^{\frac{2\pi i j k}{N}} |k\rangle = \sum_k \frac{1}{\sqrt{N}} \sum_j e^{\frac{2\pi i j k}{N} x_j} |k\rangle = \sum_k y_k |k\rangle \tag{13.27}$$

Na wyjściu mam amplitudy takie, jakie zrobiła by na x_j klasyczna transformata. Należy jednak pamiętać, że tak samo jak w przypadku algorytmu Deustcha nie będzie tak łatwo. To, że wyniki tam są, nie znaczy, że możemy je w łatwy sposób wykorzystać. W pierwszej kolejności, żeby powiedzieć, że tutaj kwantowo jest cośkolwiek lepiej, to musimy powiedzieć z ilu operacji (bramek) składa się QFT (na ćwiczeniach pokazane zostanie, że implementacja

QFT na n qubitach, potrzebować będziemy tylko $\sim n^2$ bramek – na tym polega przyspieszenie kwantowe).

Trzeba teraz zrozumieć co ma ta transformata do rozkładu liczby na czynniki pierwsze. Transformata Fouriera pozwala znajdować okres funkcji, który ma coś wspólnego z interesującym nas rozkładem. Zajmijmy się zatem w pierwszej kolejności tym. Rozważmy funkcję

$$f : \mathbb{Z}_N \to \mathbb{Z}_N$$

$$\exists_{r>0}\ \forall_x f(x+r) = f(x)$$

gdzie $\mathbb{Z}_N$ to ciało liczb od 0 do N. Dla ułatwienia przyjmijmy ponadto, żę jest różnowartościowa na okresie. Wyobraźmy sobie, że taką funkcję mamy w wersji kwantowej (transformację) U_f.

$$U_f \left|x\right\rangle \left|y\right\rangle = \left|x\right\rangle \left|y \oplus_N\ f(x)\right\rangle$$

Działamy na superpozycję wszystkich danych wejściowych

$$U_f \left(\frac{1}{\sqrt{N}} \sum_x \left|x\right\rangle \right) \otimes \left|0\right\rangle = \frac{1}{\sqrt{N}} \left|x\right\rangle \left|f(x)\right\rangle \tag{13.28}$$

Wykonujemy teraz pomiar rejestru wynikowego y. Załóżmy, że otrzymujemy wynik y_0. Oznacza to, że *przeżyją* tylko te x, które dają y_0. Stan zredukuje się do:

$$\frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} \left|x_0 + kr\right\rangle \left|y_0\right\rangle =: \left|\psi_{x_0,r}\right\rangle$$

$$f(x_0) = y_0$$

$$N = Kr$$

czyli K oznacza liczbę okresów funkcji f. Następnie można policzyć:

$$U_F \left|\psi_{x_0,r}\right\rangle = \frac{1}{\sqrt{N}} \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} \sum_{l=0}^{N-1} e^{\frac{2\pi i x_0 + kr}{N}} \left|l\right\rangle \otimes \left|y_0\right\rangle = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} e^{\frac{2\pi i l x_0}{N}} \left|l\right\rangle \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} e^{\frac{2\pi i k r L}{N}} = ... \tag{13.29}$$

Kiedy ta druga suma będzie nie zerowa? Sprawdźmy

$$\frac{1}{\sqrt{K}} \frac{1 - \left(e^{\frac{2\pi i r l}{Kr}}\right)^K}{1 - e^{\frac{2\pi i r l}{Kr}}} = \begin{cases} 0 & \frac{l}{K} \neq s \in \mathbb{Z} \\ \sqrt{K} & \frac{l}{K} = s \implies l = s\frac{N}{r} \end{cases} \tag{13.30}$$

Skoro tak, to kontynuując QFT

$$U_F \left|\psi_{x_0,r}\right\rangle = ... = \sqrt{\frac{K}{N}} \sum_{s=0}^{r-1} e^{\frac{2\pi i s \frac{N}{r} x_0}{N}} \left|s\frac{N}{r}\right\rangle = \frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{\frac{2\pi i s x_0}{r}} \left|s\frac{N}{r}\right\rangle \tag{13.31}$$

Wykonując pomiar, dostaniemy zatem $\lambda = s\frac{N}{r}$ dla jakiegoś przypadkowego $s = 0, ..., r-1$. Wynika stąd, że

$$\frac{\lambda}{N} = \frac{s}{r}.$$

Czy możemy na tej podstawie stwierdzić jakie jest r? Jeżeli s jest względnie pierwszy z r, to owszem. Trzeba się teraz odwołać do faktu z teorii liczb, żeby teoretyczna skuteczność tego algorytmu była zachowana. Okazuje się, że $NWD(s, r) = 1$ jest wystarczająco częste (prawdopodobieństwo zajścia tej sytuacji nie maleje ze wzrostem r).

Fakt

Prawdopodobieństwo, że dwie, duże przypadkowe liczby są względnie pierwsze jest dane przez

$$\sum_{p=2}^{\infty}\left(1-\frac{1}{p^2}\right)=\frac{1}{\zeta(2)}=\frac{6}{\pi^2}\simeq 0.4.$$

□

Mamy zatem okres funkcji. Co ma okres funkcji do czynników pierwszych? Niech N będzie liczbą, której czynników pierwszych szukamy. Losujemy $a < N$. Sprawdzamy (algorytmem Euklidesa, ze złożonością $\sim n^3$). Jeżeli $NWD(N,a) \neq 1$, to mamy czynnik pierwszy i koniec. Jeżeli jednak mamy równość, to na mocy tw. Eulera istnieje

$$a^r = 1 \bmod N. \tag{13.32}$$

W teorii liczb r nazywa się *rzędem a*. Oznacza to, że $a^r = kN + 1$. Jeżeli r jest parzyste, to mogę napisać

$$(a^{\frac{r}{2}}-1)(a^{\frac{r}{2}}+1)=\alpha\beta=kN \tag{13.33}$$

Skoro tak, to albo jedna z tych liczb jest wielokrotnością N – czyli nic ciekawego – albo któraś z nich musi mieć wspólny dzielnik z N (w szczególności, jeśli tak się zdarzy, że będzie N-em, to będzie miała z nim wspólne wszystkie dzielniki). Pomijając przypadek, gdy α lub β jest wielokrotnością N, to

$$NWD(\alpha,N)\neq 1 \tag{13.34}$$
$$NWD(\beta,N)\neq 1. \tag{13.35}$$

Dostaniemy zatem czynnik pierwszy N. Potrzebny jest kolejny fakt.

Fakt (z teorii liczb)

Prawdopodobieństwo, że $NWD(a,N) = 1$ i r jest parzyste i $a^{\frac{r}{2}} \pm 1$ nie są wielokrotnością N jest większe niż 0.5.

□

Potrzebujemy teraz znaleźć r. Definiujemy funkcję

$$f_a(x) = a^x mod N \tag{13.36}$$

Wtedy

$$f_a(x+r) = a^x a^r mod\ N, \tag{13.37}$$

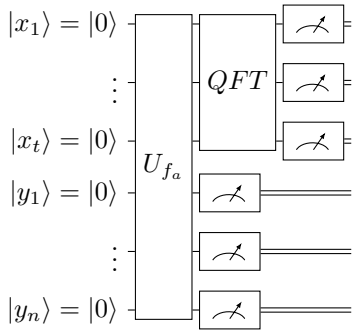
gdy $a^r = 1 mod\ N$. Wyznaczenie r to zatem znalezienie $f(x)$. Definiujemy funkcję

$$U_{f_a}\left|x\right\rangle\left|y\right\rangle = \left|x\right\rangle\left|a^x mod\ N\right\rangle. \tag{13.38}$$

(Złożoność takiego tworu nie jest większa niż n^3.) Dalej już wiadomo co robić – było na początku wykładu.

Algorytm Shora

Mamy $a < N$ i $N < 2^n$. Na wejściu biorę t qubitów, a na wyjściu n.



Rysunek 13.15: Obwód kwantowy realizujący algorytm Shora.

Na wyjściu otrzymujemy t-bitowe przybliżenie ułamka $\frac{s}{r}$. Korzystając później z metod ułamków łańcuchowych aby znaleźć ułamek prosty i znamy r.

14 Kwantowa korekcja błędów

W praktyce nie ma bramek idealnych. Klasycznie radzimy sobie metodami korekcji błędów. Przykładowo bit 0 można kodować przy pomocy trzech bitów mówiąc, że 000 to logiczne 0. Analogicznie 111 to logiczne 1. W przypadku wystąpienia błędu, 010, wiemy, że logicznie to wciąż jest 0 (z pewną dozą prawdopodobieństwa zależną od długości nadmiernego kodowania).

Należy się zastanowić, czy takie same operacje można robić kwantowo. Zauważmy, że celem tutaj jest nie tylko ochrona przed operacją typu bit-flip, ale także przed wszelakim rodzajem ogólniejszych zmian stanów (obrotów qubitu na sferze Blocha). Naturalnym wydaje się pytanie – czy możemy chronić wszystkie superpozycje? Rozpatrzmy najpierw klasyczny odpowiednik bit-flit tj. niepożądane działanie bramki σ_x. Mamy

$$|0\rangle \to |1\rangle\,, \quad |1\rangle \to |0\rangle\,.$$

Kodujemy

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle \to a\,|000\rangle + b\,|111\rangle\,. \tag{14.1}$$

Załóżmy, że wystąpił błąd typu bit-flip na jednym z qubitów. Nazwijmy $C = \{|000\rangle\,, |111\rangle\}$. Skoro błąd występuje potencjalnie tylko na jednym qubicie, to mamy tylko 4 przypadki – brak błędu, błąd na pierwszym, drugim lub trzecim. Oznacza to:

$$\begin{aligned}
a\,|000\rangle + b\,|111\rangle \to \mathbb{1}\otimes\mathbb{1}\otimes\mathbb{1} &\implies C = C_0 \implies \text{brak błędu},\\
a\,|100\rangle + b\,|011\rangle \to \sigma_x\otimes\mathbb{1}\otimes\mathbb{1} &\implies C = C_1 \implies \text{błąd na 1 qubicie},\\
a\,|010\rangle + b\,|101\rangle \to \mathbb{1}\otimes\sigma_x\otimes\mathbb{1} &\implies C = C_2 \implies \text{błąd na 2 qubicie},\\
a\,|001\rangle + b\,|110\rangle \to \mathbb{1}\otimes\mathbb{1}\otimes\sigma_x &\implies C = C_3 \implies \text{błąd na 3 qubicie}.
\end{aligned}$$

Znajdujemy się w jednej z 4 podprzestrzeni C_i. Mamy wybrać pomiar identyfikujący tę podprzestrzeń i naprawiający błąd, z którym jest ona związana.

Fakt

Jeżeli potrafimy naprawić jakiś rodzaj błędu, to potrafimy naprawić również liniowe kombinacje takich błędów (zwykle robi się to przez wprowadzenie ancilli, która zapisuje rodzaj błędu).
W przypadku informacji kwantowej mamy też do czynienia z innym rodzajem błędu tzw. *phase-flip*. W tym przypadku pozostajemy w tej samej podprzestrzeni C. Phase-flip działa w następujący sposób (na pojedynczym qubicie)

$$|0\rangle \to |0\rangle\,, \quad |1\rangle \to |-1\rangle\,.$$

Przyjmijmy następujące kodowanie

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \to \alpha\,|+++\rangle + \beta\,|---\rangle\,. \tag{14.2}$$

Stan $|\alpha_\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ będzie chronić przed błędem typu phase-flip.

Mając te dwa kodowania, chcielibyśmy teraz połączyć je w jakiś sposób tak, aby otrzymać kodowanie odporne zarówno na bit jak i phase flip. Weźmy

$$\begin{aligned}
|0\rangle &\to \frac{1}{\sqrt{2}}\left(|000\rangle + |111\rangle\right)^{\otimes 3}\\
|1\rangle &\to \frac{1}{\sqrt{2}}\left(|000\rangle - |111\rangle\right)^{\otimes 3}
\end{aligned} \tag{14.3}$$

Takie kodowanie potrafi jednocześnie naprawić phase i bit flip ($\sigma_y = -i\sigma_z\sigma_x$). Okazuje się, że takie kodowanie jest wystarczające, aby naprawić **dowolny** błąd na pojedynczym qubicie!

Rozpatrzmy teraz następujący przypadek. Niech Λ będzie kanałem opisującym dekoherencję jednego qubitu.

$$\Lambda(|\psi\rangle\langle\psi|) = \sum_i K_i |\psi\rangle\langle\psi| K_i^\dagger \tag{14.4}$$

Każdy K_i mogę zapisać w bazie σ_i.

$$K_i = \sum_j \epsilon_{ij}\sigma_j \tag{14.5}$$

Oznacza to, że na wyjściu dowolnego obwodu dostaniemy stany

$$|\psi_i\rangle = \sum_i \epsilon_{ij}\sigma_j |\psi\rangle, \tag{14.6}$$

czyli kombinacje bit / phase flip, a te potrafimy naprawić!

Tolerancja błędu (Fault tolerance)

Jeśli poziom błędów jest odpowiednio niski możemy efektywnie je naprawiać (mimo, że używamy też do tego niedoskonałych elementów). Ogólnie przyjmowany próg tolerancji błędu to $F \geqslant 99,9\%$. To ważne, żeby błędy były nieskorelowane!

Ogólne warunki dla obwodów z korekcją błędów

Niech E_a będzie bazą op. reprezentującą błędy.

$$E = \sum_a c_a E_a \tag{14.7}$$

Dodatkowo niech $|\psi_i\rangle$ oznacza stany logiczne. Wtedy

$$|\psi\rangle = \sum_i a_i |\psi_i\rangle. \tag{14.8}$$

Twierdzenie

$$\forall_{E,|\psi\rangle} \quad \langle\psi| E^\dagger E |\psi\rangle = C(E) \iff \text{mamy } E - C \text{ code.} \tag{14.9}$$

przy czym C jest niezależne od $|\psi\rangle$.

Dowód

Zaczniemy od implikacji w prawo. Mamy

$$\begin{gathered}
\langle\psi_i| E_a^\dagger E_a |\psi_i\rangle = C(E_a) \\
\langle\psi_j| E_a^\dagger |\psi_j\rangle = C(E_a) \\
\frac{4}{2}(\langle\psi_i| + \langle\psi_j|)E_a^\dagger E_a(|\psi_i\rangle + |\psi_j\rangle) = C(E_a) \\
\langle\psi_i| E_a^\dagger E_a |\psi_j\rangle = C(E_a)\delta_{ij} \\
\vdots \\
\langle\psi_i| E_a^\dagger E_b |\psi_j\rangle = C_{ab}\delta_{ij},
\end{gathered}$$

przy czym C_{ab} jest hermitowskie. Diagonalizując C_{ab} przejdziemy do nowej bazy. Nazwijmy ją F_a. Mamy

$$\langle\psi_i| F_a^\dagger F_b |\psi_j\rangle = f_a\delta_{ij}\delta_{ab} \tag{14.10}$$

Pozwala nam to na omijanie stanów po zadziałaniu błędu i możemy stwierdzić jaki błąd się zadział. Widzimy też, że

$$\langle\psi_i| F_a^\dagger F_a |\psi_j\rangle = f_a \delta_{ij}, \tag{14.11}$$

czyli, że działa proporcjonalnie do operacji unitarnej. Na przestrzeni kodowej jest to odwracalne. (????)

Dowód w drugą stronę. Jeżeli mamy $E - C$, to musi zachodzić

$$\langle\psi| E^\dagger E |\psi\rangle = \langle\phi| E^\dagger E |\phi\rangle \tag{14.12}$$

inaczej E zmieni (?) stan.

$$E(|\psi\rangle + |\phi\rangle) = N\left(|\psi\rangle + \frac{\langle\psi| E^\dagger E |\psi\rangle}{\langle\phi| E^\dagger E |\phi\rangle} |\phi\rangle\right) \tag{14.13}$$

a to też ????? stan kodujemy, ale inny.