
TRASCRIZIONE DEGLI APPUNTI DELLE LEZIONI DI

Introduzione alla teoria quantistica dell'informazione

TENUTE DAL PROF. *Simone Montangero*
PRESSO L'UNIVERSITÀ DI PADOVA

A CURA DI: *Francesco Manzali, Mattia Morgavi*

ANNO ACCADEMICO 2018-2019

Compilato il 2 ottobre 2020

Sorgente: <https://github.com/Einlar/InfoQuantistica>

This work is licensed under a Creative Commons “Attribution-NonCommercial-ShareAlike 4.0 International” license.

Indice

1	Circuiti quantistici	7
1.1	Informazioni generali del corso	7
1.2	Argomenti e motivazione del corso	7
1.3	Concetti di informazione quantistica	8
1.3.1	La natura dell'informazione	8
1.3.2	Definizioni di base	10
1.4	Modello di calcolo a circuiti	12
1.4.1	Calcolo classico deterministico	12
1.4.2	Calcolo quantistico	13
1.5	Porte logiche quantistiche	16
1.5.1	Porte a 1 qubit	16
1.5.2	Porte a 2 qubit	18
1.5.3	Funzioni di qubit	21
1.5.4	Esempio di funzione quantistica	23
1.6	No cloning theorem	24
1.7	Classi di complessità algoritmica	25
1.8	Esercizio 1	28
1.9	Esercizio 2	28
1.10	Esercizio 3	30
2	Effetti quantistici	32
2.1	Teletrasporto quantistico	32
2.2	Misure quantistiche	36
2.3	Misura senza interazione	36
2.3.1	Il Beam Splitter quantistico	36
2.3.2	Interferometro di Mach-Zehnder quantistico	38
2.3.3	Il test della bomba di Elitzur–Vaidman	39
2.4	Effetto Zenone quantistico	41
2.4.1	I paradossi di Zenone	41
2.4.2	Introduzione	42

2.4.3	Derivazione dell'Effetto Zenone	42
2.4.4	Esempio di effetto Zenone	46
2.4.5	Effetto Zenone per interazione	48
2.4.6	Matrici "Hamiltoniane" non hermitiane	49
2.4.7	Effetto Zenone senza proiezioni	50
2.4.8	L'origine delle "Hamiltoniane" non hermitiane	54
2.5	Implementazione di porte logiche	57
2.6	Implementazione di un gate CNOT	62
3	Stati e misure	64
3.1	Matrici densità	64
3.1.1	Misure statistiche	64
3.1.2	Evoluzione temporale	66
3.1.3	Proprietà della matrice densità	67
3.1.4	Stati puri e misti	68
3.1.5	Il sistema da 1 qubit	69
3.1.6	Sistemi composti	71
3.2	Matrici densità - parte 2	75
3.2.1	Correlazioni tra stati	75
3.2.2	Correlazioni ed entanglement	77
3.2.3	Schmidt decomposition	77
3.2.4	Schmidt Rank	82
3.2.5	Purificazione	83
3.2.6	Rappresentazione di Kraus	85
3.2.7	Kraus Representation Theorem	88
3.3	Misure generalizzate	90
3.3.1	Teorema di Neumark	90
3.3.2	Rappresentazione unitaria di Operatori di Kraus	91
3.3.3	Motivazione delle misure generalizzate	93
3.3.4	Weak Measurement	94
3.3.5	POVM Measurement	97
3.4	Canali quantistici	98
3.4.1	Decoerenza di un qubit	99
3.4.2	Canale bit-flip	102
3.4.3	Canale phase-flip	103
3.4.4	Canale bit-phase flip	105
3.4.5	Depolarizing channel	105
3.4.6	Amplitude damping	106
3.4.7	Phase Damping	108
3.4.8	Canale di de-entanglement	109
3.5	Master equation	111
3.6	Crittografia e meccanica quantistica	113
3.6.1	Crittografia classica	113
3.6.2	Quantum Key Distribution (QKD)	116
3.6.3	Il protocollo BB84	116
3.7	Dense coding	119

4	Correlazioni quantistiche	122
4.1	Bell Inequalities	122
4.2	Tipologie di correlazioni	130
4.3	Misura sperimentale della violazione delle disuguaglianze CHSH . .	134
4.4	Quantificare l'informazione	138
4.4.1	Entropia di Shannon	138
4.4.2	Noiseless Coding	140
4.5	Entropia di Von Neumann	141
4.5.1	Quantum Noiseless Coding	144
4.6	Caratterizzazione dell'Entanglement	145
4.6.1	Stati classicamente correlati	145
4.6.2	Misura di Entanglement	146
4.6.3	Dense Coding e correlazioni classiche	152
4.6.4	Esercizio 4	153
5	Algoritmi quantistici	156
5.0.1	Algoritmo di Deutsch	156
5.1	Algoritmo di Grover	161
5.2	Quantum Fourier Transform	168
5.3	Computer quantistici e crittografia classica	171
5.3.1	La crittografia RSA	171
5.3.2	Algoritmo di Shor - parte classica	175
5.3.3	Algoritmo di Shor - parte quantistica	176
5.3.4	Esempio	179
5.4	Phase Estimation Algorithm	180
5.5	Eigensolver	183
5.5.1	Algoritmo classico	184
5.5.2	Algoritmo quantistico	185
5.6	Error correction	187
5.6.1	Classical error correction	187
5.6.2	Quantum error correction	188
6	Complementi	192
6.1	Teoria delle perturbazioni dipendenti dal tempo	192
6.1.1	Perturbazione sinusoidale	196
6.1.2	Fermi Golden Rule	197
7	Hardware quantistico	200
7.0.1	Ioni intrappolati	201
7.0.2	Superconducting qubit	202

Introduzione

Buonsalve!

In questo documento ho cercato di riordinare gli appunti del corso di Introduzione alla teoria quantistica dell'informazione tenuto dal professor Simone Montangero presso il Dipartimento di Fisica dell'Università di Padova nel corso del secondo semestre del 2018-19.

Potrebbero esserci errori di formattazione, parentesi saltate, o peggio, coefficienti/esponenti/segni errati in giro (ma non dovrebbero essere tanti). Se ne sgamate qualcuno, fatemi sapere. Ditemi anche (se avete tempo e non vi scoccia) se ci sono passaggi non chiari.

Disclaimer: questi appunti non sono da intendere come sostituzione delle lezioni, o di altre dispense già presenti.

Buon viaggio! :)

Francesco Manzali, 20/02/2019

Aggiornamenti

Data	Aggiunte	Errata corrige	Commenti
30/6/2019	Prima pubblica- zione		

Tabella (1) – Cronologia di modifiche/aggiornamenti agli appunti

Circuiti quantistici

1.1 Informazioni generali del corso

(Lezione 1 • del
27/2/2019)

1.2 Argomenti e motivazione del corso

Nel 1899 la Meccanica Classica (MC) fu messa in crisi dal problema dello spettro del corpo nero. Nel 1900 Planck riuscì a spiegare le nuove misurazioni con l'**ipotesi quantistica**, che fu seguita anche nel 1905 da Einstein, per spiegare l'**effetto fotoelettrico**. Dopo pochi anni, la nascente Meccanica Quantistica (MQ) fu elaborata e fondata matematicamente: nel 1926 fu scritta l'equazione di Schrödinger. Nel 1935 il paradosso EPR pose un problema di *interpretazione* dei risultati finora ottenuti, di natura prettamente filosofica.

*Prima rivoluzione
quantistica*

Nel 1964 Bell scrisse delle disuguaglianze con cui rese sperimentalmente verificabili le conseguenze dell'EPR, misurando alcune osservabili di correlazione di sistemi in stati *entangled*. Un esempio è dato da misure di correlazione di spin di due particelle di $s = 1/2$ in uno stato di singoletto:

*Informazione
quantistica
(seconda
rivoluzione
quantistica)*

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$$

Ciò fu verificato nel 1982 dall'esperimento di Aspect, che mostrò la **violazione** delle disuguaglianze di Bell, aprendo un'analisi scientifica dei cosiddetti *fondamenti della MQ*, ossia i suoi aspetti più "filosofici".

Sempre nel 1982 Feynman elaborò la prima idea di **computer quantistico**, per risolvere il problema di simulazioni numerici di sistemi quantistici. Tale problema si rivela infatti *estremamente complesso*, dato che un qualsiasi insieme di molte particelle in genere si trova in stati altamente correlati, che sono particolarmente difficili da descrivere. L'idea fu allora quella di usare direttamente sistemi quantistici per tali simulazioni. Ciò portò, nel 1995, alla nascita della *teoria dell'informazione quantistica*, con la nozione di qubit. Nello stesso anno, Peter Shor, descrivendo il comportamento di un computer quantistico tramite le equazioni della MQ, trovò l'esistenza di un algoritmo per la fattorizzazione in numeri primi in tempo polinomiale (cosa che non si può fare con nessun algoritmo classico finora scoperto).

Tale algoritmo permetterebbe di violare le principali cifrature utilizzate nel mondo odierno (es. RSA). Una prospettiva del genere portò grandi finanziamenti per la teoria quantistica dell'informazione, e un intensificarsi della ricerca. Già nel 1997 fu realizzato sperimentalmente il primo **teletrasporto quantistico**.

Dagli anni 2000 in poi è stata realizzata sperimentalmente la **crittografia quantistica**, che *non può essere violata a priori* come conseguenza delle leggi fisiche per come le conosciamo ora (addirittura, un qualsiasi tentativo di intercettazione provoca l'autodistruzione del messaggio).

Ingegneria
quantistica

Sono poi stati realizzati alcuni primi piccoli computer quantistici, e anche *sensori quantistici* (apparati di misurazione che fanno uso di sistemi quantistici). Al giorno d'oggi disponiamo di *computer quantistici con rumore di taglia media*, che potrebbero essere disponibili tra qualche anno nei centri di supercomputing. Tali apparati potrebbero consentire di raggiungere la **quantum supremacy**, ossia la nascita di sistemi **più potenti** di qualsiasi supercomputer classico, potendo eseguire sia algoritmi quantistici senza corrispettivi classici, sia algoritmi classici in maniera più performante (per esempio l'algoritmo di Grover che, come vedremo, permette un guadagno di una radice quadrata nella complessità di cercare una *entry* in un database non organizzato).

Un utilizzo molto importante per i computer quantistici è quello delle **simulazioni** (in linea con la prima idea di Feynman), cosa che si sta già iniziando a fare in alcuni laboratori.

1.3 Concetti di informazione quantistica

1.3.1 La natura dell'informazione

L'informazione è fisica. Ciò può sembrar ovvio: un testo è fatto di molecole d'inchiostro su un foglio di *materia*, i *bit* sono elettroni confinati in cellette. Tuttavia, la conseguenza di ciò è *che l'informazione stessa deve obbedire alle leggi fisiche*. Manipolare l'informazione è infatti strettamente un *processo fisico*: è impossibile separare le *operazioni logiche* dalla loro effettiva *implementazione*. Per cambiare lo "stato logico di un sistema", ossia l'*informazione* che esso contiene, è per forza necessario utilizzare dei *processi fisici*.

Manipolazioni di
informazione

In particolare la **termodinamica**, e specialmente il suo **secondo principio**, deve valere anche per l'informazione. Ci chiediamo: è possibile estrarre energia utile da un sistema semplicemente osservandone lo stato, ossia avendo *informazione* su di esso? [3] Possiamo collegare *informazione* e *entropia*?

Risponderemo a tal domanda con un esperimento mentale, che è pittorescamente nominato **diavoleto di Maxwell**. In particolare, considereremo una sua variante, detta **motore di Szilard**, che vediamo rappresentato in figura 1.1.

Maxwell's demon

Consideriamo in 1.1.(a) un sistema costituito da due camere comunicanti *A* e *B*, che si trovano in un bagno termico a temperatura *T* fissata. Ipotizziamo che vi sia

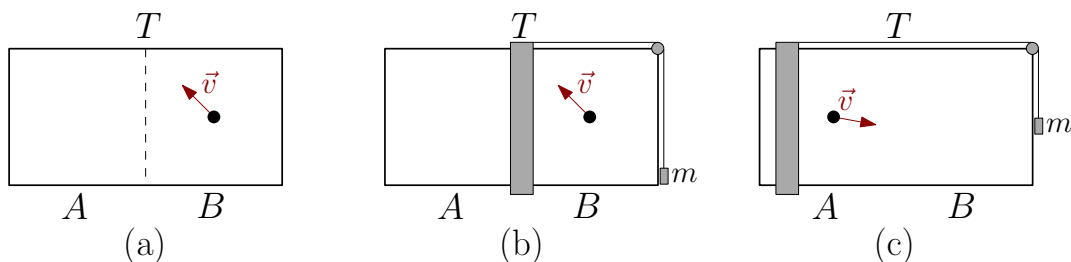


Figura (1.1) – Schema del funzionamento del motore di Szilard. In (a) si ha la configurazione iniziale, a temperatura T fissata. Si pone un pistone a dividere A e B , e dopo aver misurato la posizione della particella, in (b) si collega una massa m al pistone, dalla parte in cui si trova la particella. In tal modo, in (c) è possibile sfruttare il moto di quest'ultima per fare lavoro. Ipotizzando che il pistone (e carrucola e massa) si muovano senza attrito, il motore ha come unico effetto la produzione di *lavoro* senza alcuna spesa.

una singola particella P , con velocità \vec{v} , che effettua urti perfettamente elastici con le pareti del contenitore.

Se sappiamo che P si trova in B , cioè a destra, possiamo posizionare (in maniera reversibile, senza spendere lavoro) un pistone a separare A e B , collegato ad una massa m anch'essa posizionata a *destra* (figura 1.1.(b)). Prima o poi la particella si scontrerà con la barriera, e potremo quindi estrarre energia da essa (dato che sappiamo in che direzione mettere un *peso* per farlo sollevare dal passaggio, figura 1.1.(c)).

Se la particella si fosse trovata in A , cioè a sinistra, avremmo potuto posizionare il peso in maniera simmetrica - di nuovo senza alcuna differenza di energia spesa rispetto al caso precedente - e stavolta sfruttare il movimento di P da sinistra a destra.

Ipotizzando che il pistone si muova senza attrito, riposizionarlo non consuma alcuna energia (potremmo pensare ad un sistema di *molle* che consentano di passare reversibilmente da una configurazione all'altra), e perciò abbiamo appena ottenuto una macchina termica che ha come unico effetto quello di generare lavoro, in completa violazione del secondo principio della termodinamica.

La situazione si risolve se consideriamo, oltre al sistema di due camere, anche la *memoria* dell'osservatore. Per far funzionare il motore, infatti, abbiamo eseguito una misura di posizione della particella. Ciò, in principio, non crea problema - possiamo pensare che tale misura avvenga in maniera passiva, senza spese - ma presuppone la possibilità di **registrarne l'esito**, dato che poi dovremo basarci su di esso per posizionare la massa m a destra o a sinistra.

Perciò, se vogliamo riportare il motore di Szilard allo stato iniziale, per poi ripetere il ciclo ed estrarre energia, dobbiamo *cancellare* l'esito della misura precedente. Si dà il caso che tale operazione **consumi per forza energia**.

L'idea è data dal **principio di Landauer**. Consideriamo un registro a 1 bit in cui "salvare" la misura fatta in 1.1.(a). Tale registro è un **oggetto fisico**, realizzato con un sistema a due stati, che possiamo immaginare come una camera bipartita

da un pistone, con volume V e temperatura T fissati. Vi sono 2 configurazioni possibili, che corrispondono ai due possibili stati (0 o 1) del bit. Per esempio 0 potrebbe essere associato alla presenza della particella nella cella a sinistra, e 1 in quella a destra. Avremo allora un'entropia iniziale $S_i = k_B \log 2$.

Per cancellare il bit eliminiamo la divisione, portando a 1 il numero di stati, e quindi l'entropia S_f a 0. Abbiamo quindi una variazione $\Delta S = -k_B \log 2$.

Applicando allora il secondo principio della termodinamica, si ha che $\Delta S_{tot} = \Delta S_{amb} + \Delta S_{syst} \geq 0$, dove per ambiente intendiamo il bagno a temperatura T , mentre la ΔS_{syst} è stata appena calcolata.

Tuttavia, per una riserva termica infinita (perciò sempre all'equilibrio) possiamo usare la formula di Clausius, e quindi:

$$\Delta S_{amb} = \frac{Q_{amb}}{T}$$

Ne deriva che *cancellare* il bit richiede uno scambio di calore con la sorgente, e quindi una *dissipazione*:

$$Q_{amb} \geq k_B T \log 2$$

Applicando allora il primo principio della termodinamica, affinché l'energia del sistema non cambi (per poter rieseguire il ciclo), si avrà $Q = W$, e perciò la *cancellazione* del bit richiede una **spesa energetica** che al **minimo** è:

*Principio di
Landauer*

$$\Delta \mathcal{E} = kT \log 2$$

In altre parole, c'è un limite *minimo* all'energia spesa per una singola computazione *irreversibile* (come la cancellazione di un bit).

1.3.2 Definizioni di base

Possiamo schematizzare un **computer** come un oggetto che dato un *input* (condizione iniziale), costituito da un sistema fisico opportunamente preparato, lo trasforma tramite un *processo fisico* cambiandone lo *stato logico* (ossia le *informazioni* in esso codificate), e generando così un sistema che, una volta misurato, produce un *output* (risultato).

Computer

Un **computer quantistico** segue lo stesso schema, usando però per gli stati delle *funzioni d'onda*. Avremo quindi uno stato iniziale $|\psi_{in}\rangle$, che subisce un'evoluzione temporale tramite l'operatore $U = \exp(-\frac{i}{\hbar}tH)$, giungendo ad un output $|\psi_{out}\rangle = U|\psi_{in}\rangle$. Per poterlo usare, tuttavia, dobbiamo effettuare una misura, che selezionerà solo uno dei possibili esiti codificati da $|\psi_{out}\rangle$. Perciò, l'algoritmo quantistico deve far sì che l'esito desiderato possa essere estratto in maniera efficiente dalla funzione d'onda finale, cosa che in genere è un problema complesso.

*Computer
quantistico*

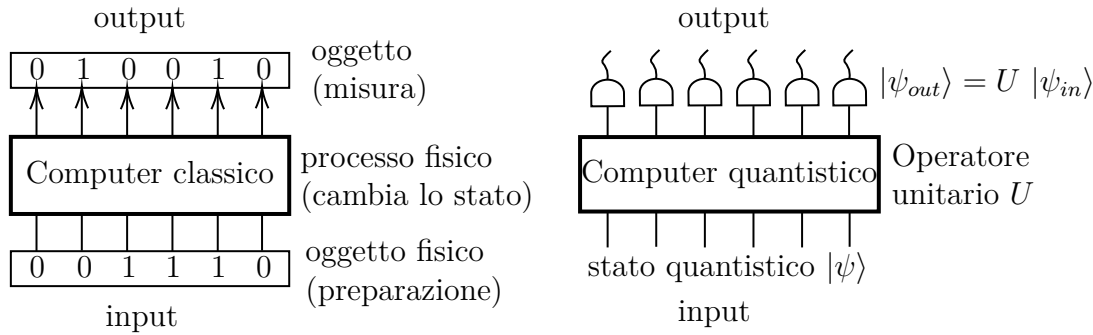


Figura (1.2) – Schema a blocchi di un computer classico o quantistico

Tuttavia, nel processo fisico possiamo usare tutte le possibilità offerte dalla MQ, come l'*entanglement*, il *principio di sovrapposizione*, il *teletrasporto*... con una **grave limitazione**: non è possibile creare una copia esatta di uno stato senza distruggere l'originale (**no cloning theorem**).

No cloning theorem

Strettamente collegato al concetto di **computazione** si trova l'idea di **comunicazione**.

Per **comunicazione classica** si intende il trasferimento di informazione da un mittente (**Alice**) e un destinatario (**Bob**), attraverso un determinato **canale**. Ciò è realizzato tramite un sistema in grado di trasferire *bit* di informazione (dato che un qualsiasi stato logico classico può essere scritto come un'opportuna sequenza di 0 e 1).

Nuovamente, la **comunicazione quantistica** segue lo stesso schema, ma questa volta si trasferisce una funzione d'onda $|\psi\rangle$ tramite un **canale quantistico**, per esempio attraverso *fotoni*. La maggiore libertà offerta dalla MQ offre possibilità più ampie rispetto a quelle permesse dalla MC, come la *crittografia quantistica*, il *trasporto di più informazione di quella classicamente permessa*, il *teletrasporto quantistico*...

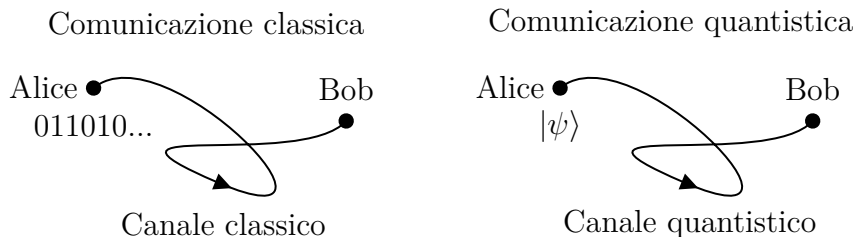


Figura (1.3) – Schema di una comunicazione classica o quantistica

Tramite **sensori quantistici** è possibile poi effettuare misure con target ben *più piccoli*, superando i limiti delle ampie medie temporali propri dei sensori classici.

1.4 Modello di calcolo a circuiti

1.4.1 Calcolo classico deterministico

Il modello di calcolo comunemente utilizzato nei computer si basa sul computare funzioni da n -bit a m -bit:

$$f : \{0, 1\}^n \mapsto \{0, 1\}^m \quad (1.1)$$

dove 0 e 1 sono rappresentati come **stati distinguibili** di un opportuno sistema classico.

Una computazione è schematizzata da *linee* che codificano lo stato di un *bit*, su cui si opera tramite **porte logiche**.

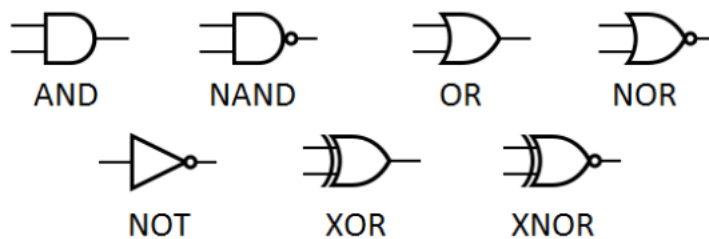


Figura (1.4) – Schemi delle principali porte logiche

Si può dimostrare che qualsiasi computazione del tipo (1.1) può essere realizzata come una combinazione di alcune **porte logiche fondamentali**, che ora specifichiamo.

Partiamo dalle porte logiche a **1 bit**. In tal caso abbiamo due sole possibilità: *identità* e *not*.

Buffer		Not	
a	a	a	\bar{a}
0	0	0	1
1	1	1	0

Tabella (1.1) – Tabella di verità per *not* (\bar{a}) e *identità* (buffer)

D'altro canto, le porte logiche fondamentali a **2 bit** sono *and* e *or*.

a	b	\wedge	\vee	\otimes
0	0	0	0	0
0	1	0	1	1
1	0	0	1	1
1	1	1	1	0

Tabella (1.2) – Tabella di verità per *and* $a \wedge b = ab$, *or* $a \vee b = a + b$ e *xor* $a \otimes b$

Due altre possibili operazioni sono il *copy* (copiare l'informazione) e lo *swap* (lo scambio di due stati):

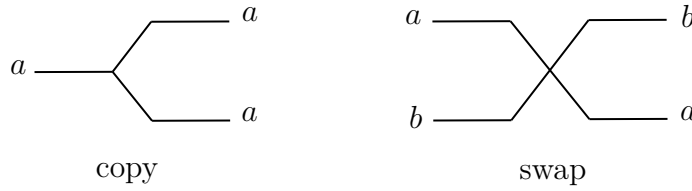


Figura (1.5) – Schema del funzionamento delle porte logiche **copy** e **swap**

Un *set univiale classico*, ossia l'insieme di operazioni logiche necessarie per eseguire una computazione logica generica, è dato da *and*, *or*, *not*, *copy*. Si può dimostrare che un set **minimale** è costituito da *copy* e una scelta tra *nand* e *nor* [6].

Notiamo che alcune di queste operazioni sono **irreversibili**, cioè dallo stato finale non è possibile risalire con univocità allo stato iniziale (es. per *or* e *and* input e output non sono in relazione biunivoca). Ciò è problematico, perché la MQ agisce in maniera unitaria e reversibile, e quindi di certo non è possibile trovare le *porte logiche quantistiche* per semplice trasposizione. Inoltre, come già accennato, non è possibile effettuare l'operazione di *copy* in MQ.

*Irreversibilità
delle porte logiche
classiche*

1.4.2 Calcolo quantistico

L'analogo quantistico di un bit classico è detto **qubit**, per cui si intende un set di due possibili stati (quantistici) distinguibili: *fondamentale* $|0\rangle$ ed *eccitato* $|1\rangle$.

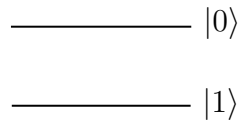


Figura (1.6) – Come nel caso classico, rappresentiamo i *qubit* come linee, che saranno opportunamente collegate agli input delle porte logiche

Dal principio di sovrapposizione della MQ deriva immediatamente la possibilità di creare stati *senza analogo classico*. Infatti, un generico qubit è dato da una

superposizione dei due stati possibili:

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle \quad \alpha_0, \alpha_1 \in \mathbb{C}, |\alpha_0|^2 + |\alpha_1|^2 = 1$$

Equivalentemente, riscrivendo α_0 e α_1 come fasi, otteniamo l'espressione:

$$|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} e^{i\varphi} |1\rangle \quad 0 \leq \theta \leq \pi, 0 \leq \varphi \leq \pi$$

Tale espressione ha un'interpretazione geometrica come un **vettore unitario** nella **sfera di Bloch** (figura 1.7).

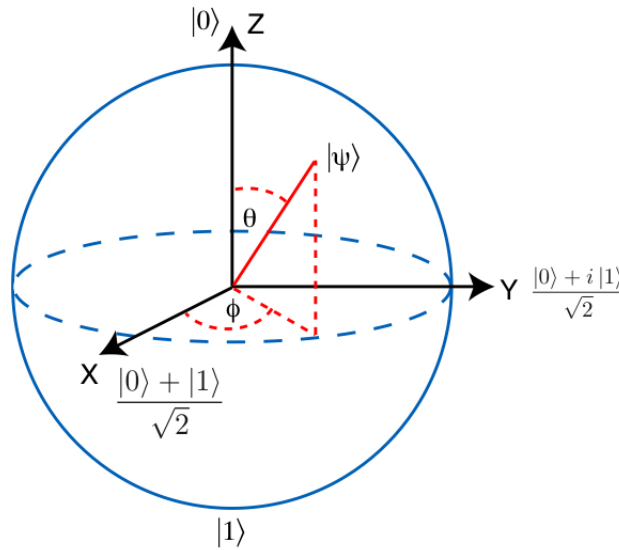


Figura (1.7) – Rappresentazione grafica della **sfera di Bloch**

Nota: Ricordiamo che moltiplicare per un numero complesso (o, nel caso normalizzato, per una **fase globale**) *non modifica lo stato*. In altre parole, $|\psi\rangle$ e $e^{i\alpha} |\psi\rangle$ rappresentano lo stesso identico stato - nel senso che misure di qualsiasi osservabile risultano in entrambi i casi negli stessi risultati.

Matematicamente ciò significa che gli stati di sistemi quantistici sono più precisamente **raggi vettori**, ossia elementi in un opportuno spazio proiettivo (\mathcal{PH}).

Ci chiediamo: quanta *informazione* (in senso classico) racchiude un qubit? Abbiamo due possibili risposte, in contraddizione tra di loro:

Informazione in un qubit

- Un qubit è l'analogo quantistico del bit, e quindi racchiude la *stessa quantità di informazione*
- Un qubit è univocamente definito da **due numeri complessi**, o da due angoli θ , φ , che hanno “infinite cifre” e quindi possono codificare *infinita informazione*. Per esempio potremmo scegliere (in rad) $\theta = 0.10010\dots$ e codificare nelle infinite cifre decimali un qualsiasi messaggio in binario.

In realtà, in pratica per produrre un output dovremo effettuare una **misura** sul qubit, che avrà due soli possibili risultati: $|0\rangle$ o $|1\rangle$. L'informazione sullo stato può essere ricavata solo da un grande insieme di qubit perfettamente identici, sui quali possiamo fare tante misure e, con tecniche di *tomografia quantistica*, ricostruire, almeno in principio, gli esatti coefficienti della $|\psi\rangle$ originaria. In particolare l'informazione estratta sarà proporzionale al numero di misure effettuate, ritornando così ad un caso equivalente ad un sistema classico di *molti bit*.

Vi è comunque un vantaggio a tutto ciò: i *numeri complessi* che definiscono la $|\psi\rangle$ sono **disponibili durante il processo quantistico** - cioè in ogni istante che precede la misura - e quindi possono influenzare la sua evoluzione. In altre parole, *computazionalmente* è “come se avessimo più informazione disponibile”, anche se, alla fine, siamo costretti a scartarne la maggior parte, estraendo un risultato “condensato”.

Un **computer quantistico** si basa allora sull'inizializzare opportunamente un certo numero di qubit, che poi si evolvono in modo unitario attraverso opportune porte logiche.

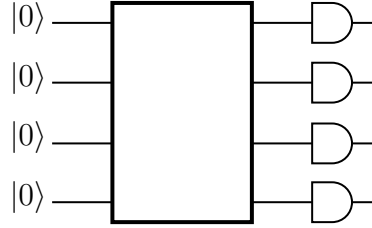


Figura (1.8) – Schema di un computer quantistico, dove l'input consiste in 4 qubit inizializzati a $|0\rangle$

Ad un certo istante dell'evoluzione, il computer quantistico si trova in un generico stato $|\psi\rangle$, che consiste nello stato di n qubit, dato da una delle possibili combinazioni lineari (con coefficienti $\Psi_{\vec{\alpha}}$) dei prodotti tensori degli stati $|\alpha_i\rangle = \{|0\rangle, |1\rangle\}$ dei singoli qubit:

$$\begin{aligned} |\psi\rangle &= \sum_{\vec{\alpha}} \Psi_{\alpha_1, \alpha_2, \dots, \alpha_N} \left(\bigotimes_{i=1}^N |\alpha_i\rangle \right) = \\ &\equiv \sum_{\vec{\alpha}} \Psi_{\alpha_1, \dots, \alpha_N} |\alpha_1, \dots, \alpha_N\rangle \end{aligned}$$

dove i coefficienti $\Psi_{\vec{\alpha}} \in \mathbb{C}$ sono opportunamente normalizzati:

$$\sum_{\vec{\alpha}} |\Psi_{\vec{\alpha}}|^2 = 1$$

Utilizzeremo spesso la notazione *sintetica*:

$$\bigotimes_{i=1}^N |\alpha_i\rangle = |\alpha_1\rangle |\alpha_2\rangle \dots |\alpha_N\rangle \equiv |\alpha_1, \dots, \alpha_N\rangle$$

Nota: in informazione quantistica supponiamo sempre di avere a che fare con qubit che sono **distinguibili** tra loro, e quindi non avremo problemi di statistica fermionica/bosonica (che comunque iniziano a essere esplorati da alcuni studi attuali).

Nella pratica non è per nulla semplice poter realizzare questi stati. $|0\rangle$ e $|1\rangle$ di un singolo qubit sono in genere separati da una piccolissima quantità di energia, che può trovarsi nel range delle particelle del bagno termico in cui si trova il sistema. In altre parole, effetti termici possono modificare lo stato di input, rendendolo inutilizzabile. Ecco perché i computer quantistici sono realizzati in sistemi controllati tramite avanzate tecniche criogeniche.

Analogamente al caso classico, si può dimostrare che **un qualsiasi operatore quantistico** O che agisce su n qubit può essere **scomposto** come l'azione combinata di opportune **porte logiche quantistiche fondamentali**. In genere, trovare tale combinazione non è però un problema banale - spesso si hanno solo risultati di esistenza, o bound sul numero di gate necessari.

Ogni operatore è la composizione di porte logiche

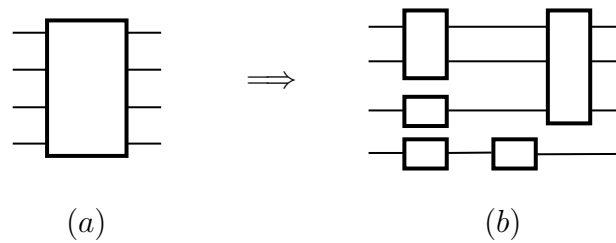


Figura (1.9) – Un qualsiasi operatore O (a) che esegue una certa operazione su n qubit può essere scomposto in una serie di porte logiche fondamentali (b) opportunamente collegate che svolgono la medesima operazione.

1.5 Porte logiche quantistiche

Nella seguente sezione ci occuperemo di esaminare le principali porte logiche utilizzate in computazione quantistica.

1.5.1 Porte a 1 qubit

Le porte a 1 qubit sono rappresentate da matrici unitarie 2×2 , di solito espresse nella **base computazionale** $\{|0\rangle, |1\rangle\}$. Nella notazione matriciale ricordiamo che:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Già nel caso a 1 bit vi sono porte *senza alcun analogo classico*, che offrono quindi una maggiore libertà nelle manipolazioni di stati logici.

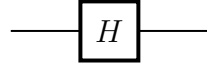
Un primo esempio è dato dalla **porta di Hadamard**, la cui rappresentazione

Porta di Hadamard

matriciale è data da:

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

e il cui elemento circuitale è rappresentato come:



Hadamard (H)

Facendola agire sugli stati della base, tramite un *prodotto di matrice per vettore*, otteniamo la sua *tavola di verità*:

$$\begin{aligned} H|0\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |+\rangle_x \\ H|1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |-\rangle_x \end{aligned}$$

L'azione di H può essere schematizzata come un passaggio dalla base degli autostati dell'operatore σ_z (che possiamo far coincidere con i $|0\rangle$ e $|1\rangle$ della base computazionale) a quella dell'operatore $\sigma_x \{|+\rangle, |-\rangle\}$ (l'equivalente di una *rotazione*, se fossimo nel comune spazio cartesiano in $d = 3$)¹.

Unitarietà. Come affermato, tutti i gate quantistici fondamentali sono **unitari**, ossia rappresentati da matrici unitarie (e hermitiane). Nel caso dell'Hadamard ciò significa che $H = H^\dagger = H^{-1}$, e quindi:

$$H|0\rangle \mapsto |+\rangle_x \Rightarrow H|+\rangle_x \mapsto |0\rangle; \quad H|1\rangle \mapsto |-\rangle_x \Rightarrow H|-\rangle_x \mapsto |1\rangle$$

Un'altra porta quantistica fondamentale (e *senza analogo classico*) è quella che “aggiunge una fase relativa”, detta di **phase-shift**:

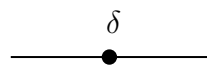
Operatore
phase-shift

$$R_z(\delta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{pmatrix}$$

che agisce come:

$$R_z(\delta)|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i(\varphi+\delta)}\sin\frac{\theta}{2}|1\rangle$$

ed è rappresentata nei circuiti mediante il simbolo:



Phase shift ($R_z(\delta)$)

¹^Qui le σ_i indicano le **matrici di Pauli**.

Si dimostra, come verificheremo più avanti, che un **generico stato** $|\psi\rangle$ di un singolo qubit può essere ottenuto partendo da $|0\rangle$ applicando una successione di porte logiche a 1-bit:

$$|\psi\rangle = R_z\left(\frac{\pi}{2} + \varphi\right) H R_z(\theta) H |0\rangle$$

Porte logiche classiche. Le porte logiche classiche a 1 bit, essendo reversibili, possono essere adattate naturalmente al caso quantistico. In particolare il **buffer** è descritto da una matrice identità, mentre il NOT si ottiene “invertendo” le colonne (e righe) di \mathbb{I} :

$$\text{BUFFER} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{I}_2 \quad \text{NOT} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_x$$

1.5.2 Porte a 2 qubit

Un generico stato di 2 qubit è una combinazione di 4 possibili stati (in generale, n qubit sono combinazioni di 2^n possibilità):

$$|\psi\rangle = \alpha_{00} |00\rangle + \alpha_{01} |01\rangle + \alpha_{10} |10\rangle + \alpha_{11} |11\rangle$$

con la normalizzazione:

$$\sum_{n=0}^3 |\alpha_n|^2 = 1 \quad \alpha_n \in \mathbb{C}$$

dove n va da 0 a 3 e “codifica” le 4 possibilità dei due indici binari ($0 = 00$, $1 = 01$, $2 = 10$, $3 = 11$). Useremo spesso questa “notazione mista” poiché la forma in **binario** è utile per riconoscere immediatamente a quale autoket si riferisca un certo coefficiente, mentre quella **decimale** è comoda per gli indici delle sommatorie.

La **base computazionale** $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, in notazione matriciale, è data da:

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

La prima porta logica che analizziamo è detta **CNOT** (*controlled not*) e corrisponde ad un’operazione su un bit *condizionata* da un altro bit. In particolare CNOT esegue un’operazione NOT sul secondo qubit solo se il primo è pari a 1, e non fa nulla altrimenti (ossia funge da “buffer” se il primo bit è 0). Si tratta di un’operazione di not “controllata”, dato che viene “accesa” solo se il bit di controllo è pari a 1.

Gate CNOT

In notazione matriciale, usando la base $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ già introdotta, CNOT è descritta da:

$$\text{CNOT} = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array} \right)$$

In questa notazione, le prime due colonne (a sinistra) descrivono il funzionamento quando il bit di controllo - che è quello *più significativo* (= più a sinistra) nella notazione binaria - è 0, e le ultime due quando è 1.

Perciò gli stati $|00\rangle$ e $|01\rangle$, dove tale bit di controllo è nullo, vengono mandati in se stessi senza variazioni, mentre per $|10\rangle$ e $|11\rangle$ il risultato si ottiene *invertendo* il secondo bit (quello *meno significativo* nella notazione binaria).

Possiamo schematizzare il comportamento usando un analogo delle tavole di verità classiche (tabella 1.3), ricordando però che lo stato iniziale (con bit di controllo indicato con c e secondo bit con a) può essere una *combinazione lineare* degli stati possibili di 2 qubit, e di conseguenza l'output sarà anch'esso una combinazione lineare. In altre parole, in una tavola di verità quantistica è possibile che “più righe siano applicate contemporaneamente”.

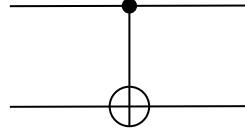
c	a	\bar{a}	out	
$ 0\rangle$	$ 0\rangle$	$ 1\rangle$	$ 0\rangle$	$ 0\rangle$
$ 0\rangle$	$ 1\rangle$	$ 0\rangle$	$ 0\rangle$	$ 1\rangle$
$ 1\rangle$	$ 0\rangle$	$ 1\rangle$	$ 1\rangle$	$ 1\rangle$
$ 1\rangle$	$ 1\rangle$	$ 0\rangle$	$ 1\rangle$	$ 0\rangle$

Tabella (1.3) – Tavola di verità per il gate CNOT

Nota: come si nota dalla tavola di verità la porta CNOT ha 2 output: uno è quello condizionale descritto, e l'altro è sempre pari al qubit c di controllo. In particolare, non è possibile realizzare una porta quantistica a 2 bit con meno di 2 output, dato che in tal caso non si potrebbe avere corrispondenza biunivoca tra stati finali e iniziali, violando l'unitarietà dell'evoluzione quantistica (torneremo su questa *reversibilità* nei prossimi paragrafi).

Porte logiche classiche a 2 bit. Poiché le porte logiche quantistiche devono essere reversibili (essendo operazioni unitarie), non vi è modo di implementare *direttamente* AND o OR (e relative), dato che associano 2 input a un solo output. Vedremo però più avanti come “reversibilizzare” una qualsiasi funzione classica, in modo da adattarla naturalmente al mondo quantistico.

Nei circuiti il gate CNOT ha la seguente rappresentazione:



Controlled NOT (CNOT)

Poiché il primo qubit in generale sarà una combinazione lineare di $|0\rangle$ e $|1\rangle$, la CNOT crea una *correlazione* tra i due bit, portando alla generazione di **stati entangled**, ossia stati che non sono rappresentabili come un prodotto tensore “puro”:

Stato entangled

$$|\psi\rangle \neq |\phi_1\rangle \otimes |\phi_2\rangle$$

Questo succede poiché uno stato generale è una combinazione lineare degli stati realizzati dai prodotti tensori, e solo alcune combinazioni sono anche *fattorizzabili*.

Vediamo esplicitamente come la CNOT possa creare stati entangled. Partiamo da uno stato iniziale dato da:

La CNOT può creare stati entangled

$$|\psi_i\rangle = (\alpha |0\rangle + \beta |1\rangle) \otimes |0\rangle$$

In cui il primo qubit (che è quello “di controllo”) è una combinazione di $|0\rangle$ e $|1\rangle$, ossia ha una probabilità $|\alpha|^2$ di essere $|0\rangle$ e $|\beta|^2$ di essere $|1\rangle$. Ciò si traduce in una certa probabilità che il qubit *a* sia *negato*, e in un'altra che invece non sia modificato - fino a che non si effettua una misura, i due percorsi sono *entrambi realizzati*. Infatti, operando su $|\psi_i\rangle$ con CNOT otteniamo:

$$\begin{aligned} \text{CNOT}(|\psi\rangle_i) &= \text{CNOT}(\alpha |00\rangle + \beta |10\rangle) \underset{(b)}{=} \text{CNOT}(\alpha |00\rangle) + \text{CNOT}(\beta |10\rangle) = \\ &= \alpha |00\rangle + \beta |11\rangle \end{aligned}$$

dove in (a) abbiamo distribuito il prodotto tensore, e in (b) abbiamo usato la linearità di CNOT (dato che è un operatore unitario, e quindi lineare - infatti ha una rappresentazione matriciale).

Lo stato finale è quindi **entangled**, dato che è **combinazione** di prodotti tensori, e non può essere scritto come **un solo** prodotto tensore “puro”. Si ha quindi una *correlazione* tra due parti diverse del sistema - ciò sarà fondamentale per la computazione quantistica, dato che permette di sfruttare la maggiore libertà dell'*entanglement*, un fenomeno senza analogo classico.

Un altro gate importante è l'analogo di CNOT per le fasi, che effettua una *rotazione* di un qubit (introducendo su di esso una fase relativa) a seconda del valore di un altro qubit. Tale gate è detto **C-PHASE** (*controlled phase*), e ha rappresentazione matriciale:

Gate C-PHASE

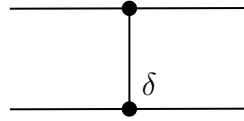
$$\text{CPHASE} = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\delta} \end{array} \right) = \begin{pmatrix} \mathbb{I}_2 & \mathbb{O}_2 \\ \mathbb{O}_2 & R_z(\delta) \end{pmatrix}$$

Analogamente al caso di CNOT, quando il bit di controllo c è 1 allora viene applicata una $R_z(\delta)$ all'altro bit a , e altrimenti non si fa nulla, come mostrato nella *tavola di verità* in tabella 1.4.

c	a	out
$ 0\rangle$	$ 0\rangle$	$ 0\rangle 0\rangle$
$ 0\rangle$	$ 1\rangle$	$ 0\rangle 1\rangle$
$ 1\rangle$	$ 0\rangle$	$ 1\rangle R_z(\delta) 0\rangle$
$ 1\rangle$	$ 1\rangle$	$ 1\rangle R_z(\delta) 1\rangle$

Tabella (1.4) – Tavola di verità per la porta CPHASE

La rappresentazione circuitale per il gate CPHASE è la seguente:



Controlled phase (CPhase)

1.5.3 Funzioni di qubit

Occupiamoci ora di rappresentare una funzione binaria, del tipo generico:

$$f : \{0, 1\}^n \rightarrow \{0, 1\}$$

Il problema è che, in generale, le $f(x)$ **non sono biunivoche**, e quindi non possono essere rappresentate da una serie di operazioni unitarie, in cui dallo stato finale si può recuperare l'informazione sullo stato iniziale in modo *univoco*.

*Computazione
quantistica è
reversibile*

Per risolvere la situazione dobbiamo estendere tali generiche funzioni a una classe di **funzioni reversibili**. Il modo più semplice per farlo è “portarsi dietro l’input”, raddoppiando il numero di gradi di libertà n . Avremo quindi un circuito che in input ha un registro di n -qubit contenente lo stato iniziale $|x\rangle$, e uno di pari dimensioni (altri n qubit) che contiene $|0\rangle$. L’output sarà allora $|x\rangle$ per i primi n qubit (ossia di nuovo l’input) e per gli ultimi $|f(x)\rangle$. Si ha quindi un circuito che, tenendo costante il primo l’input, modifica il secondo tramite opportune funzioni. In realtà lo schema si può generalizzare, partendo per il secondo registro da un generico $|y\rangle$, che si fa evolvere a $|f(x) + y\rangle$ (figura 1.10).

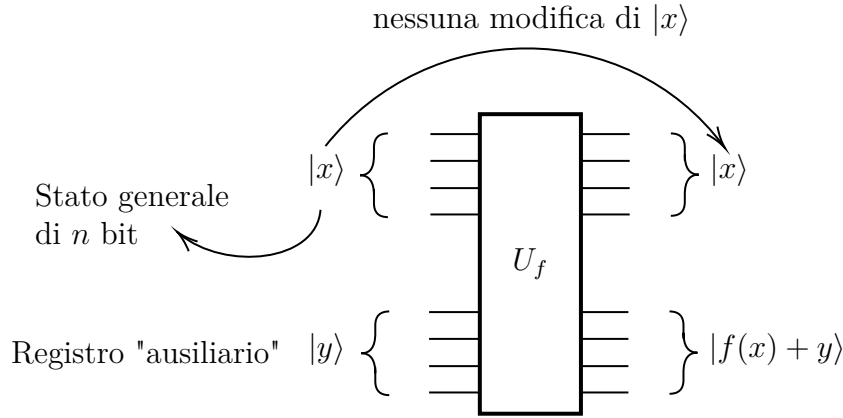


Figura (1.10) – Si può estendere una qualsiasi funzione binaria $f : \{0, 1\}^n \rightarrow \{0, 1\}^m$ a una **funzione reversibile** $f' : \{0, 1\}^n \times \{0, 1\}^m \rightarrow \{0, 1\}^n \times \{0, 1\}^m$ dove i primi n qubit dell’output sono esattamente pari ai primi n dell’input.

Avremo cioè:

$$|\psi_f\rangle = U_f |\psi_i\rangle = U_f |x\rangle |y\rangle = |x\rangle |y + f(x)\rangle$$

In generale, potremmo usare come stato iniziale una combinazione dei $2^n - 1$ possibili stati del registro $|x\rangle$ di n -qubit (fissando $|y\rangle$ che funge da “offset” dell’output), con opportuni coefficienti α_x :

$$U_f \sum_{x=0}^{2^n-1} \alpha_x |x\rangle |y\rangle = \sum_{x=0}^{2^n-1} c_x |x\rangle |y + f(x)\rangle$$

Così facendo stiamo calcolando *contemporaneamente* U_f su *ognuna* delle combinazioni. Questo è il fenomeno del **parallelismo quantistico**.

*Quantum
parallelism*

Unico problema: per leggere il risultato dobbiamo fare una misura, e quindi effettivamente avremo accesso a solo *uno dei valori*. Per sapere tutti i risultati dovremo ripetere l’esperimento un numero di volte di ordine 2^n e perciò non si hanno, per ora, particolari guadagni rispetto al caso classico.

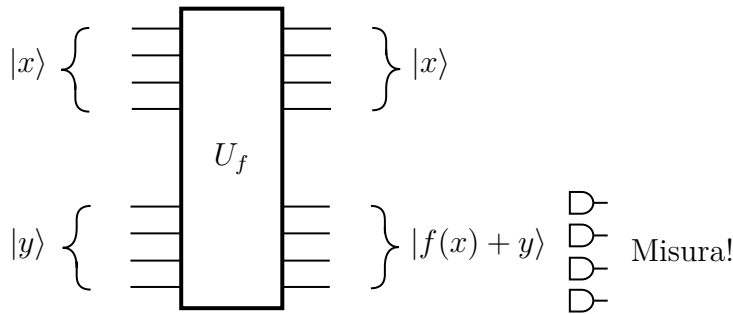


Figura (1.11) – Seppure una computazione quantistica possa avvenire in “parallelo” su tutte le singole configurazioni di cui lo stato iniziale è combinazione lineare, per poter “usare” l’output è necessario effettuare una misura, che collassa lo stato finale $|\psi_o\rangle = \sum_{x=0}^{2^n-1} c_x |x\rangle |y + f(x)\rangle$ in **uno solo** dei 2^n esiti possibili, con probabilità $|c_x|^2$

Si prospetta però la possibilità di regolare opportunamente lo stato finale in modo che l'esito desiderato sia quello dominante, ossia con la maggiore probabilità di presentarsi. In tal modo si può ottenere (con alta probabilità) la risposta voluta.

1.5.4 Esempio di funzione quantistica

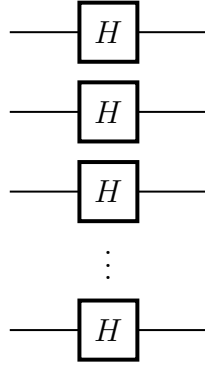
Assumiamo di poter configurare n qubit nello stato $|0000 \dots 0\rangle$. Vogliamo trovare un'operazione per farlo evolvere a uno stato del tipo:

$$\underbrace{|0000 \dots 0\rangle}_n \mapsto \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle$$

Per esempio, nel caso $n = 2$ avremo:

$$|00\rangle \mapsto \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$$

Stiamo cioè mappando uno stato *nullo* in una combinazione lineare di *tutti gli stati possibili* a pari coefficienti. Un modo per far ciò consiste nell'usare n porte Hadamard:



Riscriviamo l'azione di H in modo compatto come:

$$H|x\rangle = \frac{1}{\sqrt{2}} \sum_{y=0,1} (-1)^{xy} |y\rangle$$

che riproduce la stessa casistica definita in precedenza. Infatti se $x = 0$, allora $(-1)^0 = 1$ per un qualsiasi valore di y , e quindi avremo $H|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, mentre per $x = 1$ avremo un (-1) quando anche $y = 1$, e quindi $H|1\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$, come richiesto.

Se abbiamo due qubit, $|x_1\rangle$ e $|x_2\rangle$, e li mandiamo ciascuno attraverso una H , otteniamo:

$$\begin{aligned} (H \otimes H) |x_1\rangle |x_2\rangle &= \frac{1}{\sqrt{2}} \left(\sum_{y_1=0}^1 (-1)^{x_1 y_1} |y_1\rangle \right) \frac{1}{\sqrt{2}} \left(\sum_{y_2=0}^1 (-1)^{x_2 y_2} |y_2\rangle \right) = \\ &= \frac{1}{2} \left(\sum_{y_1, y_2=0}^1 (-1)^{x_1 y_1 + x_2 y_2} |y_1\rangle |y_2\rangle \right) \end{aligned}$$

Generalizzando a n qubit, e schematizzando la somma dei prodotti di bit corrispondenti con un *prodotto scalare* $\vec{x} \cdot \vec{y}$, con \vec{x} e \vec{y} vettori di bit:

$$(H \otimes \cdots \otimes H) |x_1\rangle \dots |x_n\rangle = \frac{1}{(\sqrt{2})^n} \sum_{\{y_i\}=0}^1 (-1)^{\vec{x} \cdot \vec{y}} |y_1, \dots, y_n\rangle = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} (-1)^{\vec{x} \cdot \vec{y}} |\vec{y}\rangle$$

Se $\vec{x} = (0, \dots, 0)$ il coefficiente $(-1)^{\vec{x} \cdot \vec{y}}$ sparisce, e quindi otteniamo l'output desiderato:

$$|\psi_f\rangle = \sum_{y=0}^{2^n-1} |\vec{y}\rangle$$

Nota: nonostante la complessità della forma finale, non si ha alcuna correlazione tra i vari qubit - infatti lo stato finale è un prodotto tensore di H applicate ciascuna a *una singola entrata*, ossia **non** è uno stato entangled. Pittorescamente: stiamo agendo sui singoli qubit *uno alla volta*, e quindi nessuno “è a conoscenza” dello stato degli altri - perciò non possono esservi effetti di interferenza.

1.6 No cloning theorem

Abbiamo accennato che in MQ non è possibile copiare esattamente uno stato. Vediamo ora esattamente perché..

Cerchiamo un'operazione che possa copiare uno stato $|\psi\rangle$, indipendentemente dal suo stato. Formalmente, vogliamo scrivere una certa U tale che:

$$U |\psi\rangle |0\rangle = |\psi\rangle |\psi\rangle \quad U \text{ unitaria, } \forall |\psi\rangle \quad (1.2)$$

(dove stiamo usando due registri di pari dimensione per garantire l'unitarietà dell'operazione).

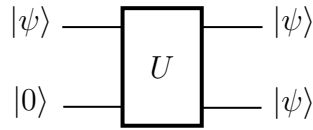


Figura (1.12) – Schema di una porta logica *copy* quantistica

Il **no cloning theorem** afferma che tale U **non esiste**.

Dimostriamolo. Supponiamo che esista una U del genere, per cui $\forall |\psi\rangle$ vale $U |\psi\rangle |0\rangle = |\psi\rangle |\psi\rangle$. Consideriamo due stati $|\Psi_1\rangle = |\psi\rangle |0\rangle$ e $|\Psi_2\rangle = |\varphi\rangle |0\rangle$. Applicando U a ciascuno di essi si ha:

Dimostrazione

$$\begin{aligned} |\Phi_1\rangle &= U |\Psi_1\rangle = U |\psi\rangle |0\rangle \stackrel{(1.2)}{=} |\psi\rangle |\psi\rangle \\ |\Phi_2\rangle &= U |\Psi_2\rangle = U |\varphi\rangle |0\rangle \stackrel{(1.2)}{=} |\varphi\rangle |\varphi\rangle \end{aligned} \quad (1.3)$$

Calcoliamo il prodotto scalare dei due stati finali:

$$P = \langle \Phi_2 | \Phi_1 \rangle$$

Abbiamo due modi per farlo:

- Usiamo per $|\Phi_1\rangle$ e $|\Phi_2\rangle$ l'espressione in termini di U unitario.

$$P = \langle \varphi | \langle 0 | U^\dagger U | \psi \rangle | 0 \rangle \underset{(a)}{=} \langle \varphi | \langle 0 | \psi \rangle | 0 \rangle = \langle \varphi | \psi \rangle$$

dove in (a) si è usata l'unitarietà di U , per cui vale $U^\dagger U = \mathbb{I}$.

- Usiamo la definizione di U , ossia la (1.3):

$$P = \langle \varphi | \langle \varphi | \psi \rangle | \psi \rangle = \langle \varphi | \psi \rangle^2$$

Ma allora deve essere:

$$\langle \varphi | \psi \rangle = \langle \varphi | \psi \rangle^2$$

cosa che **non è vera** in generale per ogni stato. Perciò *non è possibile* clonare un generico stato senza conoscerlo.

Dal risultato trovato, tuttavia, notiamo che stati ortogonali, per cui $\langle \varphi | \psi \rangle = 0$, possono però essere clonati senza problemi: possiamo quindi clonare tutti gli stati *di una base ortonormale*, ma non le loro generiche combinazioni (che formano stati generici).

Dimostrazione alternativa: proviamo a clonare una sovrapposizione, e in un caso otteniamo uno stato separabile $U(|x\rangle + |y\rangle) \otimes |0\rangle = (|x\rangle + |y\rangle)^{\otimes 2}$. Ma per linearità di U possiamo anche guardarla come $U(|x\rangle + |y\rangle) \otimes |0\rangle = |x\rangle |x\rangle + |y\rangle |y\rangle$, che in generale sarà uno stato entangled.

1.7 Classi di complessità algoritmica

I computer quantistici permettono di raggiungere una maggiore efficienza rispetto ai computer classici. Per quantificare tale fenomeno si utilizza una definizione di **classi di complessità**, che permettono di quantificare “quanto un algoritmo è complesso”. L'idea è quella di esaminare come la quantità di tempo/risorse necessarie a risolvere un problema *scali* al crescere della dimensione del problema. Per esempio, se consideriamo un algoritmo per scomporre un numero in numeri primi, la sua complessità è legata a *quanto velocemente* il tempo o le risorse necessarie alla computazione crescano all'aumentare del *numero di cifre* del numero dato in input.

Si trova che, in generale, *tempo* e *risorse* (la “memoria” utilizzata dal computer) sono correlate tra loro: è possibile usare più di una per compensare l'altra. Tuttavia, un problema di una certa complessità richiede in ogni caso una certa quantità

minima *globale* di risorse.

In particolare, consideriamo la funzione $f(n)$ che, data la grandezza n dell'input, ossia il *numero di bit* in ingresso, computa la **quantità di risorse necessarie** alla computazione. Abbiamo due possibilità:

Problemi facili e difficili

- $f(n)$ ha un andamento **polinomiale** in n (P), cioè è maggiorabile con un polinomio: $\exists N \in \mathbb{N}$ t.c. $f(n) \leq n^N$ definitivamente.
I problemi in cui le risorse scalano in modo polinomiale sono considerati **facili**.
- Se invece la funzione è **superpolinomiale**, cioè cresce più velocemente di ogni polinomio, il problema è ritenuto **difficile**.

Troviamo quindi tre classi:

Classi P, NP, NPC

- **P**: problemi risolvibili in tempo **polinomiale**.
- **NP**: *Non-deterministic Polynomial time*. Si tratta di problemi che una macchina di Turing non deterministica (che specifichiamo più avanti) può risolvere in tempo polinomiale. Si trova che i problemi NP corrispondono anche ai problemi la cui soluzione può essere **verificata** in tempo polinomiale da una macchina di Turing deterministica (cioè si ha che fare con “soluzioni ovvie a posteriori ma difficili da trovare in primo luogo”).
- **NPC**: *NP-Complete*, è composta da problemi la cui soluzione può essere riadattata *in tempo polinomiale* alla soluzione di un problema NP. In altre parole, gli NPC sono almeno tanto difficili quanto il più difficile degli NP, e trovando un algoritmo polinomiale che risolva un solo NPC è automaticamente possibile risolvere tutti gli NP sempre in tempo polinomiale, semplicemente riadattando opportunamente la soluzione trovata.

La **macchina di Turing** è l'oggetto astratto che *schematizza* l'azione di ogni computer. Consideriamo un *nastro infinito*, diviso in *cellette* che fungono da **registri**, su cui una macchina può scrivere con un certo *alfabeto* fissato (che può essere per esempio quello binario).

Macchina di Turing

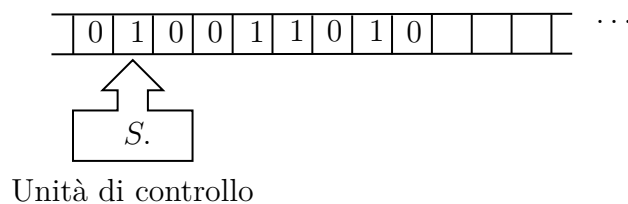


Figura (1.13) – Schema della macchina di Turing

Sul nastro agisce, tramite operazioni di **lettura/scrittura**, una **unità di controllo** S , che consiste in una **macchina a stati finiti**. Ciò significa che S può

trovarsi, in un dato istante, in uno di $N < \infty$ stati possibili. Ogni stato definisce il comportamento della macchina, e vi sono *regole* per passare da uno stato all'altro. Un esempio di comportamento associato ad uno stato potrebbe essere “leggi il bit corrente”, oppure “fai scorrere di 1 il nastro in avanti”. Tra questi stati possibili vi è lo stato **halt** (H), che indica il **termine dell'esecuzione** del programma.

In particolare, una macchina di Turing **deterministica** è tale che le regole di passaggio da uno stato all'altro dell'unità di controllo sono deterministiche, ossia conoscendo lo stato corrente A e il valore di registro a cui la macchina ha accesso si sa automaticamente il valore del prossimo stato B che assumerà S . Se invece è possibile solo dare una *probabilità* per i passaggi di stato, la macchina è **non deterministica**. Avendo una maggiore libertà d'azione, una macchina non deterministica è *più potente* di una deterministica.

Macchine di Turing (non) deterministiche

Notiamo infine che non sappiamo la **gerarchia** delle classi di complessità, ossia non è chiaro se **NP** e **P** siano in realtà lo stesso insieme, cioè se esista o meno un algoritmo *efficiente* per risolvere *problemi difficili*.

$P=NP?$

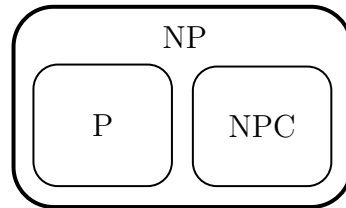


Figura (1.14) – Attualmente sappiamo che $NPC \subset NP$, e $P \subset NP$, ma non è chiaro se valgano (eventualmente) inclusioni inverse.

Considerando macchine di Turing non deterministiche, consideriamo altre classi di complessità:

- **BPP**: *Bounded Error Probabilistic Polynomial*. Un problema è in BPP se esiste un algoritmo polinomiale che dà il risultato giusto con una probabilità migliore del caso, ossia con $p = \frac{1}{2} + \delta$, con $\delta > 0$. Basta infatti tale *bias* per poter ripetere (in tempo polinomiale) l'algoritmo, migliorando “quanto si vuole” la probabilità di successo.
- **BQP**: *Bounded error Quantum Probabilistic Polynomial*. Un problema è in BQP se è risolvibile da un *computer quantistico* con una probabilità leggermente biased ($p = \frac{1}{2} + \delta$, $\delta > 0$). Per esempio, l'algoritmo di Shor per la fattorizzazione di numeri primi è di questo tipo.

Anche qui non sappiamo se tali insiemi siano gli stessi. Di sicuro vale:

$$P \subseteq BPP \subseteq BQP$$

Vi sono attualmente algoritmi quantistici *più performanti* di ogni algoritmo classico, ma nulla impedisce che in futuro si trovino nuovi algoritmi classici ancora più efficienti. Non è quindi ancora certo che i computer quantistici siano *a priori* in grado di superare i computer classici.

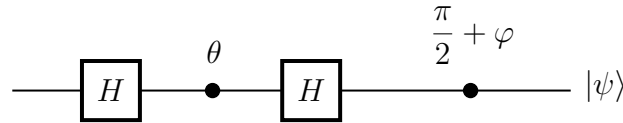
1.8 Esercizio 1

Vogliamo verificare che uno stato generico di un qubit possa essere realizzato partendo da $|0\rangle$ e applicando una precisa sequenza di gate:

$$\begin{aligned} |\psi\rangle &= \cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle = \\ &= R_z\left(\frac{\pi}{2} + \varphi\right) H R_z(\theta) H |0\rangle \end{aligned}$$

Graficamente ciò corrisponde a realizzare il circuito:

[Controllare il segno di φ]



Svolgiamo il conto, per esempio usando le rappresentazioni matriciali:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad R_z(\delta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{pmatrix}$$

Nella base computazionale ricordiamo che:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Applichiamo allora le porte logiche una dopo l'altra, svolgendo di volta in volta le moltiplicazioni *matrice per vettore*:

$$\begin{aligned} |0\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \xrightarrow{H} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \xrightarrow{R_z(\theta)} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{i\theta} \end{pmatrix} \xrightarrow{H} \begin{pmatrix} \frac{1+e^{i\theta}}{2} \\ \frac{1-e^{i\theta}}{2} \end{pmatrix} = e^{i\frac{\theta}{2}} \begin{pmatrix} \frac{e^{i\theta/2}+e^{-i\theta/2}}{2} \\ \frac{-e^{i\theta/2}+e^{-i\theta/2}}{2i} \end{pmatrix} = \\ &\stackrel{(a)}{=} e^{i\frac{\theta}{2}} \begin{pmatrix} \cos \frac{\theta}{2} \\ \underbrace{-i}_{e^{-i\pi/2}} \sin \frac{\theta}{2} \end{pmatrix} \xrightarrow{R_z(\frac{\pi}{2}+\varphi)} \left[\cos \frac{\theta}{2} |0\rangle + \exp\left(i\frac{\pi}{2} + i\varphi - i\frac{\pi}{2}\right) \sin \frac{\theta}{2} |1\rangle \right] = \\ &= \cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle \end{aligned}$$

In (a) rimuoviamo la fase globale (che non cambia lo stato).

1.9 Esercizio 2

Vogliamo verificare che l'operatore δ (C-PHASE) definito dalla matrice:

$$\delta = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\delta} \end{array} \right)$$

può essere ottenuto dalla combinazione di 5 porte logiche fondamentali (CNOT e phase-shift) come schematizzato in figura:

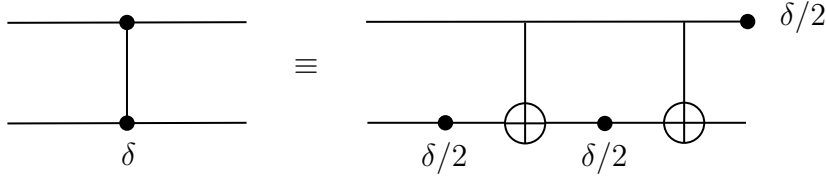


Figura (1.15) – Realizzazione dell’operatore δ tramite la composizione di 5 porte logiche quantistiche

Nota: per applicare una porta logica ad un solo bit in casi in cui si usano registri a n bit bisogna convertirla in una porta logica a n bit che effettua l’operazione solo sull’ i -esimo bit d’interesse e lascia invariati tutti gli altri. Ciò si effettua moltiplicando tensorialmente per l’identità \mathbb{I} . Per esempio il *phase-shift* sul primo di due bit si esprime come:

$$R_z^{(1)}(\delta) = R_z(\delta) \otimes \mathbb{I}$$

In notazione matriciale il prodotto tensore si ottiene partendo da una “matrice di matrici”, le cui entrate sono i prodotti tra il rispettivo membro della prima matrice e l’intera altra matrice. Per esempio, per X e Y matrici 2×2 :

$$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad Y = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

$$X \otimes Y = \begin{pmatrix} a \cdot Y & b \cdot Y \\ c \cdot Y & d \cdot Y \end{pmatrix} = \left(\begin{array}{cc|cc} a\alpha & a\beta & b\alpha & b\beta \\ a\gamma & a\delta & b\gamma & b\delta \\ \hline c\alpha & c\beta & d\alpha & d\beta \\ c\gamma & c\delta & d\gamma & d\delta \end{array} \right)$$

Nel caso della $R_z^{(1)}(\delta)$ di sopra otteniamo:

$$R_z^{(1)}(\delta) = \begin{pmatrix} 1 \cdot \mathbb{I} & 0 \cdot \mathbb{I} \\ 0 \cdot \mathbb{I} & e^{i\delta} \cdot \mathbb{I} \end{pmatrix} = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & e^{i\delta} & 0 \\ 0 & 0 & 0 & e^{i\delta} \end{array} \right)$$

Un modo è quindi procedere svolgendo la moltiplicazione matriciale:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\delta} \end{pmatrix} \stackrel{?}{=} \left[R_z\left(\frac{\delta}{2}\right) \otimes \mathbb{I} \right] \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array} \right) \left[\mathbb{I} \otimes \left(-\frac{\delta}{2}\right) \right] \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array} \right) \left[\mathbb{I} \otimes R_z\left(\frac{\delta}{2}\right) \right]$$

Nota: per le regole di composizione di applicazioni lineari, la prima matrice rappresenta l'**ultima** operazione che si effettua (quindi l'ordine delle matrici è "invertito" rispetto a quello in cui le porte logiche appaiono nel circuito. Ciò è fondamentale, poiché le matrici possono non commutare, e quindi i due ordini possono produrre risultati differenti (in questo caso specifico le matrici commutano).

Alternativamente (più veloce) si possono applicare in sequenza le varie porte logiche agli elementi della base computazionale.

Notiamo prima qualche "regola di conto veloce". Indicando con $(i)\delta$ il gate *phase-shift* che agisce sull' i -esimo qubit. Avremo che la δ aggiunge una fase di $e^{i\delta}$ solo se l' i -esimo bit è nello stato $|1\rangle$. La CNOT, d'altro canto, inverte $|1\rangle$ e $|0\rangle$ del secondo qubit **solo** se il primo qubit è $|1\rangle$.

Partiamo da $|00\rangle$, che si presenta come il caso più semplice, dato che la CNOT non fa nulla (il primo qubit è a $|0\rangle$) e nemmeno la *phase-shift* (sfasa solo il $|1\rangle$):

$$|00\rangle \xrightarrow{(2)\frac{\delta}{2}} |00\rangle \xrightarrow{\text{CNOT}} |00\rangle \xrightarrow{(2)-\frac{\delta}{2}} |00\rangle \xrightarrow{\text{CNOT}} |00\rangle \xrightarrow{(1)\frac{\delta}{2}} |00\rangle$$

Per semplicità, indicheremo un risultato con (id) se è esattamente pari a quello immediatamente precedente.

Nel caso di $|01\rangle$, invece, abbiamo una doppia fase aggiunta dalla δ :

$$|01\rangle \xrightarrow{(2)\frac{\delta}{2}} |0\rangle \otimes e^{i\frac{\delta}{2}} |1\rangle \xrightarrow{\text{CNOT}} (\text{id}) \xrightarrow{(2)-\frac{\delta}{2}} |01\rangle \xrightarrow{\text{CNOT}} (\text{id}) \xrightarrow{(1)\frac{\delta}{2}} |01\rangle$$

Analogamente ricaviamo gli ultimi due casi:

$$\begin{aligned} |10\rangle &\xrightarrow{(2)\frac{\delta}{2}} |10\rangle \xrightarrow{\text{CNOT}} |11\rangle \xrightarrow{(2)-\frac{\delta}{2}} |1\rangle \otimes e^{-i\frac{\delta}{2}} |1\rangle \xrightarrow{\text{CNOT}} |1\rangle \otimes e^{-i\frac{\delta}{2}} |0\rangle \\ &\xrightarrow{(1)\frac{\delta}{2}} e^{i\frac{\delta}{2}} |1\rangle \otimes e^{-i\frac{\delta}{2}} |0\rangle = |10\rangle \\ |11\rangle &\xrightarrow{(2)\frac{\delta}{2}} |1\rangle \otimes e^{i\frac{\delta}{2}} |1\rangle \xrightarrow{\text{CNOT}} |1\rangle \otimes e^{i\frac{\delta}{2}} |0\rangle \xrightarrow{(2)-\frac{\delta}{2}} (\text{id}) \xrightarrow{\text{CNOT}} |1\rangle \otimes e^{i\frac{\delta}{2}} |1\rangle \\ &\xrightarrow{(1)\frac{\delta}{2}} e^{i\frac{\delta}{2}} |1\rangle \otimes e^{i\frac{\delta}{2}} |1\rangle = e^{i\delta} |11\rangle \end{aligned}$$

Basta poi tornare in notazione matriciale per dimostrare il risultato.

1.10 Esercizio 3

Introduciamo la misura di **fidelity**, che valuta la "bontà" di uno stato rispetto ad un *target*. Per farlo si prende la probabilità di transizione tra i due stati: *Fidelity*

$$F = |\langle \psi_1 | \psi_2 \rangle|^2 \quad 0 \leq F \leq 1$$

che possiamo usare come una "misura di distanza" tra due stati.

1. Mostrare che F è una funzione monotona della norma (L_2) della differenza $\| |\psi_1\rangle - |\psi_2\rangle \|_2$ dei due stati (e quindi è una distanza).

2. Mostrare che $F = \cos^2 \frac{\theta}{2}$, dove θ è l'angolo fra i due vettori.

Risoluzione.

2. Partiamo dal punto 2. Potremmo calcolare il prodotto scalare prendendo due ψ generici e svolgere il conto, o notare che:

$$\langle \psi_1 | \psi_2 \rangle = \langle \psi_1 | U^\dagger U | \psi_2 \rangle$$

Possiamo scegliere U in maniera arbitraria. In particolare, per semplificare i conti, poniamo:

$$\begin{aligned} U | \psi_2 \rangle &\stackrel{!}{=} | 0 \rangle \\ \langle \psi_1 | U^\dagger &\equiv \langle \varphi | \end{aligned} \quad (1.4)$$

Dove φ è un generico stato:

$$| \varphi \rangle = \cos \frac{\theta}{2} | 0 \rangle + e^{i\varphi} \sin \frac{\theta}{2} | 1 \rangle$$

Così facendo il conto si semplifica in:

$$F = | \langle \psi_1 | \psi_2 \rangle |^2 = | \langle \psi_1 | U^\dagger U | \psi_2 \rangle |^2 \stackrel{(1.4)}{=} | \langle \varphi | 0 \rangle |^2 = \cos^2 \frac{\theta}{2} \quad (1.5)$$

1. Vogliamo mostrare che F è una funzione monotona della norma L_2 della differenza tra i vettori dei due stati $\| | \psi_1 \rangle - | \psi_2 \rangle \|_2$. Ricordiamo che tale norma è definita da:

$$\| | \phi \rangle \|_2 = \sqrt{ | \langle \phi | \phi \rangle | }$$

Procediamo calcolando direttamente la norma (quadra) della differenza:

$$\begin{aligned} \| | \psi_1 \rangle - | \psi_2 \rangle \|^2 &= | (\langle \psi_1 | - \langle \psi_2 |) (| \psi_1 \rangle - | \psi_2 \rangle) | = \\ &= | \underbrace{\langle \psi_1 | \psi_1 \rangle}_1 - \langle \psi_1 | \psi_2 \rangle - \langle \psi_2 | \psi_1 \rangle + \underbrace{\langle \psi_2 | \psi_2 \rangle}_1 | = \\ &= | 2 - (\langle \psi_1 | \psi_2 \rangle + \langle \psi_2 | \psi_1 \rangle) | \stackrel{(a)}{=} | 2 - 2 \operatorname{Re}(\langle \psi_1 | \psi_2 \rangle) | = \\ &\stackrel{(1.5)}{=} \left| 2 \left(1 - \cos \frac{\alpha}{2} \right) \right| = 2 (1 - \sqrt{F}) \end{aligned}$$

dove in (a) si è usata la definizione di parte reale di un numero complesso $a \in \mathbb{C}$:

$$\operatorname{Re} a = \frac{a + a^*}{2} \Rightarrow a + a^* = 2 \operatorname{Re} a$$

Dato che $\langle \psi_1 | \psi_2 \rangle$ e $\langle \psi_2 | \psi_1 \rangle$ sono complessi coniugati. Otteniamo perciò:

$$\| | \psi_1 \rangle - | \psi_2 \rangle \|_2 = \sqrt{2(1 - \sqrt{F})}$$

Effetti quantistici

2.1 Teletrasporto quantistico

(Lezione 3 • del
6/3/2019)

L'idea del teletrasporto quantistico consiste nel partire da un **qubit**, definito da una funzione d'onda generica $|\psi\rangle$ **non nota**, e trasmettere l'informazione in esso contenuta ad un qubit di un altro laboratorio. Dato che non è possibile **clonare** $|\psi\rangle$, il teletrasporto ha come effetto la distruzione del qubit originale.

Nota: stiamo comunque parlando di un trasporto di **informazione**, non di materia. Nel laboratorio di arrivo deve essere presente una particella a cui “applicare” il qubit trasportato.

Il protocollo di teletrasporto quantistico che esamineremo richiede:

- Due qubit in uno stato EPR (cioè in stato *massimamente entangled*), che vengono separati e portati uno al laboratorio di Alice (A) e l'altro a quello di Bob (B).
- Un qubit (C) nel laboratorio di Alice il cui contenuto sarà *teletrasportato* nel qubit B di Bob.
- Due bit classici che permettano ad Alice di registrare il risultato di due misurazioni, e trasmettere tale esito a Bob mediante un *canale di comunicazione classico*. Poiché tale passaggio di informazione può avvenire solo a velocità $v < c$, si ha che il teletrasporto quantistico non può essere un processo superluminale, e quindi non viola la relatività speciale.

Consideriamo quindi il setup di figura 2.1.

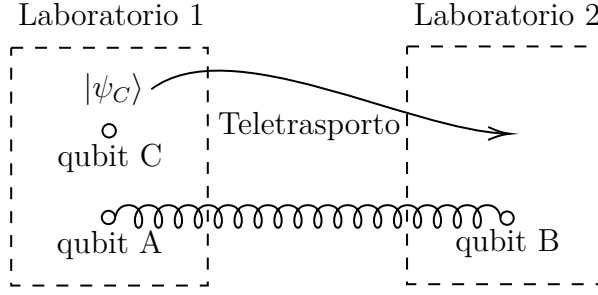


Figura (2.1) – Setup sperimentale per il teletrasporto quantistico

Presso i laboratori di Alice e Bob sono disponibili i qubit A e B , nello stato *entangled* dato da:

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})$$

Perciò se una misura di A trova un risultato 0 o 1, una qualsiasi successiva misura di B troverà lo stesso esito.

Alice ha poi anche il qubit C , nello stato generico $|\psi_c\rangle$ non noto, che vogliamo trasferire a Bob.

Mettendo tutto insieme, si ha che lo stato iniziale $|\Phi\rangle_{ABC}$ del sistema è dato da:

$$|\Phi_0\rangle_{ABC} = |\psi\rangle_C \otimes \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}} \right)_{AB} \quad |\psi\rangle_C = \alpha|0\rangle + \beta|1\rangle$$

Il **protocollo** di teletrasporto quantistico consiste in una serie di operazioni che possiamo schematizzare come un **circuito quantistico** (figura 2.2).

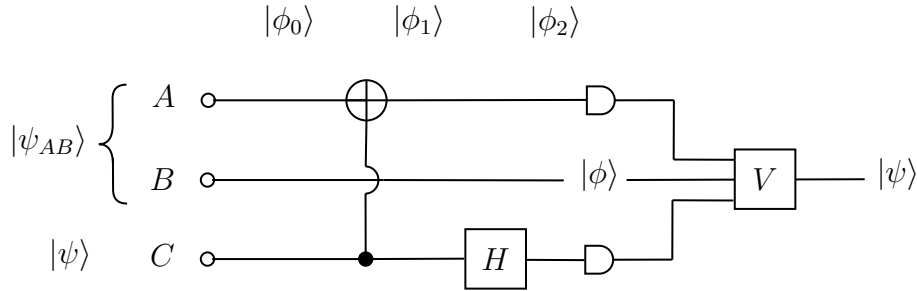


Figura (2.2) – Schema circuitale del protocollo per il teletrasporto quantistico

Alla fine del circuito Alice misura i qubit A e C , registra i risultati in una coppia di bit classici e li trasmette a Bob che, a seconda del loro valore, esegue una certa operazione V sul qubit B che possiede. L'algoritmo fa sì che, dopo aver applicato V , il qubit B sia esattamente pari al C iniziale, il cui contenuto è stato quindi

“teletrasportato” da Alice a Bob.

Esaminiamo perciò passo per passo l'azione del circuito, indicando con $|\Phi_0\rangle$ lo stato iniziale, $|\Phi_1\rangle$ lo stato a seguito del primo CNOT, e $|\Phi_2\rangle$ lo stato prima della misura di Alice. Si ha che:

$$\begin{aligned} |\Phi_1\rangle &= U_{\text{CNOT}} |\Phi_0\rangle = U_{\text{CNOT}} \left[\alpha |0\rangle_C \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}} \right)_{AB} + \beta |1\rangle_C \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}} \right)_{AB} \right] = \\ &= \alpha |0\rangle_C \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}} \right)_{AB} + \beta |1\rangle_C \left(\frac{|10\rangle + |01\rangle}{\sqrt{2}} \right)_{AB} \end{aligned} \quad (2.1)$$

$$\begin{aligned} |\Phi_2\rangle &= H_C |\Phi_1\rangle = \frac{1}{2} [\alpha(|0\rangle + |1\rangle)_C(|00\rangle + |11\rangle)_{AB} + \beta(|0\rangle - |1\rangle)_C(|10\rangle + |01\rangle)_{AB}] = \\ &= \frac{1}{2} \left[|00\rangle_{AC} (\alpha |0\rangle + \beta |1\rangle)_B + \right. \\ &+ |01\rangle_{AC} (\alpha |1\rangle + \beta |0\rangle)_B + \\ &+ |10\rangle_{AC} (\alpha |0\rangle - \beta |1\rangle)_B + \\ &\left. + |11\rangle_{AC} (\alpha |1\rangle - \beta |0\rangle)_B \right] \end{aligned} \quad (2.2)$$

A questo punto Alice misura i qubit A e C , collassando lo stato $|\Phi_2\rangle$ in **uno solo** dei 4 termini di cui è formato, per poi comunicare i risultati a Bob, che ora sa esattamente in qualche combinazione di $|0\rangle$ e $|1\rangle$ si trova lo stato di B , e può *manipolarlo* per ricondurlo allo stato di C .

Per esempio, se Alice misura 00 per AC (con $p = 1/4$), allora lo stato di B è dato dalla prima riga di (2.2), ossia: $|\phi\rangle_B = \alpha |0\rangle + \beta |1\rangle = |\psi\rangle_C$. Troviamo quindi che, in questo caso, il qubit B posseduto da Bob ha assunto *lo stesso stato* del qubit C che aveva Alice.

Negli altri casi B si trova in uno stato diverso, che però, se si è a conoscenza della misura di Alice, si può ricondurre a quello di C effettuando di conseguenza un'opportuna operazione V , come mostrato in tabella 2.1.

Alice (AC)	Bob ($ \phi\rangle_B$)	V
00	$\alpha 0\rangle + \beta 1\rangle = \psi\rangle$	\mathbb{I}_B
01	$\alpha 1\rangle + \beta 0\rangle$	$\hat{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
10	$\alpha 0\rangle - \beta 1\rangle$	\hat{Z}
11	$\alpha 1\rangle - \beta 0\rangle$	$\hat{Z}\hat{X}$

Tabella (2.1) – Le prime due colonne riportano i possibili risultati della misura di Alice e il conseguente stato di B . Per ricondurre B allo stato del qubit C che si voleva teletrasportare, Bob deve eseguire l'operazione V indicata, dove indichiamo con \hat{X} e \hat{Z} gli operatori dati dalle rispettive matrici di Pauli σ_x e σ_z

Per esempio, se Alice misura 01 per AC , troviamo che $|\phi\rangle_B = \alpha|1\rangle + \beta|0\rangle$, che è corrisponde allo stato di C $|\psi\rangle_C = \alpha|0\rangle + \beta|1\rangle$ ottenuto *scambiando* $|0\rangle \leftrightarrow |1\rangle$. Basta allora applicare un NOT quantistico a B - che in notazione matriciale corrisponde alla σ_x di Pauli. Avremo quindi:

$$|\phi'\rangle_B = \text{NOT}(|\phi\rangle_B) = \text{NOT}(\alpha|1\rangle + \beta|0\rangle) = \alpha|0\rangle + \beta|1\rangle = |\psi\rangle_C$$

Quando Alice misura 10 per AC , invece, $|\phi\rangle_B = \alpha|0\rangle - \beta|1\rangle$. Per ricondursi a $|\psi\rangle_C$ basta correggerne la fase relativa, aggiungendo un π tramite la porta logica di *phase-shift* - la cui forma matriciale diviene pari, in questo caso, a quella della σ_z di Pauli:

$$\delta(\pi) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \hat{X}$$

Per l'ultimo caso (Alice che misura 11 per AC) basta combinare le due manipolazioni appena esaminate.

Notiamo due cose:

- Bob ottiene in B esattamente lo stato di C solo in un caso su 4. Per ricondursi allo stato di C è necessario effettuare una certa operazione, che è determinata dagli esiti delle misure di Alice. Perciò Bob **deve conoscere** tali risultati, che devono essere comunicati tramite un canale *classico* (soggetto al limite di velocità c).
- Dato che Alice effettua una misura su C , lo stato originale $|\psi\rangle_C$ viene distrutto nel processo. Ritroviamo quindi, come già dimostrato, che **non è possibile clonare** stati quantistici arbitrari.

Lo stato EPR può essere creato partendo da $|00\rangle_{AB}$ tramite:

$$\text{CNOT}(H \otimes \mathbb{I})|01\rangle = \text{CNOT} \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle)_{AB} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{AB}$$

Perciò, uno schema del circuito che comprenda anche la generazione dello stato è dato da:

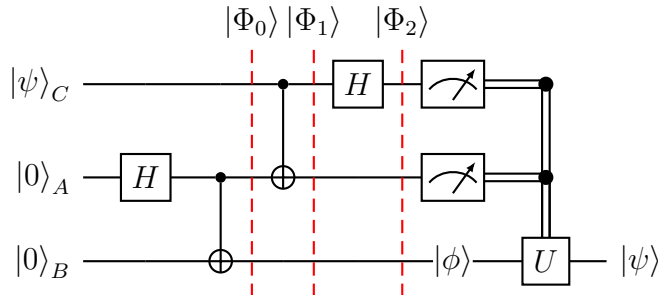


Figura (2.3) – Schema a porte logiche quantistiche del teletrasporto quantistico - con anche la generazione dello stato EPR

2.2 Misure quantistiche

Consideriamo uno stato $|\psi\rangle$ e un'osservabile \hat{A} , che consideriamo a spettro discreto e non degenerare. Si ha quindi che gli autoket $|a_n\rangle$ di \hat{A} , tali da soddisfare l'equazione agli autovalori:

$$\hat{A}|a_n\rangle = a_n|a_n\rangle$$

formano una base ortonormale di \mathcal{H} , e perciò possiamo rappresentare (per il teorema spettrale) l'azione di \hat{A} hermitiana su un generico vettore come la combinazione lineare di proiettori su tale base:

$$\hat{A} = \sum_n a_n \underbrace{|a_n\rangle\langle a_n|}_{\hat{P}_n} = \sum_n a_n \hat{P}_n$$

Effettuando una misura (ideale di prima specie) di \hat{A} sullo stato $|\psi\rangle$ si ottiene uno stato $|\phi\rangle$ dato, per il **postulato di proiezione di von Neumann**, da:

$$|\phi\rangle = \frac{\hat{P}_n|\psi\rangle}{\sqrt{\langle\psi|\hat{P}_n|\psi\rangle}}$$

In altre parole, se una misura di \hat{A} sullo stato iniziale $|\psi\rangle$ è pari a a_n , lo stato finale $|\psi\rangle$ è pari alla componente di $|\psi\rangle$ *parallela* all'autoket $|a_n\rangle$, opportunamente normalizzata.

Poiché abbiamo supposto \hat{A} a spettro discreto non degenerare, vale infine la completezza di Dirac nella sua versione più semplice:

$$\sum_n \hat{P}_n = \mathbb{I} \quad \hat{P}_n \hat{P}_m = \delta_{mn} \hat{P}_m \quad P^2 = P$$

Sperimentalmente abbiamo accesso **solo ai valor medi** ottenuti da misure (riperate). Risulta quindi utile definire p_n come la media di \hat{P}_n nello stato $|\psi\rangle$:

$$p_n = \langle\psi|\hat{P}_n|\psi\rangle$$

In questi termini, il valor medio di A e la sua fluttuazione ΔA sono dati da:

$$\langle A \rangle = \sum_n a_n p_n \quad \langle \Delta A \rangle = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}$$

2.3 Misura senza interazione

2.3.1 Il Beam Splitter quantistico

Normalmente possiamo pensare che il limite minimo per una misurazione si ottenga con una *singola interazione*, per esempio lanciando un solo fotone contro un ostacolo. In realtà è possibile, con una variazione del setup delle due fenditure, effettuare

III postulato della MQ

una misura *senza alcuna possibile interazione* con l'oggetto che si vuole misurare.

Per poter comprendere il protocollo di misura senza interazione introduciamo un elemento di **ottica quantistica** (che non staremo a discutere, dato che esula dagli obiettivi del corso):

- **Specchio semiriflettente** (o *Beam-splitter* [5]): si tratta di un elemento ottico in grado di riflettere il 50% della radiazione incidente, lasciando passare il restante 50% ($R = 1/2, T = 1/2$). Possiamo interpretare ciò anche a livello di singoli fotoni, pensando che ciascuno di essi abbia una probabilità pari a $1/2$ di essere riflesso o trasmesso (figura 2.4).

Beam-splitter

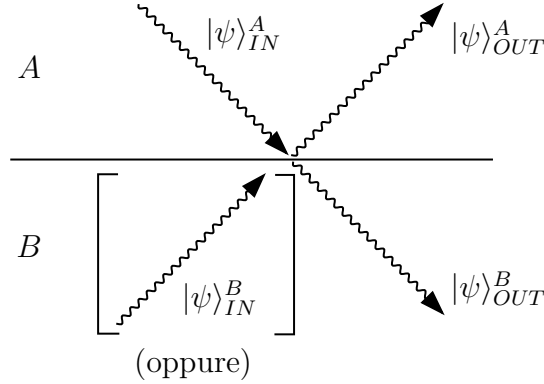


Figura (2.4) – Schema del funzionamento del Beam Splitter.

L'effetto di un beam-splitter può essere schematizzato come quello di un opportuno **gate quantistico** $U_{B.S.}$, tale che:

$$|1\rangle_{in}^A |0\rangle_{in}^B \xrightarrow{U_{B.S.}} \frac{1}{\sqrt{2}} |1\rangle_{out}^A |0\rangle_{out}^B + \frac{i}{\sqrt{2}} |0\rangle_{out}^A |1\rangle_{out}^B$$

Per cui un fotone che arriva da *fuori* (cioè da *A*) può essere rimandato fuori (di nuovo in *A*) o dentro (in *B*), con ugual probabilità. Nel caso di **trasmissione** il fotone acquisisce una fase i (cioè di $\pi/2$).

Vale la relazione analoga nel caso un fotone arrivi “da dentro”, ossia da *B*:

$$|0\rangle_{in}^A |1\rangle_{in}^B \xrightarrow{U_{B.S.}} \frac{i}{\sqrt{2}} |1\rangle_{out}^A |0\rangle_{out}^B + \frac{1}{\sqrt{2}} |0\rangle_{out}^A |1\rangle_{out}^B$$

Possiamo sintetizzare queste due relazioni scrivendo U_{BS} in notazione matriciale nella base $\{|1\rangle^A |0\rangle^B, |0\rangle^A |1\rangle^B\}$:

$$|\psi_{out}\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} |\psi_{in}\rangle$$

2.3.2 Interferometro di Mach-Zehnder quantistico

Un modo per effettuare una misura senza interazione avviene tramite l'uso di un **interferometro di Mach-Zehnder**, rappresentato in figura 2.5.

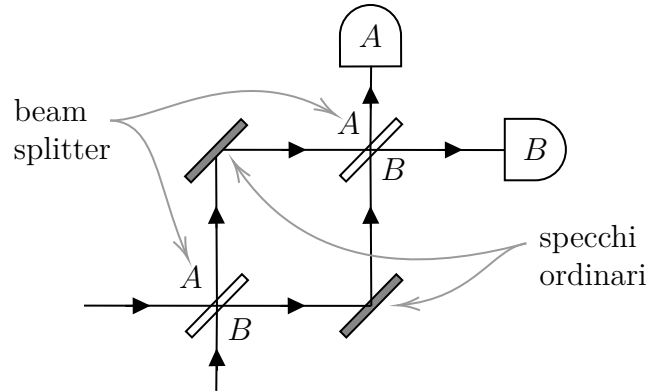


Figura (2.5) – Schema dell'interferometro di Mach-Zehnder.

Consideriamo un fotone che parte da A . Al primo passaggio attraverso il *beam-splitter* avremo con $p = 1/2$ una riflessione verso l'alto e con $p = 1/2$ una trasmissione verso destra B . In entrambi i casi il fotone incide poi contro un normale specchio, e viene riflesso verso il secondo *beam-splitter*. In MQ entrambi i percorsi sono esplorati contemporaneamente, e interferiscono tra loro al ricongiungimento nel secondo BS .

Innanzitutto, in entrambi i percorsi si ha una riflessione su uno specchio normale, che introduce in ciascun caso una fase¹ π , senza però variare la differenza tra le fasi delle due “versioni” del fotone: possiamo perciò trascurare questo dettaglio. Percorriamo allora il circuito, seguendo i due possibili percorsi, partendo da un fotone che inizialmente è “in fase con se stesso”:

- Se il fotone viene inizialmente riflesso dal primo BS, allora non acquisisce nessuna fase. Dopo aver oltrepassato il secondo BS, avremo quindi una possibilità con fase 1 ($\theta = 0$) appena prima del detector in A , dovuta ad una seconda riflessione sul BS, e una con fase ancora i prima del detector B , prodotta dalla trasmissione attraverso il BS.
- Se il fotone viene trasmesso dal primo BS, allora arriva al secondo con una fase i . A questo punto, per raggiungere il detector A deve essere nuovamente trasmesso attraverso il BS, acquisendo un'ulteriore fase i e arrivando ad una fase totale i^2 . Se invece viene riflesso verso il detector B mantiene la sua fase di i .

Notiamo allora che prima del detector A , due versioni dello stesso fotone con *fase opposta* ($i^2 = -1$, e 1) interagiscono tra loro in modo *distruttivo*, e quindi il

¹ΛDato che il materiale riflettente ha un indice di rifrazione per forza maggiore di quello del vuoto

rivelatore in A non troverà alcun segnale.

D'altro canto, davanti a B le “due versioni del fotone” interagiscono con la *stessa fase* i , e quindi avremo *interferenza costruttiva* e il rivelatore B osserverà **sempre** la presenza del fotone.

In maniera analoga si può ricavare il funzionamento nel caso il fotone parta da B e non da A , ottenendo una $p = 1$ di rivelazione al detector A .

Un modo più semplice per capire tutto ciò è schematizzare l'interferometro come l'applicazione consecutiva di due gate U_{BS} , come mostrato in figura 2.6.

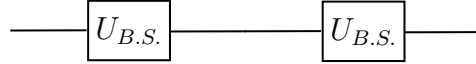


Figura (2.6) – Schema con porte logiche quantistiche dell'interferometro di Mach-Zehnder.

Effettuando allora la computazione passo a passo:

$$|0\rangle \xrightarrow{U_{BS}} \frac{|0\rangle + i|1\rangle}{\sqrt{2}} \xrightarrow{U_{BS}} \frac{1}{2} [|0\rangle + i|1\rangle + i(i|0\rangle + |1\rangle)] = i|1\rangle$$

Alternativamente si ottiene lo stesso risultato calcolando il quadrato della matrice di U_{BS} :

$$U_{BS}^2 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

2.3.3 Il test della bomba di Elitzur–Vaidman

Siamo partiti dicendo che, tramite l'interferometro di Mach-Zehnder, risulta possibile effettuare una misura senza interazione. Vediamo allora come fare, con un esempio *fortemente drammatico*.

Un terrorista ha costruito una *bomba*, dotandola di un detonatore che può essere attivato dall'interazione con un *singolo fotone*. Tale bomba è nascosta in una scatola trasparente, che viene posta lungo il cammino dell'interferometro di Mach-Zehnder (figura 2.7), ma a priori non sappiamo se si tratti di un *bluff* (cioè di una scatola vuota) o di un effettivo pericolo (la scatola contiene la bomba). Scopo dell'esperimento è distinguere i due casi *senza attivare la bomba*.

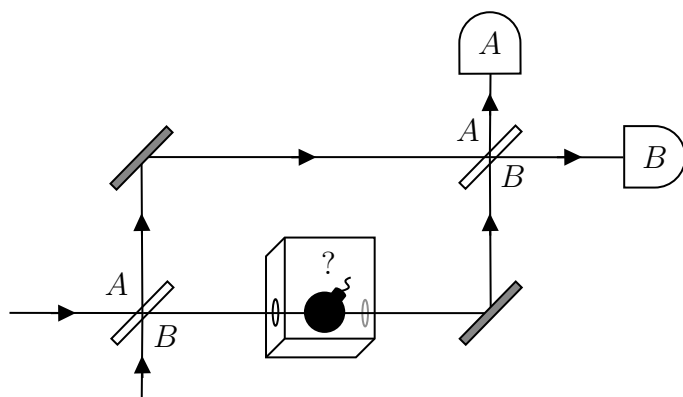


Figura (2.7) – Interferometro di Mach-Zehnder utilizzato per misure senza interazione.

Se proviamo ad iniettare in A della luce classica, il 50% del raggio viene trasmesso verso la bomba, che quindi esplode. Chiaramente, in MC il problema non è risolvibile.

Proviamo allora ad effettuare lo stesso esperimento in regime quantistico, utilizzando un singolo fotone alla volta.

Abbiamo due possibilità:

- Se la **scatola è vuota**, allora il fotone arriva all'uscita percorrendo entrambe le vie, e valgono le considerazioni fatte in precedenza per lo schema di figura 2.5. Per il detector all'uscita A avremo interferenza distruttiva, e quindi $p_A = 0$, mentre in B si ha interferenza costruttiva, e $p_B = 1$. Perciò il fotone viene rivelato sempre dal detector B .
- Se la **scatola è piena**, allora al 50% il fotone raggiunge la bomba e la fa saltare in aria. Nell'altro 50%, tuttavia, il fotone viene riflesso nel primo BS, giungendo al secondo, dove stavolta non vi è alcuna interferenza (figura 2.8). Avremo quindi due possibilità: un 25% che si trasmetta in B , e un 25% che si rifletta in A .

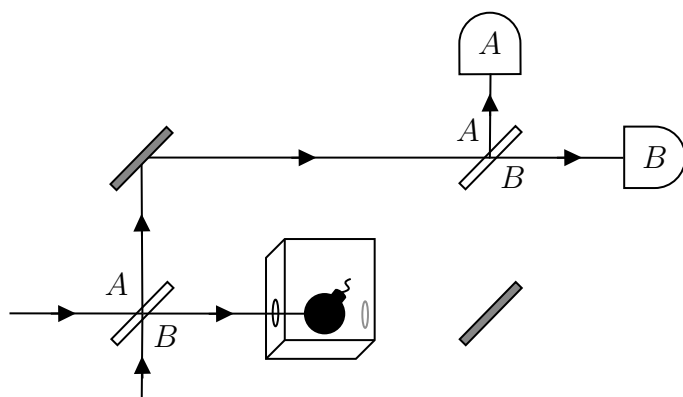


Figura (2.8) – Nel caso in cui vi sia effettivamente una bomba, solo uno dei due percorsi è possibile, e perciò non si ha alcuna interferenza nel secondo beam-splitter.

Notiamo che se la scatola è vuota, il rivelatore A non trova mai il fotone. Perciò, se troviamo il fotone in A sappiamo immediatamente che la bomba è presente. Ciò succede in un caso su 4: abbiamo allora un modo per misurare la presenza o meno di un oggetto **senza** interagire con esso in alcun modo.

Per quanto ciò sia paradossale, tale setup - il cosiddetto *detector di bombe* di Elitzur-Vaidman - è stato realizzato sperimentalmente (sostituendo la bomba con un ostacolo meno pericoloso).

Per di più si trova ([1]) che è possibile concatenare tanti circuiti simili per aumentare (arbitrariamente) la probabilità di rivelazione senza interazione.

Interpretazione del tester Elitzur-Vaidman. Un modo di pensare gli esiti dell'esperimento sta nel considerare la bomba come un *sistema interagente*. Se è presente, allora interagisce con la funzione d'onda del fotone, e la fa collassare su uno dei due percorsi possibili. Se non è presente, invece, non avviene alcun collasso, e la funzione d'onda è libera di interferire con se stessa nel secondo BS. In un certo senso, perciò, l'apparato non sta facendo altro che determinare se una interazione sia avvenuta o meno. Interpretando tale interazione non unitaria come una *misurazione*, pittorescamente si potrebbe affermare che il tester Elitzur-Vaidman “misura l'avvenire o meno di un'altra misura”.

Possibilità del genere hanno applicazione, per esempio, per ridurre la quantità di radiazioni assorbite da un tessuto durante una radiografia, dato che parte dei fotoni utilizzati portano *informazione utile* senza interazione con il paziente.

2.4 Effetto Zenone quantistico

2.4.1 I paradossi di Zenone

I **paradossi di Zenone** furono una serie di paradossi ideati nell'antica Grecia.

Il più conosciuto è quello di *Achille e la tartaruga*, che percorre il seguente ragionamento:

1. Achille (detto *pie' veloce* per la sua formidabile rapidità) viene sfidato da una tartaruga ad una gara di corsa. Per rendere un minimo più equa la competizione, alla tartaruga è concesso partire con una certa distanza di vantaggio.
2. Achille raggiunge la tartaruga in un certo tempo t_1 , percorrendo la distanza x a velocità v_A . Ma nel frattempo la tartaruga si è spostata di un ulteriore $x_2 = v_t t_1$.
3. Achille copre la distanza x_2 in un tempo t_2 , ma nel frattempo la tartaruga si posta ancora di una distanza $x_3 = v_t t_2$, e così via.

Poiché la serie *continua all'infinito*, si potrebbe pensare che Achille non possa mai raggiungere la tartaruga. Serviranno i concetti di analisi matematica, per cui una

serie infinita può convergere a un valore finito, per risolvere il problema.

Un altro paradosso è quello della **freccia**, secondo cui un dardo non può mai raggiungere nessuna destinazione, poiché ad ogni istante di tempo è *fermo*, e la somma di *infiniti istanti stazionaria* non può - almeno in principio - costituire un moto.

L'esperienza comune dimostra come il moto avvenga in realtà senza alcun problema. Il vero paradosso sorge allora quando, in MQ, si scopre che il *paradosso della freccia* è davvero possibile: misurando *tante* volte una particella è possibile *arrestare* la sua evoluzione temporale!

2.4.2 Introduzione

(Lezione 4 • del
7/3/2019)

L'idea dell'effetto Zenone quantistico sta nel fatto che, osservando *ripetutamente e frequentemente* una particella quantistica la si può *bloccare* nello stesso stato di partenza.

In questa sezione ci occuperemo di derivare matematicamente l'effetto, usando solo l'operatore di evoluzione temporale unitaria e il postulato di proiezione di von Neumann, seguendo lo schema riportato in [4]. Mostreremo quindi un esempio di sistema che esibisce tale fenomeno, e daremo un cenno a come ciò possa essere esito di una *interazione* con l'ambiente, dando un significato “più fisico” alle proiezioni geometriche finora usate.

Schema dei
paragrafi

Introdurremo poi il concetto di “Hamiltoniana” non-hermitiana, che useremo per modellizzare le proiezioni alla von Neumann. Ricondurremo infine tale matrice alla normale Hamiltoniana hermitiana di un sistema considerato “non nella sua interezza” (di cui per esempio solo alcuni stati sono di interesse o accessibili - come nel caso di un qubit posto in un apparato complesso).

2.4.3 Derivazione dell'Effetto Zenone

L'evoluzione di un sistema quantistico **instabile** segue qualitativamente l'andamento riportato in figura 2.9.

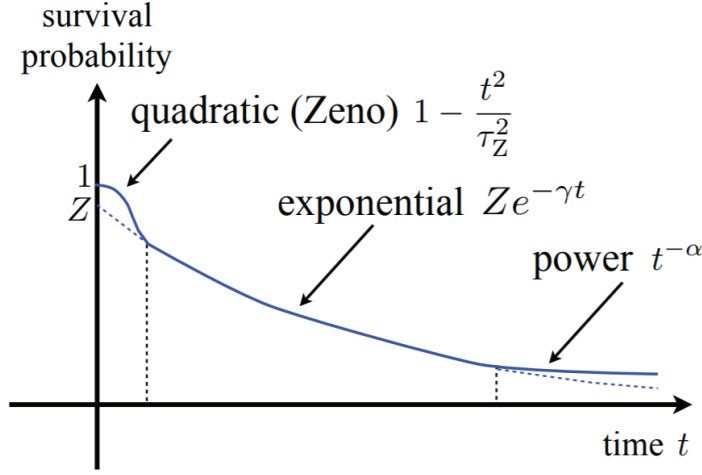


Figura (2.9) – Evoluzione temporale di un sistema quantistico instabile. Sull’asse y è riportata la *survival probability*, ossia la probabilità che il sistema rimanga nello stato iniziale (instabile)

La probabilità che il sistema rimanga nello stato iniziale esibisce una decrescita quadratica per tempi brevi, che diviene poi esponenziale per tempi medi, e con un andamento in potenza per tempi lunghi. Per l’effetto Zeno ci interessa esaminare nel dettaglio l’andamento quadratico iniziale.

Consideriamo allora un sistema isolato inizialmente nello stato $|\psi_0\rangle$ a $t = 0$. Poniamo, da qui in avanti, $\hbar = 1$ per alleggerire la notazione. AD un certo tempo t , il sistema si è evoluto ad uno stato $|\psi(t)\rangle$, che si ottiene dalla formula per l’**evoluzione unitaria**:

$$|\psi(t)\rangle = e^{-iHt} |\psi_0\rangle \quad (2.3)$$

Ci interessa quantificare quanto $|\psi(t)\rangle$ sia “vicino” allo stato iniziale $|\psi_0\rangle$. Consideriamo allora l’*ampiezza* della funzione d’onda che è rimasta *parallela* a $|\psi_0\rangle$, e che è data dal prodotto scalare:

$$\mathcal{A}(t) = \langle \psi_0 | \psi(t) \rangle = \langle \psi_0 | e^{-iHt} | \psi_0 \rangle \quad (2.4)$$

*Survival
amplitude*

Prendendo il modulo quadro otteniamo la *probabilità* $p(t)$ che lo stato al tempo t sia ancora pari a quello iniziale, detta **survival probability**:

$$p(t) = |\mathcal{A}(t)|^2 = |\langle \psi_0 | e^{-iHt} | \psi_0 \rangle|^2 \quad (2.5)$$

*Survival
probability*

che è pari a 0 se a t il sistema si trova in uno stato $|\psi(t)\rangle$ *ortogonale* a quello iniziale.

Esaminiamo l’evoluzione temporale per un piccolo intervallo di tempo δt . Sviluppando in serie di MacLaurin la (2.3) attorno a $\delta t = 0$ si ottiene:

$$|\psi(\delta t)\rangle = e^{-iH\delta t} |\psi_0\rangle \approx \left(\mathbb{I} - iH\delta t - \frac{H^2}{2}(\delta t)^2 + O((\delta t)^3) \right) |\psi_0\rangle \equiv |\psi_0\rangle + |\delta\psi\rangle$$

Tronchiamo questa espressione al secondo ordine, e calcoliamo la “survival amplitude” $\mathcal{A}(t)$ prendendo il prodotto scalare con $|\psi_0\rangle$, e riconoscendo i valori medi $\langle H \rangle_{\psi_0} = \langle \psi_0 | H | \psi_0 \rangle$:

$$\mathcal{A}(\delta t) = 1 - i\langle H \rangle_{\psi_0} \delta t - \frac{1}{2} \langle H^2 \rangle_{\psi_0} (\delta t)^2 \quad (2.6)$$

E a sua volta calcoliamo la survival probability prendendo il modulo quadro, sfruttando il fatto che, poiché H è hermitiano, i suoi valori medi sono reali:

*Survival
probability per
tempi piccoli*

$$\begin{aligned} p(\delta t) &= \left(1 + i\langle H \rangle_{\psi_0} \delta t - \frac{1}{2} \langle H^2 \rangle_{\psi_0} (\delta t)^2 \right) \left(1 - i\langle H \rangle_{\psi_0} \delta t - \frac{1}{2} \langle H^2 \rangle_{\psi_0} (\delta t)^2 \right) \\ &\stackrel{(a)}{\approx} 1 - \cancel{i\langle H \rangle_{\psi_0} \delta t} - \frac{1}{2} \langle H^2 \rangle_{\psi_0} (\delta t)^2 + \cancel{i\langle H \rangle_{\psi_0} \delta t} + \langle H \rangle_{\psi_0}^2 (\delta t)^2 - \frac{1}{2} \langle H^2 \rangle_{\psi_0} (\delta t)^2 = \\ &= 1 - (\langle H^2 \rangle_{\psi_0} (\delta t)^2 - \langle H \rangle_{\psi_0}^2 (\delta t)^2) = 1 - (\delta t)^2 \underbrace{(\langle H^2 \rangle_{\psi_0} - \langle H \rangle_{\psi_0}^2)}_{\tau_z^{-2}} \equiv 1 - \frac{(\delta t)^2}{\tau_z^2} \end{aligned} \quad (2.7)$$

dove in (a) trascuriamo tutti i termini di ordine superiore a 2. Nell’ultimo passaggio abbiamo introdotto la grandezza τ_z , detta **tempo di Zenone**, definita da:

$$\tau_z^{-2} \equiv \langle H^2 \rangle_{\psi_0} - \langle H \rangle_{\psi_0}^2 \quad (2.8)$$

Dalle espressioni appena ottenute abbiamo che la funzione d’onda evolve **linearmente** per tempi piccoli², mentre la survival probability **quadraticamente** in δt . Abbiamo allora trovato che, almeno per tempi piccoli, un sistema quantistico evolve “lentamente”, cioè in modo polinomiale. Ci aspettiamo quindi che, entro un δt , il sistema sia in uno stato *molto simile* al $|\psi_0\rangle$ iniziale.

Dal **postulato di proiezione di von Neumann** sappiamo che una misura (ideale di prima specie) *proietta* il sistema nell’autostato compatibile con l’esito ottenuto dalla misura. A seguito di tale processo, strettamente **non unitario**, il nuovo stato riprende ad evolversi in modo unitario. In un certo senso, la misura “resetta” l’evoluzione del sistema - dato che in quell’attimo il sistema *non è isolato*.

Consideriamo allora una misura che verifichi se il sistema si trova ancora nello stato iniziale $|\psi\rangle$. Ripetiamo tale misura N volte in un tempo t , in modo che l’intervallo τ tra ogni misura sia *molto piccolo*:

$$\tau = \frac{t}{N} = \delta t \quad (2.9)$$

Dopo ogni misura il sistema ha una probabilità $p(\tau)$ (survival probability valutata in $t = \tau$) di rimanere nello stato iniziale, e una probabilità $1 - p(\tau)$ di muoversi allo stato $|\psi_0^\perp\rangle$ ortogonale a quello iniziale. Se τ è piccolo, poiché $p(\tau) = 1 - \tau^2/\tau_z^2$,

²^Su tempi lunghi si osserva invece un’evoluzione esponenziale, che poi si stabilizza in una potenza $t^{-\alpha}$.

avremo $p(\tau) \approx 1$, e perciò il sistema *tende* a rimanere nello stato iniziale.

Dopo N misure le probabilità si combinano tra loro³, e si ha $|\psi(t)\rangle = |\psi_0\rangle$ con una probabilità:

$$p^{(N)}(t) = [p(\tau)]^N \stackrel{(2.7)}{=} \left[1 - \frac{\tau^2}{\tau_z^2}\right]^N \stackrel{(2.9)}{=} \left(1 - \left(\frac{t}{N\tau_z}\right)^2\right)^N \xrightarrow[N \gg 1]{(b)} \exp\left(-\frac{t^2}{N\tau_z^2}\right) \xrightarrow[N \rightarrow \infty]{} 1 \quad (2.10)$$

dove in (b) si è usata la definizione dell'esponenziale:

$$e^x = \lim_{n \rightarrow +\infty} \left(1 + \frac{x}{n}\right)^n$$

Perciò, al limite di *misure infinitamente frequenti*, l'evoluzione del sistema in esame è **completamente arrestata**.

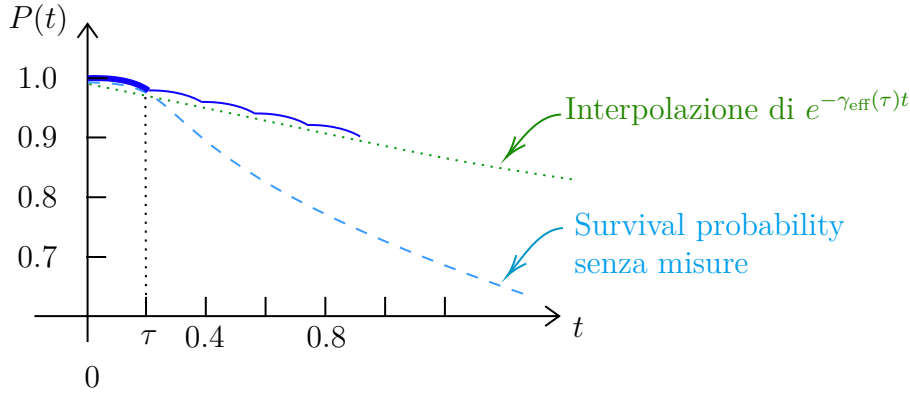


Figura (2.10) – Effetto di Zeno quantistico per $N = 5$ misure “alla Von Neumann”. La linea blu continua mostra la *survival probability* nel caso di misure ripetute, mentre quella azzurra tratteggiata nel caso *senza misure*. La linea in verde puntinata è l'esponenziale che *interpola* l'effetto Zeno definito in (2.11).

La *survival probability* dopo N misure può essere interpolata da un'esponenziale:

$$p^{(N)}(t) = p(\tau)^N = \exp(N \log p(\tau)) \stackrel{N=t/\tau}{=} \exp\left(t \frac{\log p(\tau)}{\tau}\right) = \exp(-\gamma_{\text{eff}}(\tau)t) \quad (2.11)$$

dove abbiamo definito opportunamente la “costante di decadimento effettiva” γ_{eff} come:

$$\gamma_{\text{eff}} = -\frac{1}{\tau} \log p(\tau)$$

Per $\tau \rightarrow 0$, che corrisponde a $N \rightarrow \infty$, vale l'espansione $p(\tau) = 1 - \tau^2/\tau_z^2$, e quindi ricaviamo (dato che $\log(1+x) \sim x$ per $x \rightarrow 0$):

$$\gamma_{\text{eff}}(\tau) \approx \frac{\tau}{\tau_z^2} \quad \tau \rightarrow 0 \quad (2.12)$$

³^Stiamo qui trascurando tutti i *percorsi molto improbabili* in cui il sistema visita stati differenti, come quello ortogonale, e poi ritorna a $|\psi\rangle_0$ nel corso delle N misure.

Equivalentemente, tale espressione si ricava confrontando (2.10) con $\exp(-t\gamma_{\text{eff}})$, dato che:

$$p^{(N)}(t) \underset{N \gg 1}{=} \exp\left(-\frac{t^2}{N\tau_z^2}\right) = \exp\left(-t \frac{\tau}{\tau_z^2}\right) \underset{(2.9)}{=} \exp(-t\gamma_{\text{eff}})$$

Riepilogando, se si esegue una misura ogni τ , la probabilità $p(t)$ che il sistema rimanga nello stato iniziale $|\psi_0\rangle$ ha l'andamento esponenziale decrescente della frazione *non decaduta* di atomi radioattivi. Il decadimento è tanto più lento quanto γ_{eff} , proporzionale a τ , è piccolo. Perciò, nell'ipotetico caso in cui si effettuasse una misura ad ogni istante, $\tau = 0$, $\gamma_{\text{eff}} = \infty$ e $p(t) \equiv 1 \forall t$.

2.4.4 Esempio di effetto Zenone

Un *caso semplice* che possiamo usare per studiare l'effetto Zenone è dato da un sistema che effettua *oscillazioni di Rabi*. Possiamo immaginarlo come un atomo con due livelli a energia diversa, che viene illuminato da un raggio di fotoni di energia (definita dalla loro *frequenza*) pari a quella di transizione tra i due livelli. Si osserva che il sistema *oscilla* tra i due stati: prima assorbe un fotone per portarsi al livello eccitato, e poi ritorna indietro per *emissione stimolata* (analogamente a quanto avviene in un laser).

Detti $|0\rangle$ e $|1\rangle$ gli stati corrispondenti ai due livelli, nella base $\{|0\rangle, |1\rangle\}$ l'Hamiltoniana di un sistema del genere diviene:

$$H = \Omega \hat{\sigma}_x = \begin{pmatrix} 0 & \Omega \\ \Omega & 0 \end{pmatrix} \quad |\psi_0\rangle = |0\rangle \quad (2.13)$$

dove i termini fuori dalla diagonale sono indicativi degli *accoppiamenti* (*coupling*) tra livelli, ossia delle probabilità di transizione tra $|0\rangle$ e $|1\rangle$ e viceversa.

Nota: poiché vogliamo descrivere un sistema che oscilla (per evoluzione unitaria) tra due stati $|0\rangle$ e $|1\rangle$, si ha che tali $|0\rangle$ e $|1\rangle$ **non** possono essere autoket di H : se lo fossero sarebbero stati stazionari, e non sarebbe possibile un'evoluzione da uno all'altro senza un disturbo esterno.

Infatti, se $|0\rangle$ è autoket di H , l'equazione di Schrödinger dipendente dal tempo ha come soluzione un'onda stazionaria che resta in $|0\rangle$:

$$i\hbar \frac{\partial}{\partial t} |0\rangle = \mathcal{E}_0 |0\rangle \Rightarrow |\psi(t)\rangle = e^{-i\mathcal{E}_0 t} |0\rangle$$

Segue che, scrivendo H in una base $\{|0\rangle, |1\rangle\}$ che **non** è costituita dai suoi autoket, H non può essere diagonale.

L'evoluzione temporale determinata da H è data da:

$$|\psi(t)\rangle = e^{-iHt} |\psi_0\rangle = \exp\left(-i\Omega t \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

L'esponenziale di matrice si calcola in modo diretto notando che:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad A^2 = \mathbb{I}$$

Da cui:

$$\begin{aligned} \exp(-i\Omega t A) &= 1 + (-i\Omega t)A + \frac{1}{2!}(-i\Omega t)^2 \mathbb{I} + \frac{1}{3!}(-i\Omega t)^3 A + \dots = \\ &= 1 + (-i)[\Omega t]A - \frac{1}{2!}[\Omega t]^2 \mathbb{I} - \frac{1}{3!}(-i)[\Omega t]^3 A + \dots = \\ &= \mathbb{I} \sum_{k=0}^{+\infty} (-1)^k \frac{(\Omega t)^{2k}}{(2k)!} + (-i)A \sum_{k=0}^{+\infty} (-1)^k \frac{(\Omega t)^{2k+1}}{(2k+1)!} = \\ &= \mathbb{I} \cos(\Omega t) - iA \sin(\Omega t) = \begin{pmatrix} \cos(\Omega t) & -i \sin(\Omega t) \\ -i \sin(\Omega t) & \cos(\Omega t) \end{pmatrix} \end{aligned}$$

Moltiplicando per $|\psi_0\rangle = |0\rangle = (1, 0)^T$ otteniamo allora:

$$|\psi(t)\rangle = \cos(\Omega t) |0\rangle - i \sin(\Omega t) |1\rangle$$

Calcoliamo quindi $\mathcal{A}(t)$ e $p(t)$ tramite (2.4) e (2.5):

$$\begin{aligned} \mathcal{A}(t) &= \langle \psi_0 | \psi(t) \rangle = \cos(\Omega t) \\ \Rightarrow p(t) &= |\mathcal{A}(t)|^2 = \cos^2(\Omega t) = 1 - [\Omega t]^2 + \frac{[\Omega t]^3}{3!} + O([\Omega t]^6) \end{aligned}$$

Per il tempo di Zenone usiamo la definizione (2.8):

$$\tau_z^{-2} = \langle H^2 \rangle_{\psi_0} - \langle H \rangle_{\psi_0}^2$$

dove il calcolo del valor medio, in notazione matriciale, diviene:

$$\begin{aligned} \langle H \rangle_{\psi_0} &= \langle 0 | H | 0 \rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & \Omega \\ \Omega & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0 \\ \langle H^2 \rangle_{\psi_0} &= \langle 0 | H^2 | 0 \rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \Omega^2 & 0 \\ 0 & \Omega^2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \Omega^2 \end{aligned}$$

E otteniamo:

$$\tau_z^{-2} = \Omega^2 \Rightarrow \tau_z = \Omega^{-1}; \quad \gamma_{\text{eff}} \stackrel{(2.12)}{=} \tau \Omega^2 \quad (2.14)$$

Notiamo in particolare che vale l'espansione da cui abbiamo ricavato il tempo di Zenone:

$$p(t) \approx 1 - \frac{t^2}{\tau_z^2} = 1 - \Omega^2 t^2$$

2.4.5 Effetto Zenone per interazione

Nel calcolare il tempo di Zenone τ_z nell'esempio abbiamo visto, si è trovato che l'unico contributo deriva da $\langle H^2 \rangle_{\psi_0}$. Questo *non* è un caso, ma è indice dell'origine dell'effetto Zenone come *interazione*.

In generale, consideriamo un sistema di Hamiltoniana H , scritta in una qualche base $\{|\psi_n\rangle\}$ che non sia quella dei suoi autostati. Perciò H è una matrice non diagonale, che può essere scomposta in una matrice diagonale H_0 e una matrice con tutti gli elementi che non stanno sulla diagonale H_{int} . Quest'ultima è detta **hamiltoniana di interazione**, in quanto i suoi termini caratterizzano i *coupling* tra i livelli $|\psi_n\rangle$.

$$H = H_0 + H_{\text{int}}$$

In questa notazione, chiaramente $\{|\psi_n\rangle\}$ sono gli autostati di H_0 (dato che è diagonale in questa base, per costruzione). Detto $|\psi_0\rangle$ uno di essi, avremo:

$$H_0 |\psi_0\rangle = \omega_0 |\psi_0\rangle$$

Calcoliamo il tempo di Zenone τ_z dalla definizione:

$$\tau_z^{-2} = \langle H^2 \rangle_{\psi_0} - \langle H \rangle_{\psi_0}^2 = \langle H_0^2 + H_0 H_{\text{int}} + H_{\text{int}} H_0 + H_{\text{int}}^2 \rangle_{\psi_0} - \langle H_0 + H_{\text{int}} \rangle_{\psi_0}^2$$

Notiamo che $\langle H_{\text{int}} \rangle_{\psi_0} = 0$, dato che il valor medio “seleziona” un termine sulla diagonale, che è nulla per come abbiamo definito H_{int} . Allo stesso modo, i prodotti tra H_0 e H_{int} hanno sempre la diagonale nulla. Infatti moltiplicando una matrice diagonale A con una B con diagonale nulla si ottiene una matrice C con diagonale nulla:

$$C_{ij} = A_{ik} B_{kj} \underset{(a)}{=} \delta_{ik} A_{ik} B_{kj} \Rightarrow C_{ii} = A_{ii} \underbrace{B_{ii}}_0 = 0$$

dove in (a) si è usata la definizione di matrice diagonale.

Abbiamo allora ricavato che i valor medi dei “termini misti” sono nulli. Inoltre, dato che $|\psi_0\rangle$ è autostato di H_0 vale (per definizione di autostato):

$$\langle H_0^2 \rangle_{\psi_0} - \langle H_0 \rangle_{\psi_0}^2 = 0$$

Perciò resta un unico termine:

$$\tau_z^{-2} = \langle H_{\text{int}}^2 \rangle_{\psi_0} = \sum_n \langle \psi_0 | H_{\text{int}} | \psi_n \rangle \langle \psi_n | H_{\text{int}} | \psi_0 \rangle$$

Perciò, in questo caso il tempo di Zenone dipende esclusivamente dall'Hamiltoniana di interazione, che contiene i *coupling* tra i livelli del sistema considerato. In effetti, questo è proprio il caso del sistema (oscillatore di Rabi) esaminato nel paragrafo precedente, dove $H = H_{\text{int}}$.

Come vedremo nei prossimi paragrafi, questa idea di *interazione* sorge esaminando nel dettaglio i “retroscena” delle proiezioni di Von Neumann.

2.4.6 Matrici “Hamiltoniane” non hermitiane

Poiché l’operatore U di evoluzione temporale è unitario, si ha che:

$$\langle \psi(t) | \psi(t) \rangle = \langle \psi_0 | \underbrace{U^\dagger U}_{\mathbb{I}} | \psi_0 \rangle = \langle \psi_0 | \psi_0 \rangle$$

Perciò l’evoluzione temporale di un qubit avviene tra punti posti sulla *superficie* della sfera di Bloch, ossia rimane nel sottospazio di Hilbert definito da $\|\psi\| = 1$ (figura 2.11.a). Un’eventuale evoluzione non unitaria (figura 2.11.b) fa sì che la funzione d’onda *abbandoni* tale sottospazio (come vedremo più in dettaglio nelle prossime sezioni).

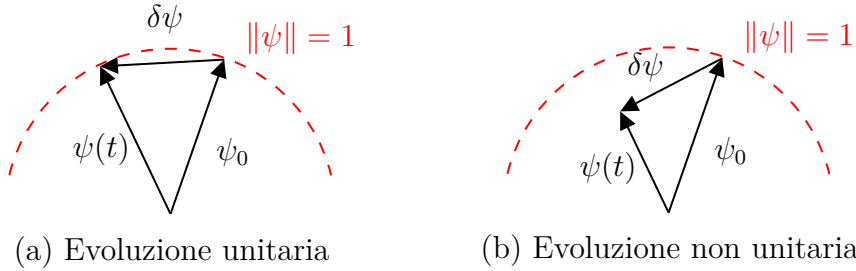


Figura (2.11) – L’evoluzione unitaria (a) fa sì che un vettore non abbandoni mai la superficie della sfera di Bloch, cosa che invece succede nel caso non unitario (b).

L’evoluzione di sistemi instabili, in cui la *survival probability* diminuisce nel tempo, può essere vista come un processo che “porta via probabilità”, ossia un’evoluzione non unitaria. In altre parole, ci stiamo muovendo nella direzione di *integrare* gli effetti delle misure ripetute precedentemente considerate direttamente nell’Hamiltoniana.

Per trattare l’evoluzione non unitaria introduciamo allora *ad hoc* una “Hamiltoniana modificata” data da:

$$H = H - iV\mathbb{I} \quad V \in \mathbb{R}^+ \quad (2.15)$$

dove V è un’opportuna grandezza detta **potenziale ottico**⁴.

Un sistema che evolve con un’Hamiltoniana del genere, che non è hermitiana, “perde probabilità”, ossia giunge a stati non normalizzati. Partendo da uno stato iniziale $|\psi_0\rangle$, avremo cioè $\|\psi(t)\| \leq \|\psi_0\|$. Fisicamente, tale “probabilità mancante” va a finire in altri stati - non contemplati nel sistema che stiamo esaminando - attraverso un *decay channel*. Pittorescamente, possiamo immaginare la probabilità come un fluido, racchiuso in un contenitore, di cui osserviamo la dinamica. L’evoluzione unitaria corrisponde allora alla *visuale completa* dell’intero contenitore, per cui la *quantità di fluido* osservata rimane sempre la stessa. Se invece ci concentriamo solo

⁴^In maniera pittoresca, si può pensare all’evoluzione della funzione d’onda come a quella della luce che si propaga in un mezzo *rifrattivo* e *dissipativo*. Da qui l’idea di aggiungere un “potenziale ottico” per tenere conto di tali “dissipazioni di probabilità”.

su una parte dell'insieme, può capitare che del fluido scorra via, muovendosi nel resto del sistema, e apparentemente scomparendo da questa visuale *parziale*.

Giustificeremo tutto ciò matematicamente tra qualche paragrafo. Per ora, partiamo analizzando l'evoluzione dettata dall'Hamiltoniana non hermitiana appena introdotta, calcolandone *survival amplitude* $\mathcal{A}(t)$ e *survival probability* $p(t)$:

$$\begin{aligned}\mathcal{A}(\delta t) &= e^{-Vt} \langle \psi_0 | e^{-iH\delta t} | \psi_0 \rangle = \\ &\stackrel{(2.6)}{=} \left(1 - V\delta t + \frac{1}{2}V^2[\delta t]^2 + O([\delta t]^3) \right) \left[1 - i\langle H \rangle_{\psi_0}\delta t - \frac{1}{2}\langle H^2 \rangle_{\psi_0}[\delta t]^2 + \dots \right] = \\ &= 1 - (V + i\langle H \rangle_{\psi_0})\delta t - \frac{1}{2}(\langle H^2 \rangle_{\psi_0} - V^2 - 2iV\langle H \rangle_{\psi_0})[\delta t]^2 + O([\delta t]^3) \\ p(\delta t) &= e^{-2Vt} |\langle \psi_0 | e^{-iHt} | \psi_0 \rangle|^2 \approx 1 - 2V\delta t + O([\delta t]^2)\end{aligned}$$

Otteniamo perciò un decadimento **lineare** per tempi piccoli sia per la *survival amplitude* $\mathcal{A}(\delta t)$ che per la *survival probability* $p(\delta t)$.

2.4.7 Effetto Zenone senza proiezioni

All'inizio della trattazione, abbiamo spiegato l'effetto Zenone quantistico in termini di *misure ripetute*, attuate come un rapido susseguirsi di proiezioni di Von Neumann. Ma qual è il *significato fisico* di ciascuna di tale proiezioni?

Secondo il parere attuale di diversi fisici, una proiezione di Von Neumann non è altro che una *notazione abbreviata* dell'effetto (idealizzato) del complicato processo che porta alla misurazione, tramite apparati macroscopici, di caratteristiche microscopiche di un sistema. Pur non essendo chiari i dettagli di tale processo (e in effetti si parla di *measurement problem*), possiamo lo stesso cercarne una *descrizione effettiva*.

In particolare, possiamo usare le “Hamiltoniane” non hermitiane appena introdotte per descrivere l'effetto delle misurazioni, che porta ad un ben preciso andamento della *survival probability*, come osservato.

Partiamo allora dall'esempio (2.13), a cui aggiungiamo un termine non hermitiano $-i2V$, giungendo all'Hamiltoniana (espressa in notazione matriciale nella base $\{|0\rangle, |1\rangle\}$):

$$H_{\text{int}} = \begin{pmatrix} 0 & \Omega \\ \Omega & -2iV \end{pmatrix} = -iV\mathbb{I} + \vec{h} \cdot \vec{\sigma} \quad \vec{\sigma} = \begin{pmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \end{pmatrix}; \quad \vec{h} = \begin{pmatrix} \Omega \\ 0 \\ iV \end{pmatrix} \quad (2.16)$$

Il sistema parte dallo stato iniziale $|0\rangle$ ed evolve verso $|1\rangle$, come visto nell'esempio iniziale. Tuttavia, nella colonna di $|1\rangle$ abbiamo inserito un termine non hermitiano $-i2V$, che non è altro che una versione asimmetrica della modifica già esaminata in (2.15), che ha quindi l'effetto di “assorbire” la probabilità **solo** dallo stato $|1\rangle$ verso cui evolve il sistema, e che funge quindi da “decay channel”. Ci interessa capire se la *survival probability* $p(t)$ calcolata con una H_{int} non hermitiana può riprodurre

l'andamento “esponenziale efficace” dell'effetto Zenone con N proiezioni di Von Neumann ripetute.

Nota: Il fattore 2 presente nel termine non hermitiano serve solo a semplificare i conti, dato che permette una scomposizione di H_{int} come somma di una matrice identità \mathbb{I} per un opportuno fattore, e di un **vettore di Pauli** $\vec{h} \cdot \vec{\sigma}$.

Partiamo allora calcolando l'evoluzione temporale del sistema a partire dallo stato $|\psi_0\rangle = |0\rangle$:

$$|\psi(t)\rangle = e^{-iH_{\text{int}}t} |0\rangle = \exp[-it(-iV\mathbb{I} + \vec{h} \cdot \vec{\sigma})] \stackrel{(a)}{=} \exp(-Vt\mathbb{I}) \exp(-it\vec{h} \cdot \vec{\sigma})$$

dove il passaggio in (a) è giustificato dal fatto che \mathbb{I} commuta con qualsiasi matrice, e in particolare con $\vec{h} \cdot \vec{\sigma}$.

Il primo esponenziale si calcola velocemente:

$$\exp(-Vt\mathbb{I}) = e^{-Vt}\mathbb{I}$$

Per il secondo, invece, usiamo il risultato:

$$e^{i\vec{a} \cdot \vec{\sigma}} = e^{ia(\hat{n} \cdot \vec{\sigma})} = \mathbb{I} \cos(a) + i(\hat{n} \cdot \vec{\sigma}) \sin(a) \quad (2.17)$$

dove $\vec{a} \in \mathbb{C}^3$ è un generico vettore, che fattorizziamo in prodotto di modulo e versore: $\vec{a} = a\hat{n}$, con $\|\hat{n}\| = 1$. Tale formula si dimostra⁵ notando che $(\hat{n} \cdot \vec{\sigma})^{2k} = \mathbb{I}$ e $(\hat{n} \cdot \vec{\sigma})^{2k+1} = \hat{n} \cdot \vec{\sigma}$ (con $k \in \mathbb{Z}$), scrivendo i primi termini dello sviluppo e riconoscendo le espansioni di sin e cos.

Nel nostro caso, partiamo scomponendo \vec{h} :

$$\|\vec{h}\|^2 = \vec{h} \cdot \vec{h} = \Omega^2 - V^2$$

Per esaminare l'effetto Zenone ci servirà un **forte coupling** con l'ambiente, ossia $V \gg \Omega$. Perciò possiamo scrivere il modulo di \vec{h} in termini di un'opportuna grandezza h , che definiamo positiva per comodità:

Strong coupling

$$\|\vec{h}\| = i\sqrt{V^2 - \Omega^2} \equiv ih \Rightarrow h = \sqrt{V^2 - \Omega^2} > 0 \quad (2.18)$$

Possiamo allora applicare la (2.17) all'esponenziale che dobbiamo calcolare:

$$\begin{aligned} \exp(-it\vec{h} \cdot \vec{\sigma}) &= \exp\left(\underbrace{i(-iht)}_a \underbrace{\frac{\vec{h}}{ih} \cdot \vec{\sigma}}_{\hat{n} \cdot \vec{\sigma}}\right) = \mathbb{I} \cos(-iht) + \underbrace{i}_{\text{color red}} \left(\frac{\vec{h} \cdot \vec{\sigma}}{ih}\right) \sin(-iht) = \\ &\stackrel{(a)}{=} \mathbb{I} \cosh(ht) - \underbrace{i}_{\text{color purple}} (-\underbrace{i}_{\text{color red}} \sin(iht)) \left(\frac{\vec{h} \cdot \vec{\sigma}}{h}\right) \stackrel{(b)}{=} \cosh(ht)\mathbb{I} - i \sinh(th) \left(\frac{\vec{h} \cdot \vec{\sigma}}{h}\right) \end{aligned}$$

⁵^Vedi esercizi del corso di Metodi Matematici

Dove in (a) e in (b) abbiamo usato rispettivamente le due relazioni:

$$\cos(ix) = \cosh(x) \quad -i \sin(ix) = \sinh(x)$$

Mettendo insieme i due pezzi otteniamo:

$$e^{-iH_{\text{int}}t} = e^{-Vt} \begin{pmatrix} \cosh(ht) + \frac{V}{h} \sinh(ht) & -i\frac{\Omega}{h} \sinh(ht) \\ -i\frac{\Omega}{h} \sinh(ht) & \cosh(ht) + \frac{V}{h} \sinh(ht) \end{pmatrix}$$

Possiamo ora finalmente calcolare la *survival amplitude* $\mathcal{A}(t)$:

$$\begin{aligned} \mathcal{A}(t) &= \langle 0 | e^{-iH_{\text{int}}t} | 0 \rangle = e^{-Vt} \left[\cosh(ht) + \frac{V}{h} \sinh(ht) \right] = \\ &\stackrel{(a)}{=} \frac{1}{2} \left(1 + \frac{V}{h} \right) e^{-(V-h)t} + \frac{1}{2} \left(1 - \frac{V}{h} \right) e^{-(V+h)t} \end{aligned}$$

dove in (a) abbiamo riscritto le funzioni iperboliche tramite esponenziali ($\sinh(x) = (e^x - e^{-x})/2$, $\cosh(x) = (e^x + e^{-x})/2$).

Poiché $V \gg \Omega$ si ha che $h = \sqrt{V^2 - \Omega^2} \approx V$. Perciò l'argomento dell'esponenziale in giallo è vicino a 0 (decadimento lento), mentre quello dell'esponenziale in azzurro è circa $-2Vt$ (decadimento veloce).

Perciò, facendo passare un piccolo tempo t , è solo il primo termine a dominare. In ogni caso, per $t \rightarrow +\infty$ l'ampiezza viene "completamente assorbita" e $\mathcal{A}(t) \rightarrow 0$ - esattamente come vogliamo.

Nell'esaminare la *survival probability* $p(t)$ concentriamoci sul caso di tempi sufficientemente lunghi, in modo da trascurare il secondo termine. Visto che $\mathcal{A}(t)$ è reale, $p(t)$ è data da:

$$p(t) \approx \left[\frac{1}{2} \left(1 + \frac{V}{h} \right) \exp(-(V-h)t) \right]^2$$

Approssimiamo tale espressione usando $V \gg \Omega$, e quindi espandendo attorno a $\Omega/V = 0$. Per l'esponenziale troviamo:

$$\begin{aligned} \exp(-2t(V-h)) &= \exp(2t(-V + \sqrt{V^2 - \Omega^2})) = \exp\left(2Vt \left(-1 + \sqrt{1 - \frac{\Omega^2}{V^2}}\right)\right) = \\ &\stackrel{(a)}{\approx} \exp\left(2tV \left(-1 + 1 - \frac{\Omega^2}{2V^2}\right)\right) = \exp\left(-\frac{\Omega^2}{V}t\right) \end{aligned}$$

dove in (a) si è usata la solita espansione:

$$(1-x)^n \approx 1 - nx \quad x \rightarrow 0$$

Per il fattore iniziale otteniamo invece:

$$\left[\frac{1}{2} \left(1 + \frac{V}{h} \right) \right]^2 = \frac{1}{4} \left(1 + \frac{1}{\sqrt{1 - \frac{\Omega^2}{V^2}}} \right)^2 \approx 1 + \frac{\Omega^2}{2V^2}$$

Mettendo tutto insieme giungiamo a:

$$p(t) \approx \left(1 + \frac{\Omega^2}{2V^2}\right) \exp\left(-\frac{\Omega^2}{V}t\right) \quad (2.19)$$

che vale per t sufficientemente alti (per $t = 0$ produce addirittura una normalizzazione > 1 , decisamente errata, ma che viene compensata velocemente).

Come ci si aspetta, la $p(t)$ decade con un andamento *esponenziale efficace* che possiamo confrontare direttamente con il γ_{eff} in (2.14), giungendo ad una relazione tra V e il tempo τ che intercorre tra una proiezione e l'altra nell'equivalente evoluzione considerata in partenza (tenendo conto del fatto che la $p(t)$ ricavata è frutto di un'approssimazione):

$$\frac{\Omega^2}{V} \approx \tau \Omega^2 \Rightarrow \tau \approx \frac{1}{V} \quad (2.20)$$

Tale risultato è decisamente *controintuitivo*. Il parametro V quantifica il coupling con l'ambiente. Aumentandolo, ci aspettiamo, pittorescamente, di *aprire un grosso varco* verso l'esterno, consentendo a una grande quantità di “fluido” (la probabilità) di scappar via più velocemente. Intuitivamente, perciò, V maggiori dovrebbero corrispondere ad una survival probability che scende molto più rapidamente.

Tuttavia, V compare al denominatore di (2.19). Ciò indica che una *forte interazione* con l'ambiente “stazionarizza” il sistema (figura 2.12). In un certo senso, V grande fa sì che l'ambiente “misuri molto frequentemente” (precisamente, una volta ogni $\tau = 1/V$, in una sorta di *continuous measurement*) il sistema in esame, producendo di conseguenza l'effetto Zenone.

Attenzione! Nel paper di riferimento, nel primo fattore V al denominatore non è al quadrato - ma ciò non concorda con i conti che ho rifatto

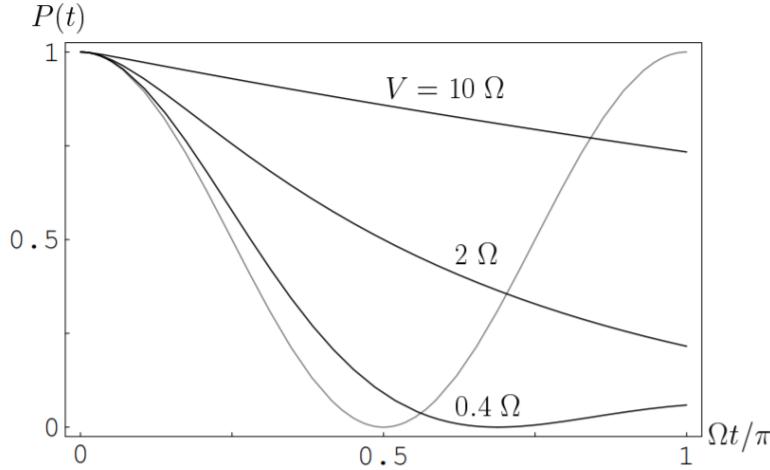


Figura (2.12) – Survival probability $p(t)$ al variare del parametro di coupling V . Coupling maggiori equivalgono ad un decadimento *più lento* dallo stato iniziale.

Perciò è possibile per un sistema rimanere *limitato allo stato iniziale* solo a seguito di una forte interazione con un sistema esterno.

2.4.8 L'origine delle “Hamiltoniane” non hermitiane

Finora abbiamo motivato fisicamente l'utilizzo di “Hamiltoniane” non hermitiane come modelli in cui il sistema interagisce con un *ambiente* più grande che non è accessibile allo sperimentatore, generando una “fuga di probabilità”.

Vediamo ora, matematicamente, come si possa giustificare tutto ciò, almeno nell'esempio *semplice* del sistema a due stati finora caratterizzato.

Partiamo dal sistema in (2.13), in cui abbiamo due stati $|0\rangle$ e $|1\rangle$ *non stazionari*, per cui il sistema può oscillare da uno all'altro. Immaginiamo che tale sistema sia immerso nell'ambiente, che modellizziamo con un *continuo* di stati $|\omega\rangle$, ciascuno dei quali avrà una certa energia. Da queste considerazioni, l'Hamiltoniana nella base ON $\{|0\rangle, |1\rangle, |\omega\rangle\}$ diviene:

$$H_0 = \underbrace{\Omega(|0\rangle\langle 1| + |1\rangle\langle 0|)}_{\text{Qubit}} + \underbrace{\int d\omega \omega |\omega\rangle\langle \omega|}_{\text{Ambiente}}$$

Aggiungiamo infine un'interazione tra qubit e ambiente. Seguendo l'idea di (2.16), denotiamo $|1\rangle$ come *decay channel* del sistema. In altre parole, un sistema che parte a $|0\rangle$ può passare, per evoluzione unitaria, solo a $|1\rangle$, da cui però può evolvere ad un qualsiasi $|\omega\rangle$ (o tornare a $|0\rangle$). Come visto in precedenza, le transizioni tra un livello e l'altro sono regolate dai *coupling* tra stati, ossia dai termini fuori dalla diagonale. Accoppiare $|1\rangle$ e $|\omega\rangle$ significa aggiungere opportuni termini $|1\rangle\langle \omega|$ e il simmetrico $|\omega\rangle\langle 1|$, giungendo (nel caso di ω continuo) ad un'Hamiltoniana:

$$H = \underbrace{\Omega(|0\rangle\langle 1| + |1\rangle\langle 0|)}_{\text{Qubit}} + \underbrace{\int d\omega \omega |\omega\rangle\langle \omega|}_{\text{Ambiente}} + \underbrace{\sqrt{\frac{\Gamma}{2\pi}} \int d\omega \overbrace{(|1\rangle\langle \omega| + |\omega\rangle\langle 1|)}^{g(\omega)}}_{\text{Interazione qubit-ambiente}} \quad (2.21)$$

La funzione $g(\omega)$ specifica l'accoppiamento (coupling) tra lo stato del qubit che funge da *decay channel* ($|1\rangle$) e l'ambiente circostante. In questo caso $g(\omega)$ è *piatta*, ossia $|1\rangle$ è accoppiato in ugual misura a tutti gli stati del continuo $|\omega\rangle$ - e infatti l'*intensità* di tale coupling è stata estratta a fattore comune nel termine $\sqrt{\Gamma}$, normalizzato a $\sqrt{2\pi}$ per convenzione sulle trasformate di Fourier.

Notazione. Una qualsiasi matrice M , fissata una base ON $\{|v_i\rangle\}$, può essere espressa in notazione di Dirac come:

$$M = \sum_{ij} \alpha_{ij} |v_i\rangle\langle v_j| \quad (2.22)$$

Dove gli α_{ij} sono gli **elementi di matrice** di M , cioè le sue entrate:

$$\langle v_i | M | v_j \rangle = \alpha_{ij} \quad (2.23)$$

Si nota infatti che l'espressione (2.23) è identicamente verificata usando per M

l'espressione vista in (2.22):

$$\langle v_i | \left(\sum_{ij} \alpha_{ij} |v_i\rangle \langle v_j| \right) |v_j\rangle \stackrel{(a)}{=} \alpha_{ij}$$

dove in (a) abbiamo usato l'ortonormalità: $\langle v_i | v_j \rangle = \delta_{ij}$.

Per esempio, la matrice di Pauli σ_x è data, nella base $\{|+\rangle, |-\rangle\}$ degli autoket di S_z , da:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = |+\rangle \langle -| + |-\rangle \langle +|$$

Se consideriamo $|0\rangle$ e $|1\rangle$ come equivalenti rispettivamente a $|+\rangle$ e $|-\rangle$, si ha:

$$H_{00} = \Omega \sigma_x = \Omega(|0\rangle \langle 1| + |1\rangle \langle 0|)$$

E otteniamo perciò l'espressione usata in (2.21) (limitatamente alla parte del sistema-qubit).

Nota: l'Hamiltoniana (2.21) compare in *Quantum Field Theory* (QFT) per descrivere l'interazione tra un sistema a due livelli e un “one-dimensional massless boson field in the rotating-wave approximation”.

Ad un'istante t , una generica funzione d'onda del sistema $|\psi(t)\rangle$ si può scrivere come combinazione lineare dei termini della base $\{|0\rangle, |1\rangle, |\omega\rangle\}$:

$$|\psi(t)\rangle = x(t) |0\rangle + y(t) |1\rangle + \int d\omega z(\omega, t) |\omega\rangle$$

dove i coefficienti sono generiche funzioni del tempo.

Scriviamo allora l'equazione di Schrödinger non stazionaria (si ricorda $\hbar = 1$):

$$i \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle \quad (2.24)$$

Calcolando $H |\psi(t)\rangle$ è dato da:

$$H |\psi(t)\rangle = \Omega(y(t) |0\rangle + x(t) |1\rangle) + \int d\omega z(\omega, t) \omega |\omega\rangle + \sqrt{\frac{\Gamma}{2\pi}} \int d\omega (z(\omega, t) |1\rangle + y(t) |\omega\rangle)$$

Lo sostituiamo in (2.24), e separiamo i vari termini in tre equazioni prendendo i prodotti scalari con $|0\rangle$, $|1\rangle$ e un generico $|\omega\rangle$ (che ci consente di “far collassare” gli integrali, dato che supponiamo che gli $|\omega\rangle$ siano ON):

$$\begin{aligned} |0\rangle : i\dot{x}(t) &= \Omega y(t) \\ |1\rangle : i\dot{y}(t) &= \Omega x(t) + \sqrt{\frac{\Gamma}{2\pi}} \int d\omega z(\omega, t) \\ |\omega\rangle : i\dot{z}(t) &= \omega z(\omega, t) + \sqrt{\frac{\Gamma}{2\pi}} y(t) \end{aligned}$$

che formano un sistema di equazioni differenziali lineari alle derivate parziali. Imponiamo, come **condizioni iniziali**, che il sistema si trovi a $t = 0$ in $|0\rangle$, e quindi che i coefficienti siano:

$$\begin{cases} x(0) = 1 \\ y(0) = 0 \\ z(\omega, 0) = 0 \end{cases}$$

L'equazione per z ha soluzione⁶:

$$z(\omega, t) = -i\sqrt{\frac{\Gamma}{2\pi}} \int_0^t d\tau e^{-i\omega(t-\tau)} y(\tau) \quad (2.25)$$

Sostituendo (2.25) nell'equazione per $y(t)$ giungiamo a:

$$i\dot{y}(t) = \Omega x(t) - i\frac{\Gamma}{2\pi} \int d\omega \int d\tau e^{-i\omega(t-\tau)} y(\tau) = \Omega x(t) - i\frac{\Gamma}{2} y(t)$$

Otteniamo perciò un sistema di sole due equazioni differenziali, che sono quelle che riguardano gli stati $|0\rangle$ e $|1\rangle$ del qubit che ci interessano:

$$\begin{cases} i\dot{x} = \Omega y(t) \\ i\dot{y} = -i\frac{\Gamma}{2} y(t) + \Omega x(t) \end{cases}$$

che possiamo esprimere in forma matriciale come:

$$i\frac{d}{dt} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & \Omega \\ \Omega & -i\frac{\Gamma}{2} \end{pmatrix}}_{H_{\text{eff}}} \begin{pmatrix} x \\ y \end{pmatrix} \quad (2.26)$$

La (2.26) ha la struttura dell'equazione di Schrödinger non stazionaria per un sistema a due livelli con una Hamiltoniana *efficace* non-hermitiana - che corrisponde a quella usata in (2.16) ponendo $\Gamma = 4V$.

Perciò, **riepilogando**, il sistema *completo* evolve con una H hermitiana (2.21) in modo unitario, mentre la parte legata al qubit si evolve in modo *non unitario* tramite una H non hermitiana, poiché le transizioni da $|1\rangle$ agli stati $|\omega\rangle$ esterni vengono qui interpretate come una *perdita di probabilità* del sistema.

Se il coupling con l'ambiente (regolato da Γ o, equivalentemente, da V) è sufficientemente forte, controintuitivamente, l'evoluzione del sistema viene “stallata” dall'effetto di interazione - e la stessa cosa succede nel caso eseguitissimo sul sistema una serie di *proiezioni di Von Neumann*. Questa è l'essenza dell'**effetto Zenone quantistico**. Perciò, in un certo senso, potremmo dire che l'ambiente effettua delle “misurazioni continue” sul sistema, con una frequenza proporzionale al coupling Γ (o V).

⁶ΛLo enunciamo solamente, è poi immediato verificarla per sostituzione

2.5 Implementazione di porte logiche

Per realizzare *fisicamente* una porta logica dobbiamo trovare un *sistema* che abbia la giusta Hamiltoniana per produrre l'evoluzione temporale desiderata. Si tratta, in altre parole, di risolvere l'equazione di Schrödinger dipendente dal tempo “al contrario”.

Un caso che può essere risolto analiticamente è quello della porta logica **NOT** quantistica, che si realizza con un sistema che effettua *oscillazioni di Rabi* analogo a quello visto in (2.13). L'idea sta nel controllare la presenza o meno di oscillazioni tra i livelli, e la loro pulsazione Ω . In maniera sintetica, il comportamento diviene:

1. Il sistema parte da un generico stato $|\psi_i\rangle$, che prendiamo pari a $|0\rangle$ per semplicità
2. Si “accende” l'oscillazione, e si aspetta un certo tempo τ affinché il sistema completi un mezzo ciclo. In questo modo, la componente lungo $|0\rangle$ di $|\psi_i\rangle$ viene *ruotata* lungo $|1\rangle$, e quella lungo $|0\rangle$ verso $|1\rangle$.
3. Il sistema finale è (a meno di fasi) $|1\rangle$, cioè l'opposto dello stato di partenza.

Vediamo, nel dettaglio, un esempio di implementazione fisica di tutto ciò.

Consideriamo una particella di spin 1/2 isolata (il qubit di partenza) in presenza di un campo magnetico B_0 stazionario lungo \hat{z} , e uno oscillante B_1 che ruota attorno a \hat{z} , mantenendosi sul piano $\hat{x}\hat{y}$.

L'Hamiltoniana di tale sistema è data da:

$$H = -\mu \{B_0 \hat{\sigma}_z + B_1 [\cos(\omega t) \hat{\sigma}_x + \sin(\omega t) \hat{\sigma}_y]\}$$

dove $\hat{\sigma}_i$ sono le matrici di Pauli.

Come **esercizio**⁷, vogliamo:

1. Studiare l'equazione di Schrödinger
2. Studiare la soluzione per il caso $\omega = -2\mu B_0/\hbar$ in cui il campo oscillante è *in risonanza* con la transizione energetica tra i due livelli.

Soluzione

1. Uno stato generico $|\psi(t)\rangle$ è dato da:

$$|\psi(t)\rangle = \alpha(t) |0\rangle + \beta(t) |1\rangle = \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} \quad (2.27)$$

dove per la forma matriciale stiamo usando la base computazionale $\{|0\rangle, |1\rangle\}$. Scriviamo allora l'equazione Schrödinger dipendente dal tempo:

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

⁷^Vedi Es. 3.22 pag. 180 di [7].

Inserendovi $|\psi(t)\rangle$ otteniamo, in forma matriciale:

$$i\hbar \frac{d}{dt} \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} = \begin{pmatrix} -\mu B_0 & -\mu B_1 [\cos(\omega t) - i \sin(\omega t)] \\ -\mu B_1 [\cos(\omega t) + i \sin(\omega t)] & +\mu B_0 \end{pmatrix} \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix}$$

Usando l'identità di Eulero $e^{ix} = \cos(x) + i \sin(x)$, giungiamo a:

$$i \begin{pmatrix} \dot{\alpha}(t) \\ \dot{\beta}(t) \end{pmatrix} = \frac{1}{\hbar} \begin{pmatrix} -\mu B_0 & -\mu B_1 e^{-i\omega t} \\ -\mu B_1 e^{i\omega t} & \mu B_0 \end{pmatrix} \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} = \begin{pmatrix} \frac{\omega_0}{2} \alpha(t) & \frac{\omega_1}{2} e^{-i\omega t} \beta(t) \\ \frac{\omega_1}{2} e^{i\omega t} \alpha(t) & -\frac{\omega_0}{2} \beta(t) \end{pmatrix}$$

dove abbiamo introdotto le pulsazioni ω_0 e ω_1 :

$$\omega_0 = -2\mu \frac{B_0}{\hbar} \quad \omega_1 = -2\mu \frac{B_1}{\hbar} \quad (2.28)$$

Nota: il fattore 2 che evidenziamo serve a prepararsi per la risoluzione del secondo punto, in cui studieremo il caso con $\omega = -2\mu B_0/\hbar = \omega_0$.

L'equazione matriciale equivale ad un sistema lineare di due equazioni differenziali:

$$\begin{cases} i\dot{\alpha}(t) = \frac{\omega_0}{2} \alpha(t) + \frac{\omega_1}{2} e^{-i\omega t} \beta(t) \\ i\dot{\beta}(t) = \frac{\omega_1}{2} e^{i\omega t} \alpha(t) - \frac{\omega_0}{2} \beta(t) \end{cases}$$

Per risolvere il sistema effettuiamo un cambio di coordinate, con lo scopo di “simmetrizzare” le esponenziali e semplificarli. Poniamo allora:

$$\begin{cases} a(t) \equiv \alpha(t) e^{i\omega t/2} \\ b(t) \equiv \beta(t) e^{-i\omega t/2} \end{cases} \Rightarrow \begin{cases} \alpha(t) = a(t) e^{-i\omega t/2} \\ \beta(t) = b(t) e^{+i\omega t/2} \end{cases}$$

E ricaviamo le sostituzioni per le derivate prime:

$$\begin{cases} \dot{\alpha}(t) = \dot{a}(t) e^{-i\omega t/2} + a(t) \left(-\frac{i\omega}{2}\right) e^{-i\omega t/2} \\ \dot{\beta}(t) = \dot{b}(t) e^{i\omega t/2} + b(t) \left(\frac{i\omega}{2}\right) e^{i\omega t/2} \end{cases}$$

Così facendo, il sistema diviene:

$$\begin{aligned} & \begin{cases} i\dot{a}(t) e^{-i\omega t/2} = \frac{\omega_0}{2} a(t) e^{-i\omega t/2} - \frac{\omega}{2} a(t) e^{-i\omega t/2} + \frac{\omega_1}{2} b(t) e^{-i\omega t/2} \\ i\dot{b}(t) e^{i\omega t/2} = \frac{\omega_1}{2} a(t) e^{i\omega t/2} + \frac{\omega}{2} b(t) e^{i\omega t/2} - \frac{\omega_0}{2} b(t) e^{i\omega t/2} \end{cases} \\ & \Rightarrow \begin{cases} i\dot{a}(t) = \left(\frac{\omega_0 - \omega}{2}\right) a(t) + \frac{\omega_1}{2} b(t) \\ i\dot{b}(t) = \frac{\omega_1}{2} a(t) - \left(\frac{\omega_0 - \omega}{2}\right) b(t) \end{cases} \Rightarrow i \frac{d}{dt} \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \omega_0 - \omega & \omega_1 \\ \omega_1 & -(\omega_0 - \omega) \end{pmatrix} \end{aligned}$$

Moltiplicando entrambi i membri della forma matriciale per \hbar ci riconduciamo all'equazione di Schrödinger dipendente dal tempo, per un sistema con un'Hamiltoniana \tilde{H} data da:

$$\tilde{H} = \frac{\hbar}{2} \begin{pmatrix} \Delta\omega_0 & \omega_1 \\ \omega_1 & -\Delta\omega_0 \end{pmatrix} \quad \Delta\omega_0 \equiv \omega_0 - \omega \quad (2.29)$$

Tale \tilde{H} è decisamente più semplice di quella da cui siamo partiti, e perciò conviene studiare l'evoluzione temporale in questo sdr. Indicando con $|\tilde{\psi}(t)\rangle$ la funzione d'onda nelle coordinate $\{a, b\}$ del sdr ruotato, avremo:

$$|\tilde{\psi}(t)\rangle = a(t)|0\rangle + b(t)|1\rangle \quad |\tilde{\psi}(t)\rangle = \exp\left(-\frac{i}{\hbar}\tilde{H}t\right)|\tilde{\psi}(0)\rangle \quad (2.30)$$

Se scriviamo il cambio di coordinate in forma matriciale:

$$\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} e^{i\omega t/2} & 0 \\ 0 & e^{-i\omega t/2} \end{pmatrix} \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} \quad (2.31)$$

ci accorgiamo che stiamo usando una matrice di rotazione di spin. Identificando per esempio $|0\rangle$ e $|1\rangle$ della base computazionale con gli autostati di S_z , ricordiamo che, fissato un versore \hat{n} ad angolo θ con \hat{z} e φ con \hat{x} , l'operatore di spin in tale direzione si scrive⁸:

$$\hat{S}(\theta, \varphi) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right)e^{-i\varphi/2} & \sin\left(\frac{\theta}{2}\right)e^{-i\varphi/2} \\ \sin\left(\frac{\theta}{2}\right)e^{i\varphi/2} & \cos\left(\frac{\theta}{2}\right)e^{i\varphi/2} \end{pmatrix}$$

Perciò la matrice in (2.31) è proprio $\hat{S}(0, -\omega t)$, che corrisponde ad una rotazione attorno a \hat{z} di un angolo $-\omega t$ per il passaggio da $\{\alpha, \beta\}$ a $\{a, b\}$. Per la trasformazione inversa basta cambiare il segno dell'angolo:

$$\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \hat{S}(0, -\omega t) \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} \quad \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} = \hat{S}(0, +\omega t) \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}$$

Nella notazione delle funzioni d'onda:

$$|\tilde{\psi}(t)\rangle = \hat{S}(0, -\omega t)|\psi(t)\rangle \quad |\psi(t)\rangle = \hat{S}(0, +\omega t)|\tilde{\psi}(t)\rangle$$

Mettendo tutto insieme, e usando la notazione sintetica $R_z(\theta)$ per indicare la rotazione $\hat{S}(0, \theta)$ attorno a \hat{z} , troviamo:

$$\begin{aligned} |\psi(t)\rangle &= R_z(+\omega t)|\tilde{\psi}(t)\rangle = R_z(\omega t) \exp\left(-\frac{i}{\hbar}\tilde{H}t\right)|\tilde{\psi}(0)\rangle = \\ &= \underbrace{R_z(\omega t) \exp\left(-\frac{i}{\hbar}t\tilde{H}\right) R_z(-\omega t)}_{\exp(-iHt/\hbar)}|\psi(0)\rangle \end{aligned}$$

⁸^Vedi pag. 241 degli appunti di Fisica Teorica

Ci manca solo calcolare l'esponenziale di \tilde{H} . Per farlo conviene riscrivere \tilde{H} nella base in cui è diagonale, costituita dai suoi autoket:

$$\tilde{H} = \mathcal{E}_1 |\mathcal{E}_1\rangle \langle \mathcal{E}_1| + \mathcal{E}_2 |\mathcal{E}_2\rangle \langle \mathcal{E}_2|$$

Partendo allora da (2.29), cerchiamo gli autovalori $\lambda_{1,2}$ della matrice senza il prefattore $\hbar/2$:

$$\begin{aligned} \det \left(\frac{2}{\hbar} \tilde{H} - \lambda \mathbb{I} \right) &= [-(\Delta\omega_0)^2 + \lambda^2 - \omega_1^2] \stackrel{!}{=} 0 \Rightarrow \lambda^2 = (\Delta\omega_0)^2 + \omega_1^2 \\ &\Rightarrow \lambda_{1,2} = \pm \sqrt{(\Delta\omega_0)^2 + \omega_1^2} \end{aligned}$$

E reinserendo il fattore $\hbar/2$:

Autovalori di \tilde{H}

$$\mathcal{E}_1 = \frac{\hbar}{2} \sqrt{(\Delta\omega_0)^2 + \omega_1^2} \quad \mathcal{E}_2 = -\frac{\hbar}{2} \sqrt{(\Delta\omega_0)^2 + \omega_1^2} \quad \Delta\omega_0 = \omega_0 - \omega \quad (2.32)$$

Per l'autovettore $|\mathcal{E}_1\rangle$ cerchiamo il ker di $\tilde{H} - \mathcal{E}_1 \mathbb{I}$, che è lo spazio dei vettori $(a, b)^T$ che soddisfano l'equazione:

$$a(\Delta\omega_0 - \sqrt{(\Delta\omega_0)^2 + \omega_1^2}) + b\omega_1 = 0 \Rightarrow -b \frac{\omega_1}{\Delta\omega_0} = a \left(1 - \sqrt{1 + \frac{\omega_1^2}{(\Delta\omega_0)^2}} \right)$$

Possiamo trovare una soluzione in forma migliore ponendo:

$$\frac{\omega_1}{\Delta\omega_0} \equiv \tan \theta$$

In questo modo all'interno della radice possiamo usare l'identità trigonometrica $1 + \tan^2 \theta = \sec^2 \theta$, e giungere a:

$$-b \tan \theta = a \left(1 - \frac{1}{\cos \theta} \right) \Rightarrow a \stackrel{(a)}{=} b \frac{\sin \theta}{\frac{1 - \cos \theta}{2}} \frac{1}{2} \stackrel{(b)}{=} b \frac{2 \sin \frac{\theta}{2} \cos \frac{\theta}{2}}{\sin^2 \frac{\theta}{2}} \frac{1}{2} = b \frac{\cos \frac{\theta}{2}}{\sin \frac{\theta}{2}}$$

dove in (a) e in (b) si sono usate rispettivamente la formula di bisezione del cos e di duplicazione del sin. Poiché b si può scegliere arbitrariamente, poniamo $b = \sin \theta/2$ per semplicità, ottenendo per il primo autovettore:

$$|\mathcal{E}_1\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix}$$

Conti analoghi (con una sola differenza di segno) portano a trovare anche $|\mathcal{E}_2\rangle$, giungendo così alla fine a:

Autovettori di \tilde{H}

$$|\mathcal{E}_1\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix} \quad |\mathcal{E}_2\rangle = \begin{pmatrix} -\sin \frac{\theta}{2} \\ \cos \frac{\theta}{2} \end{pmatrix} \quad \tan \theta = \frac{\omega_1}{\Delta\omega_0}$$

Ponendo allora $|\psi(0)\rangle = |0\rangle$ possiamo calcolare l'esponenziale, e quindi l'**evoluzione temporale** cercata:

$$\begin{aligned} |\tilde{\psi}(t)\rangle &= \sum_{i=1}^2 c_i \exp\left(-\frac{i}{\hbar}t\mathcal{E}_i\right) |\mathcal{E}_i\rangle \quad c_i = \langle \mathcal{E}_i | 0 \rangle \\ &= \cos \frac{\theta}{2} \exp\left(-\frac{i}{\hbar}t\mathcal{E}_1\right) |\mathcal{E}_1\rangle - \sin \frac{\theta}{2} \exp\left(-\frac{i}{\hbar}t\mathcal{E}_2\right) |\mathcal{E}_2\rangle = \\ &= \begin{pmatrix} \cos^2 \frac{\theta}{2} \exp\left(-\frac{i}{\hbar}t\mathcal{E}_1\right) + \sin^2 \frac{\theta}{2} \exp\left(-\frac{i}{\hbar}t\mathcal{E}_2\right) \\ \cos \frac{\theta}{2} \sin \frac{\theta}{2} [\exp\left(-\frac{i}{\hbar}t\mathcal{E}_1\right) - \exp\left(-\frac{i}{\hbar}t\mathcal{E}_2\right)] \end{pmatrix} \end{aligned}$$

Possiamo ora calcolare la survival probability per lo stato $|1\rangle$:

$$\begin{aligned} p_1(t) &= |\langle 1 | \psi(t) \rangle|^2 = |\beta(t)|^2 \stackrel{(a)}{=} |\langle \tilde{1} | \tilde{\psi}(t) \rangle|^2 = |b(t)|^2 = \\ &= \left| \cos \frac{\theta}{2} \sin \frac{\theta}{2} \left[\exp\left(-\frac{i}{\hbar}t\mathcal{E}_1\right) - \exp\left(-\frac{i}{\hbar}t\mathcal{E}_2\right) \right] \right|^2 \end{aligned}$$

In (a) si è usato il fatto che la trasformazione **unitaria** $R_z(\omega t)$ che collega le due autofunzioni *preserva* i prodotti scalari (geometricamente si intuisce dal fatto che una rotazione preserva le lunghezze).

Semplifichiamo questa espressione. Partiamo da:

$$\left| \cos \frac{\theta}{2} \sin \frac{\theta}{2} \right|^2 = \frac{1}{4} \sin^2 \theta = \frac{1}{4} \sin^2 \arctan \frac{\omega_1}{\Delta\omega_0} \stackrel{(b)}{=} \frac{1}{4} \frac{\omega_1^2}{\Delta\omega_0^2 + \omega_1^2}$$

dove in (b) si è usato:

$$\sin(\arctan x) = \frac{x}{\sqrt{1+x^2}}$$

D'altro canto, nella differenza delle due esponenziali riconosciamo un sin. Sostituendo le espressioni per \mathcal{E}_1 e \mathcal{E}_2 :

$$\begin{aligned} &\left| \frac{2i}{2i} \left(\exp\left(-\frac{it}{2}\sqrt{\Delta\omega_0^2 + \omega_1^2}\right) - \exp\left(\frac{it}{2}\sqrt{\Delta\omega_0^2 + \omega_1^2}\right) \right) \right|^2 = \\ &= \left| -2i \sin\left(\sqrt{\Delta\omega_0^2 + \omega_1^2} \frac{t}{2}\right) \right|^2 = 4 \sin^2\left(\sqrt{\Delta\omega_0^2 + \omega_1^2} \frac{t}{2}\right) \end{aligned}$$

Mettendo tutto insieme otteniamo:

$$p_1(t) = \frac{\omega_1^2}{\Delta\omega_0^2 + \omega_1^2} \sin^2\left(\underbrace{\sqrt{(\omega_0 - \omega)^2 + \omega_1^2}}_{\Omega} \frac{t}{2}\right)$$

dove Ω è detta **frequenza di Rabi**.

La probabilità di transizione a $|1\rangle$, perciò, oscilla nel tempo.

2. Esaminiamo ora il **caso risonante**, ossia quello in cui $\omega = -2\mu B_0/\hbar = \omega_0$. Ne deriva che $\Delta\omega_0 = 0$, e quindi:

$$p_1(t) = \sin^2\left(\frac{\omega_1 t}{2}\right)$$

In generale, l'ampiezza delle oscillazioni di probabilità è < 1 . Solo nel caso risonante, con $\omega = \omega_0$, si ottengono “oscillazioni complete” - ossia tali che, dopo π/Ω , si abbia probabilità di transizione a $|1\rangle$ (partendo da $|0\rangle$) pari a 1 (certezza). In tal caso, inoltre $\Omega = \omega_1$ (figura 2.13).

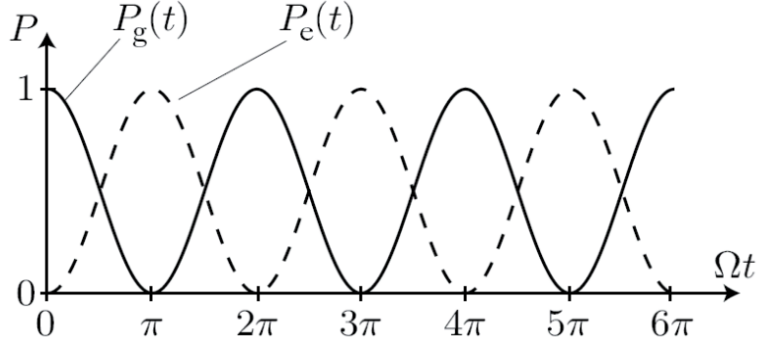


Figura (2.13) – Probabilità di transizione a $|0\rangle$ per uno stato di partenza $|0\rangle$ ($P_g(t)$) o $|1\rangle$ ($P_e(t)$). Si nota che dopo un tempo π/Ω lo stato finale è *invertito* rispetto a quello iniziale.

Si può sfruttare tale dinamica per costruire un gate NOT. L'idea è di inizializzare lo stato del qubit a $|\psi\rangle \in \mathbb{C}^2$, accendere il campo magnetico rotante ponendo $\omega = \omega_0$, e aspettare $\Delta t = \pi/\Omega = \pi/\omega_1$. A questo punto, lo stato finale è $\sigma_x |\psi\rangle$, ossia NOT $|\psi\rangle$, come desiderato.

A livello fisico, ciò corrisponde ad *eccitare* la transizione $|0\rangle \leftrightarrow |1\rangle$ con una pulsazione $\omega = (\mathcal{E}_1 - \mathcal{E}_2)/\hbar = \omega_1$ (come si verifica dalle espressioni (2.32)).

2.6 Implementazione di un gate CNOT

Per poter realizzare gate a 2 qubit è necessario considerare *interazioni* tra i singoli qubit. A scopo dimostrativo, esaminiamo ora un semplice modello di gate CNOT. Partiamo da un'Hamiltoniana $H(t)$:

$$H(t) = H_s + H_p(t); \quad H_s = -(\mu_1 \sigma_z^{(1)} + \mu_2 \sigma_z^{(2)}) B_0 + J \sigma_z^{(1)} \sigma_z^{(2)}$$

H_s è l'Hamiltoniana che descrive due particelle di spin 1/2 sottoposte ad un campo magnetico B_0 . Le direzioni di spin delle due particelle non sono indipendenti, ma interagiscono tra loro mediante il termine $J \sigma_z^{(1)} \sigma_z^{(2)}$, che ha l'effetto di aggiungere una *correlazione*⁹ tra le due. Nello specifico, per $J > 0$ gli spin tenderanno a

⁹ L'idea è quella del modello di Ising, che caratterizza il ferromagnetismo in meccanica statistica.

disporsi antiparallelamente (per minimizzare l'energia H_s), o parallelamente per $J < 0$.

$H_p(t)$, d'altro canto, è il potenziale perturbativo che consente di controllare dall'esterno il sistema (es. impulso laser).

Nota: H_s è una matrice 4×4 , dato che agisce su stati di 2 qubit. Perciò $\sigma_z^{(1)}\sigma_z^{(2)}$ sottintende in realtà un prodotto tensore: $\sigma_z^{(1)} \otimes \sigma_z^{(2)}$. Analogamente, quando un operatore agisce su un solo qubit, va inteso moltiplicato per un'opportuna matrice identità. In questo caso:

$$\mu_1 \sigma_z^{(1)} = \mu_1 \sigma_z^{(1)} \otimes \mathbb{I}_2$$

Nella base computazionale $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, dove $|0\rangle$ e $|1\rangle$ sono autostati di σ_z , troviamo che H_s è diagonale:

$$H_s = \text{diag} \left[-(\mu_1 + \mu_2)B_0 + J, -(\mu_1 - \mu_2)B_0 - J, (\mu_1 - \mu_2)B_0 - J, (\mu_1 + \mu_2)B_0 + J \right]$$

Esaminiamo le energie necessarie per le transizioni $|00\rangle \rightarrow |01\rangle$ e $|10\rangle \rightarrow |11\rangle$ (spin flip della seconda particella):

$$\begin{aligned} \Delta\mathcal{E}_1 &= \mathcal{E}_{01} - \mathcal{E}_{00} = 2(B_0\mu_2 - J) \\ \Delta\mathcal{E}_2 &= \mathcal{E}_{11} - \mathcal{E}_{10} = 2(\mu_2 B_0 + J) \end{aligned}$$

Notiamo che, se $J = 0$ (qubit non interagenti), $\Delta\mathcal{E}_1 = \Delta\mathcal{E}_2$. In tal caso, un laser di pulsazione $\omega = \Delta\mathcal{E}/\hbar$, ossia:

$$\omega = \frac{2\mu_2 B_0}{\hbar}$$

realizza entrambe le transizioni. In altre parole, per $J = 0$, l'azione di *invertire* il secondo qubit non dipende in alcun modo dallo stato del primo.

Nel caso $J \neq 0$, tuttavia, $\Delta\mathcal{E}_1 \neq \Delta\mathcal{E}_2$. Perciò, usando un laser con $\omega = \mathcal{E}_2/\hbar$, si può selezionare la sola transizione $|11\rangle \leftrightarrow |10\rangle$. In altre parole, il secondo qubit viene *invertito* solamente nel caso il primo sia nello stato $|1\rangle$ - replicando così il comportamento di un gate CNOT.

(Lezione 5 • del
13/3/2019)

3.1 Matrici densità

In questa sezione ci occuperemo di introdurre una notazione più avanzata e flessibile per trattare gli **stati** di un sistema quantistico nella loro più ricca generalità. Non sempre, infatti, è possibile caratterizzare un sistema con una precisa funzione d'onda, dato che ciò comporta possedere la *massima informazione* possibile (stato puro). Spesso si ha a che fare con incertezze o informazioni parziali (stati misti), e saper trattare tali frequenti situazioni permette di accedere ad un insieme di possibilità computazionali molto più ampio.

3.1.1 Misture statistiche

Per **mistura statistica** si intende un sistema S che è descritto da una funzione d'onda $|\psi_k\rangle$ che *non conosciamo*, scelta in un insieme di N possibilità $\{|\psi_i\rangle\}_{i=1\dots N}$, ciascuna con probabilità (classica) p_i (con $p_i > 0 \forall i$ e $\sum_{i=1}^N p_i = 1$).

Mistura statistica

Nota: Una situazione di questo tipo è fondamentalmente differente da quella in cui il sistema si trova in una *sovrapposizione* delle N funzioni d'onda (ossia in un $|\Psi\rangle = \sum_{i=1}^N \alpha_i |\psi_i\rangle$), poiché nel caso di mistura statistica le singole $|\psi_i\rangle$ **non interferiscono** tra loro, cosa che invece succede nel caso di una sovrapposizione.

Come possiamo descrivere efficacemente una mistura statistica, in modo da calcolarne valor medi ed evoluzione temporale?

Partiamo dalla definizione di **osservabile** \hat{A} in termini di proiettori \hat{P}_i , e cerchiamo di riadattarla alla nuova situazione:

$$\hat{A} = \sum_i a_i \hat{P}_i \quad \hat{P}_i = |a_i\rangle \langle a_i| \quad (3.1)$$

Se S fosse nello stato $|\psi_k\rangle$ potremmo usare direttamente la formula per il valor medio di \hat{A} , pari alla combinazione lineare dei valor medi $q_k(i)$ dei singoli proiettori:

$$\langle \hat{A} \rangle_{\psi_k} = \sum_i a_i \langle \hat{P}_i \rangle_{\psi_k} = \sum_i a_i q_k(i); \quad q_k(i) = \langle \psi_k | \hat{P}_i | \psi_k \rangle \quad (3.2)$$

Non sapendo però in quale $|\psi_k\rangle$ si trovi il sistema, dovremo considerare tutte le N possibilità $\{|\psi_k\rangle\}$, pesandole con le rispettive probabilità di occorrenza p_k . La media pesata dei valori di aspettazione dell' i -esimo proiettore \hat{P}_i nell'*ensemble* dei possibili stati $\{|\psi_k\rangle\}$ viene denotata con $\tilde{P}(i)$.

$$\tilde{P}(i) = \sum_{k=1}^N p_k q_k(i) = \sum_{k=1}^N p_k \langle \psi_k | \hat{P}_i | \psi_k \rangle$$

Sommando allora su tutti i proiettori giungiamo alla formula per il valor medio di \hat{A} della mistura statistica $\{(|\psi_k\rangle, p_k)\}$:

$$\langle \hat{A} \rangle = \sum_i a_i \tilde{P}(i) = \sum_{k=1}^N p_k \sum_i a_i \langle \psi_k | \hat{P}_i | \psi_k \rangle = \sum_{k=1}^N p_k \langle \psi_k | \hat{A} | \psi_k \rangle = \text{Tr}(\rho \hat{A}) \quad (3.3)$$

Dove ρ è detta **matrice di densità** ed è definita da:

Matrice densità

$$\rho \equiv \sum_{k=1}^N p_k |\psi_k\rangle \langle \psi_k| \quad (3.4)$$

Ricordiamo che l'operazione di **traccia** consiste nella somma sugli elementi sulla diagonale della rappresentazione matriciale di ρA in una qualsiasi base ON $\{|\chi_j\rangle\}$ di \mathcal{H} (il risultato è indipendente dalla scelta della base):

$$\text{Tr}(B) = \sum_{j=1}^{\dim \mathcal{H}} \langle \chi_j | B | \chi_j \rangle$$

Verifichiamo la formula (3.3). Nel caso $N = 0$ avremo una sola possibilità per lo stato $|\psi_0\rangle$, da cui $\rho = p_0 |\psi_0\rangle \langle \psi_0|$ (dove $p_0 = 1$). Considerando allora una base con $|\chi_1\rangle = |\psi_0\rangle$ si trova:

$$\text{Tr}(\rho A) = \text{Tr}(p_0 |\psi_0\rangle \langle \psi_0| \hat{A}) = \sum_{i=1}^{\dim \mathcal{H}} p_0 \langle \chi_i | \psi_0 \rangle \langle \psi_0 | \hat{A} | \chi_i \rangle = p_0 \langle \psi_0 | \hat{A} | \psi_0 \rangle$$

E il risultato è uguale per ogni altra base ON per le proprietà della traccia. Considerando al posto di $|\psi_0\rangle$ un $|\psi_k\rangle$ generico vale quindi:

$$\text{Tr}(p_k |\psi_k\rangle \langle \psi_k| \hat{A}) = p_k \langle \psi_k | \hat{A} | \psi_k \rangle$$

Perciò possiamo sommare su k per riottenere la (3.3):

$$\sum_{k=0}^N p_k \langle \psi_k | \hat{A} | \psi_k \rangle = \sum_{k=0}^N \text{Tr}(p_k |\psi_k\rangle \langle \psi_k| \hat{A}) \stackrel{(a)}{=} \text{Tr} \left(\sum_{k=0}^N p_k |\psi_k\rangle \langle \psi_k| \hat{A} \right) = \text{Tr}(\rho \hat{A})$$

dove in (a) abbiamo infine utilizzato la **linearità** della traccia.

3.1.2 Evoluzione temporale

Le singole $|\psi_k\rangle$ evolvono in maniera unitaria:

$$|\psi_k(t)\rangle = U(t - t_0) |\psi_k(t_0)\rangle \leftrightarrow \langle\psi_k(t)| = \langle\psi_k(t_0)| U^\dagger(t - t_0)$$

Sostituendo nella definizione (3.4) otteniamo perciò l'**evoluzione** della matrice densità:

$$\rho(t) = \sum_{k=1}^N p_k |\psi_k(t)\rangle \langle\psi_k(t)| = \sum_{k=1}^N p_k [U |\psi_k(t_0)\rangle] [\langle\psi_k(t_0)| U^\dagger] = U(t - t_0) \rho_0 U^\dagger(t - t_0)$$

Per gli stati puri $|\psi\rangle$ (che siano nel dominio dell'Hamiltoniana \hat{H} del sistema) vale l'equazione di Schrödinger dipendente dal tempo:

$$i\hbar \frac{\partial \psi(t)}{\partial t} = \hat{H} \psi(t) \quad (3.5)$$

In notazione braket:

$$i\hbar \frac{d}{dt} |\psi\rangle = H |\psi\rangle \leftrightarrow -i\hbar \frac{d}{dt} \langle\psi| = \langle\psi| H \quad (3.6)$$

Un'equazione analoga vale per le misture statistiche (dette anche **stati misti**), e prende il nome di equazione di **Liouville-von Neumann**. La si ricava calcolando la derivata temporale della matrice densità:

*Equazione di
Liouville-von
Neumann*

$$\begin{aligned} \dot{\rho}(t) &= \frac{d}{dt} \sum_{k=1}^N p_k |\psi_k\rangle \langle\psi_k| = \sum_{k=1}^N p_k \left[\left(\frac{d}{dt} |\psi_k\rangle \right) \langle\psi_k| + |\psi_k\rangle \left(\frac{d}{dt} \langle\psi_k| \right) \right] = \\ &\stackrel{(3.6)}{=} \sum_{k=1}^N p_k \left(\frac{H |\psi_k\rangle}{i\hbar} \langle\psi_k| - |\psi_k\rangle \frac{\langle\psi_k| H}{i\hbar} \right) = \\ &= \frac{1}{i\hbar} (H\rho - \rho H) = \frac{1}{i\hbar} [H, \rho] \end{aligned}$$

Riarrangiando giungiamo a:

$$i\hbar \frac{d}{dt} \rho(t) = [H, \rho(t)]$$

Nel dettaglio, questa equazione è valida nella **visuale di Schrödinger**, dove lo stato (rappresentato dalla matrice densità ρ) evolve nel tempo, mentre gli operatori che agiscono su di esso no.

La natura matriciale di $\rho(t)$, tuttavia, mostra una forte analogia con l'equazione di evoluzione temporale degli operatori $A^{(H)}(t)$ nella **visuale di Heisenberg** (dove, invece, ρ è indipendente dal tempo):

$$i\hbar \frac{dA^{(H)}}{dt} = -[H, A^{(H)}(t)]$$

a meno di una differenza di segno, data dalla diversa visuale.

3.1.3 Proprietà della matrice densità

La matrice densità ρ gode delle seguenti proprietà:

1. ρ è **hermitiana**.

Hermitiana

Dim. Una matrice è hermitiana se vale la seguente relazione tra i suoi elementi:

$$\rho_{ij} = (\rho_{ji})^*$$

Verifichiamolo direttamente. Sia $\{|i\rangle\}$ una base ON. Scriviamo le $|\psi_k\rangle$ in questa base:

$$|\psi_k\rangle = \sum_{i=1}^{\dim \mathcal{H}} c_i^k |i\rangle \quad c_i^k = \langle i | \psi_k \rangle \in \mathbb{C} \quad (3.7)$$

Gli elementi di matrice di ρ sono quindi dati da:

$$\begin{aligned} \rho_{ij} &= \langle i | \rho | j \rangle = \sum_{k=1}^N p_k \langle i | \psi_k \rangle \langle \psi_k | j \rangle \stackrel{(a)}{=} \sum_{k=1}^N p_k \sum_{l,m}^{\dim \mathcal{H}} c_l^k (c_m^k)^* \underbrace{\langle i | l \rangle}_{\delta_{il}} \underbrace{\langle m | j \rangle}_{\delta_{mj}} = \\ &= \sum_{k=1}^N p_k c_i^k (c_j^k)^* \end{aligned} \quad (3.8)$$

dove in (a) abbiamo usato l'espansione delle $|\psi_k\rangle$ nella base (3.7). In altre parole, gli elementi di matrice ρ_{ij} sono prodotti dei coefficienti i e j della rappresentazione di $|\psi_k\rangle$ nella base $|i\rangle$ in cui ρ è scritta, mediati sull'intero ensemble delle $\{|\psi_k\rangle\}$ possibili, con i pesi dati dalle loro probabilità p_k .

Invertendo i e j e coniugando otteniamo lo stesso risultato:

$$\rho_{ji}^* = \sum_{k=1}^N p_k (c_j^k)^* c_i^k = \rho_{ij}$$

Ne deriva allora che ρ è hermitiana.

2. $\text{Tr } \rho = 1$.

Traccia unitaria

Dim. Basta fare il conto diretto usando l'espressione (3.8) per gli elementi di matrice:

$$\text{Tr } \rho = \sum_{i=1}^{\dim \mathcal{H}} \rho_{ii} = \sum_{k=1}^N p_k \left(\sum_{i=1}^{\dim \mathcal{H}} |c_i^k|^2 \right) = \sum_{k=1}^N p_k \sum_{i=1}^{\dim \mathcal{H}} \langle \psi_k | i \rangle \langle i | \psi_k \rangle \stackrel{(a)}{=} \sum_{k=1}^N p_k \stackrel{(b)}{=} 1$$

dove in (a) si è usata la completezza di Dirac, e la normalizzazione $\langle \psi_k | \psi_k \rangle = 1$. In (b) usiamo poi la *convessità* delle p_k , che quindi “esauriscono tutte le possibilità”, dato che sappiamo con certezza il sistema si trova in uno stato in $\{|\psi_k\rangle\}$ (per come abbiamo definito la mistura statistica).

3. ρ è un operatore **non-negativo**, cioè $\forall |\psi\rangle$ vale $\langle \psi | \rho | \psi \rangle \geq 0$.

Non negativa

Dim. Per calcolo diretto con una generica $|\varphi\rangle$, usando di nuovo la (3.8):

$$\langle \varphi | \rho | \varphi \rangle = \langle \varphi | \left(\sum_{k=1}^N p_k |\psi_k\rangle \langle \psi_k| \right) | \varphi \rangle = \sum_{k=1}^N p_k |\langle \varphi | \psi_k \rangle|^2 \geq 0$$

3.1.4 Stati puri e misti

La matrice di densità ρ può essere usata sia per rappresentare **stati puri** che **misure statistiche (stati misti)**.

Riepilogando, uno stato **puro** è dato da ρ prodotta da un unico termine:

$$\rho_p = |\psi\rangle \langle\psi|$$

Mentre per uno stato **misto** si ha:

$$\rho_m = \sum_{k=1}^N p_k |\psi_k\rangle \langle\psi_k|$$

Si possono differenziare i due casi notando che il primo è un **proiettore**, e quindi $\rho_p^2 = \rho_p$. Da ciò segue che:

Differenza tra stati misti e puri

$$\text{Tr } \rho_m^2 < 1 \quad \text{Tr } \rho_p^2 = 1$$

Dim. Eseguiamo il calcolo diretto, usando la base $\{|j\rangle\}$ in cui ρ è diagonale (ossia scrivendo ρ nella sua rappresentazione spettrale):

$$\rho = \sum_{j=1}^N \lambda_j |j\rangle \langle j| \Rightarrow \rho^2 = \sum_{j,k=1}^N \lambda_j \lambda_k |k\rangle \underbrace{\langle k | j \rangle}_{\delta_{kj}} \langle j| = \sum_{j=1}^N \lambda_j^2 |j\rangle \langle j|$$

Dalle proprietà di ρ deriva che:

$$\begin{aligned} \text{Tr } \rho = 1 &\Rightarrow \sum_{j=1}^N \lambda_j = 1 \\ \rho \geq 0 &\Rightarrow \lambda_j \geq 0 \end{aligned}$$

e perciò $0 \leq \lambda_j \leq 1$. Vale allora la disuguaglianza:

$$\text{Tr } \rho^2 = \sum_{j=1}^N \lambda_j^2 \leq \sum_{j=1}^N \lambda_j = 1$$

E si ha l'uguaglianza solo nel caso specifico in cui $\lambda_{j^*} = 1$ per un certo $0 \leq j^* \leq N$, e tutti gli altri $\lambda_{j \neq j^*} = 0$, che corrisponde a quello in cui ρ rappresenta uno **stato puro**. In tutti gli altri casi (**stati misti**), $\text{Tr}(\rho^2)$ sarà quindi *strettamente* minore di 1.

Per questò la grandezza (misurabile sperimentalmente, dato che ρ è hermitiana, e quindi è un'osservabile) **Tr**(ρ^2) è detta **purezza** dello stato ρ : è massima (pari a 1) per stati puri, e $0 \leq \text{Tr}(\rho^2) < 1$ per stati misti.

Purezza di uno stato ρ

Nota: l'evoluzione **unitaria** **conserva** la purezza. Segue che uno stato puro può evolvere unitariamente esclusivamente a stati puri, e allo stesso modo uno stato misto resterà sempre misto (almeno finché il sistema rimane isolato).

Nella base $\{|j\rangle\}$ che abbiamo appena considerato, la matrice ρ è diagonale, ma generalmente potremmo scriverla in una qualsiasi altra base ON.

Distinguiamo perciò tra gli elementi $\lambda_1 \dots \lambda_N$ sulla diagonale di ρ , detti **termini di popolazione**, e quelli fuori dalla diagonale, detti **termini di coerenza**. Come vedremo nelle prossime sezioni, questi ultimi sono indice delle *correlazioni* tra i singoli stati (o meglio, di superposizioni quantistiche che possono portare a correlazioni).

Termini di popolazione/coerenza

Nota: la “misurabilità” delle correlazioni dipende quindi dalla base utilizzata. Se la base computazionale è quella che diagonalizza ρ , avremo solo termini di popolazione, e nessuna correlazione:

$$\rho = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_N \end{pmatrix}$$

Tuttavia, come già visto negli esempi di algoritmi, le correlazioni sono molto importanti per la computazione quantistica.

3.1.5 Il sistema da 1 qubit

Consideriamo un generico qubit nello **stato puro** $|\psi\rangle$, che, nella base computazionale $\{|0\rangle, |1\rangle\}$, è dato da:

$$|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i\varphi} \end{pmatrix}$$

La matrice densità ad esso associata è data da:

$$\rho(\theta, \varphi) = |\psi\rangle \langle \psi| \underset{(a)}{=} \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i\varphi} \end{pmatrix} \begin{pmatrix} \cos \frac{\theta}{2} & \sin \frac{\theta}{2} e^{-i\varphi} \end{pmatrix} = \begin{pmatrix} \cos^2 \frac{\theta}{2} & \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{-i\varphi} \\ \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{i\varphi} & \sin^2 \frac{\theta}{2} \end{pmatrix} \quad (3.9)$$

Nota: il prodotto tra un ket e un bra in (a) si realizza in notazione matriciale come il prodotto tra un vettore colonna (il ket) e un vettore riga con componenti coniugate (il bra). Questo si ha poiché bra e ket sono collegati da una coniugazione, che corrisponde alla *trasposta coniugata* nel linguaggio delle matrici.

Le coordinate (θ, φ) sono le coordinate polari nella sfera di Bloch, e individuano un vettore (**unitario**) di coordinate cartesiane:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix} \quad x^2 + y^2 + z^2 = 1 \quad (3.10)$$

Possiamo allora riscrivere la (3.9) in questa forma, usando formule di bisezione e duplicazione per \sin/\cos :

$$\begin{aligned}\rho &= \begin{pmatrix} \cos^2 \frac{\theta}{2} & \frac{1}{2} \sin \theta e^{-i\varphi} \\ \frac{1}{2} \sin \theta e^{i\varphi} & \sin^2 \frac{\theta}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + \cos \theta & \sin \theta (\cos \varphi - i \sin \varphi) \\ \sin \theta (\cos \varphi + i \sin \varphi) & 1 - \cos \theta \end{pmatrix} = \\ &= \frac{1}{2} \begin{pmatrix} 1 + z & x - iy \\ x + iy & 1 - z \end{pmatrix} = \frac{1}{2} (\mathbb{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z) \end{aligned} \quad (3.11)$$

Nell'ultimo passaggio si è ottenuta un'espansione nella base delle matrici hermitiane 2×2 , data da identità e matrici di Pauli $\{\mathbb{I}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$, che è molto comoda per i conti (dato che esiste una formula per l'esponenziale di un vettore di Pauli).

Ponendo $\vec{r} = (x, y, z)$, il determinante di ρ è dato da:

$$\det \rho = \frac{1}{4} [(1+z)(1-z) - (x+iy)(x-iy)] = \frac{1}{4} (1 - \|\vec{r}\|^2) \quad (3.12)$$

Abbiamo visto che una matrice densità ρ che descrive uno stato puro può essere rappresentata come un vettore \vec{r} che “punta” sulla superficie della sfera di Bloch.

Estendiamo ora tale notazione, partendo da un vettore \vec{r} generico.

Poiché ρ è hermitiana e non-negativa, ha autovalori reali positivi (al più nulli) e quindi il suo determinante (prodotto degli autovalori) è non negativo:

$$\det \rho \geq 0$$

Tale condizione, unita alla (3.12), porta a $\|\vec{r}\| \leq 1$.

Mostriamo ora che l'uguaglianza $\|\vec{r}\| = 1$ avviene solo quando \vec{r} rappresenta uno **stato puro**. Sappiamo già che le ρ di stati puri generano vettori \vec{r} unitari (abbiamo infatti mostrato che ogni $\rho = |\psi\rangle\langle\psi|$ individua (3.11) un vettore unitario (3.10)), e manca solo da mostrare il viceversa. Se $\|\vec{r}\| = 1$, si ha per la (3.12) che $\det \rho = 0$, ma vale sempre $\text{Tr}(\rho) = 1$. Detti λ_1 e λ_2 gli autovalori di ρ , si ha quindi:

$$\begin{cases} \det \rho = \lambda_1 \lambda_2 = 0 \\ \text{Tr} \rho = \lambda_1 + \lambda_2 = 1 \end{cases}$$

Perciò uno dei due (es. λ_1) è identicamente 1, e l'altro è nullo. Ma allora, se $|u\rangle$ è l'autovettore di autovalore 1, si ha $\rho = |u\rangle\langle u|$, che rappresenta uno stato puro.

Perciò, si ha che tutte le scelte di $\|\vec{r}\| < 1$, che corrispondono ad altrettante possibilità per ρ , codificano **stati misti**.

In definitiva, tra tutte le ρ possibili per un sistema a **1 qubit** (date da generici vettori nella sfera di Bloch) la maggior parte è prodotta da stati misti (l'interno). Considerare misture statistiche (nella notazione della matrici di densità) permette di sfruttare tale numero molto più ampio di possibilità.

*Gli stati puri \leftrightarrow
superficie della
sfera di Bloch
($\|\vec{r}\| = 1$)*

*Stati misti \leftrightarrow
 $\|\vec{r}\| < 1$*

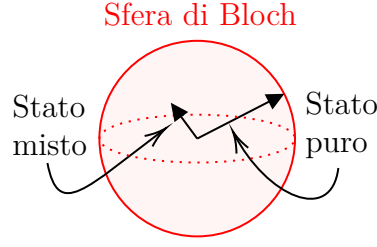


Figura (3.1) – La sfera di Bloch comprende tutti gli stati possibili per un sistema a 1 qubit. Gli stati *interni* sono **stati misti**, mentre quelli *sulla superficie* sono **stati puri**

In questa notazione vettoriale possiamo visualizzare la genesi di uno stato misto come *somma* di vettori di stati puri opportunamente pesati da probabilità. Il caso più semplice è quello in cui \vec{r} è distribuito uniformemente su tutta la superficie della sfera unitaria, e corrisponde alla **mistura** di tutti gli stati puri possibili, ossia allo **stato massimamente misto**. Possiamo calcolare analiticamente la ρ risultante integrando sulla superficie della sfera di Bloch (dove l'integrazione si intende su ciascun elemento della matrice (3.9)):

$$\rho = \frac{1}{4\pi} \int_0^{2\pi} d\varphi \int_0^\pi (d\theta \sin \theta) \rho(\theta, \varphi) = \frac{1}{2} \mathbb{I} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = \frac{1}{2} |0\rangle \langle 0| + \frac{1}{2} |1\rangle \langle 1|$$

3.1.6 Sistemi composti

Nell'ultima sezione abbiamo esaminato un sistema formato da un singolo qubit, notando l'utilità della nuova notazione introdotta.

Il principale utilizzo delle matrici densità, tuttavia, si realizza nel considerare *sistemi composti*, tali da codificare 2 o più qubit.

Procedendo con ordine, partiamo dal caso più semplice di **2 qubit**.

Sappiamo che se gli stati del sistema 1 sono elementi di uno spazio di Hilbert \mathcal{H}_1 e gli stati del sistema 2 in \mathcal{H}_2 , gli stati del sistema composto si trovano nel prodotto tensore dei due spazi, cioè in:

$$\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2$$

Una generica funzione d'onda in \mathcal{H}_{12} si scrive come:

$$|\psi\rangle = \sum_{i,\alpha}^{\dim \mathcal{H}_{1,2}} c_{i\alpha} |i\rangle_1 \otimes |\alpha\rangle_2 \quad \sum_{i,\alpha}^{\dim \mathcal{H}_{1,2}} |c_{i,\alpha}|^2 = 1$$

dove $\{|i\rangle_1\}$ e $\{|\alpha\rangle_2\}$ sono basi ON rispettivamente di \mathcal{H}_1 e \mathcal{H}_2 .

Convenzioni di notazione. Spesso useremo notazioni abbreviate per il prodotto tensore, come:

$$|a, b\rangle \equiv |a\rangle |b\rangle \equiv |a\rangle \otimes |b\rangle$$

Analogamente, per sommatorie su più indici:

$$\sum_{ij}^{\dim \mathcal{H}_{1,2}} \equiv \sum_i^{\dim \mathcal{H}_1} \sum_j^{\dim \mathcal{H}_2}$$

Quando i limiti delle sommatorie sono chiari dal contesto, tenderemo ad ometterli. Infine, impropriamente, useremo direttamente gli indici come ket, ossia $|\alpha_i\rangle \equiv |i\rangle$.

Riadattando allora la definizione di matrice di densità giungiamo a scrivere¹:

Matrice densità di sistemi composti

$$\rho_{12} = |\psi\rangle \langle\psi| = \sum_{i,\alpha} \sum_{j,\beta} c_{i\alpha} c_{j\beta}^* |i\rangle \langle\alpha| \langle j| \langle\beta| = \sum_{i,\alpha} \sum_{j,\beta} \rho_{i\alpha}^{j\beta} |i, \alpha\rangle \langle j, \beta| \quad (3.13)$$

Consideriamo allora un'osservabile \hat{A} del sistema 1:

$$\hat{A} = \hat{A}_1 \otimes \mathbb{I}_2$$

Il suo valore atteso nello stato ρ si ottiene dalla formula (3.3), con l'unica differenza di usare per la traccia una base ON $\{|\kappa\gamma\rangle\}$ del sistema composto. Esprimendo anche ρ in questa base, possiamo sfruttare l'ortonormalità e calcolare:

$$\begin{aligned} \langle \hat{A} \rangle_\rho &= \text{Tr}(\rho \hat{A}) = \sum_{\kappa\gamma} \langle \kappa\gamma | \rho \hat{A} | \kappa\gamma \rangle \stackrel{(3.13)}{=} \sum_{\kappa\gamma} \langle \kappa\gamma | \left(\sum_{i,\alpha} \sum_{j,\beta} \rho_{i\alpha}^{j\beta} |i\rangle \langle\alpha| \langle j| \langle\beta| \right) (\hat{A}_1 \otimes \mathbb{I}_2) | \kappa\gamma \rangle = \\ &= \sum_{\kappa\gamma} \sum_{i\alpha} \sum_{j\beta} \rho_{i\alpha}^{j\beta} \underbrace{\langle \kappa | i \rangle}_{\delta_{\kappa i}} \underbrace{\langle \gamma | \alpha \rangle}_{\delta_{\gamma \alpha}} \langle j | \hat{A}_1 | \kappa \rangle \underbrace{\langle \beta | \mathbb{I}_2 | \gamma \rangle}_{\delta_{\beta \gamma}} \stackrel{(a)}{=} \sum_{ij\alpha} \rho_{i\alpha}^{j\alpha} \langle j | \hat{A}_1 | i \rangle \end{aligned}$$

dove in (a) usiamo le δ di Kronecker per “collassare” le sommatorie.

Se nel calcolare la traccia usassimo la base ON di uno solo dei due sistemi, otterremo una **matrice densità ridotta**. Per esempio, se $\{|\alpha\rangle\}$ è base ON di \mathcal{H}_2 , allora la traccia parziale risulta nella matrice ridotta del primo sistema, i cui elementi sono dati da:

Matrice densità ridotta

$$(\rho_1)_{ij} = \left(\text{Tr}_2 \rho_{12} \right)_{ij} = \sum_{\alpha}^{\dim \mathcal{H}_2} \langle \alpha |_2 \rho_{12} | \alpha \rangle_2 = \sum_{\alpha}^{\dim \mathcal{H}_2} \rho_{i\alpha}^{j\alpha} |i\rangle \langle j|$$

Poiché non stiamo usando una base del sistema composto, l'operazione di traccia si dice **traccia parziale**, ed è analoga al processo di **marginalizzazione** di distribuzioni statistiche, per cui si integra (o nel nostro caso, si somma) sulle variabili che non interessano (nell'esempio quelle del sistema 2).

La matrice ridotta appena calcolata è comoda per calcolare il valor medio di un'osservabile \hat{A} che agisce solo sul primo sistema, poiché si ha:

$$\langle \hat{A} \rangle_1 = \text{Tr}(\rho_1 \hat{A}_1)$$

In maniera simmetrica, si può *ridurre* la matrice di densità composta ρ_{12} sul sistema 2:

$$(\rho_2)_{\alpha\beta} = \left(\text{Tr}_1 \rho_{12} \right)_{\alpha\beta} = \sum_i \langle i |_1 \rho_{12} | i \rangle_1 = \sum_i \rho_{i\alpha}^{i\beta} |\alpha\rangle \langle\beta| \Rightarrow \langle \mathbb{I}_1 \otimes B_2 \rangle_2 = \text{Tr}(\rho_2 B_2)$$

Tuttavia, ρ_1 e ρ_2 **non** conservano la **purezza** dello stato. Cioè, se ρ_{12} descrive uno stato puro, ossia con $\rho_{12}^2 = \rho_{12}$, non è detto che la stessa condizione valga per le singole ρ_1 e ρ_2 ottenute da essa per riduzione.

Matrici ridotte non conservano la purezza

Esempio

Proviamo a calcolare le matrici ridotte per un sistema di 2 qubit che si trova nello **stato di Bell** dato da:

Calcolo di matrici ridotte

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$

La matrice densità del sistema composto si ottiene come:

$$\begin{aligned}\rho_{12} &= |\psi^-\rangle \langle\psi^-| = \frac{1}{2}(|01\rangle - |10\rangle)(\langle 01| - \langle 10|) = \\ &= \frac{1}{2}(|01\rangle \langle 01| - |01\rangle \langle 10| - |10\rangle \langle 01| + |10\rangle \langle 10|) = \\ &= \frac{1}{2} \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ \hline 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)\end{aligned}\quad (3.14)$$

dove la matrice è espressa nella base computazionale $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$.

Calcoliamo le matrici ridotte:

$$\rho_1 = \text{Tr}_2 \rho_{12} = \langle 0|_2 \rho_{12} |0\rangle_2 + \langle 1|_2 \rho_{12} |1\rangle_2$$

Il termine in giallo *seleziona* tutti i termini di (3.14) che hanno il secondo qubit a 0 sia nel bra che nel ket, ossia solo il termine $|10\rangle \langle 10|$.

Si ottiene allora:

$$2 \langle 0|_2 \rho_{12} |0\rangle_2 = \langle 0|_2 (|1\rangle_1 \otimes |0\rangle_2 \langle 1|_1 \otimes \langle 0|_2) |0\rangle_2 = |1\rangle_1 \langle 1|_1$$

D'altro canto, il termine in azzurro verifica la stessa condizione, ma con 1 al posto di 0, cosa che si ottiene solo per $|01\rangle \langle 01|$, producendo allora:

$$\langle 1|_2 \rho_{12} |1\rangle_2 = \frac{1}{2} |0\rangle_1 \langle 0|_1$$

E quindi otteniamo:

$$\rho_1 = \frac{1}{2}(|0\rangle \langle 0| + |1\rangle \langle 1|) = \frac{1}{2}\mathbb{I}$$

E si ha che $\text{Tr}(\rho_1^2) = 1/2 < 1$, per cui ρ_1 rappresenta uno stato misto - ma siamo partiti da ρ_{12} che rappresenta uno stato puro. Perciò abbiamo verificato che le

¹La posizione degli indici non indica nessun tipo di covarianza/controvarianza, è solo un fatto di comodità stilistica.

matrici ridotte **non** conservano la purezza dello stato.

Analogamente, possiamo calcolare ρ_2 :

$$\rho_2 = \text{Tr}_1 \rho_{12} = \frac{1}{2} \mathbb{I}$$

Tracce parziali in notazione matriciale. Il calcolo di ρ_1 e ρ_2 può essere effettuato direttamente in notazione matriciale, esaminando nel dettaglio l'operazione di traccia parziale.

Per ρ_1 la traccia seleziona gli elementi di matrice in cui il secondo qubit è **0** o **1**, ossia quelli qui evidenziati:

$$\rho_{12} = \begin{array}{c} |00\rangle \\ |01\rangle \\ |10\rangle \\ |11\rangle \end{array} \left(\begin{array}{cc|cc} |00\rangle & |01\rangle & |10\rangle & |11\rangle \\ \hline \text{a} & b & e & f \\ c & \text{d} & g & h \\ \hline i & j & m & n \\ k & l & o & p \end{array} \right)$$

Il conto porta a:

$$\begin{aligned} \rho_1 &= \text{Tr}_2 \rho_{12} = (a + d) |0\rangle \langle 0| + (e + h) |1\rangle \langle 0| + (i + l) |0\rangle \langle 1| + (m + p) |1\rangle \langle 1| = \\ &= \begin{pmatrix} a + d & e + h \\ i + l & m + p \end{pmatrix} \end{aligned}$$

Basta allora considerare la matrice i cui termini sono le somme sulle diagonali dei singoli blocchi 2×2 della matrice ρ_{12} originaria. Ciò non sorprende, perché ρ_1 racchiude l'informazione necessaria a calcolare valor medi di osservabili del solo primo sistema, che hanno la forma:

$$\hat{A} = \hat{A}_1 \otimes \mathbb{I}_2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \mathbb{I}_2 = \left(\begin{array}{cc|cc} a & 0 & b & 0 \\ 0 & a & 0 & b \\ \hline c & 0 & d & 0 \\ 0 & c & 0 & d \end{array} \right)$$

D'altro canto, per ρ_2 i termini selezionati riguardano il primo qubit, e quindi avremo:

$$\rho_{12} = \begin{array}{c} |00\rangle \\ |01\rangle \\ |10\rangle \\ |11\rangle \end{array} \left(\begin{array}{cc|cc} |00\rangle & |01\rangle & |10\rangle & |11\rangle \\ \hline \text{a} & b & e & f \\ c & d & g & h \\ \hline i & j & m & n \\ k & l & o & p \end{array} \right)$$

E si giunge a:

$$\rho_2 = \text{Tr}_1 \rho_{12} = \begin{pmatrix} a+m & b+n \\ c+o & d+p \end{pmatrix}$$

Gli elementi di ρ_2 sono quindi somme degli elementi corrispondenti dei blocchi 2×2 sulla diagonale di ρ_{12} .

Di nuovo, tutto ciò è compatibile con il fatto che ρ_2 codifica l'informazione per calcolare il valor medio di un'osservabile del sistema 2, data da:

$$\hat{B} = \mathbb{I}_2 \otimes \hat{B}_2 = \mathbb{I}_2 \otimes \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \left(\begin{array}{cc|cc} \alpha & \beta & 0 & 0 \\ \gamma & \delta & 0 & 0 \\ \hline 0 & 0 & \alpha & \beta \\ 0 & 0 & \gamma & \delta \end{array} \right)$$

(E infatti i termini sommati sono quelli “pesati allo stesso modo” dall'azione di \hat{B}).

Nota: il prodotto tensore delle matrici ρ_1 e ρ_2 non riproduce, in generale, la matrice densità ρ_{12} del sistema composto:

$$\rho_{12} \neq \rho_1 \otimes \rho_2 \quad (3.15)$$

(Non)
Fattorizzabilità
delle matrici
densità

Del resto, se valesse l'uguaglianza in generale non avremmo il problema della perdita di informazione sulla *purezza* dello stato originario.

Nell'esempio appena visto possiamo verificare la (3.15) per calcolo diretto:

$$\rho_1 \otimes \rho_2 = \frac{1}{4} \mathbb{I}_1 \otimes \mathbb{I}_2 = \frac{1}{4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \neq \rho_{12}$$

La scomposizione $\rho_{12} = \rho_1 \otimes \rho_2$ è analoga alla **fattorizzabilità** degli stati puri in singoli prodotti tensori, e quindi vale solo negli specifici casi in cui ρ_1 e ρ_2 codificano sistemi “separati”. Interpretare tale condizione dal punto di vista matriciale non è tuttavia per nulla banale, e per questo non ce ne occuperemo.

3.2 Matrici densità - parte 2

3.2.1 Correlazioni tra stati

(Lezione 6 • del
14/3/2019)

Nella sezione precedente abbiamo notato che $\rho_{12} = \rho_1 \otimes \rho_2$ vale solo se ρ_1 e ρ_2 sono **indipendenti**, cioè **scorrelati**. Se così non è, come nella maggior parte dei casi, ρ_1 e ρ_2 codificano solo una parte delle informazioni del sistema composto ρ_{12} , dato che mancano i termini di *interazione* tra i due sottosistemi, che effettivamente in

tal caso non hanno una individualità ben definita.

Esaminiamo allora, nello specifico, cosa significhi **correlazione**.

Partiamo dal caso più semplice di due variabili casuali classiche x_i e y_i con opportune distribuzioni di probabilità. Supponiamo che x_i e y_i assumano valori binari $\{1, -1\}$, corrispondenti per esempio all'esito di una misura di una certa osservabile sui due sistemi. Possiamo campionare N volte le due variabili costruendo due vettori \vec{x} e \vec{y} :

$$\begin{aligned}\vec{x} &= \{1, -1, -1, -1, \dots, -1\} \\ \vec{y} &= \{-1, 1, +1, +1, \dots, +1\}\end{aligned}$$

I vettori \vec{x} e \vec{y} corrispondono, fisicamente, a misure ripetute di osservabili con risultati binari (sì/no) eseguite sui due sistemi.

Supponiamo che le x_i e y_i si distribuiscano simmetricamente attorno a 0, e cioè abbiano media nulla:

$$\bar{x} = 0; \quad \bar{y} = 0$$

Un modo per quantificare la **correlazione** tra le due è dato dal calcolare la media del prodotto:

$$\overline{xy} = \frac{1}{N} \sum_{i=1}^N x_i y_i$$

Ci si potrebbe aspettare che, poiché $\bar{x} = 0$ e $\bar{y} = 0$, anche $\overline{xy} = 0$. In realtà ciò succede solo se le due variabili sono *indipendenti*. Se esiste una qualche relazione tra le due, per cui y tende ad assumere valori che dipendono da quelli di x (e viceversa), non è detto che i singoli prodotti $x_i y_i$ si distribuiscano ancora in modo simmetrico.

Per esempio:

- Se $x_i = y_i$, avremo $x_i y_i = 1 \ \forall i$, e quindi $\overline{xy} = 1$. x e y si dicono allora **completamente correlate**
- Se $x_i = -y_i$ si ha $x_i y_i = -1 \ \forall i$, da cui $\overline{xy} = -1$, e x e y sono **completamente anticorrelate**
- Se $x_i y_i$ si distribuiscono simmetricamente attorno a 0, $\overline{xy} = 0$ e le due variabili si dicono **scorrelate**.

Nel caso generale avremo una correlazione *parziale*, che si valuta tramite il **coefficiente di Pearson**:

Coefficiente di Pearson

$$\sigma_{xy} = \frac{E[(x - \mu_x)(y - \mu_y)]}{\sigma_x \sigma_y} \quad \begin{cases} \mu_x = \bar{x} \\ \mu_y = \bar{y} \end{cases} \quad \begin{cases} \sigma_x^2 = \overline{(x - \mu_x)^2} \\ \sigma_y^2 = \overline{(y - \mu_y)^2} \end{cases}$$

dove $E(x)$ indica il valore di aspettazione (speranza matematica) di una certa variabile casuale x .

Notiamo che il coefficiente di Pearson corrisponde alla media di *infiniti* prodotti (nel senso di $N \rightarrow \infty$ misure) $(x_i - \mu_x)(y_i - \mu_y)$ (per cui le x_i e y_i possono non essere simmetriche attorno a 0), normalizzata alle loro dispersioni attorno alle loro medie.

Tale formula si adatta immediatamente al caso quantistico, usando operatori al posto di variabili, e applicando la linearità del valore di aspettazione:

$$\begin{aligned} E[(X - \mu_x)(Y - \mu_y)] &= \langle (\hat{X} - \langle \hat{X} \rangle)(\hat{Y} - \langle \hat{Y} \rangle) \rangle = \\ &= \langle \hat{X}\hat{Y} \rangle - \langle \langle \hat{X} \rangle \hat{Y} \rangle - \langle \hat{Y} \langle \hat{X} \rangle \rangle + \langle \langle \hat{X} \rangle \langle \hat{Y} \rangle \rangle = \\ &= \langle \hat{X}\hat{Y} \rangle - \langle \hat{X} \rangle \langle \hat{Y} \rangle - \langle \hat{Y} \rangle \langle \hat{X} \rangle + \langle \hat{X} \rangle \langle \hat{Y} \rangle = \langle \hat{X}\hat{Y} \rangle - \langle \hat{X} \rangle \langle \hat{Y} \rangle \end{aligned}$$

3.2.2 Correlazioni ed entanglement

La presenza di correlazioni in **stati puri** è propria² degli **stati entangled**, ossia stati composti $|\phi\rangle$ non decomponibili in un prodotto tensore:

$$|\phi\rangle \neq |\phi_1\rangle \otimes |\phi_2\rangle$$

Mostriamo infatti che **non** vi sono **correlazioni** tra le componenti di un sistema **separabile**. Consideriamo uno stato $|\psi\rangle$ **non entangled** di un sistema composto S , per cui quindi vale:

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$$

Stati non entangled non presentano correlazioni

Siano \hat{X}_A e \hat{X}_B due generiche osservabili relative, rispettivamente, ai sottosistemi A e B che compongono S . Calcoliamone la correlazione σ_{xy} a partire dalla definizione:

$$\begin{aligned} \sigma_{xy} &= \langle \hat{X}_A \hat{Y}_B \rangle - \langle \hat{X}_A \rangle \langle \hat{Y}_B \rangle = \\ &= \langle \psi_A \psi_B | \hat{X}_A \hat{X}_B | \psi_A \psi_B \rangle - \langle \psi_A \psi_B | \hat{X}_A | \psi_A \psi_B \rangle \langle \psi_A \psi_B | \hat{Y}_B | \psi_A \psi_B \rangle = \\ &= \langle \psi_A | \hat{X}_A | \psi_A \rangle \langle \psi_B | \hat{X}_B | \psi_B \rangle - \langle \psi_A | \hat{X}_A | \psi_A \rangle \langle \psi_B | \hat{Y}_B | \psi_B \rangle = 0 \end{aligned}$$

D'altro canto, per uno stato $|\phi\rangle$ non separabile (entangled), in generale σ_{xy} può assumere valori non nulli: si hanno quindi **correlazioni** tra misure di osservabili eseguite sulle **diverse componenti** del sistema composto.

3.2.3 Schmidt decomposition

Come possiamo riconoscere uno **stato puro entangled**?

Per esempio, dato un sistema S composto di due qubit A e B , con $\mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B$, si nota immediatamente che:

$$|\psi_1\rangle = |0\rangle_A |0\rangle_B$$

²^Ed esclusiva, nel caso specifico dei soli *stati puri* stiamo trattando.

è uno stato non correlato, mentre

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A |1\rangle_B + |1\rangle_A |0\rangle_B) \quad (3.16)$$

è entangled (infatti è uno stato di Bell).

Tuttavia uno stato come:

$$|\psi_3\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A |1\rangle_B + |0\rangle_A |0\rangle_B)$$

non è entangled, dato che lo si può fattorizzare come:

$$|\psi_3\rangle = |0\rangle_A \otimes \frac{1}{\sqrt{2}}(|0\rangle_B + |1\rangle_B) \quad (3.17)$$

In questo caso semplice (due soli fattori) la decomposizione si trova abbastanza in fretta, ma considerando più termini, oppure basi differenti da quella computazionale, il problema diviene sempre più complesso. Per esempio, il seguente stato $|\psi_4\rangle$ è entangled?

$$|\psi_4\rangle = |0\rangle_A |0\rangle_B - |0\rangle_A |1\rangle_B + |0\rangle_A |1\rangle_B - |1\rangle_A |1\rangle_B$$

Per rispondere a domande del genere conviene introdurre un risultato di algebra lineare, che permette di riscrivere un generico stato in una “forma minima”, dove stati fattorizzati sono scritti come un unico prodotto (come in 3.17), mentre gli stati entangled sono dati dalla somma di più prodotti (3.16), per cui la distinzione tra i due casi risulta evidente.

Tale processo prende il nome di **decomposizione di Schmidt**. Consideriamo un’autofunzione $|\psi\rangle$ in uno spazio di Hilbert (separabile) di un sistema composto **bipartito**³:

Decomposizione di Schmidt

$$|\psi\rangle \in \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$$

Allora esistono due **basi** ON $\{|u_i\rangle_A\}_{i=1}^{\dim \mathcal{H}_A}$ e $\{|v_j\rangle_B\}_{j=1}^{\dim \mathcal{H}_B}$ per cui vale la decomposizione:

$$|\psi\rangle = \sum_{i=1}^k \sqrt{p_i} |u_i\rangle_A |v_i\rangle_B \quad p_i \geq 0; \quad \sum_{i=1}^k p_i = 1; \quad k \leq \min\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\} \quad (3.18)$$

Nota: In (3.18) stiamo sommando sugli indici di una **sola** base. La decomposizione di Schmidt è quindi molto più semplice della normale decomposizione di $|\psi\rangle$ su due **generiche** basi $\{|\alpha\rangle_A\}$ e $\{|\beta\rangle_B\}$ di \mathcal{H}_A e \mathcal{H}_B :

$$|\psi\rangle = \sum_{ij}^{\dim \mathcal{H}_{A,B}} \gamma_{ij} |\alpha_i\rangle_A |\beta_j\rangle_B \quad (3.19)$$

³Non è necessario che A e B siano singoli qubit. Per esempio potremmo considerare un sistema di 5 qubit, separato in parte A data dai primi 3 e parte B con i restanti 2, e applicare lo stesso i risultati del teorema.

dove la somma avviene su **entrambi** gli indici i e j . Il risultato di Schmidt dimostra che è possibile *rimuovere* uno dei due indici con una scelta opportuna delle due basi.

Nota 2: In informazione quantistica, si tende a far coincidere le basi $\{|u_i\rangle_A\}$ e $\{|v_i\rangle_B\}$ con opportune **basi computazionali** $\{|i\rangle_A\}, \{|i'\rangle_B\}$, in cui i ket sono denotati “in modo numerico” direttamente dall’indice i . In questa notazione la decomposizione di Schmidt diviene:

$$|\psi\rangle_{AB} = \sum_{i=1}^k \sqrt{p_i} |i\rangle_A |i'\rangle_B$$

Per esempio, per $k = 2$ avremo:

$$|\psi\rangle_{AB} = \sqrt{p_1} |0\rangle_A |0'\rangle_B + \sqrt{p_2} |1\rangle_A |1'\rangle_B$$

dove gli apici mostrano che il qubit del sistema B può in generale utilizzare “livelli diversi” rispetto a quelli del qubit A (ossia in un’*altra base*).

Dimostrazione. Partiamo dalla scrittura (3.19) di $|\psi\rangle$ in due basi ON generiche $\{|\alpha\rangle_A\}$ e $\{|\beta\rangle_B\}$ di \mathcal{H}_A e \mathcal{H}_B : *Dimostrazione*

$$|\psi\rangle = \sum_{ij}^{\dim \mathcal{H}_{A,B}} \gamma_{ij} |\alpha_i\rangle_A |\beta_j\rangle_B = \sum_i^{\dim \mathcal{H}_A} |\alpha_i\rangle_A \underbrace{\left(\sum_j^{\dim \mathcal{H}_B} \gamma_{ij} |\beta_j\rangle_B \right)}_{|v_i\rangle} = \sum_i^{\dim \mathcal{H}_A} |\alpha_i\rangle |\tilde{v}_i\rangle \quad (3.20)$$

Il problema è che $\{|\tilde{v}_i\rangle\}$ **non** è, in generale, una base ON di \mathcal{H}_B . Poiché i $|\tilde{v}_i\rangle$ dipendono dai fattori γ_{ij} , che a loro volta dipendono dalla scelta dei $|\alpha_i\rangle$, possiamo giocare su questi ultimi per avere la proprietà desiderata.

In effetti, si trova che se invece di partire da una base ON $\{|\alpha_i\rangle\}$ generica per \mathcal{H}_A , consideriamo la base $\{|u_i\rangle\}$ che *diagonalizza* la matrice densità ridotta ρ_1 , automaticamente $\{|\tilde{v}_i\rangle\}$ è ortogonale. Vediamo come.

Sia quindi $\{|u_i\rangle\}$ la base di \mathcal{H}_A in cui la matrice ρ_1 è diagonale:

$$\rho_1 = \sum_{i=1}^k p_i |u_i\rangle_A \langle u_i|_A \quad (3.21)$$

dove k è il numero delle popolazioni di ρ_1 .

Calcolando ρ_1 a partire dalla definizione otteniamo:

$$\begin{aligned}
\rho_1 &= \text{Tr}_2 \rho_{12} = \text{Tr}_2 |\psi\rangle \langle \psi| \stackrel{(3.20)}{=} \text{Tr}_2 \left(\sum_{ij}^{\dim \mathcal{H}_{A,B}} |\alpha_i\rangle_A |\tilde{v}_i\rangle_B \langle \alpha_j|_A \langle \tilde{v}_j|_B \right) = \\
&\stackrel{(a)}{=} \sum_k^{\dim \mathcal{H}_B} \langle k|_B \left(\sum_{ij}^{\dim \mathcal{H}_{A,B}} |\alpha_i\rangle_A |\tilde{v}_i\rangle_B \langle \alpha_j|_A \langle \tilde{v}_j|_B \right) |k\rangle_B = \\
&= \sum_{ij}^{\dim \mathcal{H}_{A,B}} |\alpha_i\rangle_A \langle \alpha_j|_A \left(\sum_k^{\dim \mathcal{H}_B} \langle k| \tilde{v}_i\rangle \langle \tilde{v}_j| k\rangle \right) = \sum_{ij}^{\dim \mathcal{H}_{A,B}} |\alpha_i\rangle_A \langle \alpha_j|_A \left(\sum_k^{\dim \mathcal{H}_B} \langle \tilde{v}_j| \underbrace{|k\rangle \langle k|}_{\mathbb{I}} |\tilde{v}_i\rangle \right) = \\
&\stackrel{(b)}{=} \sum_{ij}^{\dim \mathcal{H}_{A,B}} \langle \tilde{v}_j | \tilde{v}_i\rangle |\alpha_i\rangle_A \langle \alpha_j|_A \tag{3.22}
\end{aligned}$$

dove in (a) usiamo una generica base ON $\{|k\rangle_B\}$ di \mathcal{H}_B per calcolare la traccia (che risulta la stessa per *qualsiasi* scelta di base) e in (b) usiamo la completezza di Dirac per la base ON $\{|k\rangle_B\}$.

Usando allora la base $\{|u_i\rangle_A\}$ al posto di $\{|\alpha_i\rangle_A\}$ in (3.22), possiamo uguagliare con l'espressione equivalente in (3.21), e giungere a:

$$\sum_{ij}^{\dim \mathcal{H}_{A,B}} \langle \tilde{v}_j | \tilde{v}_i\rangle |u_i\rangle_A \langle u_i|_A = \sum_{i=1}^k p_i |u_i\rangle_A \langle u_i|_A \Leftrightarrow \langle \tilde{v}_j | \tilde{v}_i\rangle = p_i \delta_{ij}$$

Perciò si ha che, usando per \mathcal{H}_A la base $\{|u_i\rangle\}$ che diagonalizza ρ_1 , la decomposizione (3.20) produce una base $\{|\tilde{v}_i\rangle\}$ che è ortogonale. Perché sia **ortonormale** basta normalizzare:

$$|v_i\rangle = \frac{1}{\sqrt{p_i}} |\tilde{v}_i\rangle \Rightarrow |\tilde{v}_i\rangle = \sqrt{p_i} |v_i\rangle \Rightarrow \langle v_i | v_j\rangle = \delta_{ij}$$

Partendo allora dalla (3.20) e ponendo $|\alpha_i\rangle = |u_i\rangle$, e di conseguenza $|\tilde{v}_i\rangle = \sqrt{p_i} |v_i\rangle$ otteniamo la tesi:

$$|\psi\rangle = \sum_{i=1}^k \sqrt{p_i} |u_i\rangle_A |v_i\rangle_B$$

Vale inoltre (3.21):

*Matrici ridotte
nelle basi di
Schmidt*

$$\rho_1 = \text{Tr}_2 |\psi\rangle \langle \psi| = \sum_{i=1}^k p_i |u_i\rangle \langle u_i|$$

E in maniera simmetrica si trova⁴:

$$\rho_2 = \text{Tr}_1 |\psi\rangle \langle \psi| = \sum_{j=1}^k p_j |v_j\rangle \langle v_j|$$

⁴ Bastava calcolare $\rho_{12} = |\psi\rangle \langle \psi|$ usando per $|\psi\rangle$ il risultato della decomposizione di Schmidt, e quindi svolgere la traccia parziale.

Dove k è il numero di coefficienti p_i non nulli nell'espansione di $|\psi\rangle$ (pari al numero di **popolazioni** - cioè termini sulla diagonale - di ρ_1 o ρ_2). Per come è definita la sommatoria in (3.20), si ha che k è al più pari al minimo tra la $\dim \mathcal{H}_A$ e $\dim \mathcal{H}_B$. Usando le basi $\{|u_i\rangle_A\}$ e $\{|v_i\rangle_B\}$, perciò, ρ_1 e ρ_2 hanno la forma:

$$\rho_1 = \begin{pmatrix} p_1 & & & 0 \\ & \ddots & & \\ & & p_k & \\ 0 & & & 0 & \ddots \end{pmatrix} \quad \rho_2 = \begin{pmatrix} p_1 & & & 0 \\ & \ddots & & \\ & & p_k & \\ 0 & & & 0 & \ddots \end{pmatrix}$$

Dimostrazione alternativa. Usando un teorema di algebra lineare, possiamo interpretare geometricamente la decomposizione in (3.20). Partiamo dalla decomposizione nelle basi ON generiche di \mathcal{H}_A e \mathcal{H}_B :

$$|\psi\rangle_{AB} = \sum_{i=1}^{\dim \mathcal{H}_A} \sum_{j=1}^{\dim \mathcal{H}_B} \gamma_{ij} |\alpha_i\rangle_A |\beta_j\rangle_B \quad (3.23)$$

Interpretiamo i coefficienti γ_{ij} come le entrate di una matrice Γ . Dall'algebra lineare si ha che è possibile scomporre ogni matrice $k \times k$ in un prodotto:

$$\Gamma = U\Lambda V$$

dove U è una matrice **unitaria** $d_A \times d_A$, V è una matrice **unitaria** $d_B \times d_B$ e Λ è una matrice $d_A \times d_B$ con k numeri positivi $p_i > 0$ lungo la diagonale principale e 0 altrimenti. Tale relazione prende il nome di **singular value decomposition**. Geometricamente, prendendo matrici reali, corrisponde al fatto che ogni trasformazione lineare può essere vista come una composizione di una rotazione V , un riscalamento Λ e una rotazione finale U .

Detto u_{ij} le entrate di U , e v_{ij} quelle di V , i singoli elementi di Γ derivano dal prodotto matriciale:

$$\gamma_{ij} = \sum_{n=1}^k u_{in} \lambda_n v_{nj}$$

Sostituendo tale risultato in (3.23) giungiamo a:

$$|\psi\rangle_{AB} = \sum_{n=1}^k \lambda_n \underbrace{\left(\sum_{i=1}^{\dim \mathcal{H}_A} u_{in} |\alpha_i\rangle_A \right)}_{|u_i\rangle_A} \otimes \underbrace{\left(\sum_{j=1}^{\dim \mathcal{H}_B} v_{nj} |\beta_j\rangle_B \right)}_{|v_j\rangle_B}$$

Dato che U e V sono unitarie, esse trasformano basi ON in basi ON. Perciò $\{|u_i\rangle_A\}$ e $\{|v_j\rangle_B\}$ sono le basi ON di \mathcal{H}_A e \mathcal{H}_B che realizzano la decomposizione di Schmidt.

3.2.4 Schmidt Rank

Consideriamo la decomposizione di Schmidt di uno stato $|\psi\rangle_{AB}$:

$$|\psi\rangle_{AB} = \sum_{i=1}^k \sqrt{p_i} |u_i\rangle_A |v_i\rangle_B$$

Il numero k di coefficienti $p_i > 0$ è detto **rango di Schmidt**, e consente di determinare se $|\psi\rangle$ è uno stato **entangled** o meno.

In particolare, vale:

$$k = 1 \Leftrightarrow |\psi\rangle \text{ **non entangled** (separabile)}$$

*$k = 1$ per stati
non entangled*

Dimostrazione. Per $k = 1$ la decomposizione di Schmidt ha un unico termine:

$$|\psi\rangle_{AB} = |u_1\rangle_A |v_1\rangle_B$$

e quindi $|\psi\rangle_{AB}$ è separabile.

D'altro canto, per verificare l'implicazione inversa partiamo da $|\psi\rangle_{AB}$ separabile, da cui:

$$|\psi\rangle = |\psi_A\rangle |\psi_B\rangle \Rightarrow \rho = |\psi_A\rangle |\psi_B\rangle \langle\psi_A| \langle\psi_B|$$

Scegliendo una base ON $|u_i\rangle$ per \mathcal{H}_A con $|u_1\rangle = |\psi_A\rangle$ e allo stesso modo una base ON $\{|v_j\rangle\}$ per \mathcal{H}_B con $|v_1\rangle = |\psi_B\rangle$, troviamo che le matrici ridotte dei due sottosistemi sono diagonali $\text{diag}(p_1, \dots, p_n)$, con $p_1 = 1$ e $p_{i \neq 1} = 0$:

$$\rho_A = \text{Tr}_B \rho = |\psi_A\rangle \langle\psi_A| = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\rho_B = \text{Tr}_A \rho = |\psi_B\rangle \langle\psi_B| = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Poiché k corrisponde al numero di popolazioni di ρ_1 o ρ_2 , ricaviamo che $k = 1$.

D'altro canto, se partiamo da uno stato entangled, come lo stato di Bell:

*$k > 1$ per stati
entangled*

$$|\psi_{Bell}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Otteniamo $k = 2$ per il rango di Schmidt. Lo si nota usando la base computazionale per A e B , per cui vale la decomposizione:

$$|\psi_{Bell}\rangle = \sum_{i=1}^2 \frac{1}{\sqrt{2}} |i\rangle_A |i\rangle_B$$

Analogamente, possiamo estendere tali risultati a dimensioni superiori, costruendo sistemi a 3 o più livelli, per cui saranno possibili ranghi di Schmidt più alti.

Per esempio, in un sistema bipartito con 3 livelli $\{|0\rangle, |1\rangle, |2\rangle\}$, il seguente stato porta a $k = 3$:

$$|\psi_a\rangle = \sqrt{1 - 2\epsilon^2} |00\rangle + \epsilon |11\rangle + \epsilon |22\rangle \quad (3.24)$$

Nota: k , essendo un numero naturale, informa solo della presenza o meno di correlazioni, ma non quantifica la loro “intensità”. Per esempio, se poniamo $\epsilon \approx 0$ in (3.24), avremo:

$$|\psi_a\rangle \approx |00\rangle$$

Cioè lo stato $|\psi_a\rangle$ è approssimativamente uno stato separabile (non entangled).

Nella pratica, per poter usare computazionalmente l’entanglement è necessario lavorare con *correlazioni forti*, e quindi un $k \neq 1$ non è per forza indice di uno stato “usabile” sperimentalmente, per cui servirà valutare parametri più sofisticati. Inoltre, il rango di Schmidt è utile per valutare la presenza di entanglement **solo per stati puri**. Più avanti mostreremo tecniche più avanzate per superare tali limitazioni.

3.2.5 Purificazione

L’evoluzione temporale di stati misti può essere complessa da calcolare, e spesso si vorrebbe operare con **stati puri**, senza però perdere la maggiore flessibilità offerta dalle misture statistiche.

Un metodo ingegnoso per realizzar ciò è dato dalla procedura di **purificazione**. Ricordiamo infatti che, dato un sistema composto nello stato ρ_{12} , le matrici ridotte ρ_1 e ρ_2 generalmente **non** conservano la purità dello stato originario. Può allora capitare che lo stato totale ρ_{12} sia puro, ma quello “singolo” ρ_1 sia misto. In tal caso, facendo evolvere ρ_{12} come uno stato puro (seppur con la complessità di operare in dimensione maggiore) è completamente determinata l’evoluzione dello stato misto ρ_1 .

La procedura di **purificazione** si occupa di trovare la ρ_{12} pura partendo da uno stato misto conosciuto ρ_1 . Vediamo come.

La procedura di purificazione

Partiamo “dalla fine”, e cioè dallo stato puro $|\psi\rangle_{AB}$ del sistema composto. Vogliamo trovare la relazione che lo lega ai termini di ρ_1 , e che ci permetterà di determinarlo a partire da questi ultimi.

Consideriamo un sistema S bipartito, con stati in $\mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B$. Dette $\{|\alpha\rangle_A\}$ e $\{|\beta\rangle_B\}$ due basi ON per \mathcal{H}_A e \mathcal{H}_B , un generico stato di S si scrive come:

$$|\psi\rangle_{AB} = \sum_{\alpha\beta}^{\dim \mathcal{H}_{A,B}} c_{\alpha\beta} |\alpha\rangle_A |\beta\rangle_B \quad (3.25)$$

La matrice densità è allora data da:

$$\rho_{AB} = |\psi\rangle\langle\psi| = \sum_{\alpha\beta}^{\dim \mathcal{H}_{A,B}} \sum_{ab}^{\dim \mathcal{H}_{A,B}} c_{\alpha\beta} (c_{ab})^* |\alpha\beta\rangle_{AB} \langle ab|_{AB}$$

E la matrice ridotta ρ_1 si ricava calcolando la traccia parziale:

$$\begin{aligned} \rho_A &= \text{Tr}_B \rho_{AB} = \sum_{\gamma}^{\dim \mathcal{H}_B} \langle \gamma |_B \rho_{AB} | \gamma \rangle_B = \sum_{\alpha\beta} \sum_{ab} c_{\alpha\beta} (c_{ab})^* |\alpha\rangle_A \langle a|_A \sum_{\gamma}^{\dim \mathcal{H}_B} \langle b | \gamma \rangle \langle \gamma | \beta \rangle = \\ &\stackrel{(a)}{=} \sum_{\alpha a}^{\dim \mathcal{H}_A} \sum_{k=1}^{\dim \mathcal{H}_B} c_{\alpha k} (c_{\alpha k})^* |\alpha\rangle_A \langle a|_A \end{aligned} \quad (3.26)$$

dove in (a) si è usata la completezza di Dirac, per cui $\sum_{\gamma} |\gamma\rangle \langle \gamma| = \mathbb{I}$, e il fatto che la base di \mathcal{H}_B , di cui $|\beta\rangle$ e $|b\rangle$ sono elementi, è ortonormale, per cui $\langle b | \beta \rangle = \delta_{b\beta}$ permette di “far collassare” una sommatoria, e di identificare gli indici b e β con un unico indice k .

Dalla (3.26) ricaviamo che l’elemento ij della matrice ρ_A è dato da:

$$(\rho_A)_{ij} = \sum_{k=1}^{\dim \mathcal{H}_B} c_{ik} (c_{jk})^* \quad (3.27)$$

Sia $a = \dim \mathcal{H}_A$ e $b = \dim \mathcal{H}_B$. Interpretando la matrice ρ_{AB} come una matrice $b \times b$ in cui ogni elemento è in realtà un blocco di dimensioni $a \times a$ (da cui ρ_{AB} ha dimensioni $(ab) \times (ab)$), la (3.27) collega le entrate di ρ_A alle somme sulle diagonali dei singoli blocchi di ρ_{AB} (come già visto nell’introdurre la notazione matriciale per le tracce parziali). Abbiamo quindi un **sistema** di a^2 **equazioni**, ciascuna delle quali riguarda un singolo blocco di una matrice $b \times b$, che ha quindi b^2 blocchi. Avremola possibilità di trovare una soluzione (unica) solo quando $a^2 = b^2$, e cioè quando:

Condizione per la risolubilità

$$a = b \Rightarrow \dim \mathcal{H}_A = \dim \mathcal{H}_B$$

Volendo possiamo scegliere $\dim \mathcal{H}_B$ maggiore, giungendo lo stesso a trovare (più di) una soluzione, seppur con uno spreco di computazione.

Esempio: purificazione di 1 qubit

Sia dato 1 qubit in un generico stato ρ_1 , che vogliamo purificare ad uno stato puro $\rho_{12} = |\psi\rangle\langle\psi|$ di un sistema a 2 qubit. Partiamo scrivendo le equazioni per le

entrate di ρ_1 , seguendo la (3.27):

$$\begin{cases} (\rho_1)_{00} &= c_{00}c_{00}^* + c_{01}c_{01}^* \\ (\rho_1)_{01} &= (\rho_1)_{10}^* = c_{00}c_{10}^* + c_{01}c_{11}^* \\ (\rho_1)_{11} &= c_{10}c_{10}^* + c_{11}c_{11}^* \end{cases} \quad (3.28)$$

dove c_{00} , c_{01} , c_{10} e c_{11} sono l'espansione di $|\psi\rangle$ nella base computazionale $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ (3.25):

$$|\psi\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle$$

A cui corrisponde la matrice densità ρ_{12} :

$$\rho_{12} = \begin{pmatrix} |c_{00}|^2 & c_{00}c_{01}^* & c_{00}c_{10}^* & c_{00}c_{11}^* \\ c_{01}c_{00}^* & |c_{01}|^2 & c_{01}c_{10}^* & c_{01}c_{11}^* \\ c_{10}c_{00}^* & c_{10}c_{01}^* & |c_{10}|^2 & c_{10}c_{11}^* \\ c_{11}c_{00}^* & c_{11}c_{01}^* & c_{11}c_{10}^* & |c_{11}|^2 \end{pmatrix}$$

Come si nota dal sistema in (3.28), abbiamo 4 incognite in sole 3 equazioni indipendenti (poiché ρ_1 , essendo una matrice 2×2 hermitiana, ha solo 3 gradi di libertà). Possiamo allora fissare arbitrariamente una di esse, per esempio $c_{10} = 0$, e trovare quindi la soluzione:

$$c_{01} = 0 \quad c_{00} = \sqrt{\rho_{00}} \quad c_{10} = \frac{\rho_{01}^*}{\sqrt{\rho_{00}}} \quad c_{11} = \sqrt{\frac{\rho_{00}\rho_{11} - |\rho_{01}|^2}{\rho_{00}}}$$

Perciò possiamo vedere un qualsiasi stato ρ di 1 qubit come la matrice ridotta di un sistema a 2 qubit in uno stato puro dato da $|\psi\rangle$:

$$|\psi\rangle = c_{00}|00\rangle + c_{10}|10\rangle + c_{11}|11\rangle$$

Si dice quindi che $|\psi\rangle$ “**purifica**” lo stato misto ρ_1 da cui siamo partiti.

3.2.6 Rappresentazione di Kraus

Consideriamo un sistema composto da due sottosistemi 1 e 2, che si trova inizialmente in un generico stato ρ_{12} (puro o misto). Se $U(t)$ è l'operatore di evoluzione temporale, lo stato al tempo t si ottiene dalla formula di evoluzione per una matrice densità:

$$\rho_{12}(t) = U(t)\rho_{12}U^\dagger(t) \quad (3.29)$$

Ci chiediamo *come appaia* tale relazione dal punto di vista del sottosistema 1. In termini matematici, ci interessa scrivere una **relazione di evoluzione** come la (3.29) per gli stati del sottosistema 1, che sono descritti dalla matrice ridotta ρ_1 . Vogliamo trovare allora una mappa K tra matrici densità:

$$K : \rho_1 \mapsto \rho_1(t)$$

Si trova che, seppur ρ_{12} evolva unitariamente, non è detto che lo faccia anche ρ_1 . In altre parole in generale **non esiste** un operatore **unitario** U_1 per cui valga:

$$\rho_1(t) = U_1(t)\rho_1 U_1^\dagger(t) \quad (3.30)$$

Vediamolo esplicitamente. Partiamo, per semplicità, facendo due supposizioni:

- Lo stato iniziale $\rho_{12}(0)$ è **separabile**, cioè vale:

$$\rho_{12}(0) = \rho_1 \otimes \rho_2$$

Fisicamente ciò corrisponde al fatto che i sistemi 1 e 2 sono inizialmente **scorrelati** (non entangled)

- Lo stato ρ_2 è puro, ossia:

$$\rho_2 = |0\rangle_2 \langle 0|_2$$

dove con $|0\rangle_2 \in \mathcal{H}_2$ indichiamo uno stato di riferimento del sottosistema 2.

Nota: tale assunzione non fa perdere generalità, poiché se ρ_2 non è inizialmente puro basta ampliare il secondo sottosistema, aumentando la dimensione di \mathcal{H}_2 , e **purificare** ρ_2 con la procedura vista nelle sezioni precedenti.

Dalle ipotesi fatte possiamo scrivere lo stato iniziale del sistema composto come:

$$\rho_{12}(0) = \rho_1 \otimes |0\rangle_2 \langle 0|_2 \quad (3.31)$$

Calcoliamone allora l'evoluzione temporale tramite la (3.29):

$$\rho_{12}(t) = U(t)\rho_{12}(0)U^\dagger(t)$$

Determiniamo infine la matrice ridotta $\rho_1(t)$ calcolandone la traccia parziale:

$$\begin{aligned} \rho_1(t) &= \text{Tr}_2(\rho_{12}(t)) = \text{Tr}_2[U(\rho_1 \otimes |0\rangle_2 \langle 0|_2)U^\dagger] \stackrel{(a)}{=} \sum_{k=1}^{\dim \mathcal{H}_2} \underbrace{\langle k|_2 U |0\rangle_2}_{E_k} \rho_1 \underbrace{\langle 0|_2 U^\dagger |k\rangle_2}_{E_k^\dagger} = \\ &= \sum_{k=1}^{\dim \mathcal{H}_2} E_k \rho_1 E_k^\dagger \end{aligned} \quad (3.32)$$

dove in (a) abbiamo introdotto una base ON $\{|k\rangle_2\}$ di \mathcal{H}_2 per poter calcolare la traccia parziale.

La (3.32) mostra come l'evoluzione di ρ_1 abbia una forma *più generale* di quella che sarebbe propria di un'evoluzione unitaria (3.30).

In altre parole, un'*evoluzione generalizzata* è composta dalla “somma di più di una evoluzione” - analogamente ad uno stato entangled, che non può essere espresso come un singolo prodotto tensore.

Come nel caso unitario, dove $U^\dagger U = \mathbb{I}$, troviamo che anche gli operatori $\{E_k : \mathcal{H}_1 \rightarrow \mathcal{H}_1\}$, detti **operatori di Kraus**, che agiscono⁵ sugli stati di \mathcal{H}_1 , soddisfano una simile *relazione di normalizzazione*:

$$\sum_{k=1}^{\dim \mathcal{H}_2} E_k^\dagger E_k = \sum_{k=1}^{\dim \mathcal{H}_2} \langle 0|_2 U^\dagger \underbrace{|k\rangle_2 \langle k|_2}_{\mathbb{I}_2} U |0\rangle_2 \stackrel{(a)}{=} \langle 0|_2 \underbrace{U^\dagger U}_{\mathbb{I}_{12}} |0\rangle_2 = \langle 0|_2 \mathbb{I}_{12} |0\rangle_2 \stackrel{(b)}{=} \mathbb{I}_1 \quad (3.33)$$

In (a) abbiamo usato la completezza di Dirac, e il fatto che $\mathbb{I}_2 U = U$. Poiché \mathbb{I}_2 e U hanno dimensioni differenti, l'espressione ha senso in ambito tensoriale (perciò $U^\dagger \mathbb{I}_2 U$ non indica un prodotto di matrici, che non si potrebbe fare). Possiamo risolvere tutto lavorando in notazione di Dirac:

$$\mathbb{I}_2 U = \left(\sum_{j=1}^{d_2} |j\rangle_2 \langle j|_2 \right) \left(\sum_{ik}^{d_1, d_2} c_{ij} |i\rangle_1 |k\rangle_2 \right) = \sum_{ij}^{d_1, d_2} c_{ij} |i\rangle_1 |j\rangle_2 = U$$

dove $d_1 = \dim \mathcal{H}_1$ e $d_2 = \dim \mathcal{H}_2$.

Analogamente, in (b) si ha:

$$\langle 0|_2 \mathbb{I}_{12} |0\rangle_2 = \langle 0|_2 \sum_{i=0}^{d_1-1} \sum_{j=0}^{d_2-1} |i\rangle_1 |j\rangle_2 \langle i|_1 \langle j|_2 |0\rangle_2 = \sum_{i=0}^{d_1-1} |i\rangle_1 \langle i|_1 \langle 0|_2 |0\rangle_2 = \mathbb{I}_1$$

La mappa $S : \rho_1 \rightarrow \rho'_1$ definita da (3.32), con la condizione (3.33), è detta **operazione quantistica**, o **superoperatore**, dato che si tratta di una *trasformazione di operatori* (matrici). In particolare l'equazione (3.32) è detta **rappresentazione di Kraus** (o in “somma di operatori”) di S .

In questo caso siamo partiti dall'evoluzione temporale per costruire S , ma potremmo considerare delle evoluzioni “più esotiche” in cui si parte da generici operatori U unitari, giungendo sempre alla stessa rappresentazione:

$$S : \rho_1 \mapsto \rho'_1 = \sum_k E_k \rho_1 E_k^\dagger$$

Un superoperatore S è una **mappa lineare** tra operatori. Di più, si dimostra che S mappa matrici densità in matrici densità, dato che:

- $S(\rho_1)$ è hermitiana se lo è ρ_1 :

$$(\rho_1)^\dagger = \left(\sum_k E_k \rho_1 E_k^\dagger \right)^\dagger \stackrel{(a)}{=} \sum_k (E_k^\dagger)^\dagger \rho_1^\dagger E_k^\dagger = \sum_k E_k \rho_1 E_k^\dagger = \rho'_1$$

dove in (a) si usa il fatto che $E_k \rho_1 E_k^\dagger$ è un prodotto di matrici, e la trasposta coniugata inverte il senso della moltiplicazione ($(AB)^\dagger = B^\dagger A^\dagger$).

Perciò S conserva l'hermitianicità.

- S conserva la traccia:

$$\text{Tr}(\rho'_1) = \text{Tr} \left(\sum_k E_k \rho_1 E_k^\dagger \right) \stackrel{(b)}{=} \sum_k \text{Tr}(\rho_1 E_k^\dagger E_k) = \text{Tr} \left(\rho_1 \sum_k E_k^\dagger E_k \right) = \text{Tr}(\rho_1)$$

⁵Infatti $U : \mathcal{H}_1 \otimes \mathcal{H}_2 \rightarrow \mathcal{H}_1 \otimes \mathcal{H}_2$, e perciò $E_k = \langle k|_2 U |0\rangle_2 : \mathcal{H}_1 \rightarrow \mathcal{H}_1$

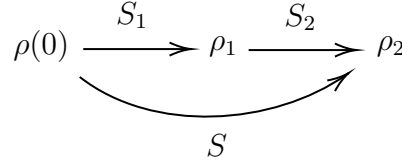
dove in (b) si è usata la ciclicità della traccia ($\text{Tr}(ABC) = \text{Tr}(BCA)$).
In particolare, per una matrice densità avremo $\text{Tr}(\rho_1) = 1$, e quindi $\text{Tr}(\rho'_1) = 1$.

- Se ρ_1 è non negativa, allora anche ρ'_1 lo è. Supponiamo allora che $\langle m | \rho_1 | m \rangle \geq 0$ per ogni $|m\rangle \in \mathcal{H}_1$. Calcolando lo stesso valor medio per ρ'_1 :

$$\langle m | \rho'_1 | m \rangle_1 = \sum_k \underbrace{\langle m |_1 E_k \rho_1 E_k^\dagger | m \rangle_1}_{\langle n |} = \sum_k \langle n |_1 \rho_1 | n \rangle_1 \geq 0$$

Inoltre S ha altre caratteristiche utili:

- **“Proprietà di gruppo”**. Si possono comporre due mappe di Kraus una dopo l'altra:



e il risultato è ancora una mappa di Kraus $S(\rho) = (S_2 \circ S_1)(\rho) = S_2(S_1(\rho))$ che gode ancora di tutte le proprietà di S .

- **Invertibilità**. Data S , S^{-1} esiste solo se è anche unitaria. Se non lo è significa, fisicamente, che vi è stata un'interazione tra A e B che non è racchiusa nella descrizione di A data da ρ_A , cioè si è persa dell'informazione nel passaggio da ρ_{A+B} alla singola ρ_A (tale processo è detto *decoerenza*). In altre parole, compare una *freccia del tempo*: la transizione $\rho \mapsto \rho'$ ha una descrizione completa, ma non quella inversa.

Nota: uno stesso superoperatore S può essere scritto come combinazione di diverse classi di operatori di Kraus, legate fra loro da una trasformazione unitaria. Per esempio $S(\rho_1) = \sum_k E_k \rho_1 E_k^\dagger$ e $S'(\rho_1) = \sum_k F_k \rho_1 F_k^\dagger$ coincidono se $F_i = \sum_j W_{ij} E_j$, con W matrice unitaria.

Geometricamente, S mappa matrici $N \times N$ in matrici $N \times N$. Una base di $\mathcal{M}_{N \times N}$ ha N^2 elementi, e perciò una generica trasformazione lineare $\mathcal{M}_{N \times N} \rightarrow \mathcal{M}_{N \times N}$ ha $(N^2)^2$ parametri. Considerando che $\sum_k E_k E_k^\dagger = \mathbb{I}_1$ forma un sistema di N^2 equazioni, avremo $N^4 - N^2$ parametri liberi per un superoperatore. Per stati di singolo qubit, le ρ sono matrici 2×2 , e quindi $N = 2$ e $2^4 - 2^2 = 12$. Per 2 qubit $N = 2^2 = 4$ e $4^4 - 4^2 = 240$ (!).

3.2.7 Kraus Representation Theorem

L'aver elencato le proprietà di S ci permette di **generalizzare** tale costruzione, considerando “evoluzioni generalizzate” relative alle varie *operazioni quantistiche* che si possono fare su un sistema. Vedremo diversi esempi di ciò nel trattare i

“canali quantistici” nella prossima sezione.

Consideriamo allora una mappa “evoluzione generalizzata” *tra stati* $S : \rho_1 \rightarrow \rho'_1$ con le seguenti proprietà (che abbiamo notato partendo dall'esempio sull'evoluzione temporale nel paragrafo precedente):

1. **Lineare:** $S(a\rho_1 + b\rho_2) = aS(\rho_1) + bS(\rho_2)$
2. Manda matrici Hermitiane in matrici **Hermitiane:** in dimensione finita vale allora $S(\rho) = [S(\rho)^T]^*$ ($S(\rho)$ descrive un operatore simmetrico)
3. Conserva la **traccia:** $\text{Tr } \rho = \text{Tr } S(\rho)$
4. **Completamente positiva.** La positività significa che se ρ è un operatore non negativo, ossia ha valori medi ≥ 0 , allora lo è anche $S(\rho)$. La *completa* positività, invece, aggiunge anche che $S \otimes \mathbb{I}_E$, ossia l'estensione di S da \mathcal{H}_1 a $\mathcal{H}_1 \otimes \mathcal{H}_E$ per un certo \mathcal{H}_E , è una mappa positiva. In altre parole, è sempre possibile estendere $S(\rho)$ ad un'operazione che agisce *localmente su un sistema*, lasciando *invariato* tutto il resto (che è il caso di interesse sperimentale, dato che non lavoreremo mai con stati dell'intero universo). La completa positività garantisce che tale operazione naturale sia sempre ben definita.

Allora si dimostra che S si può scrivere in **rappresentazione di Kraus**, cioè nella forma “decomposta” come somma di M termini:

$$S : \rho \mapsto \rho'_1 = \sum_{k=1}^M E_k \rho E_k^\dagger \quad \sum_{k=1}^M E_k^\dagger E_k = \mathbb{I}$$

Esempio

Sia ρ_1 la matrice densità di un generico qubit. Consideriamo il sistema di 2 qubit nello stato ρ **separabile**:

$$\rho = \begin{pmatrix} \rho_1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \rho_1 = |0\rangle_2 \langle 0|_2 \otimes \rho_1$$

Facciamo evolvere unitariamente la ρ , ed esaminiamo quanto accade per ρ_1 :

$$\rho'_1 = \text{Tr}_2(U\rho U^\dagger) \stackrel{(a)}{=} \text{Tr}_2 \left(\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \rho_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A^\dagger & C^\dagger \\ B^\dagger & D^\dagger \end{pmatrix} \right) = A\rho_1 A^\dagger + C\rho_1 C^\dagger$$

dove in (a) rappresentiamo le matrici U e U^\dagger in una forma a blocchi 2×2 .

Confrontando con la formula generale in (3.32), troviamo che $E_k = \{A, C\}$. Verifichiamo che la traccia è conservata:

$$\begin{aligned} \text{Tr}(\rho'_1) &= \text{Tr}(A\rho_1 A^\dagger + C\rho_1 C^\dagger) \stackrel{(a)}{=} \text{Tr}(\rho_1 A A^\dagger) + \text{Tr}(\rho_1 C C^\dagger) = \\ &= \text{Tr}(\rho \underbrace{A A^\dagger + C C^\dagger}_{\mathbb{I}}) = \text{Tr}(\rho_1) \end{aligned}$$

dove in (a) abbiamo usato la ciclicità della traccia, cioè l'invarianza per permutazioni cicliche dell'argomento:

$$\text{Tr}(ABC) = \text{Tr}(BCA) = \text{Tr}(CAB)$$

3.3 Misure generalizzate

(Lezione 7 ● del
20/3/2019)

Il formalismo delle “evoluzioni generalizzate” ci permette di trattare una classe più ampia di misurazioni.

Finora, infatti, abbiamo identificato la misura di uno stato quantistico come una *misura ideale di prima specie*, che corrisponde ad una proiezione di Von Neumann. Matematicamente, consideriamo una funzione d'onda $|\psi(t)\rangle$, e una misura dell'osservabile A a $t = 0$, che ha esito $a \in \sigma(A)$. Immediatamente dopo la misura, il sistema si trova in uno stato in cui il valore di A è ben definito, ossia in un autostato di A dell'autospazio con autovalore a , ottenuto proiettando $|\psi(0)\rangle$:

$$|\psi(0^+)\rangle = \frac{P_a^A |\psi(0)\rangle}{\sqrt{\langle \psi(0) | P_a^A | \psi(0) \rangle}}$$

Perciò una qualsiasi misura *immediatamente successiva* trova per A sempre il valore a , con certezza.

Potremmo però considerare un “processo di misura” con un risultato **incerto**, per cui lo stato finale è un elemento $\{|\psi_i\rangle\}_{i=1,\dots,N}$ scelto con **probabilità** p_i . Al posto del proiettore usiamo allora una classe più ampia di operatori $\{M_i\}$, *non necessariamente autoaggiunti*, che mappano $|\psi\rangle$ nelle varie possibilità $|\psi_i\rangle$:

Misura
generalizzata

$$|\psi_i\rangle = \frac{M_i |\psi\rangle}{\sqrt{\langle \psi | M_i^\dagger M_i | \psi \rangle}} \quad p_i = \langle \psi | M_i^\dagger M_i | \psi \rangle \quad i = 1, \dots, N$$

Poiché vogliamo che le $|\psi_i\rangle$ esauriscano tutte le possibili evoluzioni di $|\psi\rangle$, richiediamo che le p_i abbiano somma 1, e quindi:

Condizione di
completezza

$$\sum_{i=1}^N p_i = 1 \Leftrightarrow \langle \psi | \sum_{i=1}^N M_i^\dagger M_i | \psi \rangle = 1 \Leftrightarrow \sum_{i=1}^N M_i^\dagger M_i = \mathbb{I} \quad (3.34)$$

Nota: se $M_i^\dagger = M_i$ (M_i autoaggiunti) e $M_i M_j = \delta_{ij} M_i \Rightarrow M_i^2 = M_i$ (M_i proiettori ortonormali), riotteniamo le **misure proiettive** (proiezione di Von Neumann), per cui la completezza è data direttamente da $\sum_i M_i = \mathbb{I}$.

3.3.1 Teorema di Neumark

Si trova che le misure generalizzate sono equivalenti a **misure proiettive** effettuate in uno **spazio più grande** dopo una certa **evoluzione unitaria**.

In altre parole, una usuale misura proiettiva su un sistema di 2 qubit può essere descritta, a livello dei singoli qubit, da una opportuna misura generalizzata.

In particolare, il **teorema di Neumark** afferma che l'azione di una misura generalizzata può essere schematizzata come una sequenza di:

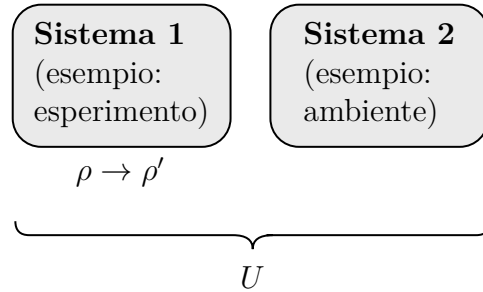
Teorema di Neumark

1. **Ampliamento del sistema:** al sistema 1 in esame si aggiunge un sistema 2 ausiliario, detto **ancilla**, con cui può interagire.
2. **Evoluzione unitaria** del sistema 1 + 2 secondo un qualche operatore U_{1+2}
3. **Misura proiettiva** sul sottosistema 2 (misura dell'ancilla), osservata dal punto di vista dello stato del sottosistema 1.

3.3.2 Rappresentazione unitaria di Operatori di Kraus

Avevamo visto che, se consideriamo una matrice densità ridotta ρ_1 , possiamo vedere ogni sua possibile **evoluzione generalizzata** come l'applicazione di opportuni *operatori di Kraus*.

Nel caso dell'evoluzione temporale, avevamo notato che la rappresentazione di Kraus dell'evoluzione del sottosistema 1 deriva dall'evoluzione unitaria del sistema composto da 1 e un sottosistema 2 (che funge da *ancilla*). Ci chiediamo se ciò sia vero in generale. Possiamo cioè affermare che *ogni* evoluzione generalizzata, descritta da una qualsiasi classe di operatori di Kraus, sia *indotta* dall'evoluzione unitaria di un sistema più grande?



Il quesito è per certi versi analogo a quello affrontato nel discutere la *purificazione* di uno stato misto, e procediamo in maniera simile.

Partendo dalla mappa $U_1 : \rho_1 \mapsto \rho'_1$ vogliamo allora trovare l'evoluzione unitaria U_{1+2} che agisce su ρ_{12} inducendo U_1 per ρ_1 . In altre parole, vogliamo interpretare una generica “evoluzione generalizzata” U_1 come un'evoluzione unitaria U_{1+2} di un sistema più ampio osservata da una delle sue parti.

Distinguiamo due casi:

- Il sottosistema 1 evolve **unitariamente** tramite U_1 . Ma allora una $U_{1+2} : \mathcal{H}_1 \otimes \mathcal{H}_2 \rightarrow \mathcal{H}_1 \otimes \mathcal{H}_2$ unitaria che agisce sul sistema 1 + 2 è data da:

$$U_{1+2} = U_1 \otimes \mathbb{I}_2$$

Dimostrazione. Supponiamo che lo stato iniziale di 1 sia $|\psi\rangle_1$ puro, e quello di 2 sia $|0\rangle_2$. U_{1+2} è il prodotto tensore tra una matrice unitaria e l'identità, e quindi è unitario. Si verifica immediatamente che tale U_{1+2} induce l'evoluzione data da U_1 per ρ_1 :

$$\begin{aligned}\rho'_1 &= \text{Tr}_2(\rho_{12}(t)) = \text{Tr}_2(U_{1+2} |\psi\rangle_1 |0\rangle_2 \langle\psi|_1 \langle 0|_2 U_{1+2}^\dagger) = \\ &= \sum_{k=1}^{\dim \mathcal{H}_2} \langle k|_2 (U_1 \otimes \mathbb{I}_2) |\psi\rangle_1 |0\rangle_2 \langle\psi|_1 \langle 0|_2 (U_1^\dagger \otimes \mathbb{I}_2) |k\rangle_2 = \\ &= \sum_{k=1}^{\dim \mathcal{H}_2} \langle k|0\rangle \langle 0|k\rangle U_1 |\psi\rangle_1 \langle\psi|_1 U_1^\dagger = U_1 \rho_1 U_1^\dagger\end{aligned}$$

Il risultato si generalizza anche considerando per ρ_1 un generico stato misto, dato che vale $\rho_1 = \sum_n p_n |\phi_n\rangle \langle\phi_n|$ e le operazioni svolte nella dimostrazione (evoluzione unitaria e traccia) sono tutte lineari. Analogamente, non è richiesto che ρ_2 corrisponda ad uno stato puro: se così non è basta estendere \mathcal{H}_2 e purificare lo stato.

- Il sottosistema 1 evolve **non unitariamente**, tramite un processo che può essere rappresentato da operatori di Kraus $\{E_k\}_{k=1,\dots,M}$, $E_k : \mathcal{H}_1 \rightarrow \mathcal{H}_1$:

$$\rho'_1 = \sum_{k=1}^M E_k \rho_1 E_k^\dagger \quad \sum_{k=1}^M E_k^\dagger E_k = \mathbb{I} \quad (3.35)$$

Questo è il caso generale in cui i sottosistemi 1 e 2 *interagiscono* durante l'evoluzione. Si trova allora che U_{1+2} in questo caso è dato da:

$$U_{1+2} |\psi\rangle_1 |0\rangle_2 = \sum_{k=1}^M (E_k \otimes \mathbb{I}_2) |\psi\rangle_1 |k\rangle_2 \quad (3.36)$$

dove $\{|k\rangle_2\}_{k=1}^M$ è una qualsiasi base ON di \mathcal{H}_2 , con $M = \dim \mathcal{H}_2$.

Dimostrazione. Supponiamo che lo stato iniziale del sistema sia $|\psi\rangle_1 |0\rangle_2 \equiv |\psi 0\rangle_{12}$.

Dimostriamo che:

1. U_{1+2} è unitario. Possiamo vederlo mostrando che preserva le norme:

$$\begin{aligned}\langle\psi 0|_{12} U_{1+2}^\dagger U_{1+2} |\psi 0\rangle_{12} &\stackrel{(3.36)}{=} \left(\sum_{k=1}^M \langle\psi|_1 \langle k|_2 (E_k^\dagger \otimes \mathbb{I}_2) \right) \left(\sum_{l=1}^M (E_l \otimes \mathbb{I}_2) |\psi\rangle_1 |l\rangle_2 \right) = \\ &\stackrel{(a)}{=} \sum_{k=1}^M \sum_{l=1}^M \underbrace{\langle k|l\rangle}_{\delta_{kl}} \langle\psi|_1 E_k^\dagger E_l |\psi\rangle_1 = \sum_{k=1}^M \langle\psi| E_k^\dagger E_k |\psi\rangle = \\ &= \langle\psi| \sum_{k=1}^M E_k^\dagger E_k |\psi\rangle \stackrel{(3.35)}{=} \langle\psi| \psi\rangle = 1\end{aligned}$$

dove in (a) usiamo il fatto che gli E_k agiscono solo su \mathcal{H}_1 , ossia sul primo sottosistema, e lasciano invariato il secondo. Tutto ciò vale anche se al posto di $|\psi\rangle$ consideriamo uno stato misto (per la linearità delle operazioni coinvolte).

2. L'evoluzione dello stato ρ_{12} del sistema composto secondo U induce l'evoluzione di ρ_1 data dagli operatori di Kraus $\{E_k\}$. Partiamo dal caso in cui ρ_1 descrive uno stato puro, ossia $\rho_1 = |\psi\rangle\langle\psi|$, da cui $\rho_{12} = |\psi_1 0\rangle\langle\psi_1 0|$. Allora si ha:

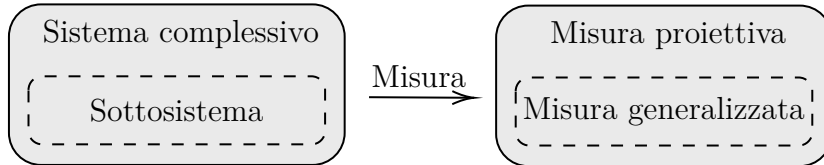
$$\begin{aligned}
\rho'_1 &= \text{Tr}_2(\rho_{12}(t)) = \text{Tr}_2(U_{1+2}\rho_{12}U_{1+2}^\dagger) = \\
&= \text{Tr}_2(U_{1+2}|\psi_1 0\rangle_{12}\langle\psi_1 0|U_{1+2}^\dagger) = \\
&= \sum_{k=1}^M \langle k|_2 \left(\sum_{l=1}^M (E_l |\psi\rangle_1) |l\rangle_2 \right) \left(\sum_{m=1}^M (\langle\psi|_1 E_m^\dagger) \langle m|_2 \right) |k\rangle_2 = \\
&= \sum_{k,l,m=1}^M \underbrace{\langle k|l\rangle}_{\delta_{kl}} \underbrace{\langle m|k\rangle}_{\delta_{km}} E_l |\psi\rangle_1 \langle\psi|_1 E_m^\dagger = \sum_{k=1}^M E_k |\psi\rangle_1 \langle\psi|_1 E_k^\dagger = \\
&= \sum_{k=1}^M E_k \rho_1 E_k^\dagger
\end{aligned}$$

Nel caso invece ρ_1 sia uno stato misto, tale risultato continua a valere: avremo infatti $\rho_1 = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, e quindi $\rho_{12} = \rho_1 \otimes |0\rangle\langle 0|$, e sfruttando la **linearità** dei passaggi svolti nella dimostrazione si arriva sempre alla tesi del teorema.

3.3.3 Motivazione delle misure generalizzate

Abbiamo visto come un'evoluzione unitaria su un sistema grande si traduce in un'evoluzione generalizzata sui sottosistemi di cui è composto (e viceversa).

Analogamente, si trova che una **misura proiettiva** sul sistema composto si traduce in una **misura generalizzata** sui suoi sottosistemi. In generale, perciò, misure proiettive su un sistema non sono più proiettive se esaminate dal punto di vista delle singole parti.



Consideriamo nello specifico un sistema $1+2$ (sistema in esame + ancilla) che viene fatto evolvere unitariamente da U_{12} , che come abbiamo visto in (3.36) è legata all'evoluzione generalizzata definita dagli operatori di Kraus $\{E_k\}$ su 1. Per un sistema che parte da $|\Psi(0)\rangle_{12} = |\psi\rangle_1 |0\rangle_2$ otteniamo:

$$|\Psi(t)\rangle_{12} = U_{12} |\psi\rangle_1 |0\rangle_2 = \sum_{k=0}^M E_k |\psi\rangle_1 |k\rangle_2 \Rightarrow \rho_{12}(t) = |\Psi(t)\rangle_{12} \langle\Psi(t)|_{12}$$

Consideriamo una **misura proiettiva** sul secondo sottosistema per verificare se si trovi o meno nello stato $|i\rangle_2$, e che quindi equivale ad un operatore che proietta la

funzione d'onda nel sottospazio generato da $|i\rangle_2$:

$$\hat{P}_i = \mathbb{I}_1 \otimes |i\rangle_2 \langle i|_2$$

Calcoliamo la probabilità di ottenere i dalla misura:

$$\begin{aligned} p_i &= \text{Tr}(\rho_{12}(t)\hat{P}_i) = \text{Tr} \left(\left[\sum_{k,k'}^M (E_k |\psi\rangle_1) |k\rangle_2 (\langle\psi|_1 E_{k'}^\dagger) \langle k'|_2 \right] (\mathbb{I}_1 \otimes |i\rangle_2 \langle i|_2) \right) = \\ &= \text{Tr} \left(\sum_{k,k'}^M E_k |\psi\rangle_1 \langle\psi|_1 E_k^\dagger \otimes \langle k'|_2 |i\rangle_2 |k\rangle_2 \langle i|_2 \right) = \\ &= \sum_{m=1}^{\dim \mathcal{H}_1} \sum_{n=1}^{\dim \mathcal{H}_2} \langle m|_1 \langle n|_2 \left(\sum_{k,k'}^M E_k |\psi\rangle_1 \langle\psi|_1 E_k^\dagger \otimes \langle k'|_2 |i\rangle_2 |k\rangle_2 \langle i|_2 \right) |m\rangle_1 |n\rangle_2 = \\ &= \sum_{m,n}^{\dim \mathcal{H}_{1,2}} \sum_{k,k'}^M \langle m|_1 E_k |\psi\rangle_1 \langle\psi|_1 E_k^\dagger |m\rangle_1 \underbrace{\langle n|_2 |k\rangle_2}_{\delta_{nk}} \underbrace{\langle k'|_2 |i\rangle_2}_{\delta_{k'i}} \underbrace{\langle i|_2 |n\rangle_2}_{\delta_{in}} = \\ &\stackrel{(a)}{=} \sum_m^{\dim \mathcal{H}_1} \langle m|_1 E_i |\psi\rangle_1 \langle\psi|_1 E_i^\dagger |m\rangle_1 \stackrel{(*)}{=} \sum_m^{\dim \mathcal{H}_1} \langle\psi|_1 E_i^\dagger \underbrace{|m\rangle_1 \langle m|_1}_{\mathbb{I}_1} E_i |\psi\rangle_1 = \langle\psi|_1 E_i^\dagger E_i |\psi\rangle_1 \end{aligned}$$

dove in (a) usiamo le δ di Kronecker per identificare $k' = k = n = i$, “collassando” tre sommatorie in un colpo solo.

Al passaggio segnato (*) possiamo anche riscrivere come:

$$p_i = \text{Tr}(E_i \rho_1 E_i^\dagger) \stackrel{(b)}{=} \text{Tr}(\rho_1 E_i^\dagger E_i) \equiv \text{Tr}(\rho_1 M_i^\dagger M_i)$$

dove in (b) usiamo la ciclicità della traccia, e nel passaggio finale riconosciamo in E_i , che generalmente non sono proiettori, gli operatori M_i che avevamo introdotto nel definire le misure generalizzate.

Riepilogando, una *evoluzione generalizzata* $S : \rho_1 \rightarrow \rho'_1$ di un certo sistema 1 è sempre indotta dall'evoluzione unitaria su un sistema più grande (e viceversa), dato dall'unione di 1 con un sottosistema 2, detto *ancilla*, con cui 1 può interagire. Possiamo ora effettuare una misura proiettiva su 1, utilizzando il formalismo di Von Neumann, o una misura proiettiva su 2, che induce una *misura generalizzata* su 1 per effetto delle correlazioni tra i due sistemi prodotte (generalmente) dall'evoluzione unitaria. Ciò apre un'intera classe di possibilità in più per esaminare lo stato di 1, che esamineremo nelle prossime due sezioni.

3.3.4 Weak Measurement

Sfruttando le correlazioni tra i sottosistemi 1 e 2 è possibile *estrarre informazioni* da 1 effettuando una misura proiettiva su 2, che si traduce - come abbiamo visto nella precedente sezione - in una misura generalizzata su 1 (teorema di Neumark). Così facendo l'idea è quella di *perturbare il meno possibile* lo stato di 1, dato che non lo si sta “direttamente proiettando” come succederebbe nel caso fosse misurato

direttamente.

Una **misura debole** consiste proprio in questo: dato un sistema 1 iniziale, lo si correla con un sistema 2 (*ancilla*), poi si separano i due sistemi e si effettua una misura su 2.

A tal proposito, ci limiteremo a dare solamente un **esempio**.

Esempio: Weak measurement di 2 qubit

Consideriamo 2 qubit, uno appartenente al *sistema* in esame S e l'altro all'*ambiente* E . Consideriamo per l'ambiente uno stato iniziale di riferimento $|0\rangle_E$, e per $|\psi_S\rangle$ prendiamo uno stato generico. Lo stato totale del sistema è quindi:

$$|\psi\rangle_{SE} = (\alpha|0\rangle + \beta|1\rangle)_S \otimes |0\rangle_E = \alpha|00\rangle_{SE} + \beta|10\rangle_{SE} \quad (3.37)$$

Consideriamo l'operazione unitaria U sul sistema:

$$U = \{(R_z(\theta))_S \otimes \mathbb{I}_E\} [(\cos\theta)\mathbb{I}_{SE} - i\sin\theta(\text{CNOT})_{SE}] \quad \theta \ll 1$$

dove la CNOT *inverte o meno* il qubit dell'ambiente $|\phi\rangle_E$ a seconda dello stato del qubit del sistema $|\psi\rangle_S$.

Nella base computazionale $\{|00\rangle_{SE}, |01\rangle_{SE}, |10\rangle_{SE}, |11\rangle_{SE}\}$, la forma matriciale di U è data da:

$$U = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & e^{i\theta} & 0 \\ 0 & 0 & 0 & e^{i\theta} \end{array} \right) \left(\begin{array}{cc|cc} e^{-i\theta} & 0 & 0 & 0 \\ 0 & e^{-i\theta} & 0 & 0 \\ \hline 0 & 0 & \cos\theta & -i\sin\theta \\ 0 & 0 & -i\sin\theta & \cos\theta \end{array} \right) = \left(\begin{array}{cc|cc} e^{-i\theta} & 0 & 0 & 0 \\ 0 & e^{-i\theta} & 0 & 0 \\ \hline 0 & 0 & e^{i\theta}\cos\theta & -ie^{i\theta}\sin\theta \\ 0 & 0 & -ie^{i\theta}\sin\theta & e^{i\theta}\cos\theta \end{array} \right)$$

Applicando U allo stato iniziale (3.37) $|\psi\rangle_{SE} = (\alpha, 0, \beta, 0)^T$ troviamo:

$$\begin{aligned} |\Psi\rangle \equiv U|\psi\rangle_{SE} &= e^{i\theta}(\alpha|00\rangle + \beta\cos\theta|10\rangle - i\beta\sin\theta|11\rangle)_{SE} = \\ &= (\alpha|0\rangle_S + \beta\cos\theta|1\rangle_S)|0\rangle_E - i\beta\sin\theta|1\rangle_S|1\rangle_E \end{aligned}$$

Consideriamo ora una **misura** sul qubit dell'**ambiente** (che possiamo pensare come parte dello strumento utilizzato per estrarre informazioni dal sistema). Abbiamo due possibili risultati:

- Il qubit dell'ambiente viene trovato nello stato $|0\rangle_E$ con probabilità p_0 :

$$\begin{aligned} p_0 &= \langle\Psi|(\mathbb{I}_S \otimes |0\rangle_E\langle 0|_E)|\Psi\rangle = |\alpha|^2 + |\beta|^2\cos^2\theta = |\alpha|^2 + |\beta|^2(1 - \theta^2 + O(\theta^4)) \\ &\approx \underbrace{|\alpha|^2 + |\beta|^2}_1 - |\beta|^2\theta^2 = 1 - |\beta|^2\theta^2 \underset{\theta \ll 1}{\sim} 1 \end{aligned}$$

In questo caso il nuovo stato (normalizzato) del sistema S è dato da:

$$\begin{aligned}
|\psi_0\rangle_S &= \frac{\alpha|0\rangle_S + \beta \cos \theta |1\rangle_S}{\sqrt{|\alpha|^2 + |\beta|^2 \cos^2 \theta}} \approx \frac{\alpha|0\rangle_S + \beta \cos \theta |1\rangle_S}{\sqrt{1 - |\beta|^2 \theta^2}} \\
&\approx (\alpha|0\rangle_S + \beta \cos \theta |1\rangle_S) \left(1 + \frac{|\beta|^2}{2} \theta^2\right) \approx \\
&\approx \alpha \left(1 + \frac{1}{2} |\beta|^2 \theta^2\right) |0\rangle_S + \beta \left(1 - \frac{\theta^2}{2}\right) \left(1 + \frac{|\beta|^2}{2} \theta^2\right) = \\
&= \alpha \left(1 + \frac{1}{2} |\beta|^2 \theta^2\right) |0\rangle_S + \beta \left(1 + \frac{|\beta|^2 - 1}{2}\right) = \\
&\stackrel{(a)}{=} \alpha \left(1 + \frac{1}{2} |\beta|^2 \theta^2\right) |0\rangle_S + \beta \left(1 - \frac{1}{2} |\alpha|^2 \theta^2\right) |1\rangle_S
\end{aligned}$$

dove in (a) abbiamo usato la normalizzazione, per cui $|\beta|^2 - 1 = |\beta|^2 - |\alpha|^2 - |\alpha|^2 = -|\alpha|^2$. L'uso delle espansioni di Taylor è giustificato dal fatto che $\theta \sim 0$.

- D'altro canto il qubit dell'ambiente viene trovato nello stato $|1\rangle_E$ con probabilità p_1 :

$$\begin{aligned}
p_1 &= 1 - p_0 = |\alpha|^2 + |\beta|^2 - |\alpha|^2 - |\beta|^2 \cos^2 \theta = |\beta|^2 (1 - \cos^2 \theta) = |\beta|^2 \sin^2 \theta \\
&\stackrel{\theta \sim 0}{\approx} |\beta|^2 \theta^2 \ll 1
\end{aligned}$$

E in tal caso lo stato finale del sistema S è $|\psi_1\rangle_S = |1\rangle_S$

Una misura debole, perciò, nella maggior parte delle volte ($p \approx 1 - |\beta|^2 \theta^2 \sim 1$) lascia il sistema in uno stato poco perturbato ($|\psi_0\rangle_S \approx |\psi\rangle$ di partenza), e raramente ($p \approx |\beta|^2 \theta^2$) lo *distrugge* proiettandolo su $|1\rangle_S$.

L'informazione estratta da un tale processo è decisamente parziale: se misure ripetute trovano sempre $|0\rangle_E$ sappiamo che $|\psi_0\rangle_S$ è probabilmente molto prossima a $|0\rangle_S$, dato che in tal caso $\beta \approx 0$. Inoltre, a seguito di ogni misura, si può *modificare* β in modo controllato.

Del resto, ripetere molte volte una misura debole produce lo stesso effetto di una misura proiettiva.

Nota. Il processo di weak measurement esemplificato può essere trattato in maniera equivalente utilizzando il formalismo delle misure generalizzate $\{M_0, M_1\}$ date da:

$$\begin{aligned}
M_0 &= |0\rangle_S \langle 0|_S + \cos \theta |1\rangle_S \langle 1|_S \\
M_1 &= \sin \theta |1\rangle_S \langle 1|_S
\end{aligned}$$

che infatti verificano:

$$\sum_i M_i^\dagger M_i = \mathbb{I}$$

Si verifica che tale scelta di $\{M_0, M_1\}$ ricostruisce le probabilità delle misurazioni e gli stati proiettati. Usando la base $\{|0\rangle_S, |1\rangle_S\}$ per la notazione matriciale:

- Per $|0\rangle_E$:

$$p_0 = \langle \psi | M_0^\dagger M_0 | \psi \rangle = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \cos \theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \cos \theta \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = |\alpha|^2 + |\beta|^2 \cos^2 \theta$$

$$|\psi_0\rangle_S = \frac{M_0 |\psi\rangle}{\sqrt{p_0}} = \frac{\alpha |0\rangle + \beta \cos \theta |1\rangle}{\sqrt{|\alpha|^2 + |\beta|^2 \cos^2 \theta}}$$

- Per $|1\rangle_E$:

$$p_1 = \langle \psi | M_1^\dagger M_1 | \psi \rangle = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \sin \theta \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \sin \theta \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = |\beta|^2 \sin^2 \theta$$

$$|\psi_1\rangle_S = \frac{M_1 |\psi\rangle}{\sqrt{p_1}} = \frac{\beta \sin \theta |1\rangle}{\sqrt{|\beta|^2 \sin^2 \theta}} = |1\rangle$$

3.3.5 POVM Measurement

Una tipologia ancora più generale di misure è data dai **POVM Measurement**, ossia i **Positive-Operator Valued Measurement**. L'idea è di considerare un sistema in cui non ci interessa il suo *stato finale*. Ciò è utile per esempio in ottica quantistica, dove la particella “viene distrutta”⁶ dopo esser stata rivelata dall'apparato sperimentale, e potrebbe anche non essere facile replicare la misura nelle stesse condizioni iniziali. Per esempio, possiamo pensare al caso di un *fotone* che incide su un detector e produce un *click*.

Possiamo trattare formalmente tale processo introducendo un set di operatori **non-negativi** $\{F_i\}_{i=1,\dots,M}$ tali che la loro somma sia l'**identità**:

$$\sum_{i=1}^M F_i = \mathbb{I} \quad (3.38)$$

Ogni F_i descrive un possibile **esito** di una misura, che avviene con probabilità data dal valor medio:

$$p_i = \langle \psi | F_i | \psi \rangle$$

Poiché gli F_i sono non negativi sappiamo che $p_i \geq 0$, e dalla condizione di completezza (3.38) vale:

$$\sum_{i=1}^M p_i = \langle \psi | \sum_{i=1}^M F_i | \psi \rangle = \langle \psi | \psi \rangle = 1$$

Perciò si ha che le F_i descrivono un insieme massimale di esiti della misurazione.

Nota: nel caso la misura sia fatta su uno stato misto ρ_1 , la probabilità dell'esito i -esimo si ottiene da $p_i = \text{Tr}(F_i \rho_1)$.

⁶Per esempio può essere riassorbita, o finire dispersa a seguito di interazioni complesse.

Nota 2: Non diamo nessuna espressione per lo stato del sistema successivamente alla misura, su cui non vengono fatte ipotesi.

Nota 3: Se vale la decomposizione $F_i = M_i^\dagger M_i$ ritroviamo il caso delle misure generalizzate (che a loro volta contengono tutte le possibili misure proiettive).

Anche in questo caso ci limitiamo a dare un **esempio**.

Consideriamo un processo che genera due possibili stati $|\psi_1\rangle$ e $|\psi_2\rangle$ dati da:

Esempio: POVM measurement

$$\begin{aligned} |\psi_{k=1}\rangle &= \sin \theta |0\rangle + \cos \theta |1\rangle \\ |\psi_{k=2}\rangle &= \sin \theta |0\rangle - \cos \theta |1\rangle \end{aligned}$$

con $0 < \theta < \pi/4$.

Abbiamo *una* sola particella, che si trova in uno dei due stati, ma non sappiamo quale. Potendo fare una sola misurazione, l'unica cosa che possiamo sperare di ottenere è *discriminare* tra i due stati possibili. Un modo ovvio per farlo è misurare nella base delle $|\psi_k\rangle$, ma ciò è generalmente difficile da fare sperimentalmente, poiché non è detto che sia possibile avere a disposizione un detector *per qualsiasi base*.

Supponiamo allora di lavorare con uno strumento che può dare tre possibili risultati $i = 0, 1, 2$. Assumiamo che le probabilità di ottenere un certo esito i siano modellizzate dal valore atteso di certe matrici F_i non-negative:

$$F_0 = \frac{1}{2} \begin{pmatrix} 1 & r \\ r & r^2 \end{pmatrix} \quad F_1 = \frac{1}{2} \begin{pmatrix} 1 & -r \\ -r & r^2 \end{pmatrix} \quad F_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 - r^2 \end{pmatrix} \quad r = \tan \theta$$

Si verifica che $\sum_i F_i = \mathbb{I}$, e quindi le $\{F_i\}$ descrivono un POVM-Measurement.

Calcoliamo allora la probabilità che il detector produca un esito i a partire dallo stato k -esimo:

$$p(i|k) = \langle \psi_k | F_i | \psi_k \rangle$$

Risulta che $p(1|1) = 0$ e $p(0|2) = 0$. Perciò se $i = 1$ escludiamo sicuramente $|\psi_1\rangle$, se $i = 0$ si esclude $|\psi_2\rangle$, e se $i = 2$ non si può escludere nessuno dei due.

3.4 Canali quantistici

(Lezione 8 ● del 21/3/2019)

Il formalismo introdotto nelle sezioni precedenti ci consente ora di caratterizzare alcuni fenomeni utili. Per esempio, nella pratica, nessun sistema S può essere totalmente isolato dall'ambiente E che lo circonda. I fenomeni di interazione introducono perciò **rumore** nell'evoluzione degli stati, che cessa di essere **unitaria**. In particolare, un fenomeno importante è dato dalla **decoerenza**, in cui si creano *correlazioni* tra sistema e ambiente che hanno l'effetto di *distruggere* le sovrapposizioni quantistiche in cui si trova S , trasformando stati puri in misture statistiche.

Ciò può risultare estremamente dannoso per la computazione quantistica, dato che stati puri *entangled* sono necessari per diversi protocolli (es. teletrasporto quantistico). D'altro canto, la *decoerenza* si instaura anche nel meccanismo che porta ad una qualsiasi misura, e gioca un ruolo fondamentale nel trasformare *stati coerenti* propri della MQ in *stati classici*. Per esempio, nel noto esperimento mentale del gatto di Schrödinger, il motivo per cui non risulta possibile osservare una sovrapposizione *non fisica* gatto vivo-gatto morto è legato a tale meccanismo⁷.

Introduciamo allora la nozione di **canale quantistico** C , con cui intendiamo in generale un qualsiasi *processo fisico* - non necessariamente unitario - che trasforma uno stato ρ_S di un certo sistema S in uno stato ρ'_S :

$$C : \rho_S \mapsto \rho'_S$$

In particolare, se ρ_S è uno stato puro, non è detto che lo sia anche ρ'_S . Esamineremo ora le caratteristiche di alcuni canali quantistici d'interesse.

3.4.1 Decoerenza di un qubit

Partiamo da un modello semplice per il fenomeno di decoerenza. Consideriamo un sistema S costituito da un **singolo qubit**, che si trova inizialmente in un generico stato puro $|\psi\rangle$:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \Rightarrow \rho_S = |\psi\rangle\langle\psi| = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix}$$

Gli elementi sulla diagonale di ρ_S sono detti **popolazioni**, e descrivono la probabilità di ottenere rispettivamente gli esiti 0 o 1 da una misura di σ_z .

D'altro canto, i termini fuori dalla diagonale sono detti *coerenze*, e appaiono solo nel caso $|\psi\rangle$ presenti sovrapposizioni quantistiche degli stati della base utilizzata.

Il fenomeno di **decoerenza** D , distruggendo le superposizioni, ha come effetto l'annullamento di tali termini.

$$\rho_S \xrightarrow{D} \rho'_S = \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}$$

Tale trasformazione può essere schematizzata dall'azione di una CNOT, che usa come qubit di controllo il ρ_S , e lo correla con lo stato $|0\rangle_E$ dell'ambiente:

⁷^Tuttavia, la decoerenza non risolve il *problema della misura* in MQ. La decoerenza spiega come mai scompaiano le sovrapposizioni quantistiche, ma non come dalle risultanti possibilità mutualmente esclusive si giunga ad una e una sola realtà, cosa che è tuttora oggetto di numerose interpretazioni.

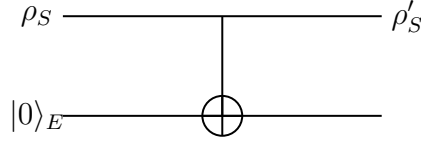


Figura (3.2) – Schema dell’azione CNOT che “distrugge” le coerenze di ρ_S

Vediamo come. Lo stato iniziale del sistema qubit-ambiente è dato da:

$$|\psi_0\rangle_{SE} = |\psi\rangle_S \otimes |0\rangle_E = (\alpha |00\rangle + \beta |10\rangle)_{SE}$$

L’azione della CNOT lo porta a $|\psi_1\rangle$:

$$|\psi_1\rangle = \alpha |00\rangle_{SE} + \beta |11\rangle_{SE} \quad (3.39)$$

In altre parole, quando il sistema è nello stato 0, lo stato dell’ambiente rimane imperturbato, ma quando il qubit S è nello stato 1, il qubit di E viene invertito. La $|\psi_1\rangle$ è uno stato non separabile, cioè entangled: la CNOT ha quindi *correlato* S ed E .

Si verifica ora che ρ'_S ottenuta eseguendo la traccia parziale di $\rho_1 = |\psi_1\rangle \langle \psi_1|$ ha i termini di coerenza nulli:

$$\rho'_S = \text{Tr}_E |\psi_1\rangle \langle \psi_1| = \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}$$

Fisicamente, lo stato finale (3.39) significa che l’ambiente *conosce* lo stato di S - infatti conoscere $|\psi\rangle_E$ equivale a conoscere $|\psi\rangle_S$, dato che a correlazione tra i due è perfetta. In altre parole, ciò significa che l’ambiente *ha misurato* il sistema S . Ciò ha l’effetto di *spostare* i coefficienti α e β da $|\psi\rangle_S$ a ρ_S . Sperimentalmente, poiché non si ha accesso allo stato globale, ma solo alla parte del sistema, ciò equivale ad una *perdita di informazione accessibile*.

Notiamo inoltre come l’evoluzione di sistema + ambiente sia unitaria, mentre quella del solo sistema non lo sia - come abbiamo notato nelle precedenti sezioni nell’esaminare l’evoluzione di una matrice ridotta.

Esaminiamo allora tale evoluzione non-unitaria “che distrugge le coerenze” focalizzandoci sul primo sottosistema. Per il teorema di rappresentazione di Kraus, esistono certi operatori $\{E_k\}$ tali che:

$$\rho_S \mapsto \rho'_S = \sum_k E_k \rho E_k^\dagger \quad \sum_k E_k^\dagger E_k = \mathbb{I} \quad (3.40)$$

Lo stato di un singolo qubit è codificato in un vettore \vec{v} nella sfera di Bloch (3.11, pag. 70), e può essere scritto nella base $\{\mathbb{I}, \hat{\sigma}_i\}$ delle matrici hermitiane 2×2 . Possiamo allora osservare l’evoluzione in quest’ottica:

$$\rho_S = \frac{1}{2}(\mathbb{I} + \vec{r} \cdot \vec{\sigma}) \Rightarrow \rho'_S = \frac{1}{2}(\mathbb{I} + \vec{r}' \cdot \vec{\sigma}) \quad (3.41)$$

dove \vec{r} e \vec{r}' sono due vettori nella sfera di Bloch.

Poiché $\rho_S \mapsto \rho'_S$ è lineare, la mappa $\vec{r} \mapsto \vec{r}'$ è una **trasformazione affine**:

$$\vec{r} \mapsto_D \vec{r}' = M\vec{r} + \vec{c} \quad (3.42)$$

per una certa matrice $M_{3 \times 3}$ e vettore $\vec{c} \in \mathbb{R}^3$.

Scrivendo gli operatori di Kraus nella stessa base:

$$E_k = \gamma_k \mathbb{I} + \sum_{l=1}^3 a_{kl} \hat{\sigma}_l$$

è possibile trovare una relazione tra le entrate di M e c e i coefficienti γ_k , a_{kl} che definiscono gli operatori di Kraus E_k . La procedura fa uso della decomposizione polare di $M = OS$, con O matrice ortogonale e S simmetrica e non negativa, e risulta nelle seguenti relazioni esplicite:

$$M_{jk} = \sum_{l=1}^3 \left\{ a_{lj} a_{lk}^* + a_{lj}^* a_{lk} + \left(|\gamma_l|^2 - \sum_{p=1}^3 |a_{lp}|^2 \right) \delta_{jk} + i \sum_{p=1}^3 \epsilon_{jkp} (\gamma_l a_{lp}^* - \gamma_l^* a_{lp}) \right\}$$

$$c_j = 2i \sum_{k,l,m=1}^3 \epsilon_{jlm} a_{kl} a_{km}^*$$

Origine della trasformazione affine e interpretazione geometrica.

We can expand the Kraus matrices as $E_k = \gamma_k \mathbb{I} + \sum_i a_{ik} \sigma_i$: our full expression becomes

$$\rho \rightarrow \rho' = \sum_k \left(\gamma_k \mathbb{I} + \sum_i a_{ik} \sigma_i \right) \frac{1}{2} (\mathbb{I} + \vec{r} \cdot \vec{\sigma}) \left(\gamma_k^* \mathbb{I} + \sum_j a_{jk}^* \sigma_j^\dagger \right) \quad (3.43)$$

and our claim is that

$$\rho' = \frac{1}{2} (\mathbb{I} + (M_i^j r_j + c_i) \sigma_i) \quad (3.44)$$

for some matrix M_i^j and vector c_i . This can be readily seen by noticing that:

1. products of Pauli matrices are linear combinations of Pauli matrices: $\sigma_a \sigma_b = \delta_{ab} \mathbb{I} + i \epsilon_{abc} \sigma_c$;
2. the Kraus transformation sends density matrices into density matrices, so the trace of ρ' will still be 1 and we will be able to separate the trace term $\mathbb{I}/2$ from the traceless Pauli matrix part (that is, there will not be any transformation-dependents coefficients multiplying the identity).

Now, to see that it is a contraction recall that $\text{Tr} \rho^2 \leq 1$. We will apply the formula:

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = (\vec{a} \cdot \vec{b}) \mathbb{I} + i(\vec{a} \wedge \vec{b}) \cdot \vec{\sigma} \quad (3.45)$$

So then:

$$\text{Tr}[(\rho')^2] = \text{Tr}\left[\frac{1}{4}(\mathbb{I} + \vec{r}' \cdot \vec{\sigma})^2\right] = \text{Tr}\left[\frac{1}{2}\left(\frac{1 + |\vec{r}'|^2}{2}\mathbb{I} + \vec{r}' \cdot \vec{\sigma}\right)\right] \leq 1 \quad (3.46)$$

Because $\text{Tr}(\vec{r}' \cdot \vec{\sigma}) = 0$, by using trace linearity we derive:

$$\text{Tr}[(\rho')^2] = \frac{1}{2} \frac{1 + |\vec{r}'|^2}{2} \text{Tr}(\mathbb{I}) = \frac{1}{2}(1 + |\vec{r}'|^2) \leq 1$$

therefore $|\vec{r}'| \leq 1$: the image of the unit sphere is contained in the unit sphere.

Geometricamente, possiamo immaginare di applicare (3.42) all'intera sfera di Bloch in una volta sola. In tal caso, l'effetto è quello di *contrarre* la sfera in un ellissoide. Il processo è in generale *non invertibile*, dato che tutte le matrici densità che differiscono per i soli termini di coerenza vengono mappate nello stesso stato finale.

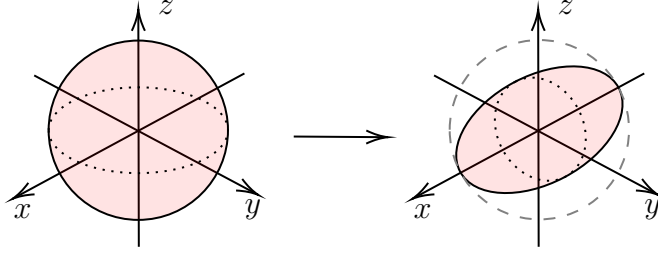


Figura (3.3) – Rappresentazione geometrica della trasformazione $\rho \mapsto \rho'$ dal punto di vista dei vettori nella sfera di Bloch. Generalmente il processo non è invertibile, e il volume può solo diminuire.

3.4.2 Canale bit-flip

Oltre a distruggere le coerenze, il *rumore* provocato dall'interazione sistema-ambiente può inserire **errori** nell'informazione contenuta nello stato trasmesso.

Per esempio, consideriamo un canale quantistico che “inverte” - tramite l'azione di σ_x (gate NOT) - un qubit con una certa probabilità $p = |\alpha|^2$, lasciandolo invariato nei restanti casi $(1 - p)$. Matematicamente, una tale trasformazione \mathcal{S} è realizzata da:

$$\rho' = \mathcal{S}(\rho) = \underbrace{|\alpha|^2}_p \sigma_x \rho \sigma_x^\dagger + (1 - |\alpha|^2) \rho$$

Riconosciamo in quest'espressione la rappresentazione di Kraus:

$$\rho' = \mathcal{S}(\rho) = \sum_{k=0}^1 E_k \rho E_k^\dagger$$

con gli operatori $\{E_0, E_1\}$ dati da:

$$E_0 = \sqrt{1 - |\alpha|^2} \mathbb{I} \quad E_1 = |\alpha| \sigma_x$$

Per esempio, \mathcal{S} mappa lo stato puro $\rho = |0\rangle\langle 0|$ a quello misto:

$$\rho = |0\rangle\langle 0| \mapsto \rho' = |\alpha|^2 |1\rangle\langle 1| + (1 - |\alpha|^2) |0\rangle\langle 0|$$

E in maniera analoga per $\rho = |1\rangle\langle 1|$:

$$\rho = |1\rangle\langle 1| \mapsto \rho' = |\alpha|^2 |0\rangle\langle 0| + (1 - |\alpha|^2) |1\rangle\langle 1|$$

Geometricamente, il bit-flip corrisponde alla trasformazione affine data da:

$$\begin{cases} x' = x \\ y' = (1 - 2|\alpha|^2)y \\ z' = (1 - 2|\alpha|^2)z \end{cases}$$

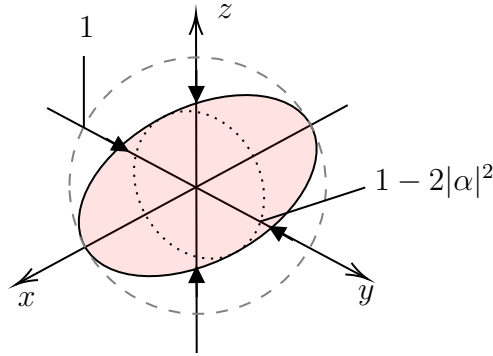


Figura (3.4) – L'operazione di bit-flip deforma la sfera di Bloch contraendola lungo le direzioni \hat{y} e \hat{z} di una quantità $1 - 2|\alpha|^2$, e lasciandola invariata lungo \hat{x}

Possiamo interpretare tale canale come l'azione di una C-NOT sullo stato iniziale ρ , condizionata da un qubit ausiliario $|\psi\rangle_c$ che *codifichi* la probabilità di inversione:

$$|\psi\rangle_c = \alpha |0\rangle + \sqrt{1 - |\alpha|^2} |1\rangle \quad (3.47)$$

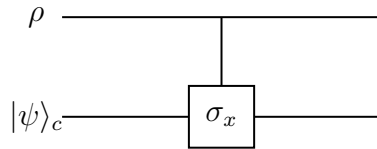


Figura (3.5) – Schema del circuito a gate quantistici equivalente all'operazione di bit-flip

3.4.3 Canale phase-flip

Analogamente al bit-flip possiamo considerare l'operazione che inverte la fase relativa di un qubit con probabilità $|\gamma|^2$. La descrizione è identica a quella del canale precedente, se non nell'uso di σ_z al posto di σ_x :

$$\rho' = \mathcal{S}(\rho) = |\gamma|^2 \sigma_z \rho \sigma_z^\dagger + (1 - |\gamma|^2) \rho$$

Esplicitando la rappresentazione di Kraus si ottiene:

$$\rho' = \sum_{k=0}^1 E_k \rho E_k^\dagger; \quad \begin{cases} E_0 = \sqrt{1 - |\gamma|^2} \mathbb{I} \\ E_1 = |\gamma| \sigma_z \end{cases} \quad (3.48)$$

Il canale quantistico mappa uno stato puro $|\varphi_+\rangle = \mu|0\rangle + \nu|1\rangle$ nella mistura statistica:

$$\rho' = \mathcal{S}(\rho) = |\gamma|^2 |\varphi_-\rangle \langle \varphi_-| + (1 - |\gamma|^2) |\varphi_+\rangle \langle \varphi_+|$$

Per un generico stato ρ , invece:

$$\rho = \begin{pmatrix} p_0 & \alpha \\ \alpha^* & 1 - p_0 \end{pmatrix} \mapsto \rho' = \begin{pmatrix} p & \alpha(1 - 2|\gamma|^2) \\ \alpha^*(1 - 2|\gamma|^2) & 1 - p \end{pmatrix} \quad (3.49)$$

Notiamo che per $|\gamma|^2 = 1/2$, ossia nel caso in cui l'inversione di fase avvenga casualmente, i termini di coerenza della matrice densità vengono annullati. Il canale di *phase-flip*, perciò, costituisce un secondo modello per il processo di **decoerenza**.

Geometricamente la sfera di Bloch viene deformata da:

$$\begin{cases} x' = (1 - 2|\gamma|^2)x \\ y' = (1 - 2|\gamma|^2)y \\ z' = z \end{cases}$$

ossia in un *ellissoide* simmetrico rispetto all'asse \hat{z} (figura 3.6) - analogamente, a meno di una rotazione, a quanto visto per il canale *bit-flip*.

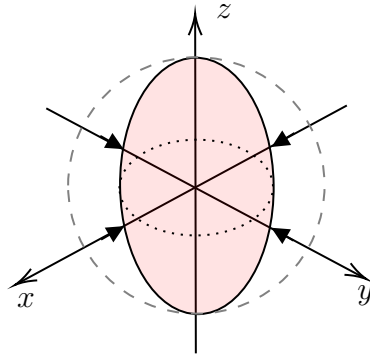


Figura (3.6) – L'operazione di phase-flip deforma la sfera di Bloch contraendola lungo le direzioni \hat{x} e \hat{y} di una quantità $1 - 2|\alpha|^2$, e lasciandola invariata lungo \hat{z}

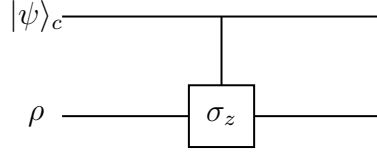


Figura (3.7) – Schema a gate quantistici dell’operazione di phase-flip (= schema del bit-flip, ma con σ_z al posto di σ_x)

3.4.4 Canale bit-phase flip

Mettendo insieme i due canali appena esaminati, consideriamo l’operazione che con una certa probabilità $|\beta|^2 = p$ effettua sia un’inversione del qubit che un’inversione della sua fase relativa:

$$\rho' = S(\rho) = |\beta|^2 \sigma_y \rho \sigma_y^\dagger + (1 - |\beta|^2) \rho$$

L’interpretazione geometrica, analogamente ai casi precedenti, porta ad una *contrazione* con simmetria attorno all’asse \hat{y} :

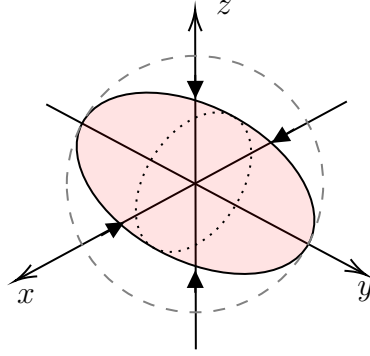


Figura (3.8) – Canale di Bit-Phase Flip

3.4.5 Depolarizing channel

Unendo tutte e tre le ultime operazioni (con una certa probabilità p) otteniamo il funzionamento del **depolarizing channel**:

$$\rho' = \frac{1}{3} [\sigma_x \rho \sigma_x^\dagger + \sigma_y \rho \sigma_y^\dagger + \sigma_z \rho \sigma_z^\dagger] + (1 - p) \rho$$

dove stiamo usando tutti e 4 gli operatori di Kraus:

$$E_k = \left\{ \sqrt{1-p} \mathbb{I}, \sqrt{\frac{p}{3}} \sigma_i \right\}$$

Geometricamente ciò equivale a una contrazione *lungo tutti gli assi*:

$$\vec{r}' = \left(1 - \frac{4}{3}p\right) \vec{r}$$

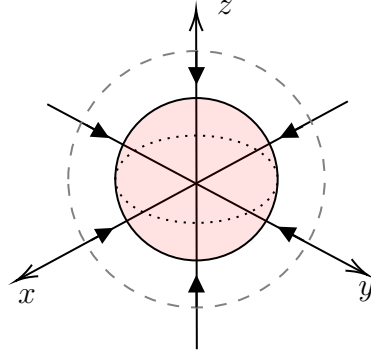


Figura (3.9) – Azione del depolarizing channel sulla Sfera di Bloch

Applicando diverse volte tale canale è possibile far collassare l'intera sfera di Bloch sull'origine (che è il punto fisso della trasformazione geometrica, e corrisponde a $\vec{r} = \vec{0} \Rightarrow \rho = \mathbb{I}/2$ - stato massimamente misto).

3.4.6 Amplitude damping

Consideriamo ora l'operazione in rappresentazione di Kraus:

$$\rho' = \mathcal{S}(\rho) = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger = \begin{pmatrix} \rho_{00} + p\rho_{11} & \sqrt{1-p}\rho_{01} \\ \sqrt{1-p}\rho_{10} & (1-p)\rho_{11} \end{pmatrix}$$

dove gli operatori di Kraus $\{E_k\}$ sono dati da:

$$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} \quad E_1 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}$$

Tale canale “trasferisce” parte della popolazione di $|1\rangle$ a $|0\rangle$. Lo si vede partendo da uno stato eccitato $|1\rangle$:

$$\rho = |1\rangle \langle 1| \mapsto \rho' = p|0\rangle \langle 0| + (1-p)|1\rangle \langle 1|$$

Notiamo che tale trasformazione mappa uno stato puro in uno stato misto (dato che è una combinazione lineare di proiettori indipendenti).

Nella sfera di Bloch, il canale di Amplitude Damping è rappresentato dalla trasformazione

$$r \mapsto \begin{pmatrix} \sqrt{1-p} & & \\ & \sqrt{1-p} & \\ & & 1-p \end{pmatrix} r + \begin{pmatrix} 0 \\ 0 \\ p \end{pmatrix} \quad (3.50)$$

che ha come punto fisso solo $(0, 0, 1)^\top$ e contrae tutte le direzioni.

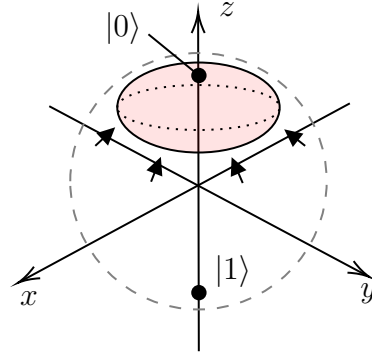


Figura (3.10) – Schema grafico del canale di amplitude damping: la sfera di Bloch viene contratta verso il punto corrispondente allo stato $|0\rangle\langle 0|$

Applicando n volte di seguito l'azione del canale, l'elemento ρ_{11} della matrice densità decade esponenzialmente:

$$\rho_{11}^{(n)} = (1 - p)^n \rho_{11}^{(0)} = \overbrace{e^{n \ln(1-p)}}^{e^{-|A|^2 n}} \rho_{11}^{(0)} \xrightarrow{n \rightarrow \infty} 0$$

In altre parole, diviene sempre più improbabile trovare il qubit nello stato $|1\rangle$. Ricordiamo infatti che una generica matrice densità di un singolo qubit, nella sua forma più generale, è data da:

$$\rho = \begin{pmatrix} p_0 & \alpha \\ \alpha^* & 1 - p_0 \end{pmatrix}$$

dove $1 - p_0 \equiv p_1$ è la probabilità di ottenere $|1\rangle$ da una misura del qubit, ed è proprio il termine che viene modificato dall'*amplitude damping* (α e α^* calano di conseguenza, dato che per normalizzazione deve valere $|\alpha| \leq \sqrt{p_0(1 - p_0)}$).

Ciò significa che lo stato limite, dopo infinite ripetizioni, è dato da:

$$\rho^{(\infty)} = |0\rangle\langle 0|$$

che è uno stato puro.

Perciò, il canale di *amplitude damping* trasforma *eventualmente* ogni stato (puro o misto) nello stato puro $|0\rangle\langle 0|$. In altre parole, l'ambiente “risucchia” l'informazione contenuta nel qubit, azzerandolo.

Nota. L'azione \mathcal{S}^∞ di infinite ripetizioni del canale *non* equivale ad effettuare un NOT su $|1\rangle$, anche se il risultato è lo stesso. Nel caso di \mathcal{S} , infatti, il processo (che può essere inteso come *continuo*) attraversa *infiniti stati misti*, ed è intrinsecamente non-coerente.

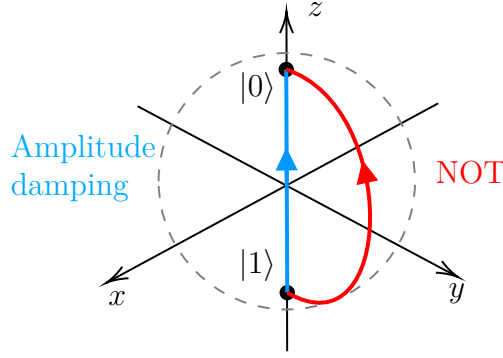


Figura (3.11) – Azione asintotica del canale di Amplitude Damping

3.4.7 Phase Damping

L'analogo per le fasi dell'amplitude damping è detto **phase-damping** - e, come vedremo ora, costituisce un terzo modello semplice per il fenomeno di **decoerenza**.

Supponiamo di partire da un generico stato ρ di un qubit, dato da:

$$\rho = \begin{pmatrix} p_0 & \alpha \\ \alpha^* & 1 - p_0 \end{pmatrix}$$

Nel canale di *phase damping* l'interazione del qubit con l'ambiente risulta in *rotazioni* $R_z(\theta)$ ad angoli casuali, che modificano i termini di coerenza α (si parla di *phase kick*):

$$\rho'(\theta) = R_z(\theta)\rho R_z^\dagger(\theta); \quad R_z(\theta) = \begin{pmatrix} \exp(-i\frac{\theta}{2}) & 0 \\ 0 & \exp(i\frac{\theta}{2}) \end{pmatrix}$$

Supponiamo che l'angolo θ sia scelto casualmente da una distribuzione gaussiana centrata in $\theta = 0$ data da:

$$p(\theta) = \frac{1}{\sqrt{4\pi\lambda}} \exp\left(-\frac{\theta^2}{4\lambda}\right)$$

dove il parametro λ ne specifica la deviazione standard.

Lo stato finale si ottiene allora integrando su tutte le possibili trasformazioni, pesate dalla loro probabilità $p(\theta)$ di essere attuate:

$$\rho' = \int_{-\infty}^{+\infty} d\theta p(\theta) R_z(\theta) \rho R_z^\dagger(\theta) \quad \rho = \begin{pmatrix} p & \alpha \\ \alpha^* & 1 - p \end{pmatrix}$$

dove gli estremi dell'integrale sono scelti in modo da rendere calcolabile analiticamente l'integrale della gaussiana.

Integrando elemento per elemento si giunge allora a:

$$\rho' = \begin{pmatrix} p & \alpha e^{-\lambda} \\ \alpha^* e^{-\lambda} & 1 - p \end{pmatrix}$$

Confrontando con (3.49), scopriamo che l'amplitude damping corrisponde al canale di *phase-flip*, ponendo $1 - 2|\gamma|^2 = e^{-\lambda}$, per cui scelta λ possiamo calcolare la probabilità di inversione di fase $|\lambda|^2$.

3.4.8 Canale di de-entanglement

Consideriamo due qubit in uno stato entangled. Se uno qualsiasi dei due subisce un fenomeno di decoerenza, l'*entanglement* viene distrutto. Vediamo esplicitamente come.

Partiamo da uno stato di Bell (massimamente entangled):

$$|\psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \Rightarrow \rho = \frac{1}{2} \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ \hline 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

dove la matrice ρ è espressa nella base computazionale $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$.

Consideriamo ora un canale che *distrugga* i termini di coerenza (evidenziati in giallo). Per esempio, usiamo il canale definito dagli operatori di Kraus F_0 e F_1 così specificati:

$$F_0 = \mathbb{I} \otimes \tilde{F}_0 = \mathbb{I} \otimes \begin{pmatrix} 1 & 0 \\ 0 & \cos \theta \end{pmatrix} = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \cos \theta \end{array} \right)$$

$$F_1 = \mathbb{I} \otimes \tilde{F}_1 = \mathbb{I} \otimes \begin{pmatrix} 0 & 0 \\ 0 & \sin \theta \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & \sin \theta & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sin \theta \end{array} \right)$$

Gli operatori di Kraus $\{\tilde{F}_0, \tilde{F}_1\}$ costituiscono una descrizione equivalente ai (3.48) per il canale di *phase-flip*.

Infatti:

$$\rho = \begin{pmatrix} p_0 & \alpha \\ \alpha^* & 1 - p_0 \end{pmatrix} \mapsto \rho' = F_0 \rho F_0^\dagger + F_1 \rho F_1^\dagger = \begin{pmatrix} p_0 & \alpha \cos \theta \\ \alpha^* \cos \theta & 1 - p_0 \end{pmatrix}$$

che è equivalente a (3.49) ponendo $\cos \theta = 1 - 2|\gamma|^2$.

In effetti, la rappresentazione di Kraus di un superoperatore non è unica (esattamente come non è unica la decomposizione in prodotti tensori di uno stato entangled), e set di operatori equivalenti sono legati da opportune matrici unitarie. In questo caso vale:

$$F_i = \sum_{j=0}^1 W_{ij} E_j; \quad W = \begin{pmatrix} \cos \frac{\theta}{2} & \sin \frac{\theta}{2} \\ \sin \frac{\theta}{2} & -\cos \frac{\theta}{2} \end{pmatrix}$$

Lo stato finale è quindi dato da:

$$\rho' = \sum_{i=0}^1 F_i \rho F_i^\dagger = \frac{1}{2} \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & \cos \theta & 0 \\ \hline 0 & \cos \theta & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Ponendo $\cos \theta = 0$ si trova:

$$\rho' = \frac{1}{2} \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) = \frac{1}{2} |01\rangle \langle 01| + \frac{1}{2} |10\rangle \langle 10|$$

che è uno **stato misto** (somma di proiettori indipendenti), e per di più **separabile** - nel senso generalizzato⁸ di somma di stati sperabili ($|01\rangle$ e $|10\rangle$, presi con uguale probabilità). In altre parole, le correlazioni presenti in ρ' sono puramente di natura classica, ossia derivano dall'ignoranza insita in una mistura statistica, e non di natura quantistica: non vi è alcuna superposizione in ρ' .

L'effetto della decoerenza è, nuovamente, quello di “effettuare una misura” su ρ - senza poterne però conoscere l'esito a priori.

Notiamo che il canale di de-entanglement *non modifica* i singoli qubit. Più precisamente, se osserviamo l'evoluzione dal punto di vista delle matrici ridotte, troviamo:

$$\begin{aligned} \rho &= \frac{1}{2}(|01\rangle \langle 01| + |10\rangle \langle 10| + |01\rangle \langle 10| + |10\rangle \langle 01|) & \Rightarrow \rho_1 = \text{Tr}_2 \rho = \frac{1}{2}(|0\rangle \langle 0| + |1\rangle \langle 1|) = \frac{1}{2}\mathbb{I} \\ \rho' &= \frac{1}{2}(|01\rangle \langle 01| + |10\rangle \langle 10|) & \Rightarrow \rho'_1 = \frac{1}{2}(|0\rangle \langle 0| + |1\rangle \langle 1|) = \frac{1}{2}\mathbb{I} \end{aligned}$$

e analogamente vale $\rho_2 = \rho'_2$ (per simmetria).

In altre parole, la decoerenza distrugge le *correlazioni* tra i due qubit, ossia la parte di informazione “senza analogo classico”. In questo caso, due qubit inizialmente *entangled* divengono alla fine completamente *indipendenti* l'uno dall'altro. In un certo senso, la decoerenza riduce qubit a *normali* bit.

⁸^Torneremo più avanti sulla caratterizzazione dell'*entanglement* in stati misti

3.5 Master equation

Cerchiamo ora un'equazione, analoga a quella di Schrödinger, che descriva l'evoluzione (generalmente non unitaria) di un *sistema aperto*, ossia interagente con l'ambiente.

Sia $\rho(t_0)$ lo stato del sistema in esame ad un istante t_0 , e $\rho_{\text{tot}}(t_0)$ quello di sistema + ambiente (generalmente non separabile). Sappiamo che ρ_{tot} evolve in maniera unitaria, mediante l'operatore $U(t_0 + \Delta t, t_0)$. Consideriamo un'evoluzione infinitesima ($\Delta t = dt$) e focalizziamoci (con una traccia parziale) sulla matrice ridotta ρ :

$$\rho(t_0 + dt) = \text{Tr}_{\text{amb.}}[\rho_{\text{tot}}(t_0 + dt)] = \quad (3.51)$$

$$= \text{Tr}_{\text{amb.}}[U(t_0 + dt, t_0)\rho_{\text{tot}}(t_0)U^\dagger(t_0 + dt, t_0)] \quad (3.52)$$

Si ha perciò che $\rho(t_0 + dt)$ dipende da $\rho_{\text{tot}}(t_0)$. Ciò porta a due problemi:

1. Non conosciamo lo stato dell'ambiente, e vorremmo una equazione che riguardi solo il sistema di interesse. Possiamo allora supporre (**approssimazione di Born**) che l'ambiente sia molto grande rispetto al sistema, e che quindi rimanga praticamente invariato a seguito delle interazioni sistema-ambiente.
2. La presenza di correlazioni tra sistema e ambiente fa sì che, in generale, $\rho_{\text{tot}}(t_0)$ dipenda dagli stati $\rho(t)$ per $t < t_0$. In altre parole, l'intera *storia* del sistema, ossia tutti i suoi precedenti stati, può influenzarne l'evoluzione. Ciò è ingestibile - nella pratica avremo a che fare con un solo dato iniziale, ossia lo stato di partenza $\rho(t_0)$. Supponiamo perciò che l'ambiente *non abbia memoria* (**approssimazione di Markov**). Ciò significa che l'informazione può solo “uscire” dal sistema, e il sistema “non può accedere” a informazioni sul suo passato registrate dall'ambiente.

Tale approssimazione è generalmente valida se ogni effetto *di memoria* dell'ambiente avviene su tempi molto minori rispetto alle dinamiche di interesse. A livello matematico, un'evoluzione Markoviana può essere espressa (localmente) come soluzione di un'equazione differenziale del primo ordine:

$$\rho(t + dt) = \rho(t) + \dot{\rho}(t)dt + O([dt]^2) \quad (3.53)$$

Come visto nelle precedenti sezioni, la (3.52) induce un'evoluzione generalizzata sul sistema, che ha una rappresentazione di Kraus:

$$\rho(t + dt) = \mathcal{S}(t, t + dt)\rho(t) = \sum_{k=0}^{M-1} E_k \rho(t) E_k^\dagger \quad (3.54)$$

con $M \leq N^2$, $N = \dim(\mathcal{H})$.

Richiediamo le seguenti (naturali) condizioni:

1. L'evoluzione per un tempo nullo equivalga all'identità: $\mathcal{S}(t; t) = \mathbb{I}$

2. $\mathcal{S}(t, t + dt)$ riproduca i risultati standard dell'equazione di Schrödinger dipendente dal tempo nel caso di un'evoluzione unitaria
3. Valga la condizione per la rappresentazione di Kraus:

$$\sum_{k=0}^{M-1} E_k^\dagger E_k = \mathbb{I}_N \quad (3.55)$$

Le prime due condizioni sono soddisfatte scegliendo gli operatori di Kraus $\{E_k\}$ dati da⁹:

$$E_0 = \mathbb{I} + \frac{1}{\hbar}(-iH + K)dt \quad (3.56)$$

$$E_k = \sqrt{dt} L_k \quad (3.57)$$

dove H e K sono operatori hermitiani, mentre L_k sono generici operatori, detti **operatori di Lindblad**.

Imponendo la (3.55):

$$\begin{aligned} & \left[\mathbb{I} + \frac{1}{\hbar}(iH + K)dt \right] \left[\mathbb{I} + \frac{1}{\hbar}(-iH + K)dt \right] + \sum_{k=1}^{M-1} L_k^\dagger L_k = \mathbb{I} \\ & \underset{(a)}{\approx} \frac{2}{\hbar} K dt + \sum_{k=1}^{M-1} L_k^\dagger L_k dt + O([dt]^2) = 0 \Rightarrow K = -\frac{\hbar}{2} \sum_{k=1}^{M-1} L_k^\dagger L_k \end{aligned} \quad (3.58)$$

dove in (a) espandiamo al primo ordine.

Sostituendo (3.58) e (3.57) nell'espressione iniziale (3.54) giungiamo a:

$$\rho(t + dt) = \rho(t) - \frac{i}{\hbar}[H, \rho(t)]dt + \sum_{k=1}^{M-1} \left(L_k \rho(t) L_k^\dagger - \frac{1}{2} L_k^\dagger L_k \rho(t) - \frac{1}{2} \rho(t) L_k^\dagger L_k \right) dt + O([dt]^2) \quad (3.59)$$

Riconosciamo nel primo termine l'evoluzione data dall'equazione di Heisenberg, come desiderato.

Nell'approssimazione di Markov l'evoluzione è data da un'equazione differenziale del primo ordine:

$$\dot{\rho}(t + dt) = \rho(t) + \dot{\rho}(t)dt + O([dt]^2) \quad (3.60)$$

Confrontando (3.60) e (3.59) giungiamo allora alla forma principale della master equation, detta anche equazione di *Gorini-Kossakowski-Sudarshan-Lindblad* (GKSL):

$$\dot{\rho} = -\frac{i}{\hbar}[H, \rho] + \sum \left(L_k \rho L_k^\dagger - \frac{1}{2} L_k^\dagger L_k \rho - \frac{1}{2} \rho L_k^\dagger L_k \right)$$

⁹Una derivazione completa della master equation è presentata in [15]. Daremo qui solo qualche cenno.

Un'altra forma (equivalente) per la stessa equazione è data da:

$$\dot{\rho} = -\frac{i}{\hbar}[H, \rho] + \frac{1}{2\hbar^2} \sum_{i,j=1}^{N^2-1} A_{ij} \{[\sigma_i, \rho\sigma_j^\dagger] + [\sigma_i\rho, \sigma_j^\dagger]\}$$

dove A è una matrice positiva, e $\{\sigma_i\}$ è la base delle matrici Hermitiane data da:

$$\sigma_0 = \frac{\mathbb{I}}{\sqrt{N}} \quad \text{Tr}(\sigma_i) = 0 \quad i > 0 \quad \text{Tr}(\sigma_i^\dagger \sigma_j) = \delta_{ij}$$

3.6 Crittografia e meccanica quantistica

(Lezione 9 ● del
27/3/2019)

Con **crittografia** si intende l'insieme di tecniche necessarie a rendere una comunicazione *sicura*, ossia inaccessibile a chiunque non sia il destinatario inteso.

Nel corso della storia si sono susseguiti metodi via via più sofisticati per ottenere ciò. Per esempio, i romani utilizzavano una striscia di cuoio da attorcigliare attorno ad un'asta di legno. Il messaggio confidenziale era quindi scritto *in verticale* sul nastro, che poi veniva srotolato, risultando così in una sequenza disordinata di caratteri. In questo modo, solo disponendo di un bastone con lo stesso diametro era possibile replicare la configurazione iniziale, e quindi leggere il messaggio. Tale asta è quindi detta **chiave** (e in questo caso è proprio un oggetto fisico), in quanto consente a chi la possiede di *accedere* al contenuto del messaggio cifrato senza difficoltà.

In crittografia si suppone che la *chiave* sia condivisa tra mittente e destinatario, e che nessun altro ne sia in possesso. Il meccanismo di cifratura, tuttavia, può essere di dominio pubblico: un cifrario si dice *sicuro* se, pur conoscendone perfettamente il meccanismo, non è possibile decifrarne i messaggi senza conoscere la chiave.

Parallelamente alla crittografia, la **crittoanalisi** si occupa di “forzare” i cifrari, ossia svelarne i contenuti cifrati senza essere in possesso delle informazioni necessarie per farlo (ossia della chiave).

Per esempio, un *cifrario* molto comune in passato è la cosiddetta **sostituzione monoalfabetica**, che consiste nel sostituire ogni carattere del messaggio con un altro. La chiave è allora data da una tabella che mostra le sostituzioni effettuate, permettendo quindi di invertirle e decifrare il messaggio.

In ogni lingua, tuttavia, certe lettere compaiono più frequentemente di altre. Da un'analisi statistica delle frequenze del messaggio cifrato (che supponiamo sufficientemente lungo) è perciò possibile riconoscere una certa parte delle sostituzioni, e ricavare *per tentativi* le altre - ricostruendo così la chiave. Perciò la *sostituzione monoalfabetica* non è, al giorno d'oggi, un cifrario sicuro.

3.6.1 Crittografia classica

Possiamo schematizzare il processo crittografico come una **trasformazione** \hat{E}_K

Definizione di
cifrario

(*encryption*) che mappa il messaggio da cifrare P (*plaintext*) nel messaggio cifrato C (*ciphertext*), e che dipende da un parametro K detto chiave (*key*):

$$P \mapsto \hat{E}_K(P) = C$$

Tale trasformazione deve essere invertibile: possedendo K si può decifrare C e riottenere P . Esiste quindi una mappa inversa \hat{D}_K (*decryption*):

$$C \mapsto \hat{D}_K(C) = P$$

Consideriamo uno schema specifico per \hat{E}_K (**cifrario di Vernam**). Partiamo associando ad ogni lettera un codice numerico:

Cifrario di Vernam

$$A \mapsto 00, B \mapsto 01, C \mapsto 02, \dots, Z \mapsto 26$$

Possiamo allora prendere un messaggio P ="Shaken not stirred", e una chiave K composta di numeri casuali, e combinarle *un numero alla volta* per formare il messaggio criptato C :

$$\hat{E}_K(P_i) = (P_i + K_i)_{\text{mod } 26} = C_i$$

La trasformazione inversa è data da:

$$\hat{D}_K(C_i) = (C_i - K_i)_{\text{mod } 26} = P_i$$

Tale schema può essere implementato facilmente in un computer convertendo P e K in una serie di *bit* (per esempio tramite il codice ASCII). Detti allora $\{p_1, p_2, \dots, p_N\}$ i bit del messaggio e $\{k_1, \dots, k_N\}$ quelli della chiave (con $k_i \in \{0, 1\}$ e $p_i \in \{0, 1\}$) si ha:

$$c_i = \hat{E}_K(p_i) = p_i \oplus k_i \quad \forall i = 1, 2, \dots, N$$

dove \oplus indica uno XOR, che equivale ad una somma in modulo 2. Per decifrare si segue il medesimo schema:

$$p_i = \hat{D}_K(c_i) = c_i \oplus k_i \quad \forall i = 1, 2, \dots, N$$

Matematicamente, si dimostra che tale processo genera messaggi sicuri, nel senso di *completamente indecifrabili*, solo se K ha le seguenti proprietà:

- K è completamente casuale
- K è usata una volta sola (*One Time Pad*)

Sicurezza del cifrario di Vernam. Se la chiave K è casuale ed è *lunga quanto il messaggio* P , anche il messaggio cifrato C sarà, effettivamente, composto di caratteri *casuali*. Procedendo per *forza bruta*, tentando tutte le chiavi K possibili, è possibile decifrare da C una qualsiasi stringa di testo: il messaggio cercato è perciò *nascosto* tra una miriade di "messaggi spuri" che, non conoscendo K , risultano

ugualmente validi. Per esempio, il messaggio originale potrebbe essere il seguente:

P: In the name of the moon I will punish you!
 K: So meb odyo nc eto ldme t hewo rldisg onn
 C: Ab flf bdkb bh xas xrar B dmhz gfqqkn mbh

Una scelta casuale di K può produrre però P completamente diversi, ma lo stesso “sensati”:

C: Abf lfb dksbh xasxrar bdmhz Gfq qknm bh
 K: Txh nrh fwykd ssfxgpt Bhmrv irw ugwi iq
 P: Hey you youre finally awake You were tr

Se la chiave K è scelta casualmente^a non vi sono scelte “più probabili” di altre, e quindi non vi è alcun modo per ricavare il messaggio P corretto, e nemmeno quelli “più probabili”.

Risulta però di estrema importanza **non riutilizzare** la stessa chiave K . In tal caso, infatti, è possibile immediatamente ricavare informazioni sui messaggi cifrati. Siano a_i e b_i i caratteri di due messaggi (con alfabeto di N caratteri), e k_i quelli della chiave usata per cifrare entrambi. Si ha allora:

$$c_i = a_i + k_i \mod N; \quad d_i = b_i + k_i \mod N$$

Sottraendo membro a membro:

$$c_i - d_i \mod N = a_i + k_i - b_i - k_i \mod N = a_i - b_i \mod N \quad (3.61)$$

Si trova così una relazione diretta tra i due messaggi. Si può ora procedere *per tentativi*, cercando di indovinare singole parole di A e B compatibili con tale relazione, in un processo detto *crib drag*. L’idea di base è che vi sono parole comuni facilmente contenute in ogni messaggio (es. *the* in inglese). Per esempio, ipotizzando che $\{b_i\}_{i=1,\dots,5} = \text{“Hello”}$, possiamo ricavare da (3.61) le prime 5 lettere di A . Se queste ultime “hanno senso” (per esempio risultano in “Helpm”) allora si ha conferma del tentativo appena fatto. Possiamo ora *espandere* quanto ricavato per A (es. “Help me please”) e di conseguenza ottenere nuovi caratteri di B , che - ammesso abbiano senso - possono portare a loro volta a nuove ipotesi, e così via. Poiché un computer può tentare migliaia di parole comuni in ogni possibile posizione dei messaggi A e B , la ripetizione di una chiave vanifica la sicurezza del cifrario di Vernam.

^aQuindi non come nell’esempio appena visto...

La sicurezza del cifrario di Vernam richiede quindi di poter condividere una chiave K che ha la stessa lunghezza del testo, e che deve pervenire unicamente al destinatario. Si presuppone quindi di avere un canale sicuro per fare questo passaggio: peccato che questo fosse proprio il problema che volevamo risolvere in partenza!

Un tale schema, detto **a cifratura simmetrica**, pone quindi il problema della **distribuzione delle chiavi**, che non è di facile risoluzione. Lo scenario peggiore si

ha se qualcuno riesce a *intercettare* la chiave senza che né mittente né destinatario se ne accorgano, poiché in tal caso il canale ritenuto sicuro è invece compromesso.

3.6.2 Quantum Key Distribution (QKD)

Il problema di distribuzione delle chiavi trova soluzioni efficaci tramite tecniche di Informazione Quantistica.

I due protocolli principali sono:

- **BB84**, che fa uso del principio di sovrapposizione e del no-cloning theorem
- **E91**, per cui invece si usa l'entanglement

In questa sezione ci concentreremo sul primo.

3.6.3 Il protocollo BB84

Supponiamo che **Alice** (A) voglia trasmettere a **Bob** (B) una chiave da utilizzare per future comunicazioni crittografate, facendo sì che nessuno possa intercettarla.

Alice può generare qubit nelle basi B_z (autoket di $\hat{\sigma}_z$) o B_x (autoket di $\hat{\sigma}_x$). Scegliendo come base computazionale gli autoket B_z avremo:

$$B_z = \{|0\rangle, |1\rangle\} \quad B_x = \left\{ \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \right\}$$

Il protocollo prosegue perciò nel seguente modo:

1. Alice genera una stringa K_1 di *bit classici* completamente casuali¹⁰:

$$K_1 = 0011100000001101$$

2. Alice codifica la stringa K_1 associando ad ogni 0 un qubit nello stato $|0\rangle$ o $|+\rangle$ in modo equiprobabile (metà degli 0 diverranno $|0\rangle$, metà diverranno $|+\rangle$). Fa lo stesso per gli 1, codificandoli o con $|1\rangle$ o con $|-\rangle$. Ottiene quindi una sequenza di qubit K_2 :

$$K_2 = |0\rangle |+\rangle |1\rangle |1\rangle |-\rangle |0\rangle |+\rangle |+\rangle \dots$$

3. Alice trasmette K_2 a Bob mediante un canale quantistico
4. Bob misura ciascun qubit nella base B_x o B_z in modo casuale, e registra i risultati in una stringa di bit classici K_3 . Avremo allora due possibilità:
 - Se Bob sceglie di misurare un qubit nella stessa base che Alice aveva usato per inviarlo, allora riottiene lo stesso bit classico che era stato usato in partenza (con una certa efficienza, limitata dalla presenza di rumore nel canale di trasmissione, e dalle efficienze degli apparati di produzione/rilevazione degli stati quantistici).

¹⁰^Ciò si può fare con un generatore di numeri casuali classico, o misurando sovrapposizioni quantistiche equiprobabili.

- Se invece Bob usa una base diversa da Alice otterrà un risultato *casuale*.

Per esempio, se Bob sceglie di misurare con $B_z, B_x, B_z, B_x, \dots$ potrebbe ottenere per K_3 :

$$K_3 = 001010 \dots$$

5. A questo punto Alice e Bob condividono le **basi** utilizzate rispettivamente per generare o misurare gli stati dei qubit:

Alice : $B_z, B_x, B_z, B_z, B_x, B_z, B_x, B_x$
 Bob : $B_z, B_x, B_z, B_x, B_x, B_x, B_z, B_x$

In questo modo Alice e Bob sanno in quali casi il bit classico è sicuramente pervenuto all'altro. Basta allora tenere tali bit “correttamente comunicati” e comporre con essi la chiave:

$$K = 00100 \dots$$

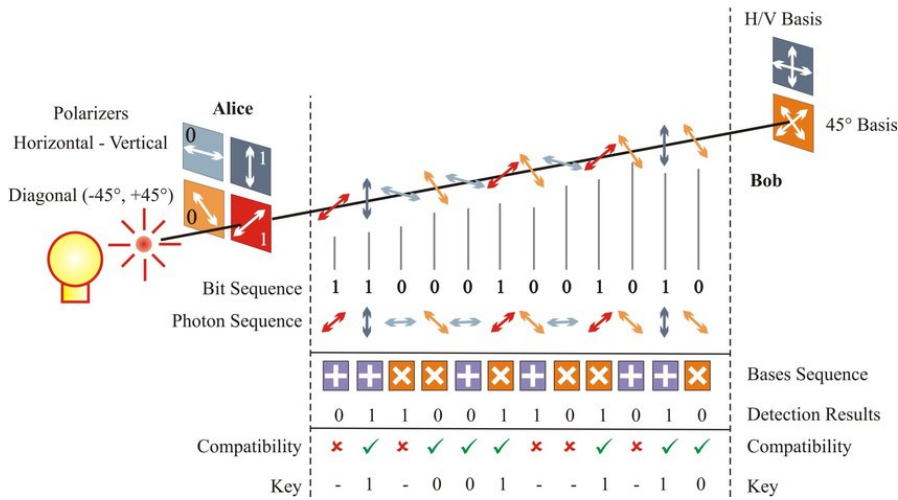


Figura (3.12) – Esempio di distribuzione quantistica di una chiave tramite il protocollo BB84. Alice parte da una sequenza di bit iniziale e la converte in una successione di qubit in 4 stati possibili. Solo quando Bob misura il qubit nella stessa base usata da Alice per generare gli stati allora il bit classico iniziale è stato correttamente ricevuto.

Immaginiamo che una terza persona, **Eve**, cerchi di intercettare il messaggio. Per farlo **non può** duplicarlo (no-cloning theorem), e quindi l'unico modo è fare direttamente delle misure, secondo una sequenza di basi nuovamente arbitraria. Poiché è molto difficile che Eve faccia al 100% le stesse scelte di Bob, quando A e B si scambiano le informazioni sulle basi utilizzate, Eve potrà al più recuperare parte della chiave.

Tuttavia, la misura di Eve condiziona i risultati delle misure di Bob: se Eve sbaglia

la base, può far collassare il qubit in uno stato *diverso* da quello generato all'inizio da Alice. Ciò produce differenze nella chiave K condivisa tra Alice e Bob, che possono essere rivelate tramite tecniche di teoria dell'informazione classica. In altre parole, non è possibile per Eve intercettare le comunicazioni senza alcuna interferenza.

Esaminiamo alcune tecniche che possono essere attuate da Alice e Bob per migliorare la sicurezza del canale (ed eventualmente determinare la presenza di ascoltatori):

1. Determinare l'**errore** R : Alice e Bob condividono parte delle loro versioni di K , calcolando la percentuale di bit differenti. Se è molto alta, ciò può essere il risultato di un canale di comunicazione fallace, o della presenza di qualcuno che sta cercando di intercettare il messaggio. In tal caso conviene effettuare i dovuti controlli e rigenerare K .
2. **Information reconciliation**: se il rate di errore R è sufficientemente basso, è possibile correggere i bit errati di K senza condividere la chiave stessa. Un modo per farlo è dato dal suddividere K in stringhe di lunghezza l , tali che ciascuna sottostringa contenga probabilmente al più un solo errore (ossia tali che $Rl \ll 1$). A tal punto Alice e Bob calcolano la **parità** di ogni stringa:

$$P = b_1 \oplus b_2 \oplus \cdots \oplus b_l$$

Le parità possono essere condivise senza problemi, poiché non permettono di ottenere informazione sui singoli bit b_i . Se le due P coincidono, sapremo che (molto probabilmente) tale stringa è stata correttamente ricevuta. Altrimenti, probabilmente vi è un errore (*bit-flip*). In tal caso si procede per bisezione, dividendo la stringa a metà e confrontando le parità delle due nuove parti:

$$P_1 = b_1 \oplus b_2 \oplus \cdots \oplus b_{(l-1)/2}; \quad P_2 = b_{(l-1)/2+1} \oplus b_{(l-1)/2+2} \oplus \cdots \oplus b_{l-1}$$

Se Alice e Bob cancellano l'ultimo bit prima della bisezione (b_l - che infatti non compare nel calcolo di P_2), allora vanificano ogni informazione ottenuta da terzi grazie alla precedente parità P condivisa.

Proseguendo per bisezione è perciò possibile (seppur con una certa perdita di bit) trovare esattamente i *bit errati* e correggerli.

3. **Privacy Amplification**. Dalla misura di R è possibile stimare il massimo numero di bit k che Eve può conoscere. Alice e Bob scelgono allora un $s \in \mathbb{N}$, e costruiscono $n - k - s$ set di bit di K (presi a caso, secondo uno schema comune), dove n è la lunghezza di K . La parità dei qubit di ciascun set forma allora un bit di una nuova chiave. Si può dimostrare (*Leftover hash theorem*) che tale processo fa sì che Eve non possa ricavare alcun bit - dato che per farlo dovrebbe possedere una certa informazione *su ogni bit di* K . Inoltre, maggiore è s (che porta a qualche bit scartato) minore è l'informazione accessibile ad Eve (che decade come $O(2^{-s})$).

D'altro canto, Eve potrebbe tentare vari approcci per cercare di capire la chiave:

- **Intercept & Resend:** Eve misura il qubit in arrivo in una base generica, e poi lo rinvia. L'effetto di una misura proiettiva, tuttavia, perturba molto lo stato iniziale, e quindi un attacco del genere è facilmente individuabile da Alice e Bob
- **Translucent Attacks:** Eve fa interagire il qubit inviato da Alice con un altro qubit “ausiliario” e rinvia il qubit iniziale senza averlo misurato. Solo quando Alice e Bob condividono le basi allora Eve misura il qubit ausiliario. Si trova che, in ogni caso, ciò o perturba lo stato iniziale, o non consente di ricavare informazione utile dalle correlazioni. Perciò, almeno finora, non si è trovato nessun modo efficace per portare a termine un attacco del genere.
- **Attacchi collettivi:** Eve manipola interi blocchi di qubit in una volta sola.

In ogni caso, è possibile rendere l'informazione che Eve ha su K arbitrariamente piccola, indipendentemente dalla strategia messa in atto da Eve (Nielsen e Chuang, 2000).

3.7 Dense coding

Tramite un opportuno canale quantistico è possibile trasmettere 2 bit classici inviando un solo qubit, supponendo che mittente e destinatario condividano due qubit in uno stato di Bell (entangled).

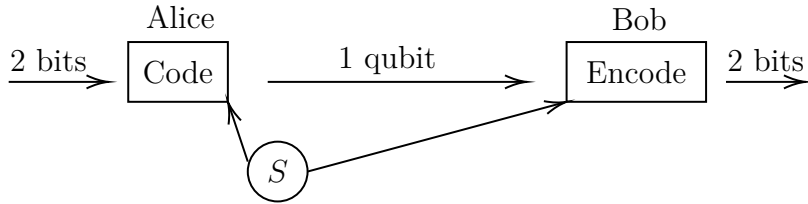


Figura (3.13) – Schema del protocollo di Dense Coding

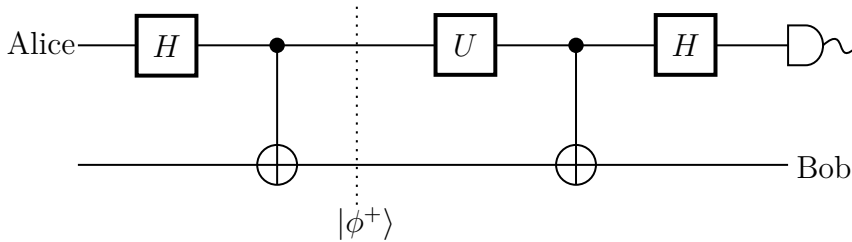


Figura (3.14) – Rappresentazione del canale come azione di gate quantistici

Il protocollo consiste nei seguenti passi:

1. **Preparazione.** Si crea uno stato entangled:

$$|\phi^+\rangle = \text{CNOT}(H \otimes \mathbb{I}) |00\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

I due qubit vengono poi lasciati uno ad A e uno a B .

2. **Messaggio.** Alice sceglie quali 2 bit classici inviare, e a seconda della scelta compie una certa operazione U sul suo qubit.

Bit	U
00	\mathbb{I}
01	$\hat{\sigma}_x$
11	$\hat{\sigma}_y$
10	$\hat{\sigma}_z$

Tabella (3.1) – Operazioni U svolte da Alice a seconda della combinazione dei 2 bit classici che vuole inviare a Bob

Avremo quindi 4 possibili stati finali per $|\phi^+\rangle$:

$$\begin{aligned}
(\mathbb{I} \otimes \mathbb{I}) |\phi^+\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = |\phi^+\rangle \\
(\hat{\sigma}_x \otimes \mathbb{I}) |\phi^+\rangle &= \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle) = |\psi^+\rangle \\
(\hat{\sigma}_y \otimes \mathbb{I}) |\phi^+\rangle &= \frac{1}{\sqrt{2}}(|10\rangle - |01\rangle) = |\psi^-\rangle \\
(\hat{\sigma}_z \otimes \mathbb{I}) |\phi^+\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) = |\phi^-\rangle
\end{aligned}$$

3. **Comunicazione.** Alice manda il suo qubit a Bob attraverso un opportuno canale quantistico.
4. **Misura.** Bob esegue l'operazione inversa di quella al punto 1 sui due qubit che ora possiede, e il cui è stato è un certo $|\psi\rangle$:

$$[\text{CNOT}(H \otimes \mathbb{I})]^{-1} |\psi\rangle = (H \otimes \mathbb{I}) \text{CNOT} B$$

dato che sia CNOT che $H \otimes \mathbb{I}$ sono hermitiane e unitarie (e quindi ciascuna è pari all'inversa di se stessa).

Ricordando che la CNOT inverte il secondo qubit solo se il primo è 1 (ossia mappa $|11\rangle \leftrightarrow |10\rangle$, e lascia invariati gli altri stati), procediamo calcolando l'azione di B sui quattro possibili stati:

$$\begin{aligned}
|\phi^+\rangle &\xrightarrow{\text{CNOT}} \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) = \frac{1}{\sqrt{2}}|+\rangle|0\rangle \xrightarrow{(H \otimes \mathbb{I})} \frac{1}{\sqrt{2}}|0\rangle|0\rangle \\
|\psi^+\rangle &\xrightarrow{\text{CNOT}} \frac{1}{\sqrt{2}}(|11\rangle + |01\rangle) = \frac{1}{\sqrt{2}}|+\rangle|1\rangle \xrightarrow{(H \otimes \mathbb{I})} \frac{1}{\sqrt{2}}|0\rangle|1\rangle \\
|\psi^-\rangle &\xrightarrow{\text{CNOT}} \frac{1}{\sqrt{2}}(|11\rangle - |01\rangle) = \frac{1}{\sqrt{2}}|-\rangle|1\rangle \xrightarrow{(H \otimes \mathbb{I})} \frac{1}{\sqrt{2}}|1\rangle|1\rangle \\
|\phi^-\rangle &\xrightarrow{\text{CNOT}} \frac{1}{\sqrt{2}}(|00\rangle - |10\rangle) = \frac{1}{\sqrt{2}}|-\rangle|0\rangle \xrightarrow{(H \otimes \mathbb{I})} \frac{1}{\sqrt{2}}|1\rangle|0\rangle
\end{aligned}$$

Misurando gli stati dei due qubit dopo l'applicazione di B , Bob può perciò ricavare i 2 bit classici scelti da Alice al punto 2. Si è così realizzata la trasmissione di 2 bit al prezzo di un solo *qubit*.

Correlazioni quantistiche

Nelle precedenti sezioni abbiamo notato come i canali quantistici offrano possibilità in apparenza più ampie dei corrispettivi classici, permettendo per esempio una crittografia sicura o di trasmettere una maggiore *densità* di informazioni. Come sarà chiaro nei prossimi paragrafi, tali opportunità sono consentite dalla presenza di *correlazioni quantistiche non-locali* che non possono essere realizzate da canali puramente classici. Partiremo quindi esaminando i comportamenti non-locali della MQ (confermati dalla violazione delle disuguaglianze di Bell), per poi cercare metriche per *quantificare* tali correlazioni, mostrando infine che senza di esse risulta del tutto impossibile ottenere risultati simili a quelli esaminati nelle precedenti sezioni.

4.1 Bell Inequalities

Nel 1964 John Stewart Bell formulò uno dei più importanti teoremi sui fondamenti della fisica, dimostrando che ogni teoria che si proponga di spiegare determinate misure sperimentali (come la Meccanica Quantistica) debba necessariamente includere comportamenti **non-locali**. Citando la risposta di Bell stesso in un'intervista del 1988 alla domanda “Cosa significa **località**?”:

(Lezione 10 ● del
28/3/2019)

“Si tratta dell’idea che ogni azione abbia conseguenze solo nelle immediate vicinanze, e che ogni conseguenza più lontana sia più debole, e arrivi a destinazione solo dopo un intervallo compatibile con il limite della velocità della luce. La località è l’idea che gli effetti si propaghino *in maniera continua*, che non *saltino* improvvisamente da un punto all’altro.”

Località

Una definizione rigorosa di località[9][10].

Una qualsiasi definizione di località è necessariamente ancorata ad una scelta di “variabili locali” che si presuppone rispettino la condizione qualitativa appena specificata. Per esempio, il fatto che alla morte della regina di Inghilterra il principe del Galles divenga immediatamente re non risulta in un nessun fenomeno non-locale, dato che una *convenzione* può “trasmettersi” a velocità arbitrarie. Analogamente,

nella teoria classica dell'elettromagnetismo non sorge alcuna preoccupazione nell'affermare che il potenziale scalare si propaghi con velocità infinita - poiché essendo una grandezza dipendente dalla scelta del *gauge* non è considerata come “fisica”, ma solo come un utile ausilio per i conti.

Per questo, Bell considera nella sua trattazione della località solo una particolare categoria di **teorie fisiche**, dette a “*beable locali*”.

Esplicitamente, una teoria fisica T è un qualsiasi insieme di **leggi** che consentono di prevedere, a partire da un set di **informazioni** sul mondo fisico, le **probabilità** di eventi futuri. Se X è una descrizione completa di un sistema, e A è un possibile evento, la teoria T permette allora di calcolare:

$$X \mapsto P(A|X)$$

Se T è “ben specificata”, deve essere chiaro quali siano gli elementi “che possono essere considerati reali” (detti *beable*) e che sono oggetto delle leggi, indipendentemente da ogni osservazione. Tali *beable* devono includere, necessariamente, le impostazioni degli apparati sperimentali, e le letture degli strumenti. In altre parole, per poter anche solo introdurre il concetto di località, è necessario che una teoria *assuma* quali grandezze “vadano prese seriamente”, ossia abbiano un “ruolo privilegiato” rispetto ad eventuali artifici matematici.

Un particolare *beable* è detto **locale** se corrisponde ad una regione dello spazio fisico (es. il momento di una particella), e **non-locale** altrimenti (es. la funzione d'onda in MQ, che appartiene ad uno spazio astratto molto più ampio di quello “fisico”). Ha senso quindi parlare delle *beable locali* contenute in una regione R dello spaziotempo, mentre *beable non-locali* non sono “ancorate” a specifiche posizioni.

Consideriamo allora una T che contenga solo *beable locali*, e ci chiediamo se possa essere considerata **locale** (o *localmente causale*). Facendo riferimento alla figura 4.1:

1. Consideriamo le *beable locali* b_1 e b_2 (es. setup di apparati ed esiti di misure) contenute rispettivamente nelle regioni 1 e 2 separate da un intervallo di tipo spazio.
2. Indichiamo con 3 una regione, appartenente al passato di 1, che separi completamente 1 dal passato di 2, e contenga tutte le *beable* B_3 necessarie affinché la teoria T possa predire le probabilità delle *beable* in 1.
3. Allora T si dice **locale** se e solo se specificare le *beable* in 2 non modifica le predizioni che si possono già dare per 1, ossia se:

$$P(b_1|B_3, b_2) = P(b_1|B_3) \quad \forall b_2 \quad (4.1)$$

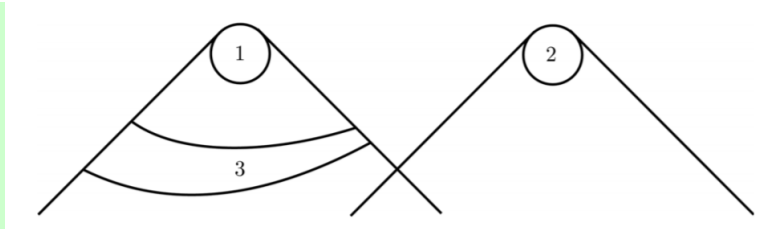


Figura (4.1) – In una teoria locale, l'informazione completa contenuta nella regione 3 del diagramma di Minkowski è sufficiente a determinare le probabilità degli eventi nella regione 1, indipendentemente da quanto accade nella regione 2.

Tale definizione permette, grazie al ruolo di B_3 , di distinguere tra i casi di semplici *correlazioni* e i *sistemi* in cui si ha un effettivo comportamento non-locale. Per esempio, sia b_1 l'evento di *avvenuta cottura di un uovo in una pentola* e b_2 lo *squillo di un timer*. Se le informazioni B_3 (es. la temperatura dell'acqua, lo stato dell'uovo) consentono di determinare che dopo pochi istanti avverrà b_1 , il realizzarsi di b_2 “non modifica la predizione”, e la teoria che consente tali previsioni è locale - nonostante b_1 e b_2 siano ovviamente fortemente correlati.

Se ciò non avviene - per esempio se per qualche motivo l'attivarsi del timer “istantaneamente cuoce l'uovo”, dovremmo per forza dedurre che vi è stato un “influsso a distanza” del timer sulla cottura dell'uovo. Notiamo che per consentire tale conclusione, 3 deve rispettare entrambe le ipotesi specificate, ossia essere *completa* e *fungere da barriera*. Se B_3 non contiene tutte le informazioni necessarie per predire b_1 , il realizzarsi di b_2 può dare alcune delle informazioni mancanti. Nell'esempio, misurare la temperatura dell'acqua in B_3 ma non lo stato dell'uovo non consente di determinare che in b_1 sarà cotto, e quindi b_2 “ha un influsso istantaneo” sulle previsioni - ma questo è dovuto all'ignoranza dello sperimentatore, e non ad un'effettiva non-località.

Analogamente, se B_3 non scherma 1 dal passato di 2, allora possono esservi eventi comuni al passato di 1 e 2 che non sono contenuti in B_3 e non sono prevedibili dalla sola conoscenza di B_3 , ma che possono creare correlazioni tra 1 e 2, rendendo quindi diverse $P(b_1|B_3, b_2)$ e $P(b_1|B_3)$ (ma solo in teorie stocastiche).

Si può estendere il criterio di località anche a teorie T che contengono *beable non-locali*, come la MQ. In questo caso, la condizione (4.1) diviene solo necessaria per la località, e le *beable* in B_3 comprendono i valori che la funzione d'onda assume su una “famiglia di superfici di Cauchy” in B_3 (ossia sui luoghi geometrici dello spaziotempo corrispondenti a determinati “istanti”).

Concentriamoci ora sulla definizione matematica di località in un preciso sistema di interesse[8].

Consideriamo una sorgente S che produce coppie di qubit entangled (per esempio coppie di particelle con *spin entangled* nell'esperimento di Stern-Gerlach, o fotoni “duplicati” da un cristallo non-lineare), che sono fornite a due osservatori, Alice (A) e Bob (B), spazialmente separati (figura 4.2).

Ciascuno dei due ha a disposizione un set di *possibili misure* (permesse dagli apparati di cui è in possesso) che può svolgere sul proprio qubit. Indichiamo con

X l'osservabile scelta da Alice, e con Y quella di Bob. A seguito della misura, A otterrà un esito $a \in \sigma(X)$, e B un esito $b \in \sigma(Y)$.

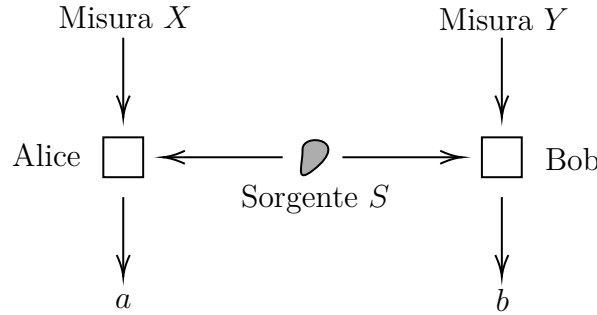


Figura (4.2) – Schema dell'esperimento di Bell per esaminare la non-località della MQ. Una sorgente S di coppie di particelle entangled invia una particella ad Alice e una a Bob, che eseguono rispettivamente una misura X e Y sulla loro particella, ottenendo come risultati a e b

In generale, ripetendo l'esperimento (partendo sempre dallo stesso stato iniziale), Alice potrà ottenere un risultato $a' \neq a$, e Bob un $b' \neq b$, a seconda della natura (potenzialmente diversa) delle misurazioni X e Y effettuate, o di incertezze intrinseche al sistema. In ogni caso, abbiamo a disposizione una teoria T che permette di determinare la probabilità $P(a, b|X, Y)$ di ottenere come esiti a e b per una specifica scelta di X e Y .

1. Non è detto che tale probabilità sia fattorizzabile:

$$P(a, b|X, Y) \neq P(a|X)P(b|Y)$$

poiché potrebbero essere presenti **correlazioni** tra esiti di misure diverse (per esempio b potrebbe dipendere dal risultato a di Alice).

La sola presenza di correlazioni, tuttavia, non significa che vi siano fenomeni non-locali in gioco: le correlazioni potrebbero essere dovute all'origine comune delle particelle misurate.

2. Supponiamo che la teoria T sia **locale**, e che le misurazioni di A e B avvengano in regioni 1 e 2 separate da un intervallo di tipo spazio (figura 4.3).

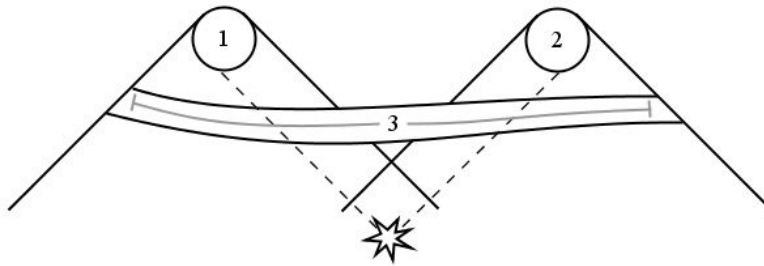


Figura (4.3) – Diagramma di Minkowski per lo scenario di Bell. Gli apparati di Alice e Bob si trovano nelle regioni 1 e 2, e lo stato del sistema è definito nella regione 3.

L'unico modo per spiegare le correlazioni tra a e b è allora dato dall'esistenza di parametri λ che *definiscono lo stato iniziale* (nella regione 3) e che permettano di calcolare le probabilità $P(a, b|X, Y, \lambda)$. In altre parole, in una teoria locale devono esistere “parametri locali” che sono già definiti quando le particelle sono separate, prima della misura, e che ne descrivono completamente il comportamento - spiegando di conseguenza ogni correlazione.

3. I parametri λ possono anche non essere misurabili (“variabili nascoste”), ma possiamo osservarne l'effetto nelle correlazioni misurate. Poiché le λ contengono già tali correlazioni, la probabilità condizionata diviene ora separabile:

$$P(a, b|X, Y, \lambda) = P(a|X, \lambda) \cdot P(b|Y, \lambda). \quad (4.2)$$

Derivazione della fattorizzabilità. Le osservabili scelte e i loro esiti a, b, X, Y sono infatti *beable* delle regioni 1 e 2. Se T è locale, allora possiamo individuare una regione 3 tali che le *beable* λ in esse contenute permettano di calcolare le probabilità degli eventi in 1 e in 2, e che separi ciascuna regione dal passato dell'altra. Si ha allora che ogni esito a, b dipende esclusivamente da λ e dal relativo setup sperimentale (X, Y) che si trova nella stessa regione:

$$\begin{aligned} P(a|X, Y, b, \lambda) &= P(a|X, \lambda) \\ P(b|X, Y, a, \lambda) &= P(b|Y, \lambda) \end{aligned}$$

In particolare, notando che anche $P(b|X, Y, \lambda) = P(b|Y, \lambda)$ e che, per la definizione di probabilità condizionata, vale:

$$P(a, b|X, Y, \lambda) = P(a|X, Y, b|\lambda) \cdot P(b|X, Y, \lambda)$$

troviamo:

$$P(a, b|X, Y, \lambda) = P(a|X, \lambda) \cdot P(b|Y, \lambda)$$

In realtà perché ciò valga è richiesta anche l'ipotesi che la scelta di X e Y non modifichi il valore di λ (la cosiddetta “free choice” o “**no conspiracy**”). Ciò è sensato, poiché la regione 3 può contenere un numero enorme di *beable*, e quelle che influenzano a e b possono essere scelte distinte da quelle che influenzano X e Y . Così facendo, una diversa scelta di X e Y comporta una variazione della “parte di λ ” che non ha alcun effetto su a e b , e quindi la fattorizzazione vale ancora.

Per esempio, consideriamo un *trial* per un nuovo farmaco, in cui misuriamo l'effetto del farmaco (X_1) o di un placebo (X_2) su pazienti diversi. L'esito (a) del test dipende da delle condizioni λ (es. salute del paziente, sensibilità al principio attivo, etc.) che si presume siano completamente diverse da quelle usate per scegliere se somministrare il farmaco o meno (es. lanciare una moneta). Perciò si può dire che X_1 e X_2 non influenzano “la parte di λ ” rilevante per l'esito a .

Rigettare tale ipotesi significa affermare l'esistenza di una qualche "cospirazione" che spinge gli sperimentatori ad assegnare il farmaco solo ai pazienti meno recettivi, cosa che è del tutto irragionevole, e mina i fondamenti stessi della possibilità di condurre esperimenti.

Nota: dato che la regione 3 interseca i passati di entrambe le regioni 1 e 2, nella definizione *generalizzata* di località possiamo considerare tra le informazioni contenute in 3 anche una funzione d'onda *non separabile* (cioè entangled) che descriva lo stato delle due particelle.

4. In generale, λ può cambiare nel tempo (es. da un esperimento all'altro). Possiamo allora pensare a λ come una variabile casuale, che assume valori $\lambda \in \Lambda$ e si distribuisce secondo la densità di probabilità $q(\lambda)$. Sperimentalmente ci interessano allora le probabilità *marginalizzate*, ossia:

$$P(a, b|X, Y) = \int_{\Lambda} d\lambda q(\lambda) p(a|X, \lambda) p(b|Y, \lambda) \quad (4.3)$$

L'introduzione di variabili nascoste permette di spiegare (per esempio) come un sistema intrinsecamente deterministico *appaia* indeterministico, non avendo accesso a *tutte* le informazioni che ne descrivono lo stato. Per esempio, potremmo pensare ad un modello come il seguente.

Consideriamo come base computazionale $\{|\uparrow\rangle_z, |\downarrow\rangle_z\}$ gli autostati di $\hat{\sigma}_z$ (spin lungo \hat{z} per un fermione). Lo stato $|\psi\rangle$ di un generico qubit è allora dato da:

$$|\psi\rangle = \cos \frac{\theta}{2} |\uparrow\rangle + \sin \frac{\theta}{2} |\downarrow\rangle \quad \theta \in [0, 2\pi)$$

Una misura di $|\psi\rangle$ è indeterminata, nel senso che possiamo calcolare a priori solamente la probabilità di un particolare esito. In particolare, si ottiene $|\uparrow\rangle$ con $p = \cos^2 \theta/2$, e $|\downarrow\rangle$ con probabilità $1 - p$.

Tuttavia, potremmo considerare un modello in cui al qubit è associato un *parametro locale* $\lambda \in [0, 1]$, uniformemente distribuito e non misurabile, che codifica l'esito delle future misure:

- $|\uparrow\rangle$ se $0 \leq \lambda \leq \cos^2 \frac{\theta}{2}$
- $|\downarrow\rangle$ se $\cos^2 \frac{\theta}{2} \leq \lambda \leq 1$

Poiché non sappiamo λ , non possiamo determinare a priori il risultato di una misura di spin, ma possiamo solo darne una **probabilità** - ritroviamo quindi la descrizione data dalla funzione d'onda in MQ, con il postulato di Bohr sulle probabilità.

Nota: il teorema di Bell non riguarda solamente le teorie alle variabili nascoste *deterministiche*, ma offre condizioni che devono essere rispettate da *ogni teoria locale*. In effetti, la condizione di località (4.2) è formulata in termini di sole probabilità.

Teorema 4.1.1. *Le predizioni $P(ab|XY)$ date dalla MQ non ammettono una decomposizione della forma (4.3). Poiché la (4.3) è una condizione necessaria perché una teoria sia considerabile locale, ogni teoria compatibile con le predizioni della MQ (verificate sperimentalmente) è quindi di carattere non-locale.*

Dimostrazione.

Consideriamo uno scenario in cui Alice e Bob hanno ciascuno a disposizione due apparati di misurazione, per cui $X \in \{x_0, x_1\}$ e $Y \in \{y_0, y_1\}$. Supponiamo che l'esito di ogni possibile misura sia binario, ossia $a, b \in \{-1, 1\}$.

Indicheremo con $\langle a_x b_y \rangle$ il valor medio del prodotti degli esiti a, b per una scelta fissata delle misure (x, y) . Per esempio, $\langle a_0 b_0 \rangle$ è il valor medio di $a \cdot b$ quando $X = x_0$ e $Y = y_0$.

Esplicitamente:

$$\langle a_x b_y \rangle \equiv \sum_{a, b \in \{-1, 1\}} a b p(a, b | X, Y) \quad (4.4)$$

L'idea è ora quella di considerare un'arbitraria funzione S delle correlazioni $\langle a_x b_y \rangle$, che ha la caratteristica di essere quella giusta per dimostrare il teorema:

$$S = \langle a_0 b_0 \rangle + \langle a_0 b_1 \rangle + \langle a_1 b_0 \rangle - \langle a_1 b_1 \rangle \quad (4.5)$$

Dimostriamo ora che in una teoria locale (dove quindi vale la fattorizzazione (4.3)) deve essere **$S \leq 2$ (disuguaglianza CHSH)**.

Tuttavia, calcolando S tramite la MQ, e valutando i valor medi per uno stato massimamente entangled (es. uno stato di Bell), troveremo che $S \leq 2$ non vale: e quindi la (4.3) non può essere compatibile con le predizioni della MQ.

*Disuguaglianza
CHSH*

Partiamo quindi supponendo di essere in una teoria locale, per cui vale la (4.3). Inserendola in (4.4) otteniamo:

$$\begin{aligned} \langle a_x b_y \rangle &\stackrel{(4.3)}{=} \sum_{a, b \in \{-1, 1\}} a b \int_{\Lambda} d\lambda q(\lambda) p(a|X, \lambda) p(b|Y, \lambda) = \\ &= \int_{\Lambda} d\lambda q(\lambda) \sum_{a, b \in \{-1, 1\}} a b p(a|X, \lambda) p(b|Y, \lambda) = \\ &= \int_{\Lambda} d\lambda q(\lambda) \underbrace{\left(\sum_{a \in \{-1, 1\}} a p(a|X, \lambda) \right)}_{\langle a_x \rangle_{\lambda}} \underbrace{\left(\sum_{b \in \{-1, 1\}} b p(b|Y, \lambda) \right)}_{\langle b_y \rangle_{\lambda}} = \int_{\Lambda} d\lambda q(\lambda) \langle a_x \rangle_{\lambda} \langle b_y \rangle_{\lambda} \end{aligned}$$

dove $\langle a_x \rangle_{\lambda}$ e $\langle b_y \rangle_{\lambda}$ sono i valori medi degli esiti a e b per una scelta di (x, y, λ) .

Sostituendo tale risultato in (4.5) troviamo:

$$\begin{aligned} S &= \int_{\Lambda} d\lambda q(\lambda) S_{\lambda} \\ S_{\lambda} &\equiv \langle a_0 \rangle_{\lambda} \langle b_0 \rangle_{\lambda} + \langle a_0 \rangle_{\lambda} \langle b_1 \rangle_{\lambda} + \langle a_1 \rangle_{\lambda} \langle b_0 \rangle_{\lambda} - \langle a_1 \rangle_{\lambda} \langle b_1 \rangle_{\lambda} \end{aligned}$$

Poiché a e b possono assumere valori tra $\{-1, +1\}$, i vari valor medi devono essere compresi tra -1 e 1 :

$$\langle a_x \rangle_{\lambda} \in [-1, 1] \quad \langle b_y \rangle_{\lambda} \in [-1, 1]$$

Ma allora si ha:

$$S_\lambda = \langle a_0 \rangle_\lambda [\langle b_0 \rangle_\lambda + \langle b_1 \rangle_\lambda] + \langle a_1 \rangle_\lambda [\langle b_0 \rangle_\lambda - \langle b_1 \rangle_\lambda] \leq |\langle b_0 \rangle_\lambda + \langle b_1 \rangle_\lambda| + |\langle b_0 \rangle_\lambda - \langle b_1 \rangle_\lambda|$$

dato che è possibile massimizzare S_λ scegliendo $\langle a_{0,1} \rangle_\lambda$ in modo che i termini tra parentesi quadre siano entrambi positivi. Precisamente, per la prima parentesi:

$$\langle a_0 \rangle (\langle b_0 \rangle + \langle b_1 \rangle) \leq \max[-1 \cdot (\langle b_0 \rangle + \langle b_1 \rangle); +1 \cdot (\langle b_0 \rangle + \langle b_1 \rangle)] = |\langle b_0 \rangle + \langle b_1 \rangle|$$

e una relazione analoga si ha anche per la seconda.

Siamo allora giunti a:

$$S_\lambda \leq |\langle b_0 \rangle + \langle b_1 \rangle| + |\langle b_0 \rangle - \langle b_1 \rangle|$$

Possiamo supporre, senza perdita di generalità, che $\langle b_0 \rangle \geq \langle b_1 \rangle \geq 0$, e quindi rimuovere i due valori assoluti e trovare:

$$S_\lambda \leq 2\langle b_0 \rangle \leq 2$$

Equivalentemente, siano $x, y \in [-1, 1]$, allora $S_\lambda = |x + y| + |x - y| \leq 2$. Infatti, prendendo il quadrato:

$$S_\lambda^2 = x^2 + y^2 + 2xy + x^2 + y^2 - 2xy + 2|x + y||x - y| = 2x^2 + 2y^2 + 2|x^2 - y^2|$$

che è uguale a $4x^2$ o a $4y^2$, e in entrambi i casi è $S_\lambda^2 \leq 4 \Rightarrow S_\lambda \leq 2$.

Si ha allora:

$$S = \int_\Lambda d\lambda q(\lambda) S_\lambda \leq 2 \int_\Lambda d\lambda q(\lambda) = 2 \Rightarrow S \leq 2$$

dato che $q(\lambda)$ è normalizzata.

Svolgiamo ora lo stesso calcolo con il formalismo della MQ. Consideriamo la base computazionale $\{|0\rangle, |1\rangle\}$ degli autostati di $\hat{\sigma}_z$. Nella MQ, l'informazione completa (λ) necessaria per calcolare le probabilità di ogni misura è contenuta nello stato $|\psi\rangle$ dei due qubit, che supponiamo essere *massimamente entangled*, ossia pari allo stato di singoletto $|\psi^-\rangle$:

*Violazione della
disuguaglianza
CHSH in MQ*

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)_{AB}$$

Supponiamo che Alice e Bob possano misurare lo spin delle rispettive particelle in arbitrarie direzioni \vec{X} e \vec{Y} ($\in \mathbb{R}^3$). Esplicitamente, (X, Y) sono quindi scelte di osservabili del tipo:

$$\hat{O}_x = \vec{X} \cdot \vec{\sigma} \quad \hat{O}_y = \vec{Y} \cdot \vec{\sigma}; \quad \vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$$

Detti $\vec{X} = (X_1, X_2, X_3)^T$ e $\vec{Y} = (Y_1, Y_2, Y_3)^T$, nella base computazionale di un qubit:

$$\vec{X} \cdot \vec{\sigma} = \begin{pmatrix} X_3 & X_1 - iX_2 \\ X_1 + iX_2 & -X_3 \end{pmatrix}; \quad \vec{Y} \cdot \vec{\sigma} = \begin{pmatrix} Y_3 & Y_1 - iY_2 \\ Y_1 + iY_2 & -Y_3 \end{pmatrix}$$

Le singole correlazioni che compaiono in S_λ sono quindi date da:

$$\langle a_x b_y \rangle = \langle \psi^- | (\vec{X} \cdot \vec{\sigma})_A \otimes (\vec{Y} \cdot \vec{\sigma})_B | \psi^- \rangle$$

Nella base computazionale per 2 qubit $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ vale:

$$\langle a_x b_y \rangle = \frac{1}{2} \begin{pmatrix} 0 & 1 & -1 & 0 \end{pmatrix} M \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = -\vec{X} \cdot \vec{Y} \quad (4.6)$$

$$M = \left(\begin{array}{cc|cc} X_3 Y_3 & X_3 (Y_1 - iY_2) & (X_1 - iX_2) Y_3 & (X_1 - iX_2) (Y_1 - iY_2) \\ X_3 (Y_1 + iY_2) & -X_3 Y_3 & (X_1 - iX_2) (Y_1 + iY_2) & -Y_3 (X_1 - iX_2) \\ \hline (X_1 + iX_2) Y_3 & (X_1 + iX_2) (Y_1 - iY_2) & -X_3 Y_3 & -X_3 (Y_1 - iY_2) \\ (X_1 + iX_2) (Y_1 + iY_2) & -Y_3 (X_1 + iX_2) & -X_3 (Y_1 + iY_2) & X_3 Y_3 \end{array} \right)$$

Restringiamoci al caso in cui Alice e Bob possano misurare solo lungo due direzioni *opportunamente predefinite* in modo da massimizzare S_λ . Nello specifico, Alice può misurare nella base “cartesiana”:

$$\vec{a}_0 = \vec{e}_0 \quad \vec{a}_1 = \vec{e}_1$$

e Bob in una base ruotata di $3\pi/4$:

$$\vec{b}_0 = \frac{-(\vec{e}_0 + \vec{e}_1)}{\sqrt{2}} \quad \vec{b}_1 = \frac{-\vec{e}_0 + \vec{e}_1}{\sqrt{2}}$$

In questo modo la (4.6) restituisce $1/\sqrt{2}$ per $\langle a_0 b_0 \rangle$, $\langle a_0 b_1 \rangle$ e $\langle a_1 b_0 \rangle$ (ossia tutti i termini positivi in S_λ) e $-1/\sqrt{2}$ per l'unico negativo. Perciò:

$$S_\lambda = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} - \left(-\frac{1}{\sqrt{2}} \right) = \frac{4}{\sqrt{2}} = 2\sqrt{2} > 2$$

E quindi abbiamo una violazione della disuguaglianza CHSH.

Tale risultato, che deriva dall'aver assunto particolari postulati all'origine della MQ, è sperimentalmente verificato.

4.2 Tipologie di correlazioni

Il teorema di Bell introduce un'importante distinzione tra le *correlazioni* che possono essere generate da fenomeni classici e quelle proprie di sistemi quantistici *entangled*. In altre parole, vi sono misure quantistiche che non possono essere

direttamente simulate con sistemi classici - nemmeno disponendo di potenza computazionale illimitata. Basta infatti considerare due sistemi in uno stato *intrecciato* ρ_{AB} , per cui si vogliono calcolare le probabilità $p(ab|xy)$ che una misura di x su A risulti in a , e una di y su B risulti in b . Poiché tale distribuzione è *non separabile*, due osservatori A e B sufficientemente lontani da non poter conoscere (località) ciascuno le misure scelte dall'altro, non hanno modo - indipendentemente dalla quantità di informazioni *condivise in precedenza* - di calcolare correttamente le $p(ab|xy)$.

Vale allora la pena di considerare, in modo generale e completamente astratto, alcune tipologie interessanti di correlazioni che possono esistere tra risultati a e b .

Si definisce **scenario di Bell** un esperimento in cui due osservatori distanti (Alice e Bob) eseguono misure su un sistema fisico condiviso (es. una coppia di particelle *entangled*), su cui non si fanno ipotesi (*una black-box*). Ciascun osservatore può scegliere tra m possibili misure, e ciascuna misura può risultare in Δ esiti possibili. Denotiamo le misurazioni scelte da A e B con $X, Y \in \{1, \dots, m\}$, e i relativi esiti con $a, b \in \{1, \dots, \Delta\}$.

Ogni scenario è completamente individuato dalla sequenza di probabilità $\vec{p} = \{p(ab|XY)\}$ corrispondenti a qualsiasi scelta degli input (le possibili misure X, Y , con m^2 scelte) e qualsiasi possibilità per l'output (gli esiti a, b , per altre Δ^2 scelte), per un totale di $\Delta^2 m^2$ parametri: $\vec{p} \in \mathbb{R}^{\Delta^2 m^2}$.

Più precisamente, \vec{p} è confinato ad un sottoinsieme proprio $\mathcal{P} \subset \mathbb{R}^{\Delta^2 m^2}$ di dimensione $(\Delta^2 - 1)m^2$, dato che valgono m^2 vincoli di normalizzazione:

$$\sum_{a,b=1}^{\Delta} p(ab|XY) \stackrel{!}{=} 1 \quad \forall (X, Y) \in \{1, \dots, m\}^2$$

Inoltre le probabilità devono essere tutte positive:

$$p(ab|XY) \geq 0 \quad \forall a, b, X, Y$$

A seconda delle proprietà della teoria che permette di calcolare le $p(ab|XY)$, vi sono altri vincoli su \vec{p} , che quindi appartiene ad un sottoinsieme più piccolo di \mathcal{P} . In altre parole, la teoria restringe le possibili correlazioni (ossia le scelte degli elementi di \vec{p}). In particolare distinguiamo tra:

- **Correlazioni No-signal**: frutto di teorie che assumono solo l'impossibilità di una comunicazione istantanea tra Alice e Bob. Si tratta di correlazioni *di forma più generale* rispetto a quelle permesse dalla sola MQ. Matematicamente si suppone che le probabilità marginali calcolate da Alice siano indipendenti dalla scelta delle misure fatta da Bob:

$$p(a|X) = p(a|XY) = \sum_b p(ab|XY) \quad \forall Y$$

In parole povere, ciò significa che le scelte di B non possono influenzare i risultati di A (e viceversa) e quindi Bob non può *segnalare* le sue scelte ad

Alice (e viceversa).

Valgono allora i seguenti vincoli su \vec{p} :

$$\begin{aligned}\sum_{b=1}^{\Delta} p(ab|XY) &= \sum_{b=1}^{\Delta} p(ab|XY') \quad \forall a, X, Y, Y' \\ \sum_{a=1}^{\Delta} p(ab|XY) &= \sum_{a=1}^{\Delta} p(ab|X'Y) \quad \forall b, Y, X, X'\end{aligned}\tag{4.7}$$

La regione di \mathcal{P} individuata da tali vincoli è denotata con \mathcal{NS} .

Per esempio, nel caso $\{\Delta = 2\}$, con $a, b \in \{-1, 1\}$, le relazioni (4.7) possono essere scritte mediante dei correlatori. In effetti, si può dimostrare che $\{\langle A_x \rangle, \langle B_y \rangle, \langle A_x B_y \rangle\}$ bastano in questo caso a definire lo scenario di Bell, dato che da essi possiamo determinare le $p(ab|XY)$ tramite:

$$\begin{aligned}\langle A_x \rangle &= \sum_a a p(a|X) \\ \langle B_y \rangle &= \sum_b b p(b|Y) \\ \langle A_x B_y \rangle &= \sum_{ab} ab p(ab|XY) \\ p(ab|XY) &= \frac{1}{4} [1 + a \langle A_x \rangle + b \langle B_y \rangle + ab \langle A_x B_y \rangle]\end{aligned}$$

La (4.7) porta allora a:

$$1 + a \langle A_x \rangle + b \langle B_y \rangle + ab \langle A_x B_y \rangle \geq 0$$

In particolare, se $\langle A_x \rangle = \langle B_y \rangle = 0$ si trova:

$$-1 \leq \langle A_x B_y \rangle \leq 1$$

2. **Correlazioni da teorie locali** (“classiche”). Come visto nelle sezioni precedenti, in una teoria locale deve valere una condizione di separabilità, “mediata” da parametri locali $\lambda \in \Lambda$ con distribuzione $q(\lambda)$:

$$p(ab|XY) = \int_{\Lambda} d\lambda q(\lambda) p(a|X, \lambda) p(b|Y, \lambda)$$

La regione di \mathcal{P} individuata da tale vincolo è denotata con \mathcal{L} . Il complemento $\mathcal{P} \setminus \mathcal{L}$ contiene perciò le scelte di \vec{p} che contengono *correlazioni non-locali*.

Si può dimostrare che tutte le correlazioni locali sono anche no-signal, ma non vale il contrario. In altre parole: $\mathcal{L} \subset \mathcal{NS}$.

3. **Correlazioni Quantistiche**. Consideriamo infine tutte le scelte di \vec{p} che sono permesse dalla MQ. Ricordiamo che una misura generalizzata x che restituisce come esito a è descritta da un operatore POVM (*Positive Operator Valued Measure*) $M_{a|x}$ tale che $M_{a|x} \geq 0$ e $\sum_{a=1}^{\Delta} M_{a|x} = \mathbb{I}$. La probabilità che si ottenga l'esito a da una tale misura in uno stato ρ_A è data da:

$$p(a|x) = \text{Tr}(\rho_A M_{a|x})$$

Estendendo al caso di un sistema composto nello stato $\rho_{ab} \in \mathcal{H}_A \otimes \mathcal{H}_B$, con generiche POVM $M_{a|x}: \mathcal{H}_A \rightarrow \mathcal{H}_A$ e $M_{b|y}: \mathcal{H}_B \rightarrow \mathcal{H}_B$, otteniamo allora la condizione:

$$p(ab|xy) = \text{Tr}(\rho_{AB} M_{a|x} \otimes M_{b|y})$$

Senza perdita di generalità, allargando opportunamente gli spazi di Hilbert, è possibile considerare solo stati puri (pari alle ρ_{AB} *purificate*) e solo misure proiettive (per cui $M_{a|x} M_{a'|x} = \delta_{aa'} M_{a|x}$ e $\sum_a M_{a|x} = \mathbb{I}_A$ e analogamente per $M_{b|y}$). In tal caso la condizione diviene:

$$p(ab|XY) = \langle \psi | P_{a|X} \otimes P_{b|Y} | \psi \rangle$$

Le probabilità calcolate in entrambi i casi sono le stesse (per costruzione), e perciò i due vincoli sono completamente equivalenti. Denotiamo con \mathcal{Q} il sottoinsieme di \mathcal{P} da essi individuato.

Nota. Vi è un modo alternativo di scrivere tali vincoli, spesso utilizzato in QFT, in cui non si fa uso di una struttura di prodotto tensore tra i sistemi di A e B . Si considera piuttosto un unico spazio di Hilbert \mathcal{H} che comprende l'intero sistema, e le misure sono descritte da “osservabili locali” che agiscono su punti distinti e indipendenti, ossia proiettori ortogonali $M_{a|x}$ e $M_{b|y}$ che commutano tra loro $[M_{a|x}, M_{b|y}] = 0$. Così facendo, si ottiene un'espressione con un prodotto matriciale al posto del prodotto tensore:

$$p(ab|XY) = \langle \psi | M_{aX} M_{bY} | \psi \rangle \quad [M_{aX}, M_{bY}] = 0$$

Denotiamo con \mathcal{Q}' il sottoinsieme di \mathcal{P} che verifica tale condizione. Poiché $[M_{a|x} \otimes \mathbb{I}_B, \mathbb{I}_A \otimes M_{b|y}] = 0$, $\mathcal{Q} \subseteq \mathcal{Q}'$. L'inclusione inversa vale nel caso finito-dimensionale (dove quindi $\mathcal{Q} = \mathcal{Q}'$), ma non si sa ancora se valga in quello infinito-dimensionale.

Si può dimostrare che:

1. Le teorie locali ammettono una descrizione quantistica: $\mathcal{L} \subseteq \mathcal{Q}$. In effetti, le correlazioni tra stati separabili (non entangled) sono di tipo “classico”.
2. Ogni comportamento quantistico soddisfa le condizioni di *no-signaling*: $\mathcal{Q} \subseteq \mathcal{NS}$. In altre parole, non è possibile utilizzare qubit entangled per trasmettere segnali a velocità superluminali.
3. Esistono comportamenti quantistici non locali (violazione delle disuguaglianze di Bell), e quindi $\mathcal{L} \subset \mathcal{Q}$ è una inclusione stretta.

Analogamente, si trova che le teorie no-signaling sono più generali della MQ, e quindi $\mathcal{Q} \subset \mathcal{NS}$ è anch'essa stretta. Mettendo tutto insieme giungiamo a:

$$\mathcal{L} \subset \mathcal{Q} \subset \mathcal{NS}$$

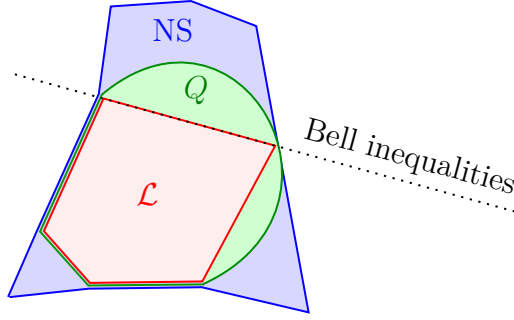


Figura (4.4) – Schema delle relazioni tra i possibili insiemi di scenari di Bell, delimitati da vincoli nello spazio \mathcal{P} (qui proiettati in $d = 2$). Si trova infatti che NS ha la struttura di un *politopo*. I piani che separano le varie classi sono detti, in generale, **disuguaglianze di Bell**.

Nel caso $\Delta = 2$, $m = 2$, la forma più generale delle disuguaglianze di Bell è data da:

$$S \cdot p = \sum_{abXY} S_{XY}^{ab} p(ab|XY) \leq S_k$$

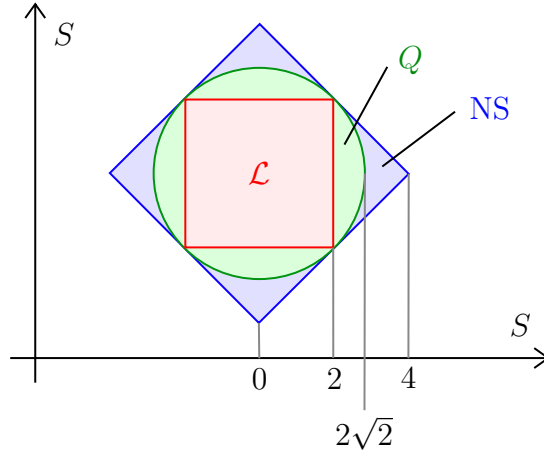


Figura (4.5) – Relazioni tra Q , \mathcal{L} e NS per $\Delta = 2$, $m = 2$.

4.3 Misura sperimentale della violazione delle disuguaglianze CHSH

Esaminiamo ora un possibile apparato sperimentale in grado di osservare la violazione delle disuguaglianze CHSH.

Utilizzeremo come qubit lo stato di polarizzazione di un fotone, nella base $\{|H\rangle, |V\rangle\}$ di polarizzazione (lineare) orizzontale (H) o verticale (V).

Per generare qubit entangled impieghiamo il fenomeno di *spontaneous parametric downconversion*. Certi cristalli (come ad esempio il borato di Bario, comunemente

detto BBO) esibiscono un comportamento ottico **non lineare**. Ciò rende possibile la conversione di un fotone in ingresso in una coppia di fotoni, rispettando conservazione di energia e momento (figura 4.6).

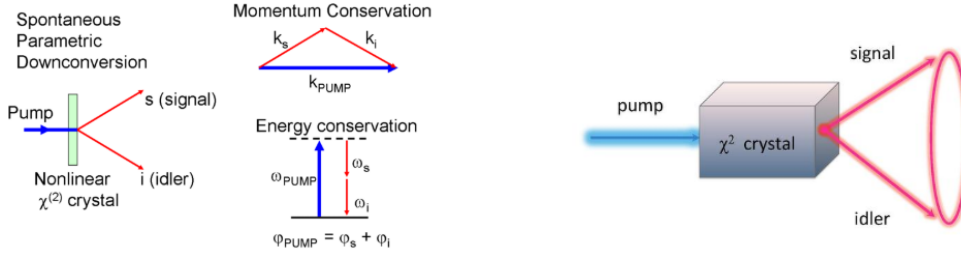


Figura (4.6) – A sinistra: schema del processo di *spontaneous parametric downconversion*. Come visibile a destra, i due fotoni risultanti sono emessi in un cono di angolo specifico.

Tale processo avviene con un'efficienza molto ridotta, e solo per certe frequenze dei fotoni in ingresso. Per poterlo sfruttare, perciò, è necessario utilizzare un laser sufficientemente potente.

A seconda della tipologia di cristallo, la polarizzazione dei due fotoni emessi è più o meno correlata a quella del fotone incidente. In particolare, in un cristallo con *phase matching* di tipo 1, se il fotone iniziale è polarizzato parallelamente ad uno specifico asse del cristallo, allora i due fotoni risultanti hanno entrambi la stessa polarizzazione, perpendicolare a quella iniziale.

Ciò permette, per esempio, di convertire un fotone $|V\rangle$ in due fotoni $|H\rangle$:

$$|V\rangle \mapsto |H\rangle_A \otimes |H\rangle_B$$

Lo stato così generato è separabile, e quindi non entangled. Possiamo rimediare a ciò sovrapponendo due cristalli non lineari, con assi perpendicolari, e illuminandoli con luce polarizzata a 45° (figura 4.7). Il fotone iniziale è quindi nello stato $|\psi_1\rangle$:

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle)$$

e perciò ha uguale probabilità di dividersi in due fotoni polarizzati orizzontalmente o verticalmente, a seconda del cristallo con cui interagisce.

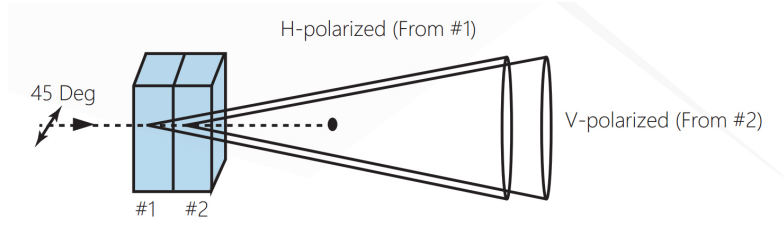


Figura (4.7) – I cristalli #1 e #2 hanno assi ottici perpendicolari: il primo verticale, il secondo orizzontale. Perciò #1 converte fotoni $|V\rangle \rightarrow |H\rangle \otimes |H\rangle$, mentre il secondo $|H\rangle \rightarrow |V\rangle \otimes |V\rangle$. Utilizzando come input fotoni polarizzati a 45° (sovrapposizione a pari coefficienti di $|H\rangle$ e $|V\rangle$) allora ogni fotone ha pari probabilità di essere splittato da #1 o da #2. Fino a quando non viene fatta una misura, entrambi i processi *sono esplorati*, e perciò lo stato finale è entangled.

L'unico problema è che, a causa dello spessore dei cristalli non lineari, i due *coni* di emissione delle coppie di fotoni non coincidono. Perciò i due percorsi possibili (split nel primo cristallo o nel secondo) non sono perfettamente sovrapposti - e quindi lo stato finale non è una superposizione coerente, ma uno stato misto:

$$|\psi'_2\rangle = \frac{1}{2}(|HH\rangle \langle HH| + |VV\rangle \langle VV|)$$

Tale fenomeno è definito *walk-off*, e può essere compensato “ritardando” uno dei due fotoni tramite un altro elemento ottico. Così facendo otteniamo, finalmente, uno stato entangled coerente:

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}(|H\rangle_A |H\rangle_B + e^{i\varphi} |V\rangle_A |V\rangle_B)$$

Non resta altro che effettuare misure di polarizzazione sui due fotoni entangled. Per farlo, esaminiamo - punto per punto - il seguente setup:

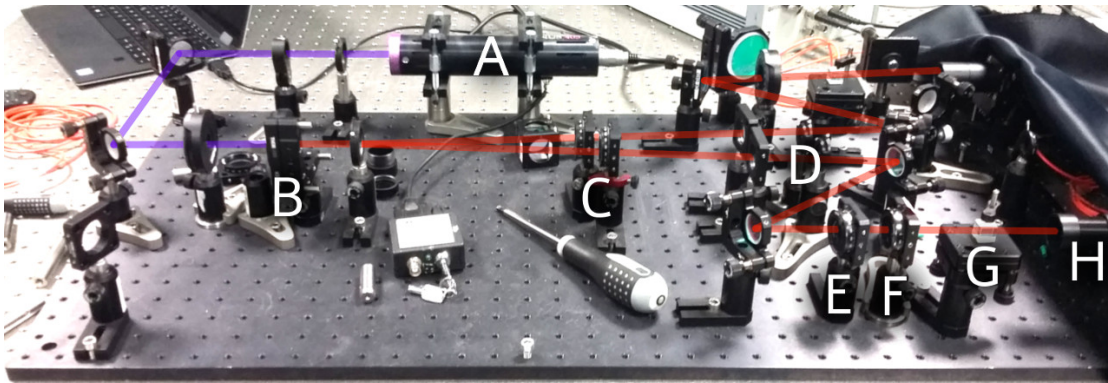


Figura (4.8) – Setup sperimentale per la violazione delle disuguaglianze CHSH

1. In **A** si ha un laser di un'opportuna lunghezza d'onda (specifica per i BBO) e sufficientemente potente. Il fascio viene polarizzato **verticalmente** da un filtro.

2. Il laser viene focalizzato e collimato da alcuni elementi ottici, e tramite una coppia di specchi viene fatto incidere sul cristallo non lineare in **B**, dove avviene il fenomeno di *parametric down-conversion*.
3. Ci concentriamo su uno specifico piano all'interno del cono di emissione delle coppie di fotoni. In **C**, tramite opportuni elementi ottici, si sovrappongono i due coni relativi ai due strati di cristallo non lineare, in modo da ottenere lo stato coerente $|\psi_2\rangle$.
4. Dopo aver attraversato un'ulteriore lente, in **D** i due fasci di fotoni entangled incidono su una coppia di *iridi*, che vengono regolati in modo da lasciar passare solo la componente effettivamente sovrapposta (dove quindi l'entanglement è massimo)
5. Ciascun fascio viene riflesso verso l'apparato di misura. Concentriamoci sul fascio inferiore. In **E** si ha una lamina $\lambda/2$, che consente di variare l'angolo θ di polarizzazione lineare della luce incidente. Nel dettaglio, la lamina "riflette" la polarizzazione della luce incidente rispetto al suo asse ottico (che può essere orientato a piacere):

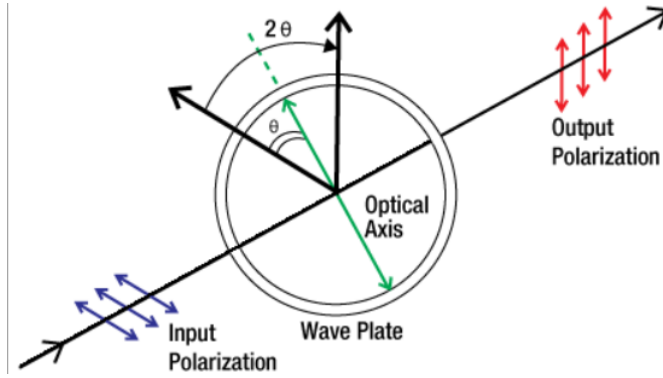


Figura (4.9) – Schema del funzionamento di una lamina $\lambda/2$. Poiché la lamina “riflette” la polarizzazione incidente, orientandola a $\theta = 45^\circ$ è possibile convertire $|H\rangle$ in $|V\rangle$ (e viceversa). Analogamente, per $\theta = 22.5^\circ$, si converte una polarizzazione a 45° ($|+\rangle$) in $|H\rangle$.

Una lamina $\lambda/4$ (in **F**) è presente in uno solo dei due fasi, e serve a fissare la fase φ che compare in $|\psi_2\rangle$. A questo punto il fascio incide su un “polarizing biscuit” in **G**, che è un normale cristallo birifrangente orientato in modo da trasmettere al 100% la componente $|H\rangle$, e da riflettere quella $|V\rangle$. Poiché possiamo convertire una qualsiasi polarizzazione in $|H\rangle$ orientando opportunamente la lamina $\lambda/2$, l'apparato consente di misurare la polarizzazione lungo un asse arbitrario.

6. Infine, in **H**, si ha un rivelatore a singolo fotone, collegato ad un'opportuna elettronica di acquisizione e ad un computer per la presa dati.

Per determinare la presenza o meno di entanglement, un modo è misurare un'osservabile detta **entanglement witness**, tale da risultare in esiti differenti per stati

separabili o stati entangled.

Un esempio è misurare le correlazioni in *basi coniugate*. Ci aspettiamo infatti di osservare coppie polarizzate HH o VV sia per $|\psi_2\rangle$ che per $|\psi'_2\rangle$, e allo stesso modo di non osservare HV e VH (se non per le incertezze sperimentali dovute ad allineamento, rumore, etc.). Tuttavia, misurando in una base *diagonale*, per $|\psi_2\rangle$ ci aspettiamo di misurare tanti fotoni polarizzati $++$, e nessuno $+-$, mentre per $|\psi'_2\rangle$ misureremo fotoni con ogni combinazione di polarizzazione ($++$ e $+-$ allo stesso modo).

In presenza di entanglement possiamo misurare la violazione delle disuguaglianze di Bell. L'idea è di misurare un fotone nelle polarizzazioni orizzontale e verticale, e l'altro in diagonale e antidiagonale, eseguendo quanto già visto in (4.5).

4.4 Quantificare l'informazione

(Lezione 11 ● del
3/4/2019)

Ci dedichiamo ora al problema di *quantificare* le correlazioni quantistiche. Per far ciò è necessario innanzitutto definire il problema, cercando un metodo preciso per *quantificare* qualcosa di astratto come “l'informazione” contenuta in un messaggio.

4.4.1 Entropia di Shannon

Partiamo allora introducendo alcuni concetti di teoria dell'informazione *classica*. Il problema di *quantificare* l'informazione contenuta in un messaggio fu risolto da Shannon nel 1948, sfruttando un'analogia tra la *complessità* di un sistema fisico (che è “codificata” dalla sua entropia) e quella di un *sequenza di caratteri*.

Partiamo da alcune definizioni di base:

- Un **messaggio** è una sequenza di lunghezza arbitraria N composta di **lettere** scelte da un certo **alfabeto** \mathcal{A} :

$$\mathcal{A} = \{a_1, \dots, a_k\}$$

Assumiamo per semplicità che le lettere nel messaggio siano **statisticamente indipendenti** l'una dall'altra. Precisamente, ciò significa che ogni lettera compare in una qualsiasi posizione nel messaggio con una probabilità p_i (conosciuta a priori) che non dipende dalle lettere che la precedono o la seguono¹. Naturalmente le p_i devono rispettare le condizioni proprie delle probabilità:

$$p_i \geq 0 \quad \forall i = 1, \dots, k; \quad \sum_{i=1}^k p_i = 1$$

¹ΛCiò è chiaramente falso nel caso di messaggi scritti in una *lingua naturale*, come l'italiano. Per esempio è molto difficile che k segua una m , ed è molto probabile che una u segua una q .

Consideriamo allora una distribuzione di probabilità fissata, ossia la sequenza delle probabilità delle k lettere dell'alfabeto: $\{p_1, p_2, \dots, p_k\}$. Si definisce **entropia di Shannon** la quantità:

$$H(p_1, \dots, p_k) = - \sum_{i=1}^k p_i \log_2 p_i$$

dove la base del logaritmo è una scelta convenzionale (in genere si usa 2 con riferimento all'alfabeto binario), e verrà spesso omessa.

Esempio. Un messaggio binario ha caratteri scelti da un alfabeto di soli due elementi: $\mathcal{A} = \{0, 1\}$. Avremo quindi $0 \leq p_1 \equiv p \leq 1$, e $p_2 = 1 - p$, da cui l'entropia di Shannon è data da:

$$H(p_1, p_2) = -p \log_2 p - (1 - p) \log_2 (1 - p) = H(p) \quad (4.8)$$

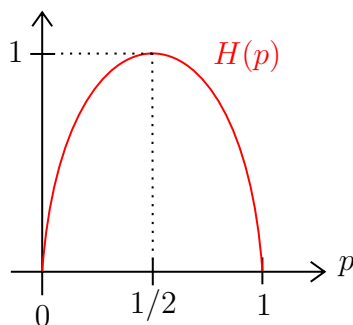


Figura (4.10) – Grafico dell'entropia di Shannon $H(p)$ per un messaggio binario.

Notiamo che $H(p) = 0$ per $p = 0$ o $p = 1$. In altre parole messaggi *costanti*, cioè formati da soli 0 o soli 1 hanno entropia *minima* (nulla). Viceversa, un messaggio in cui 0 e 1 compaiono con la stessa probabilità ha entropia *massima*.

Ciò è consistente con l'interpretare $H(p)$ come “l'**informazione media** contenuta in ciascuna lettera del messaggio”, o equivalentemente come una misura di “**ignoranza a priori**”. Immaginiamo infatti di ricevere un messaggio costante, interamente composto da 0 (come succede per $p = 1$), lettera per lettera. Poiché sappiamo già *a priori* che ogni lettera sarà uno 0, la nostra ignoranza a priori è minima (sappiamo già tutto), così così come è minima l'informazione ancorata a ciascuna nuova lettera, dato che non aggiunge nulla a quanto già sappiamo (e infatti $H(0) = 0$).

D'altro canto, se $p = 1/2$ abbiamo ignoranza massima sul messaggio (che appare completamente casuale), e ogni nuova lettera aggiunge $H(1/2) = 1$ bit di informazione “nuova” alla nostra conoscenza.

Un'altra interpretazione lega un valore basso di $H(p)$ ad un'elevata **ridondanza** all'interno del messaggio, che può essere espresso utilizzando un numero molto più basso di caratteri, ossia può essere **compressso** con un grande guadagno di spazio. D'altro canto, una $H(p)$ vicina al massimo significa che il messaggio è

Interpretazione dell'entropia di Shannon

“molto casuale”, e quindi non si può pensare di poterlo comprimere in un qualche messaggio di lunghezza molto inferiore.

Un modo “naturale” [12] per ricavare l’entropia di Shannon parte dall’associare l’informazione contenuta in un messaggio all’ignoranza *a priori*, e quindi all’*incertezza* presente nella distribuzione di probabilità delle lettere nel messaggio. Concentriamoci su una di esse, che compare con probabilità p . Potremmo stimare l’*incertezza* quantificando la “sorpresa” S legata all’apparire di tale lettera, che può essere presa come proporzionale a $1/p$ - più piccola è p , più alta è la sorpresa. Vorremmo però che tale “incertezza” (o sorpresa) sia additiva, ma $S(p \cdot q) = 1/(pq) \neq S(p) \cdot S(q)$. Possiamo recuperare l’uguaglianza definendo invece $S(p) = \log(1/p) = -\log(p)$. A questo punto, passando al caso di una distribuzione di probabilità $\{p_1, \dots, p_k\}$, definiamo $S(\vec{p})$ come la *media pesata* delle singole “sorprese” p_i , ossia esattamente come $-\sum_i p_i \log p_i$, riottenendo così l’entropia di Shannon.

4.4.2 Noiseless Coding

L’ultima interpretazione dell’entropia di Shannon porta a pensare ad un collegamento tra la *quantità di informazione* contenuta in un messaggio e la *compressibilità* del messaggio stesso. Ci si aspetta, infatti, di poter *codificare* un messaggio ridondante (ossia con poca informazione per lettera) in un messaggio significativamente più corto.

In effetti, questo è proprio quello che succede, e un messaggio di n lettere può essere *compresso* ad uno di $nH(p_1, \dots, p_k)$ lettere, come dimostrato dal teorema del *noiseless coding* di Shannon.

Teorema 4.4.1. *Dato un messaggio di lunghezza n le cui lettere sono state scelte indipendentemente tra loro da un alfabeto $\mathcal{A} = \{a_1, \dots, a_k\}$ con probabilità a priori $\{p_1, \dots, p_k\}$, esiste asintoticamente (ossia per messaggi di lunghezza grande $n \rightarrow \infty$) una codifica ottimale per comprimere il messaggio usando $H(p_1, \dots, p_k)$ bit per lettera, senza perdita di informazione.*

Teorema del
noiseless coding di
Shannon

Dimostrazione omessa.

Esempio. Forniamo solamente un esempio di applicazione del teorema del noiseless coding (mediante il codice di Huffman). Consideriamo un generico alfabeto di 4 lettere $\mathcal{A} = \{a_1, a_2, a_3, a_4\}$. Ciascuna di esse può essere *codificata* da 2 bit:

$$a_1 \leftrightarrow 00 \quad a_2 \leftrightarrow 01 \quad a_3 \leftrightarrow 10 \quad a_4 \leftrightarrow 11$$

Per il teorema del *noiseless coding* è possibile trovare una rappresentazione migliore, ossia trasmettere lo stesso messaggio usando *meno* di 2 bit per lettera. Per farlo dobbiamo però partire da delle probabilità:

$$p = \left\{ \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8} \right\}$$

Consideriamo la seguente associazione:

$$a_1 \leftrightarrow 0 \quad a_2 \leftrightarrow 10 \quad a_3 \leftrightarrow 110 \quad a_4 \leftrightarrow 111$$

Notiamo che ogni lettera a_i è ora rappresentata da una stringa di bit di *differente lunghezza* l_i , con la particolarità che la stringa più corta è associata alla lettera più probabile, e quelle più lunghe alle lettere meno probabili.

In effetti, calcolando il numero medio \bar{L} di bit necessari per inviare una lettera del messaggio otteniamo:

$$\bar{L} = \sum_{i=1}^4 p_i l_i = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 3 = \frac{7}{4} < 2$$

Nota: le stringhe non possono essere scelte casualmente, dato che è necessario poter *distinguere* le lettere che rappresentano quando sono disposte una dopo l'altra. Per esempio, non si può usare $a_2 \leftrightarrow 1$, dato che in tal caso la sequenza 110 ha traduzioni ambigue: $a_2 a_2 a_1$ e a_3 .

Nota 2: il limite $n \rightarrow \infty$ è necessario per poter interpretare situazioni estreme. Per esempio, un messaggio costante in alfabeto binario può essere codificato per 0 bit per lettera - cosa assurda per messaggi finiti, ma sensata nel limite infinito. Infatti, un qualsiasi messaggio costante di n bit può essere codificato *sempre* con un solo bit (dato che sono tutti uguali). Perciò il numero di bit necessari *per lettera* è $1/n$, e:

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0$$

4.5 Entropia di Von Neumann

Possiamo adattare i risultati del teorema di Shannon al caso quantistico, pensando ad uno stato come a un *messaggio* scritto nell'alfabeto della *base computazionale*. Esplicitamente, sia $\{|\psi_0\rangle, \dots, |\psi_k\rangle\}$ una base computazionale (di stati puri). Consideriamo la relativa base per le matrici densità $\{\rho_0, \dots, \rho_k\}$, con $\rho_i = |\psi_i\rangle\langle\psi_i|$. Una generica matrice densità ρ è una mistura statistica delle *lettere* ρ_i con probabilità a priori p_i :

$$\rho = \sum_{i=1}^k p_i \rho_i$$

Con questa notazione, possiamo direttamente adattare la definizione di entropia di Shannon, ottenendo l'**entropia di Von Neumann** S_V :

$$S_V(\rho) = -\text{Tr}(\rho \log \rho)$$

Poiché la traccia non dipende dalla scelta della base, possiamo calcolarla nella base che diagonalizza ρ , ossia quella in cui $\rho = \text{diag}(\lambda_1, \dots, \lambda_k)$, con λ_i autovalori di ρ . Tale base esiste sempre, in quanto ρ è una matrice hermitiana (dato che è una

matrice densità) e quindi è diagonalizzabile². Ricordando allora che applicare una funzione ad una matrice diagonale consiste nell'applicare la funzione a ciascun elemento della diagonale, si ottiene:

$$\begin{aligned} S_V(\rho) &= -\text{Tr} \left[\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{pmatrix} \begin{pmatrix} \log(\lambda_1) & & \\ & \ddots & \\ & & \log(\lambda_k) \end{pmatrix} \right] = \\ &= -\sum_{i=1}^k \lambda_i \log \lambda_i = H(\lambda_1, \dots, \lambda_k) \end{aligned}$$

Perciò l'entropia di Von Neumann di una matrice densità ρ non è altro che l'entropia di Shannon dei suoi autovalori λ_i .

Esaminiamo le **proprietà** di S_V :

1. L'entropia di Von Neumann di **stati puri** è nulla:

$$\text{Stati puri} \Leftrightarrow S(\rho) = 0$$

Infatti la ρ di uno stato puro può essere scritta sempre come il prodotto esterno di un solo termine:

$$\rho = |\psi\rangle \langle \psi|$$

Da cui un autovalore $\lambda_1 = 1$ e tutti gli altri $\lambda_i = 0$ per $i > 1$. Perciò:

$$S(\rho) = -\sum \lambda_{i=1}^k \log \lambda_i = -\lambda_1 \log \lambda_1 = -\log 1 = 0$$

2. S_V è invariante sotto cambi di base (trasformazioni U unitarie):

$$S(U\rho U^\dagger) = S(\rho)$$

Ciò è dovuto al fatto che $S(\rho)$ dipende dagli autovalori, che sono invarianti per trasformazioni unitarie (in effetti sono definiti anche senza scegliere una base). Una conseguenza importante è che l'entropia di Von Neumann non varia a seguito dell'evoluzione unitaria del sistema.

3. Se $\dim(\mathcal{H}) = N \Rightarrow 0 \leq S(\rho) \leq \log N$.

Infatti $S(\rho) \geq 0$ poiché $0 \leq \lambda_i \leq 1 \forall i$, e quindi $-\lambda_i \log \lambda_i \geq 0$. Usando poi il fatto che $S(\rho) = H(\lambda_1, \dots, \lambda_N)$, e ricordando che l'entropia di Shannon è massima quando tutte le lettere sono equiprobabili, ossia quando vale $\lambda_1 = \dots = \lambda_N = 1/N$, si ha:

$$S(\rho)_{\max} = -\frac{1}{N} \sum_{i=1}^N \log \frac{1}{N} = \log N$$

²^Tutto ciò è frutto dei postulati che definiscono le osservabili, introdotto proprio per permettere di calcolare *funzioni* di osservabili.

La maggiore libertà offerta dalla MQ si traduce in alcune differenze tra entropia di Von Neumann ed entropia di Shannon, che esaminiamo nei seguenti due esempi.

Esempio 1.

Scegliamo la base $\{|0\rangle, |1\rangle\}$, da cui ricaviamo l'*alfabeto* di stati puri $\{\rho_0, \rho_1\}$, con $\rho_0 = |0\rangle\langle 0|$, $\rho_1 = |1\rangle\langle 1|$, con probabilità $p_0 = p$ e $p_1 = 1 - p$. Una generica matrice densità ρ è data da:

$$\rho = \sum p_i^2 \rho_i = p |0\rangle\langle 0| + (1 - p) |1\rangle\langle 1| = \begin{pmatrix} p & 0 \\ 0 & 1 - p \end{pmatrix}$$

L'entropia di Von Neumann è allora data da:

$$S(\rho) = -\text{Tr}(\rho \log \rho) = -\sum p_i^2 \log p_i = H(p)$$

Otteniamo perciò che la mistura statistica di stati quantistici *ortogonali tra loro* ($\langle 0|1\rangle = 0$) è analoga ad una mistura di *stati classici* (4.8).

Esempio 2.

Scegliamo ora due stati *senza analogo classico*:

$$\begin{aligned} |a\rangle &= \cos \theta |0\rangle + \sin \theta |1\rangle = \begin{pmatrix} c \\ s \end{pmatrix} \\ |b\rangle &= \sin \theta |0\rangle + \cos \theta |1\rangle = \begin{pmatrix} s \\ c \end{pmatrix} \end{aligned}$$

con $0 \leq \theta \leq \pi/4$, $c = \cos \theta$, $s = \sin \theta$. Notiamo che $\langle a|b\rangle = \sin 2\theta$, quindi in generale non possiamo rappresentare $|a\rangle$ e $|b\rangle$ con “equivalenti classici” ortogonali. Le matrici densità sono date da:

$$\begin{aligned} \rho_a &= |a\rangle\langle a| = \begin{pmatrix} c^2 & cs \\ cs & s^2 \end{pmatrix} & p_a &= p \\ \rho_b &= |b\rangle\langle b| = \begin{pmatrix} s^2 & cs \\ cs & c^2 \end{pmatrix} & p_b &= 1 - p \end{aligned}$$

Una generica mistura statistica ρ dei due stati è data da:

$$\rho = p\rho_a + (1 - p)\rho_b = \begin{pmatrix} s^2 + p \cos 2\theta & cs \\ cs & c^2 - p \cos 2\theta \end{pmatrix}$$

I cui autovalori sono dati da:

$$\lambda_{\pm} = \frac{1}{2}(1 \pm \sqrt{1 + 4p(p - 1) \cos^2 2\theta})$$

Un grafico di $\lambda_{\pm}(\theta)$ per alcuni valori di θ è riportato in figura 4.11a. Notiamo che per $\theta = 0$ (per cui $\langle a|b\rangle = 0$) gli autovalori sono p e $1 - p$, e si recupera il senso

classico. Per $\theta \neq 0$, tuttavia, i grafici di λ_- e λ_+ si “respingono” a vicenda, fino ad arrivare a $\theta = \pi/4$, dove gli autovalori sono costanti per ogni p (e pari a 1 o 0).

L'entropia di Von Neumann è dunque data da:

$$S_V(\rho) = -\lambda_+ \log \lambda_+ - \lambda_- \log \lambda_-$$

dove ρ è determinata dai parametri p e θ . Un grafico di $S_V(\rho)$ in funzione di p per alcuni valori di θ è riportato in figura 4.11b. Notiamo che per $\theta = 0$ si ottiene (come ci si aspettava) lo stesso andamento che ha $H(p)$ nel caso classico. Per $\theta = \pi/4$ si ha $S(\rho) \equiv 0$, dato che in tal caso $|a\rangle = |b\rangle$ e quindi non si ha trasmissione di informazione.

Per θ qualsiasi troviamo perciò che $S(\rho) \leq H(p)$. Si può dimostrare che tale disuguaglianza vale in generale, anche in sistemi più complicati. Possiamo interpretarla notando che la “somiglianza” tra $|a\rangle$ e $|b\rangle$ è proporzionale al loro prodotto scalare $\langle a | b \rangle = \sin 2\theta$, e ricevere stati “più simili tra loro” riduce, in un certo senso, l'ignoranza a priori che si può avere per il sistema. In effetti, al caso limite $\theta = \pi/4$ entrambi gli stati sono uguali, e quindi l'ignoranza è nulla: ad ogni istante non si ha alcun dubbio su quale potrà essere il prossimo stato ricevuto. Come visto precedentemente, l'ignoranza *a priori* è correlata con la quantità di informazione trasmessa da ogni qubit.

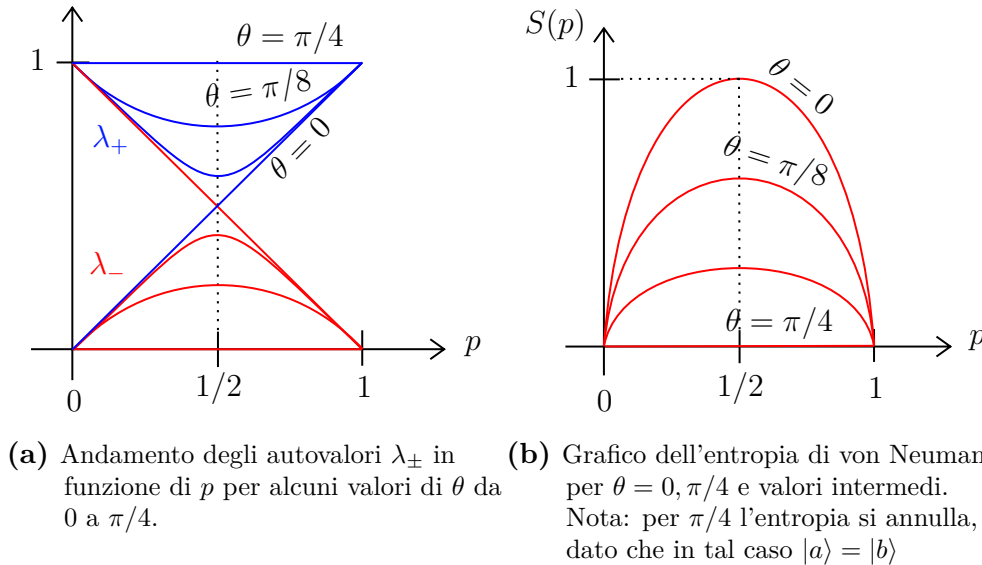


Figura (4.11) – Grafici relativi ad una mistura statistica di stati quantistici senza analogo classico (ossia non ortogonali)

4.5.1 Quantum Noiseless Coding

Il teorema di noiseless coding può essere generalizzato al caso quantistico facendo uso dell'entropia di Von Neumann.

(Lezione 12 ● del 4/4/2019)

Supponiamo che Alice trasmetta un **messaggio** di n lettere a Bob, scelte indipendentemente l'una dall'altra da un alfabeto $\mathcal{A} = \{|\psi_1\rangle, \dots, |\psi_k\rangle\}$, con $|\psi_i\rangle$ stati puri. Ciascuna lettera compare in una qualsiasi posizione del messaggio con probabilità p_i conosciuta a priori, tale che $p_i \geq 0$ e $\sum_i p_i = 1$. Il messaggio è quindi caratterizzato da un set di k probabilità $\{p_1, \dots, p_k\}$, e ciascuna lettera è descritta da una mistura statistica ρ dei k stati possibili con tali probabilità p_i :

$$\rho = \sum_{i=1}^k p_i |\psi_i\rangle \langle \psi_i|$$

Poiché abbiamo assunto che tutte le lettere siano indipendenti l'una dall'altra, l'intero messaggio ρ^n non è altro che il prodotto tensore di n stati di lettera singola, tutti pari a ρ :

$$\rho^n = \rho^{\otimes n} = \underbrace{\rho \otimes \rho \otimes \dots \otimes \rho}_{n \text{ volte}}$$

Teorema 4.5.1. *Dato un messaggio di n lettere costituite da stati puri scelti indipendentemente dall'alfabeto $\mathcal{A} = \{|\psi_1\rangle, \dots, |\psi_k\rangle\}$ con probabilità a priori $\{p_1, \dots, p_k\}$, esiste, asintoticamente nella lunghezza del messaggio ($n \rightarrow \infty$), un codice ottimale che comprime il messaggio a $S(\rho)$ qubit per lettera senza perdita di informazione, con $\rho = \sum_{i=1}^k p_i |\psi_i\rangle \langle \psi_i|$.*

Teorema del
quantum noiseless
coding di
Schumacher

Dimostrazione omessa.

In completa analogia con il caso classico, perciò, è possibile comprimere un messaggio (sufficientemente lungo) usando solo $S(\rho)$ qubit per lettera. Come visto nel precedente paragrafo, tuttavia, in generale $S(\rho) \leq H(p)$ - e perciò la MQ consente *rate di compressione maggiori* rispetto agli analoghi classici.

4.6 Caratterizzazione dell'Entanglement

Facendo uso degli strumenti appena introdotti, possiamo finalmente passare a *quantificare le correlazioni quantistiche*. Ci limiteremo al caso di correlazioni generate dall'*entanglement*, dato che sono quelle che compaiono nelle *applicazioni interessanti* come il teletrasporto quantistico o il dense coding. Oltre che dal punto di vista teorico, *quantificare l'entanglement* risulta importante anche sperimentalmente, dato che i limiti di misurazione rendono necessario lavorare con stati le cui correlazioni sono "sufficientemente ampie" da poter essere rilevate agevolmente.

4.6.1 Stati classicamente correlati

Consideriamo due sistemi A e B , in uno stato ρ_{AB} dato da una mistura statistica di stati separabili:

$$\rho_{AB}^S = \sum_i p_i \rho_i^A \otimes \rho_i^B \quad (4.9)$$

Matrici densità che hanno tale forma sono ovviamente *non entangled*, e presentano solamente *correlazioni classiche*. Si trova[11] infatti che una qualsiasi ρ_{AB} di questo tipo può essere prodotta da LOCC, ossia *Local operations & Classical Communications*. Ciò significa che, se il qubit A è presso Alice e il qubit B è presso Bob, il sistema AB può essere configurato in un qualsiasi ρ_{AB} del tipo (4.9) mediante sequenze di operazioni locali (ossia compiute da Alice e Bob sui rispettivi qubit) con la possibilità di comunicazioni classiche tra i due (Alice può comunicare a Bob l'esito di una sua misura).

Matrici densità non entangled: forma generale

Esempio: Alice con probabilità 50% prepara il suo qubit a $|0\rangle$, e nei casi restanti a $|1\rangle$. Comunica poi a Bob la sua scelta, e Bob prepara il suo stato allo stesso modo. Lo stato del sistema è allora dato da:

$$\rho = \frac{1}{2} |00\rangle \langle 00| + \frac{1}{2} |11\rangle \langle 11|$$

Analogamente, se Alice e Bob preparano *in uno stato casuale* i rispettivi qubit si otterrà uno stato finale pari a:

$$\rho = \frac{1}{4} (|00\rangle \langle 00| + |01\rangle \langle 01| + |10\rangle \langle 10| + |11\rangle \langle 11|)$$

In entrambi i casi la ρ ottenuta è del tipo (4.9).

Nota: Esistono stati ρ che **non** possono essere scritti in tale forma:

$$\rho \neq \rho_{AB}^S$$

e che quindi presentano correlazioni non classiche.

4.6.2 Misura di Entanglement

Partiamo elencando alcune proprietà che vorremmo da una qualsiasi **misura di entanglement** $\rho \mapsto E(\rho)$ [12] per un sistema bipartito AB :

Caratteristiche di una misura di entanglement

1. Se ρ è **separabile** l'entanglement deve essere nullo:

$$\forall \rho \text{ separabile} \Rightarrow E(\rho) = 0$$

In altre parole, sistemi completamente indipendenti non sono entangled.

2. **Invarianza** per trasformazioni locali unitarie:

$$E(\rho) = E[(U_A \otimes U_B) \rho (U_A^\dagger \otimes U_B^\dagger)]$$

Ciò significa che operazioni unitarie sui singoli qubit (es. una rotazione) non possono variare la “quantità di entanglement” associata alla coppia di qubit.

3. L'entanglement **non può aumentare** a seguito di LOCC e operazioni di *sub-selection* (ossia operazioni che “scartano” alcuni degli stati). Tali operazioni, infatti, possono al più creare correlazioni classiche tra A e B , e perciò possono solo ridurre (o al limite mantenere costante) la “quantità di entanglement” associata alla coppia di qubit.

Più precisamente, l'azione di LOCC è schematizzabile come una trasformazione *separabile* $A_i \otimes B_i$ sullo stato iniziale ρ con probabilità p_i . In altre parole, Alice e Bob possono operare trasformazioni locali che sono correlate da probabilità comuni p_i . Dopo una di queste operazioni, si ottiene uno stato σ_i :

$$\sigma_i = \frac{1}{p_i} (A_i \otimes B_i) \rho (A_i^\dagger \otimes B_i^\dagger)$$

con una probabilità p_i data da:

$$p_i = \text{Tr}((A_i \otimes B_i) \rho (A_i^\dagger \otimes B_i^\dagger))$$

Perciò, uno stato iniziale ρ a seguito di LOCC diviene una mistura statistica degli stati σ_i , ciascuno pesato dalla relativa probabilità p_i :

$$\rho' = \sum_i p_i \sigma_i$$

Poiché le LOCC non generano entanglement, deve quindi essere:

$$E(\rho) \geq E(\rho')$$

Se ρ_{AB} è pura, una buona misura di entanglement è data direttamente dall'entropia di Von Neumann della matrice densità ridotta di uno (qualsiasi) dei due sottosistemi³:

*Entanglement per
stati puri*

$$E(\rho) = S_V(\rho_A)$$

Ciò è conseguenza del fatto che per uno stato puro, ossia $\rho_{AB} = |\psi\rangle\langle\psi|$ per un qualche $|\psi\rangle \in \mathcal{H}_{AB}$, è definita la decomposizione di Schmidt, e perciò le matrici ridotte dei due sottosistemi hanno una “struttura” precisa e ben definita.

Ciò non succede in generale se ρ_{AB} è una mistura statistica, ossia se ammette una scrittura del tipo:

$$\rho_{AB} = \sum_{j=1}^k p_j |\psi_j\rangle\langle\psi_j| \quad |\psi_j\rangle \in \mathcal{H}_{AB}$$

In questo caso non è nemmeno ben definita la famiglia di stati puri $\{|\psi_j\rangle\}$ che *compongono* ρ_{AB} . Per esempio, si può sempre diagonalizzare ρ_{AB} , trovando un'espressione equivalente:

$$\rho_{AB} = \sum_{n=1}^k \lambda_n |\lambda_n\rangle\langle\lambda_n| \quad |\lambda_n\rangle \in \mathcal{H}_{AB}$$

³ΛBennett, Bernstein, et. al., 1996

con $|\lambda_n\rangle$ autovettori di ρ_{AB} di autovalore λ_n . In generale, perciò, esistono *famiglie diverse di stati puri* che compongono lo stesso stato misto. Si hanno perciò diverse possibilità - completamente equivalenti - per calcolare le matrici densità ridotte, che potrebbero portare a differenti entropie di Von Neumann.

Nella misura dell'**entanglement of formation**, perciò, l'idea è di considerare *tutte* le possibili realizzazioni di ρ_{AB} , ossia la famiglia di tutte le possibili matrici ridotte che si possono ottenere scegliendo diversi stati per la composizione di ρ_{AB} . Procediamo per gradi. Scegliamo una generica famiglia $\{|\psi_j\rangle\langle\psi_j|\}$ di stati puri che forma ρ_{AB} . Allora con probabilità p_i il sistema si trova nello stato puro $|\psi_i\rangle\langle\psi_i|$, che ha matrice ridotta:

$$\rho_A^i = \text{Tr}_B(|\psi_i\rangle\langle\psi_i|)$$

e di cui è ben definita l'entropia di Von Neumann. Per la famiglia scelta, calcoliamo l'entanglement pesando le varie entropie con le loro probabilità:

$$E_F(\rho)_{\{|\psi_j\rangle\}} = \sum_i p_i S(\rho_A^i)$$

Ripetiamo questo conto per *tutte* le possibili famiglie che compongono ρ_{AB} , e definiamo l'entanglement di ρ_{AB} come il **minimo** valore ottenuto:

Entanglement per stati misti

$$E_F(\rho) \equiv \min_{\text{famiglie } |\psi_j\rangle} \sum_{i=1}^k p_i S(\rho_A^i)$$

Tale calcolo è in genere molto complesso. Fortunatamente, nel caso di un sistema a 2 qubit esiste un algoritmo, detto **concurrence**, che consente di ottenere un risultato per stati generici - come esaminiamo nel seguente esempio.

Esempio. Consideriamo il caso di Alice e Bob che producono uno stato usando solamente LOCC, cercando di creare una qualche correlazione. Per esempio supponiamo che A disponga di un generatore di numeri casuali (puramente classico) e produca un qubit $|0\rangle$ con $p = 0.5$ e $|1\rangle$ altrimenti, comunicando a B di fare lo stesso. Lo stato finale dei due qubit di Alice e Bob è una mistura statistica *classicamente correlata*:

$$\rho^c = \frac{1}{2} |00\rangle\langle 00| + \frac{1}{2} |11\rangle\langle 11| = \frac{1}{2} \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right)$$

dove la matrice è scritta nella base computazionale $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$.

Confrontiamo tale sistema con quello di uno **stato di Bell**, che sappiamo essere entangled:

$$\begin{aligned}
 |\psi\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\
 \Rightarrow \rho^b &= |\psi\rangle\langle\psi| = \frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|) = \\
 &= \frac{1}{2} \left(\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{array} \right)
 \end{aligned}$$

Notiamo che:

Caratteristiche di ρ^b e ρ^c

- In ρ^b compaiono termini fuori dalla diagonale (detti *coerenze quantistiche*) che non sono presenti in ρ^c . Ciò significa che nella base computazionale scelta sono *visibili* sovrapposizioni quantistiche di stati - ma ciò non consente di stabilire la presenza di correlazioni quantistiche (del resto ogni ρ è diagonale in un'opportuna base).
- ρ^b è uno **stato puro**, dato che per costruzione $\rho^b = |\psi\rangle\langle\psi|$. Lo si può notare anche dal fatto che $(\rho^b)^2 = \rho^b$ (e quindi la purità $\text{Tr}((\rho^b)^2) = 1$), oppure da $\text{rk } \rho^b = 1$, dato che tutti gli stati puri sono proiettori unidimensionali. D'altro canto, ρ^c è uno **stato misto**. Lo si nota dalla costruzione, poiché si è usato un generatore di numeri casuali per produrre *probabilità classiche*, oppure dal fatto che ha $\text{rk} > 1$ - essendo una matrice diagonale con due termini non nulli.

Esaminiamo le **correlazioni** delle misure dell'osservabile σ^z :

Correlazioni di un'osservabile diagonale

$$\text{corr}(\sigma^z) = \langle \sigma_A^z \sigma_B^z \rangle - \langle \sigma_A^z \rangle \langle \sigma_B^z \rangle$$

Se le misure sono completamente scorrelate, avremo $\langle \sigma_A^z \sigma_B^z \rangle = \langle \sigma_A^z \rangle \langle \sigma_B^z \rangle$ e quindi $\text{corr}(\sigma^z) = 0$.

Partiamo svolgendo i conti per lo stato ρ^b .

$$\langle \sigma_A^z \rangle = \text{Tr}(\sigma_A^z \rho_A^b)$$

dove ρ_A è la matrice ridotta:

$$\begin{aligned}
 \rho_A^b &= \text{Tr}_B \rho_{AB}^b = \sum_{i=1}^{\dim \mathcal{H}_B} \langle i |_B \rho_{AB}^b | i \rangle_B = \langle 0 |_B \rho_{AB}^b | 0 \rangle_B + \langle 1 |_B \rho_{AB}^b | 1 \rangle_B = \\
 &= \frac{1}{2} |0\rangle_A \langle 0|_A + \frac{1}{2} |1\rangle_A \langle 1|_A = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
 \end{aligned}$$

che corrisponde ad uno stato misto (dato che la purità $\text{Tr}(\rho_A^2) = 1/2 < 1$). Possiamo ora calcolare il valor medio cercato:

$$\langle \sigma_A^z \rangle = \text{Tr} \left[\frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] = \frac{1}{2} \text{Tr} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 0$$

Ciò ha senso, poiché ρ_A descrive uno stato che la metà delle volte ha “spin negativo” e per l'altra metà ha “spin positivo”, e perciò, mediamente, ha “magnetizzazione nulla”.

In maniera simmetrica otteniamo $\langle \sigma_B^z \rangle = \text{Tr}(\sigma_B^z \rho_B^b) = 0$.

D'altro canto, il valor medio delle misure simultanee è dato da:

$$\langle \sigma_A^z \sigma_B^z \rangle = \text{Tr}((\sigma_A^z \otimes \sigma_B^z) \rho_{AB}^b) = \text{Tr} \left[\left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right) \rho_{AB}^b \right] = 1$$

Si trova perciò che lo stato di Bell è completamente correlato:

$$\text{corr}(\sigma^z, \sigma^z)_{\rho^b} = \langle \sigma_A^z \sigma_B^z \rangle_{\rho^b} - \langle \sigma_A^z \rangle_{\rho^b} \langle \sigma_B^z \rangle_{\rho^b} = 1$$

Ripetiamo ora gli stessi calcoli nel caso dello stato classicamente correlato ρ^c .

La matrice densità ridotta ρ_A^c è data da:

$$\rho_A^c = \text{Tr}_B(\rho_{AB}^c) = \langle 0|_B \rho_{AB}^c |0\rangle_B + \langle 1|_B \rho_{AB}^c |1\rangle_B = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Otteniamo la stessa matrice densità ridotta di prima: $\rho_A^c = \rho_A^b$. In altre parole, misure su un singolo qubit nei due stati sono completamente equivalenti: non è possibile distinguere correlazioni classiche da correlazioni quantistiche con misure singole.

Allo stesso modo, la correlazione tra i due qubit è non nulla:

$$\langle \sigma_A^z \sigma_B^z \rangle_{\rho^c} = \text{Tr}(\sigma_A^z \sigma_B^z \rho_{AB}^c) = \text{Tr} \left[\left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right) \left(\begin{array}{cc|cc} 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 \end{array} \right) \right] = 1$$

E perciò anche qui:

$$\text{corr}(\sigma^z)_{\rho^c} = \langle \sigma_A^z \sigma_B^z \rangle_{\rho^c} - \langle \sigma_A^z \rangle_{\rho^c} \langle \sigma_B^z \rangle_{\rho^c} = 1$$

Ciò non sorprende: $\sigma_A^z \otimes \sigma_B^z$ è una matrice diagonale, e gli stati ρ^b e ρ^c hanno gli stessi valori sulla diagonale, per cui il valor medio $\langle \sigma_A^z \sigma_B^z \rangle$ è lo stesso in entrambi i casi.

Chiaramente non tutte le osservabili hanno questa caratteristica. Per esempio, esaminando le correlazioni di σ^x (misura di *spin* lungo un altro asse), di nuovo $\langle \sigma_{A,B}^x \rangle_{\rho_{b,c}} = 0$, ma stavolta $\langle \sigma_A^x \sigma_B^x \rangle_{\rho_b} = 1$, mentre $\langle \sigma_A^x \sigma_B^x \rangle_{\rho_c} = 0$. Perciò:

Correlazioni di osservabili non diagonali

$$\text{corr}(\sigma^x)_{\rho_b} = 1 \neq \text{corr}(\sigma^x)_{\rho_c} = 0$$

Tali correlazioni quantistiche sono quantificate dalla misura di entanglement.

Nel caso di ρ^b , che è uno stato puro, essa coincide con l'entropia di Von Neumann di una delle matrici ridotte:

Entanglement per ρ^b e ρ^c

$$S_V(\rho_A^b) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = 1$$

che è quello che ci si aspetta, dato che ρ_A^b è una mistura statistica equiprobabile, a cui quindi è associata una *massima incertezza* (ossia una massima informazione).

Caratterizzare l'entanglement di ρ^c , uno stato misto, è invece molto più complesso. Non si può procedere calcolando l'entropia di Von Neumann di una matrice ridotta - cosa che del resto porterebbe allo stesso risultato ottenuto per ρ^b , visto che $\rho_A^b = \rho_A^c$. Utilizzare l'*entanglement of formation* risulta poi molto complesso, dato che bisogna considerare tutte le possibili decomposizioni di ρ_{AB} . Introduciamo perciò il metodo di **concurrence**, che ci permette di arrivare al risultato nel caso di sistemi a soli 2 qubit. Ne esamineremo solo il processo, senza dimostrarlo.

Concurrence

Definiamo:

$$C(\rho) = \max(0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4)$$

dove i λ_i sono gli autovalori in ordine decrescente della matrice R data da:

$$R = \rho_{AB} \tilde{\rho}_{AB} \quad \tilde{\rho}_{AB} = (\sigma_y \otimes \sigma_y) \rho (\sigma_y \otimes \sigma_y)$$

con

$$(\sigma^y \otimes \sigma^y) = \left(\begin{array}{cc|cc} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{array} \right)$$

Concretizzando tutto ciò per ρ^c notiamo che $\tilde{\rho}^c = \rho^c$, da cui R è data da:

$$R = (\rho^c)^2 = \left(\begin{array}{cc|cc} 1/4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/4 \end{array} \right)$$

Da cui $\lambda_1 = \lambda_2 = 1/4$ e $\lambda_3 = \lambda_4 = 0$, e di conseguenza $C(\rho^c) = 0$.

D'altro canto, lo stesso calcolo per ρ_{AB}^b restituisce lo stesso risultato che abbiamo ottenuto precedentemente con l'entropia di Von Neumann della matrice densità ridotta:

$$C(\rho_{AB}^b) = 1$$

come ci si aspetta, dato che la *concurrence* è una misura di entanglement.

Infine, può essere interessante confrontare l'entropia di Von Neumann degli stati ρ^b e ρ^c :

Entropia di ρ^b e ρ^c

$$S_V(\rho^c) = -\sum p_i \log_2 p_i = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = \log_2 2 = 1$$

Ciò ha senso, dato che ρ^c è una mistura statistica di due stati, e quindi si ha un certo grado di *ignoranza a priori* sulle “lettere” di cui è composta. Del resto non si tratta di uno stato ad entropia massima, che si ottiene per una mistura equiprobabile di tutti e 4 i vettori della base computazionale.

D'altro canto, poiché gli autovalori di ρ_{AB}^b sono tutti nulli tranne uno, che è pari a 1, avremo:

$$S_V(\rho_{AB}^b) = 0$$

Infatti qui abbiamo **certezza** sullo stato in cui si trova il sistema (che infatti è puro).

4.6.3 Dense Coding e correlazioni classiche

Mostriamo con un esempio come sia impossibile implementare algoritmi che necessitano di correlazione quantistiche utilizzando solo correlazioni classiche.

Nello specifico, proviamo a implementare l'algoritmo di Dense Coding usando lo stato creato da una mistura statistica equiprobabile ρ^c degli stati $|00\rangle$, $|11\rangle$:

$$\rho^c = \left(\begin{array}{cc|cc} 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 \end{array} \right)$$

A seconda della scelta dei due bit da inviare, Alice effettua una certa operazione sul suo qubit. Per esempio, per trasmettere 00 applica l'identità (per cui ρ non

cambia), mentre per 01 applica σ^z . Lo stato prodotto è quindi:

$$\begin{aligned}\rho' &= (\sigma^z \otimes \mathbb{I})\rho^c(\sigma^z \otimes \mathbb{I})^\dagger = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{array} \right) \rho^c \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{array} \right) = \\ &= \left(\begin{array}{cc|cc} 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 \end{array} \right)\end{aligned}$$

che però è esattamente pari a ρ^c iniziale!

D'altro canto, per trasmettere 10 Alice applica σ^x , ottenendo:

$$\rho'' = (\sigma^x \otimes \mathbb{I})\rho^c(\sigma^x \otimes \mathbb{I})^\dagger = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10|)$$

Infine, per trasmettere 11 Alice applica σ^y , ma anche qui riottiene $\rho''' = \rho''$.

Perciò, nonostante Alice possa applicare 4 possibili operazioni, gli stati finali saranno solo 2, e perciò può essere trasmesso un solo bit classico - esattamente come il caso di una qualsiasi comunicazione classica.

Deduciamo perciò che i vantaggi offerti dal protocollo di dense coding sono dati dalla possibilità di utilizzare i maggiori gradi di libertà offerti dalle correlazioni **quantistiche**, che non hanno alcun analogo classico.

4.6.4 Esercizio 4

1. Data una matrice densità della forma $\rho = \rho_A \otimes \rho_B$ calcolare la matrice densità ridotta di A e B .
2. Partendo dallo stato $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2} \otimes |0\rangle$ applicare un CNOT e mostrare che $\rho_{12} \neq \rho_1 \otimes \rho_2$.
3. Calcolare le correlazioni $C = \langle xy \rangle - \langle x \rangle \langle y \rangle$ presenti in un sistema descritto da $\rho = \rho_A \otimes \rho_B$.

Soluzione

1. Consideriamo due generiche matrici densità ρ_A e ρ_B , che - nella base dei loro autostati - hanno espressione:

$$\rho_A = \sum_{j=1}^{\dim \mathcal{H}_A} p_j^A |j\rangle \langle j|; \quad \rho_B = \sum_{k=1}^{\dim \mathcal{H}_B} p_k^B |k\rangle \langle k|$$

Calcoliamo una matrice ridotta di $\rho = \rho_A \otimes \rho_B$ tracciando sugli stati di B :

$$\text{Tr}_B(\rho) = \sum_{i=1}^{\dim \mathcal{H}_B} |i\rangle_B \rho_A \otimes \rho_B |i\rangle_B$$

Dato che la traccia è la stessa comunque venga scelta la base, possiamo scegliere $\{|i\rangle_B\}$ base ON di \mathcal{H}_B che coincida con quella $\{|k\rangle_B\}$ degli autostati di ρ_B . Così facendo:

$$\begin{aligned} \text{Tr}_B(\rho) &= \sum_{i=1}^{\dim \mathcal{H}_B} \langle i|_B \left(\sum_{j=1}^{\dim \mathcal{H}_A} p_j^A |j\rangle_A \langle j|_A \right) \left(\sum_{k=1}^{\dim \mathcal{H}_B} p_k^B |k\rangle_B \langle k|_B \right) |i\rangle_B = \\ &= \sum_{i=1}^{\dim \mathcal{H}_B} \left(\sum_{j=1}^{\dim \mathcal{H}_A} p_j^A |j\rangle_A \langle j|_A \right) \left(\sum_{k=1}^{\dim \mathcal{H}_B} p_k^B \underbrace{|\langle i|k\rangle|^2}_{\delta_{ki}} \right) = \\ &= \left(\sum_{j=1}^{\dim \mathcal{H}_A} p_j^A |j\rangle_A \langle j|_A \right) \underbrace{\left(\sum_{i=1}^{\dim \mathcal{H}_B} p_i^B \right)}_{=1} \stackrel{(a)}{=} \rho_A \end{aligned}$$

dove in (a) abbiamo usato $\text{Tr} \rho_B = 1$.

Analogamente, per simmetria, si ottiene:

$$\text{Tr}_A(\rho) = \rho_B$$

Come ci si potrebbe aspettare, perciò, le matrici densità ridotte di uno stato separabile non sono altro che i due stati separati.

2. Ricordiamo che un gate CNOT inverte il secondo qubit se lo stato del primo è $|1\rangle$. In questo caso:

$$|\psi'\rangle = \text{CNOT} |\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Otteniamo perciò uno stato di Bell, massimamente entangled e non separabile. Sia ρ_{12} la matrice densità dello stato finale, con:

$$\rho_{12} = |\psi'\rangle \langle \psi'| = \frac{1}{2}(|00\rangle \langle 00| + |11\rangle \langle 11| + |00\rangle \langle 11| + |11\rangle \langle 00|)$$

Le matrici ridotte sono date da:

$$\rho_1 = \text{Tr}_2 \rho_{12} = \langle 0|_2 \rho_{12} |0\rangle_2 + \langle 1|_2 \rho_{12} |1\rangle_2 = \frac{1}{2}(|0\rangle \langle 0| + |1\rangle \langle 1|) = \frac{1}{2} \mathbb{I}_1$$

E analogamente:

$$\rho_2 = \text{Tr}_1 \rho_{12} = \frac{1}{2} \mathbb{I}_2$$

Evidentemente:

$$\rho_1 \otimes \rho_2 = \frac{1}{4} \mathbb{I} \neq \rho_{12} = \frac{1}{2} \left(\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{array} \right)$$

3. Siano X_A e Y_B una qualsiasi coppia di osservabili che agiscono rispettivamente sul sistema A o B , lasciando invariato l'altro. Calcoliamo allora la correlazione C data da:

$$C = \text{Tr}(\rho_{AB} X_A Y_B) - \text{Tr}(\rho_A X_A) \text{Tr}(\rho_B Y_B)$$

Focalizziamoci sul primo termine:

$$C_1 = \text{Tr}(\rho_A \otimes \rho_B X_A Y_B) = \text{Tr}(\rho_A X_A \otimes \rho_B Y_B)$$

Siano $\{|j\rangle_A\}$ e $\{|k\rangle_B\}$ basi ON rispettivamente di \mathcal{H}_A e \mathcal{H}_B . Possiamo quindi calcolare la traccia come:

$$\begin{aligned} C_1 &= \sum_{j=1}^{\dim \mathcal{H}_A} \sum_{k=1}^{\dim \mathcal{H}_B} \langle j|_A \langle k|_B (\rho_A X_A \otimes \rho_B Y_B) |j\rangle_A |k\rangle_B = \\ &= \left(\sum_{j=1}^{\dim \mathcal{H}_A} \langle j|_A \rho_A X_A |j\rangle_A \right) \left(\sum_{k=1}^{\dim \mathcal{H}_B} \langle k|_B \rho_B Y_B |k\rangle_B \right) = \\ &= \text{Tr}(\rho_A X_A) \text{Tr}(\rho_B Y_B) \end{aligned}$$

e perciò $C = 0$. Come ci si aspetterebbe le misure di osservabili su stati separabili sono del tutto scorrelate (chiaramente, l'esito di una misura su A non può influenzare B).

Algoritmi quantistici

(Lezione 13 ● del
17/4/2019)

Supponendo di riuscire a realizzare fisicamente un computer quantistico, è possibile attuare algoritmi *più efficienti* degli analoghi classici, almeno per certi problemi. Esamineremo alcuni di essi in questa sezione.

Al giorno d'oggi, tuttavia, non esiste un “compilatore quantistico”, ossia un sistema in grado di implementare automaticamente un problema classico in un sistema quantistico con un certo guadagno computazionale.

5.0.1 Algoritmo di Deutsch

Consideriamo una **scatola nera**, ossia un *modulo*, di contenuto ignoto, che prenda un certo input x (n bit) e restituisca un output $f(x)$ (a un solo bit), dove la funzione $f: \{0, 1\}^n \rightarrow \{0, 1\}$ - intesa come *deterministica* - appartiene a una delle due classi seguenti:

- **Bilanciata:** f mappa metà dei valori possibili di x in 0 e l'altra metà in 1. In altre parole, per un x scelto a caso, $f(x)$ è pari a 0 o a 1 con la stessa probabilità.
- **Costante:** l'output $f(x)$ è sempre pari a un certo valore $k \in \{0, 1\}$. In altre parole, $f(x)$ restituisce soli 0 (o soli 1) indipendentemente dall'input x inserito.

Senza fare ipotesi su come tale scatola nera sia effettivamente realizzata, vogliamo trovare un modo, il più efficiente possibile, per determinare a *quale* delle due classi appartenga f .

Nota: una generica funzione $f: \{0, 1\}^n \rightarrow \{0, 1\}$ per $n > 1$ può non appartenere ad una delle due classi. Perciò funzioni bilanciate o costanti sono classi “speciali” di funzioni.

Tuttavia, nel caso unidimensionale, con $f: \{0, 1\} \rightarrow \{0, 1\}$, vi sono solo 4 possibilità, 2 bilanciate e 2 costanti (tabella 5.1).

x	f_0	f_1	f_2	f_3
0	0	0	1	1
1	0	1	0	1

Tabella (5.1) – Vi sono solo 4 funzioni a 1-bit, 2 bilanciate (f_1 e f_2) e 2 costanti (f_0 e f_3)

Classicamente, una possibile strategia consiste semplicemente nell'esaminare $f(x)$ per vari valori di x . Sono necessari almeno 2 input per stabilire che f sia bilanciata, dato che basta trovare x_1 e x_2 t.c. $f(x_1) \neq f(x_2)$. Nel caso f sia costante, tuttavia, servono ben $2^{n-1} + 1$ tentativi per averne certezza, dato che una f bilanciata ammette solo 2^{n-1} scelte di input che corrispondano allo stesso output.

Nel caso quantistico, d'altro canto, l'algoritmo di Deutsch permette di accertare la classe di f esaminando in ogni caso *un solo* output. Vediamo come.

Partiamo dal caso (più semplice) di $f : \{0, 1\} \rightarrow \{0, 1\}$, per cui abbiamo già elencato tutte le possibilità in tabella 5.1. Consideriamo allora una versione quantistica dell'oracolo, ottenuta sostituendo ai bit 0, 1 i qubit $|0\rangle$ e $|1\rangle$ di una opportuna base computazionale (figura 5.1a).

Ogni computazione quantistica deve essere *reversibile*. Dobbiamo perciò *portarci dietro* l'input iniziale, come schematizzato in figura 5.1b.

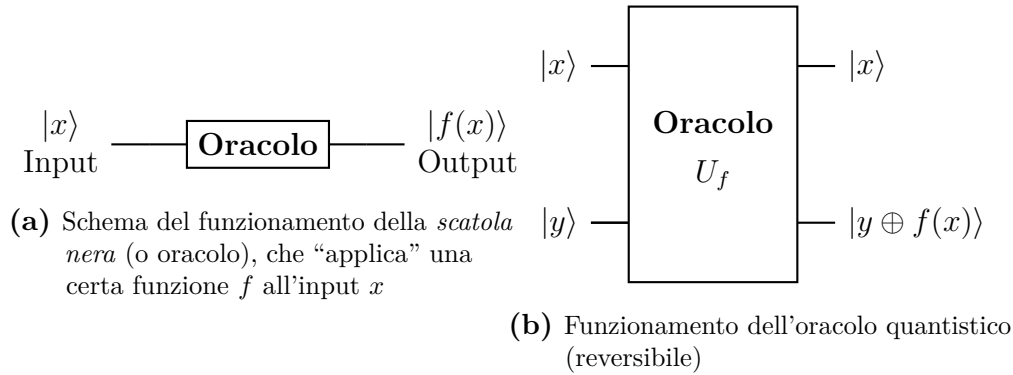


Figura (5.1) – Scatola nera: versione *classica* e *quantistica*

L'algoritmo di Deutsch consiste allora nel seguente circuito quantistico:

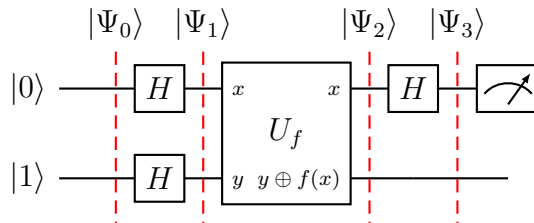


Figura (5.2) – Schema a porte logiche quantistiche dell'algoritmo di Deutsch ($n = 1$)

Esaminiamone il funzionamento. Consideriamo come input iniziale $|\Psi_0\rangle = |0\rangle |1\rangle$. Dopo l'applicazione delle Hadamard otteniamo:

$$\begin{aligned} |\Psi_1\rangle &= (H \otimes H) |\Psi_0\rangle = \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) = \\ &= \frac{1}{2}|0\rangle[|0\rangle - |1\rangle] + \frac{1}{2}|1\rangle[|0\rangle - |1\rangle] \end{aligned}$$

Ogni Hadamard trasforma un qubit in una sovrapposizione coerente dei due stati possibili. Come vedremo, è proprio questo passaggio che permette al computer quantistico di *esaminare tutti gli input* in una volta sola.

Calcoliamo l'azione di U_f su uno stato $|x\rangle[|0\rangle - |1\rangle]$. Ricordando la definizione (figura 5.1b), per linearità:

$$\begin{aligned} U_f |x\rangle[|0\rangle - |1\rangle] &= U_f |x\rangle|0\rangle - U_f |x\rangle|1\rangle = |x\rangle|f(x) + 0\rangle - |x\rangle|f(x) + 1\rangle = \\ &= |x\rangle[|f(x) + 0\rangle - |f(x) + 1\rangle] \end{aligned}$$

dove la somma binaria (+) è equivalente all'operazione XOR (\oplus).

Abbiamo due possibilità $f(x) = \pm 1$. Calcoliamole separatamente:

- Per $f(x) = 0$:

$$|f(x) + 0\rangle - |f(x) + 1\rangle = |0\rangle - |1\rangle$$

- Per $f(x) = 1$:

$$|f(x) + 0\rangle - |f(x) + 1\rangle = |1\rangle - |0\rangle = -(|0\rangle - |1\rangle)$$

Mettendo tutto insieme:

$$|x\rangle[|0\rangle - |1\rangle] \xrightarrow{U_f} (-1)^{f(x)} |x\rangle[|0\rangle - |1\rangle]$$

Possiamo ora calcolare $U_f |\Psi_1\rangle$, sovrapponendo i casi per i due valori di $|x\rangle$:

$$\begin{aligned} |\Psi_2\rangle &= \frac{1}{2}[(-1)^{f(0)}|0\rangle[|0\rangle - |1\rangle] + (-1)^{f(1)}|1\rangle[|0\rangle - |1\rangle]] = \\ &= \frac{1}{2}[(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle]_1 \otimes [|0\rangle - |1\rangle]_2 \end{aligned}$$

E infine resta da calcolare l'ultimo Hadamard:

$$\begin{aligned} |\Psi_3\rangle &= (H \otimes \mathbb{I}) |\Psi_2\rangle = \frac{1}{2\sqrt{2}}[(-1)^{f(0)}(|0\rangle + |1\rangle) + (-1)^{f(1)}(|0\rangle - |1\rangle)]_1 \otimes [|0\rangle - |1\rangle]_2 = \\ &= \frac{1}{2\sqrt{2}}[((-1)^{f(0)} + (-1)^{f(1)})|0\rangle + ((-1)^{f(0)} - (-1)^{f(1)})|1\rangle]_1 \otimes [|0\rangle - |1\rangle]_2 \end{aligned}$$

A questo punto misuriamo lo stato del primo qubit, che sarà $|0\rangle$ o $|1\rangle$ con probabilità date dal modulo quadro dei rispettivi coefficienti. Ma se $f(0) = f(1)$ (funzione

costante) si ha che $|1\rangle$ ha probabilità nulla, mentre se $f(0) \neq f(1)$ (funzione bilanciata) stavolta $|0\rangle$ ha probabilità nulla. Perciò, in un caso o nell'altro, troveremo **un solo risultato** con certezza - cosa che ci permette di distinguere da *una sola misura* la classe della funzione f .

Generalizziamo tutto ciò al caso di f (bilanciata o costante) che prende in input n qubit:

$$f : \{0, 1\}^n \rightarrow \{0, 1\}$$

Lo stato di partenza è, in questo caso, pari a:

$$|\Psi_0\rangle = \underbrace{|0\rangle |0\rangle \cdots |0\rangle}_n |1\rangle$$

Estendendo il circuito visto in figura 5.2 (che ora è detto algoritmo di Deutsch-Jorza), ciascuno degli $n + 1$ input viene mandato in una Hadamard, dopo le quali otteniamo un nuovo stato:

$$|\Psi_1\rangle = (H^{\otimes n} \otimes H) |\Psi_0\rangle = \frac{1}{\sqrt{2^n}} \left(\sum_{x=0}^{2^n-1} |x\rangle_n \right) \frac{(|0\rangle - |1\rangle)_{n+1}}{\sqrt{2}}$$

dove con $|x\rangle_n$ si intende lo stato di n qubit $|x_0\rangle |x_1\rangle |x_2\rangle \otimes \cdots \otimes |x_{n-1}\rangle$, con x_i i -esima cifra binaria di x .

L'estensione naturale dell'azione dell'oracolo a più qubit risulta in:

$$|\Psi_2\rangle = U_f |\Psi_1\rangle = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} (-1)^{f(x)} |x\rangle_n (|0\rangle - |1\rangle)_{n+1}$$

dove con $|x\rangle_n$ si intende lo stato dei primi n qubit, con la conversione di x da decimale a binario. Per esempio, per $n = 3$, $|3\rangle = |011\rangle_3$. L'ultimo ket, d'altro canto, riguarda solo l' $n + 1$ -esimo qubit.

Applichiamo ora n Hadamard a tutti gli input tranne l' $n + 1$ -esimo, ottenendo:

$$|\Psi_3\rangle = (H^{\otimes n} \otimes \mathbb{I}) |\Psi_2\rangle = \frac{1}{2^n} \left[\sum_{x,y=0}^{2^n-1} (-1)^{f(x)+x \cdot y} |y\rangle \right]_n \otimes \frac{1}{\sqrt{2}} [|0\rangle - |1\rangle]_{n+1}$$

dato che:

$$H^{\otimes n} |x\rangle = \prod_{i=0}^{n-1} \left(\frac{1}{\sqrt{2}} \sum_{y_i=0}^1 (-1)^{x_i y_i} |y_i\rangle \right) = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} (-1)^{x \cdot y} |y\rangle$$

dove $x \cdot y$ è il prodotto scalare tra i bit di x e y (che ha parità ben definita).

Misuriamo ora i primi n qubit (l' $n + 1$ -esimo può essere scartato). Di nuovo, avremo due casi:

- Per $f(x) = \text{costante}$ vi è un solo stato finale possibile:

$$|\Psi_4\rangle = \underbrace{|0\rangle \cdots |0\rangle}_n (|0\rangle - |1\rangle)_{n+1}$$

Per esempio, supponiamo $f(x) \equiv 0$. Allora lo stato dei primi n qubit prima della misura è:

$$|\Psi_3\rangle_n = \frac{1}{2^n} \sum_{x,y=0}^{2^n-1} (-1)^{x \cdot y} |y\rangle$$

La probabilità che $|0 \cdots 0\rangle_n$ venga selezionato è data dal modulo quadro del braket:

$$\langle \Psi_3 | 0 \cdots 0 \rangle = \frac{1}{2^n} \sum_{x=0}^{2^n-1} (-1)^0 \langle 0 \cdots 0 | 0 \cdots 0 \rangle = 1$$

dove abbiamo usato l'ortonormalità dei vettori della base computazionale. Poiché la probabilità di tale stato *esaurisce* tutte le possibilità, tale stato è l'unico possibile. Lo stesso risultato si ottiene nel caso sia $f(x) \equiv 1$, dove il braket è -1 , ma ciò non cambia il suo modulo quadro (la probabilità), che è sempre pari a 1 (certezza).

- Per $f(x)$ bilanciata, invece, la probabilità di ottenere $|0 \cdots 0\rangle_n$ è identicamente nulla, dato che:

$$\langle \Psi_3 | 0 \cdots 0 \rangle = \frac{1}{2^n} \sum_{x=0}^{2^n-1} (-1)^{f(x)} = \frac{1}{2^n} [(-1)^{f(0)} + (-1)^{f(1)} + \cdots + (-1)^{f(2^n-1)}]$$

e $f(x) = 0$ esattamente metà delle volte, e pari a 1 altrimenti. Si ottiene perciò una somma di $2^n/2$ termini $+1$ e altrettanti termini -1 , che porta chiaramente a 0.

Ogni altro stato $|\phi_n\rangle \perp |0 \cdots 0\rangle$ è possibile, e perciò lo stato dopo la misura sarà uno di questi:

$$|\Psi_4\rangle = |\phi\rangle_n (|0\rangle - |1\rangle)_{n+1}$$

Perciò un computer quantistico è in grado, con una sola misura, di distinguere la classe della funzione.

Nota: il problema appena esaminato richiede $2^{n-1} + 1$ misure classiche, o una sola misura quantistica. Il guadagno computazionale, nel caso quantistico, sembrerebbe esponenziale.

In realtà stiamo ignorando un fatto importante: $2^{n-1} + 1$ misure classiche sono richieste nel caso *peggiore*. In generale, se immaginiamo che gli input x siano mappati “uniformemente” a 0 e 1, le prime 2^{n-1} misure producono lo stesso output solo in due casi su 2^n , ossia con probabilità 2^{-n+1} . Mediamente, perciò, sono richieste *molte meno* misure classiche per poter discriminare la classe di f , ma comunque almeno 2. In ogni caso, perciò, l'algoritmo di Deutsch permette un significativo guadagno computazionale (nell'ipotesi che la *black-box* sia implementata efficientemente).

5.1 Algoritmo di Grover

Un altro problema che può essere risolto in maniera efficiente da un computer quantistico è la ricerca in un database non strutturato.

Possiamo schematizzare un semplice **database** come un set A di *stringhe* indicizzate dalla loro *posizione*, ossia una mappa $g: \mathbb{N} \supset U \rightarrow A$. Il database si dice *strutturato* se gli indici rispettano una relazione d'ordine su A , ossia se esiste un modo per *comparare* le stringhe e vale:

$$x_1 \leq x_2 \Rightarrow U, g(x_1) \leq g(x_2) \quad \forall x_1, x_2 \in U$$

In un database non strutturato, d'altro canto, non esiste alcuna relazione di questo tipo, e l'associazione $x \mapsto g(x)$ è completamente arbitraria.

Data una stringa $a \in A$, vorremmo trovare un modo efficiente per trovare il suo indice $x_a \in U$ tale che $g(x_a) = a$.

In un **database strutturato** ciò si può fare con l'algoritmo classico di *bisezione*:

1. Si esamina la stringa $a_1 = g(x_1)$ nella posizione x_1 centrale del database. Se $a_1 = a$ la ricerca è terminata.
2. Se $a \neq a_1$, si esamina l'ordine relativo tra a e a_1 . Se a è a destra della stringa centrale a_1 , allora sappiamo per certo che l'indice cercato è $> x_1$, e quindi compreso nella metà destra del database. Analogamente, se $a < a_1$, l'indice sarà nella metà sinistra.
3. Si esamina la stringa corrispondente all'indice intermedio della metà precedentemente considerata. Reiterando gli ultimi due punti è possibile *dimezzare* di volta in volta la regione da esaminare, e dopo pochi passaggi ($O(\log_2 N)$) si giunge alla soluzione.

D'altro canto, la ricerca in un **database non strutturato** è decisamente meno efficiente. L'unico modo è infatti procedere *sequenzialmente*, e in media dovremo esaminare $N/2$ casi. L'algoritmo è quindi di ordine $O(N/2)$.

Quantisticamente, l'algoritmo di Grover permette un guadagno *quadratico*, passando da $O(N)$ a $O(\sqrt{N})$. Vediamo come.

Ridefiniamo il problema in termini di qubit. Sia \bar{x} la posizione cercata (vista come stringa di n bit). Definiamo una funzione $f(x)$ (detta *oracolo*) che indichi se una posizione x esaminata è proprio quella \bar{x} della stringa cercata:

$$f(x) : \{0, 1\}^n \rightarrow \{0, 1\} \quad f(x) = \begin{cases} 1 & x = \bar{x} \\ 0 & x \neq \bar{x} \end{cases}$$

La ricerca di una stringa è quindi equivalente a trovare l'unico valore \bar{x} (tra tutti gli indici x possibili) tale che $f(\bar{x}) = 1$.

Supponiamo, per semplicità, di avere indici di 2 qubit, per un totale di $2^2 = 4$ possibili entrate nel database. In tal caso, l'algoritmo di Grover è implementato dal circuito di figura 5.3.

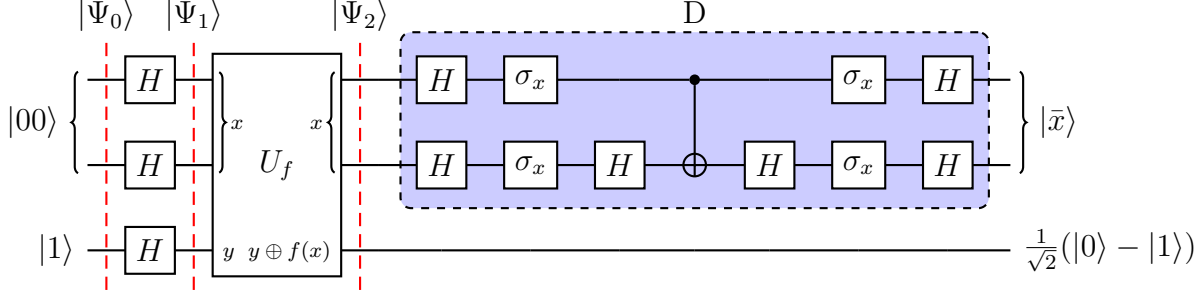


Figura (5.3) – Schema in porte logiche quantistiche dell'algoritmo di Grover per indici a 2 qubit

Partiamo dallo stato iniziale:

$$|\Psi_0\rangle = |00\rangle_{12} |1\rangle_3$$

dove i primi 2 qubit codificano il primo indice del database, e il terzo è un *qubit ancilla* necessario per la computazione reversibile.

Analogamente a quanto visto per l'algoritmo di Deutsch, applichiamo ad ogni qubit un gate Hadamard per realizzare una sovrapposizione di tutti gli stati possibili, giungendo a:

$$|\Psi_1\rangle = H^{\otimes 3} |\Psi_0\rangle = \frac{1}{2} \left[|00\rangle + |01\rangle + |10\rangle + |11\rangle \right]_{12} \frac{1}{\sqrt{2}} \otimes \left[|0\rangle - |1\rangle \right]_3$$

L'oracolo $f(x)$ è implementato dal gate U_f , che per risultare reversibile agisce come:

$$|x\rangle_{12} |y\rangle_3 \xrightarrow{U_f} |x\rangle_{12} |y \oplus f(x)\rangle_3$$

Nota: Poiché l'oracolo è parte della definizione del problema, non affrontiamo il discorso di come possa essere realizzato nella pratica.

Si hanno ora due casi. Per $f(x) = 0$ (ossia per la maggior parte delle x):

$$U_f |x\rangle_{12} \otimes \left[|0\rangle - |1\rangle \right]_3 = |x\rangle_{12} \otimes \left[|0\rangle - |1\rangle \right]_3$$

Mentre per l'unico \bar{x} per cui $f(x) = 1$:

$$U_f |\bar{x}\rangle_{12} \otimes \left[|0\rangle - |1\rangle \right]_3 = -|x\rangle_{12} \otimes \left[|0\rangle - |1\rangle \right]_3$$

Mettendo tutto insieme:

$$U_f |x\rangle_{12} \otimes \left[|0\rangle - |1\rangle \right]_3 = (-1)^{f(x)} |x\rangle_{12} \otimes \left[|0\rangle - |1\rangle \right]_3$$

Perciò lo stato dopo l'applicazione dell'oracolo è dato da:

$$\begin{aligned} |\Psi_2\rangle &= \frac{1}{2} \sum_{x=0}^{2^n-1} (-1)^{f(x)} |x\rangle_{12} \otimes \frac{1}{\sqrt{2}} [|0\rangle - |1\rangle]_3 = \\ &\stackrel{(a)}{=} \frac{1}{2} [|00\rangle + |01\rangle - |10\rangle + |11\rangle]_{12} \otimes [|0\rangle - |1\rangle]_3 \end{aligned}$$

dove in (a) abbiamo ipotizzato (senza perdita di generalità) che $x_0 = 10$.

L'applicazione di U_f , perciò, ha il solo effetto di aggiungere una fase al termine della sovrapposizione che corrisponde all'indice cercato (il terzo qubit può essere scartato). Notiamo che una misura di $|\Psi_2\rangle_{12}$ non porta alla soluzione, dato che i primi 2 qubit si trovano in una sovrapposizione equiprobabile di stati. Possiamo però manipolare $|\Psi\rangle_{12}$ in modo da “amplificare” la probabilità che una misura selezioni il solo stato con fase differente dagli altri. A tal proposito, esiste una trasformazione unitaria che mappa (a meno di un fattore $1/2$ di normalizzazione):

$$(-1, 1, 1, 1)^T \mapsto \hat{e}_1; \quad (1, -1, 1, 1)^T \mapsto \hat{e}_2; \quad (1, 1, -1, 1)^T \mapsto \hat{e}_3; \quad (1, 1, 1, -1)^T \mapsto \hat{e}_4$$

che è proprio data dalla matrice D che ha per colonne i vettori da trasformare¹:

$$D = \frac{1}{2} \left(\begin{array}{cc|cc} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ \hline 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{array} \right)$$

Scrivendo la $|\Psi_2\rangle$ in forma vettoriale possiamo calcolare lo stato finale applicando la matrice al vettore:

(Lezione 14 ● del 18/4/2019)

$$|\Psi_3\rangle = D |\Psi_2\rangle = D \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = |10\rangle$$

A questo punto, una misura dei 2 qubit rivela l'indice cercato.

Nel dettaglio, osserviamo che la matrice D può essere decomposta in un blocco D' che agisce sui 2 qubit, delimitato da due applicazioni di Hadamard:

$$D = H^{\otimes 2} D' H^{\otimes 2}$$

dove D' , come rappresentato nel circuito, è pari, a meno di una fase globale, a:

$$D' = \sigma_x^{\otimes 2} (\mathbb{I} \otimes H) \text{CNOT} (\mathbb{I} \otimes H) \sigma_x^{\otimes 2}$$

¹ΛIn generale, una matrice ha come colonne le immagini dei vettori della base canonica - per cui D mappa $\hat{e}_1 \mapsto (-1, 1, 1, 1)^T/2$ (e così via), che è proprio l'inverso della trasformazione che stiamo cercando. Tuttavia si tratta di una matrice unitaria ($DD^\dagger = \mathbb{I}$) ed hermitiana ($D = D^\dagger$), per cui $D^{-1} = D$.

con $\sigma_x^{\otimes 2}$ che corrisponde all'operazione NOT effettuata su entrambi i qubit. In D' , il blocco centrale rappresenta una CPHASE(π) (anche detta CMINUS):

$$(\mathbb{I} \otimes H) \text{CNOT}(\mathbb{I} \otimes H) = \frac{1}{2} \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{array} \right) \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array} \right) \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{array} \right) = \\ = \text{diag}(1, 1, 1, -1)$$

Nota: dato che si tratta di matrici suddivise in blocchi *conformi*, i prodotti matriciali si possono svolgere rapidamente moltiplicando tra loro blocchi 2×2 corrispettivi.

La CPHASE(π) aggiunge una fase -1 al solo stato $|11\rangle$. Trasformando il blocco con due NOT si ottiene D' . A livello di matrici, $(\sigma_x \otimes \sigma_x) \text{CMINUS}(\sigma_x \otimes \sigma_x)$ equivale a invertire l'ordine sia delle righe che delle colonne, giungendo a:

$$D' = \left(\begin{array}{cc|cc} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right) \stackrel{(a)}{=} \text{diag}(1, -1, -1, -1) = -\mathbb{I} + 2|00\rangle\langle 00|$$

dove in (a) abbiamo evidenziato e rimosso una fase globale $e^{i\pi}$ per semplicità di calcolo. La D' così ottenuta, perciò, aggiunge una fase -1 solo² allo stato $|00\rangle$ (=NOT $|11\rangle$).

La matrice D' è trasformata a sua volta da $H^{\otimes 2}$:

$$D = (H^{\otimes 2})(-\mathbb{I} + 2|00\rangle\langle 00|)H^{\otimes 2}$$

Notando che $H^{\otimes 2}H^{\otimes 2} = \mathbb{I}$ (dato che è una matrice hermitiana unitaria), e che:

$$H^{\otimes 2}|00\rangle = |S\rangle; \quad \langle 00|(H^{\otimes 2})^\dagger = \langle 00|$$

dove $|S\rangle$ è la sovrapposizione equiprobabile di tutti gli stati della base computazionale:

$$|S\rangle = H^{\otimes 2}|00\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle$$

otteniamo:

$$D = -\mathbb{I} + 2|S\rangle\langle S| = -(\mathbb{I} - 2|S\rangle\langle S|) \equiv -R_{|S\rangle}$$

²ΛO, equivalentemente, a tutti gli stati che non sono $|00\rangle$

$R_{|S\rangle}$ è la **riflessione** rispetto all'iperpiano perpendicolare a $|S\rangle$, ossia la trasformazione (unitaria) che cambia il segno della componente di un vettore parallela a $|S\rangle$.

In effetti è possibile interpretare in modo simile anche l'azione dell'oracolo U_f . Dato che:

$$U_f: |x\rangle \mapsto (-1)^{f(x)} |x\rangle$$

e $f(x) = 1 \Leftrightarrow x = \bar{x}$, U_f agisce come l'identità su tutte le $|x\rangle$, eccetto che su $|\bar{x}\rangle$, a cui aggiunge una fase:

$$U_f = \mathbb{I} - 2 |\bar{x}\rangle \langle \bar{x}| \equiv R_{|\bar{x}\rangle}$$

$R_{|\bar{x}\rangle}$ è la riflessione rispetto all'iperpiano perpendicolare a $|\bar{x}\rangle$ (figura 5.4).

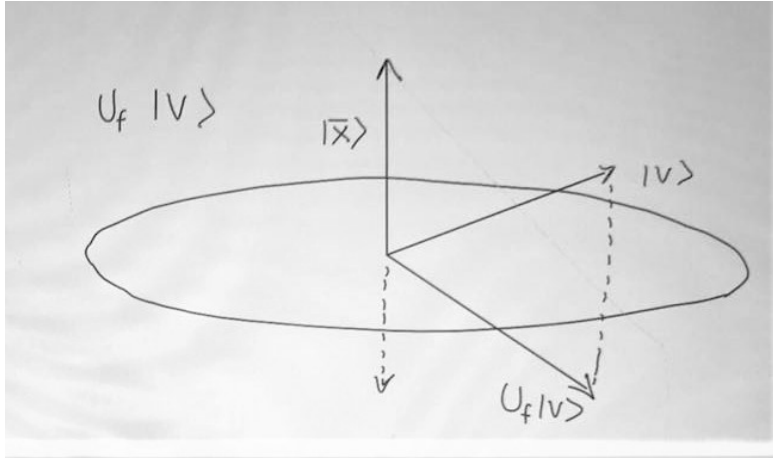


Figura (5.4) – Interpretazione geometrica dell'azione di U_f su un generico vettore $|v\rangle$

Ciò suggerisce un'interpretazione geometrica per l'algoritmo di Grover. Partiamo notando che $-R_{|S\rangle} = R_{|S^\perp\rangle}$. Infatti, un qualsiasi vettore $|v\rangle$ nell'iperpiano generato da $\{|S\rangle, |S^\perp\rangle\}$ si può scrivere come combinazione lineare:

$$|v\rangle = \alpha |S\rangle + \beta |S^\perp\rangle$$

Applicando la riflessione:

$$\begin{aligned} R_{|S\rangle} |v\rangle &= -\alpha |S\rangle + \beta |S^\perp\rangle \\ R_{|S^\perp\rangle} |v\rangle &= \alpha |S\rangle - \beta |S^\perp\rangle = -R_{|S\rangle} |v\rangle \end{aligned}$$

Mettendo tutto insieme, l'algoritmo di Grover è dato da³:

$$G = DU_f = R_{|S^\perp\rangle} R_{|\bar{x}\rangle}$$

³^Si ricorda che l'ordine di applicazione delle matrici è da destra a sinistra

Poniamoci allora nell'iperpiano generato da $\{|\bar{x}\rangle, |S\rangle\}$ (figura 5.5) su cui rappresentiamo anche i vettori (unitari) perpendicolari $|\bar{x}^\perp\rangle$ e $|S^\perp\rangle$. Consideriamo uno stato iniziale $|\psi\rangle$ che appartiene al piano, e si trova ad un angolo α rispetto a $|\bar{x}^\perp\rangle$. Applichiamo l'algoritmo di Grover:

1. L'azione dell'oracolo U_f è data da $R_{|\bar{x}\rangle}$, che inverte la componente di $|\psi\rangle$ lungo $|\bar{x}\rangle$ (riflessione rispetto a $|\bar{x}^\perp\rangle$).
2. L'azione di D è la riflessione $R_{|S^\perp\rangle}$, che porta $U_f |\psi\rangle$ a $G |\psi\rangle$, riflettendolo rispetto a $|S\rangle$.

Detto β l'angolo tra $|S\rangle$ e $G |\psi\rangle$, l'angolo tra $U_f |\psi\rangle$ e $G |\psi\rangle$ è 2β , o equivalentemente a $\beta + \theta + \alpha$:

$$2\beta = \beta + \theta + \alpha \Rightarrow \beta = \theta + \alpha$$

L'angolo tra $|\psi\rangle$ e $G |\psi\rangle$ è pari a $2\beta - 2\alpha$, ossia a 2θ . Perciò l'azione di D corrisponde ad una rotazione di $|\psi\rangle$ di un angolo 2θ .

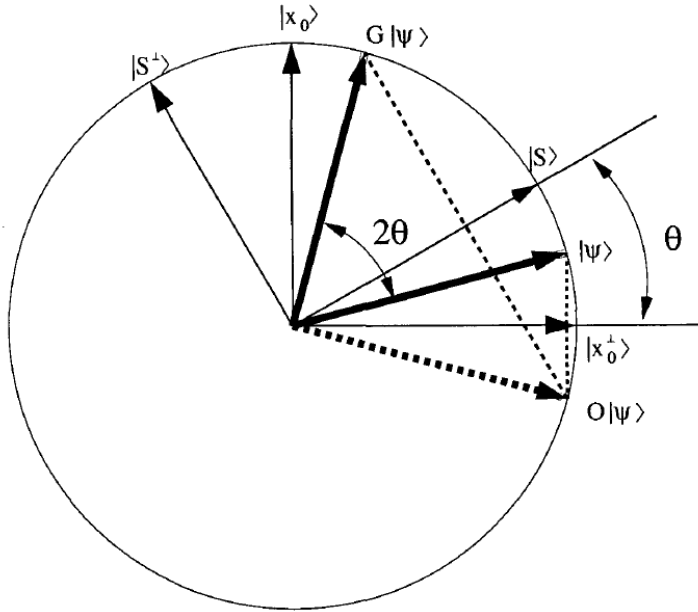


Figura (5.5) – Interpretazione grafica dell'algoritmo di Grover.

Poiché lo stato iniziale $|\psi_0\rangle$ è pari a $|S\rangle$ (è la sovrapposizione creata dalle due Hadamard iniziali) si ha:

$$|\psi_0\rangle = |S\rangle = \sin \theta |\bar{x}\rangle + \cos \theta |\bar{x}^\perp\rangle$$

dato che θ è, per definizione, l'angolo tra $|S\rangle$ e $|\bar{x}^\perp\rangle$. Poiché stiamo lavorando con 2 soli qubit, possiamo calcolare direttamente θ :

$$\cos\left(\frac{\pi}{2} - \theta\right) = \sin \theta = \langle S | \bar{x} \rangle = \frac{1}{\sqrt{2^2}} = \frac{1}{2} \Rightarrow \theta = 30^\circ$$

per qualsiasi scelta di \bar{x} , dato che $|S\rangle$ le comprende tutte con ugual probabilità. Ma allora, dopo una rotazione di $2\theta = 60^\circ$, il vettore $D|S\rangle$ ha un angolo $\theta = 90^\circ$ rispetto a $|\bar{x}^\perp\rangle$, e quindi $D|S\rangle = |\bar{x}\rangle$ - esattamente come previsto dal calcolo matriciale.

Notiamo allora che tutti i ragionamenti fatti possono essere **generalizzati** al caso di indici con n qubit, applicando n Hadamard ai primi n qubit (il ruolo dell'ancilla resta invariato):

*Generalizzazione
a n qubit*

$$|\Psi_1\rangle = H^{\otimes n} |0 \dots 0\rangle_n = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle_n \equiv |S\rangle$$

In tal caso, tuttavia, l'angolo θ_0 di partenza (ossia quello tra $|S\rangle$ e $|\bar{x}^\perp\rangle$) è dato da:

$$\sin \theta_0 = \frac{1}{\sqrt{2^n}} \leq \frac{1}{2} \Rightarrow \theta_0 \leq 30^\circ$$

Dopo l'applicazione dell'algoritmo di Grover, $|\Psi_1\rangle$ viene ruotato di $2\theta_0$ verso \bar{x} , ma per $n > 2$ ciò non è sufficiente a far sì che $|\Psi_1\rangle = |\bar{x}\rangle$. Risulta quindi necessario **ripetere** più volte l'algoritmo. In generale, dopo j ripetizioni di G lo stato del sistema è:

$$|\Psi_j\rangle \equiv G^j |S\rangle = \sin((2j+1)\theta_0) |\bar{x}\rangle + \cos((2j+1)\theta_0) |\bar{x}^\perp\rangle \quad (5.1)$$

Possiamo fermare l'algoritmo dopo un certo numero k di passi per cui:

$$(2k+1)\theta_0 \approx \frac{\pi}{2}$$

ossia tale che $|\psi_k\rangle \approx |\bar{x}\rangle$. Invertendo la relazione si trova:

$$k = \left\lfloor \frac{\pi}{4\theta_0} - \frac{1}{2} \right\rfloor$$

dove $\lfloor \dots \rfloor$ indica l'arrotondamento all'intero più vicino (poiché chiaramente non si può effettuare un numero frazionario di passi).

Per n sufficientemente grande:

$$\sin \theta_0 = \langle S | \bar{x} \rangle = \frac{1}{\sqrt{2^n}} \Rightarrow \theta_0 \approx \frac{1}{\sqrt{2^n}} \Rightarrow k = \left\lfloor \frac{\pi}{4} \sqrt{N} - \frac{1}{2} \right\rfloor = O(\sqrt{N})$$

dove $N = 2^n$ è la dimensione del database.

Possibilità di errore. Lo stato finale, prodotto dall'algoritmo di Grover dopo k passaggi, è molto vicino alla soluzione, ma non coincide con essa. Vi è quindi una certa probabilità p_{fail} che la misura finale non risulti nel valore desiderato \bar{x} . Dimostriamo ora che p_{fail} decade come $1/N$, ed è perciò in genere trascurabile. p_{fail} è pari al modulo quadro del coefficiente di $|\bar{x}^\perp\rangle$ in (5.1) dopo k iterazioni di G :

$$p_{\text{fail}} = |\cos((2k+1)\theta)|^2 \quad (5.2)$$

Per $N \gg 1$ si ha:

$$k = \left\lfloor \frac{\pi}{4} \sqrt{N} - \frac{1}{2} \right\rfloor \approx \frac{\pi}{4} \sqrt{N}; \quad \theta \approx \frac{1}{\sqrt{N}}$$

Sostituendo in (5.2) giungiamo a:

$$p_{\text{fail}} \approx \left| \cos \left(\frac{\pi}{2} + \frac{1}{\sqrt{N}} \right) \right|^2 \underset{(a)}{\approx} \left| \frac{\pi}{2} - \left(\frac{\pi}{2} - \frac{1}{\sqrt{N}} \right) \right|^2 = \frac{1}{N}$$

dove in (a) si è usata l'espansione di $\cos(x)$ attorno a $x = \pi/2$:

$$\cos(x) \underset{x \approx \pi/2}{=} \left(\frac{\pi}{2} - x \right) + O \left(x - \frac{\pi}{2} \right)^3$$

Al giorno d'oggi, tuttavia, l'algoritmo di Grover è solo una *proof of concept*. Il problema è implementare fisicamente l'oracolo U_f . L'idea è che U_f è determinata dal database in questione e dalla stringa cercata. Non esistendo un sistema generale per “convertire” un database classico e una stringa da cercare in un'implementazione quantistica di U_f , l'algoritmo di Grover non è facile da utilizzare.

Del resto, è possibile che l'implementazione fisica di U_f vari drasticamente a seconda della stringa cercata - e perciò risulti realizzabile solo “sapendo a priori la soluzione”, cosa che ne rimuove l'utilità pratica. Un altro problema è immagazzinare le entrate del database in una memoria quantistica - cosa che richiede di manipolare molti più qubit di quanto non sia possibile fare al giorno d'oggi.

5.2 Quantum Fourier Transform

Si può rappresentare una stringa di (qu)bit come combinazione lineare di una certa base di stringhe - la base *di Fourier*. Esplicitamente, data una sequenza di N termini complessi $\{f(0), f(1), \dots, f(N-1)\}$, la **trasformata di Fourier discreta** è una nuova sequenza di N numeri complessi $\{\tilde{f}(0), \tilde{f}(1), \dots, \tilde{f}(N_1)\}$ dati da:

$$\tilde{f}(k) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \exp \left(\frac{2\pi i}{N} jk \right) f(j) \quad (5.3)$$

Tale formula - a meno di fattori di normalizzazione convenzionali - non è altro che una versione *discreta* della trasformata di Fourier di una funzione.

Poiché si tratta di un operatore unitario, è possibile adattarla direttamente al caso di un registro di n qubit ($N = 2^n$):

$$\text{QFT}(|j\rangle) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \exp \left(2\pi i \frac{jk}{N} \right) |k\rangle \quad \forall |j\rangle \in \mathbb{C}^N \quad (5.4)$$

Ricaviamo un'espressione in termini di gate quantistici per la QFT.
Partiamo scrivendo k in rappresentazione binaria:

$$k = k_{n-1}2^{n-1} + \dots + k_02^0 = \sum_{t=0}^{n-1} k_t 2^t$$

dove k_i è l' i -esima cifra, con $i = 0$ che corrisponde alla cifra meno significativa (LSB) e $i = n - 1$ a quella più significativa (MSB).

Sostituendo in (5.4), scomponiamo la somma tra $k = 0$ e $k = N - 1$ in una serie di n somme annidate su ciascun bit:

$$\text{QFT}(|j\rangle) = \frac{1}{\sqrt{N}} \sum_{k_{n-1}=0}^1 \dots \sum_{k_0=0}^1 \exp\left(\frac{2\pi i j}{2^n} \sum_{t=0}^{n-1} 2^t k_t\right) |k_{n-1} \dots k_0\rangle \quad (5.5)$$

Manipoliamo la sommatoria all'interno dell'esponenziale:

$$\sum_{t=0}^{n-1} \frac{2^t k_t}{2^n} = \sum_{t=0}^{n-1} \frac{k_t}{2^{n-t}} \stackrel{(a)}{=} \sum_{l=1}^n \frac{k_{n-l}}{2^l}$$

dove in (a) scegliamo di usare come indice della sommatoria l'esponente del 2 al denominatore, ponendo $l = n - t$, da cui $t = n - l$, cosa che permetterà ad un passaggio successivo di introdurre la notazione binaria frazionale. Quando $t = 0$ si ha $l = n$, mentre quando $t = n - 1$, $l = n - (n - 1) = 1$, e perciò i nuovi estremi sono $l = 1$ e $l = n$. Nel sostituire in (5.5) trasformiamo una sommatoria ad esponente in un prodotto di esponenziali:

$$\text{QFT}(|j\rangle) = \frac{1}{\sqrt{N}} \sum_{k_{n-1}=0}^1 \dots \sum_{k_0=0}^1 \left[\prod_{l=1}^n \exp\left(2\pi i j \frac{k_{n-l}}{2^l}\right) \right] |k_{n-1} \dots k_0\rangle$$

Notando che:

$$|k_{n-1} \dots k_0\rangle = \bigotimes_{l=1}^n |k_{n-l}\rangle$$

Possiamo riscrivere:

$$\text{QFT}(|j\rangle) = \frac{1}{\sqrt{N}} \sum_{k_{n-1}=0}^1 \dots \sum_{k_0=0}^1 \bigotimes_{l=1}^n \exp\left(2\pi i j \frac{k_{n-l}}{2^l}\right) |k_{n-l}\rangle$$

Analogamente, le sommatorie annidate non sono altro che un prodotto di sommatorie, per cui:

$$\text{QFT}(|j\rangle) = \frac{1}{\sqrt{N}} \bigotimes_{l=1}^n \left[\sum_{k_{n-l}=0}^1 \exp\left(2\pi i j \frac{k_{n-l}}{2^l}\right) |k_{n-l}\rangle \right]$$

Non resta che svolgere esplicitamente l'unica sommatoria rimasta:

$$= \frac{1}{\sqrt{N}} \bigotimes_{l=1}^n \left[|0\rangle + \exp\left(2\pi i j \frac{1}{2^l}\right) |1\rangle \right]$$

Scriviamo infine j in notazione binaria:

$$j = j_{n-1}j_{n-2} \dots j_0 = \sum_{t=0}^{n-1} j_t 2^t$$

e introduciamo la notazione binaria frazionale:

$$0.j_l j_{l+1} \dots j_m = \frac{1}{2} j_l + \frac{1}{4} j_{l+1} + \dots + \frac{1}{2^{m-l+1}} j_m$$

In questo modo possiamo calcolare $j/2^l$:

$$\frac{j}{2^l} = j_{n-1}j_{n-2} \dots j_{l+1}.j_l \dots j_0$$

Fattorizzando l'esponenziale, notiamo che la parte **intera** non è importante:

$$\exp\left(2\pi i \frac{j}{2^l}\right) = \underbrace{\exp(2\pi i j_{n-1} \dots j_{l+1})}_{=1} \exp(2\pi i 0.j_l \dots j_0)$$

Possiamo così espandere il prodotto tensore:

$$\text{QFT}(|j\rangle) = \frac{1}{\sqrt{N}} \left[|0\rangle + \exp(2\pi i 0.j_0) |1\rangle \right]_{n-1} \otimes \left[|0\rangle + \exp(2\pi i 0.j_1 j_0) |1\rangle \right]_{n-2} \otimes \dots \otimes \quad (5.6)$$

$$\otimes \left[|0\rangle + \exp(2\pi i 0.j_{n-1} j_{n-2} \dots j_0) |1\rangle \right]_0 \quad (5.7)$$

Notiamo che gli stati dei singoli qubit sono ancora separabili: in altre parole, la QFT non produce *entanglement*.

L'espressione (5.7) può essere ora realizzata mediante gate elementari.

Partiamo notando che l' $n-1$ -esimo qubit della trasformata dipende solo dallo stato del primo qubit - $|j_0\rangle$ - e può essere ottenuto mediante un'Hadamard:

$$\text{QFT}(|j\rangle)_{n-1} \equiv |\tilde{j}\rangle_{n-1} = \frac{1}{\sqrt{2}} \left[|0\rangle + \exp(2\pi i 0.j_0) |1\rangle \right] = H |j_0\rangle$$

Infatti, j_0 può essere 1 o 0:

$$\begin{aligned} j_0 = 0 : \quad & |\tilde{j}\rangle_{n-1} = \frac{1}{\sqrt{2}} \left[|0\rangle + |1\rangle \right] = H |0\rangle \\ j_0 = 1 : \quad & |\tilde{j}\rangle_{n-1} = \frac{1}{\sqrt{2}} \left[|0\rangle + \exp(\pi i) |1\rangle \right] = \frac{1}{\sqrt{2}} \left[|0\rangle - |1\rangle \right] = H |1\rangle \end{aligned}$$

Il qubit successivo, l' $n-2$ -esimo, ha una fase più complessa. La parte $0.j_1$ si può ottenere applicando un'Hadamard a $|j_1\rangle$, ma resta un contributo $0.0j_0$ che richiede una C-PHASE controllata da $|k_0\rangle$. Definiamo, in generale, la C-PHASE di ordine k come:

$$R_k = \begin{pmatrix} \mathbb{I}_2 & \mathbb{O} \\ \mathbb{O} & R_z(2\pi i/2^k) \end{pmatrix}$$

Ne risulta che lo stato del qubit $n - 2$ -esimo è dato da:

$$|\tilde{j}\rangle_{n-2} = R_2^{[j_0]} H |j_1\rangle$$

Generalizzando, per $m \geq 1$:

$$|\tilde{j}\rangle_m = R_{m+1}^{[j_0]} \cdots R_3^{[j_{m-2}]} R_2^{[j_{m-1}]} H |j_m\rangle = \left(\prod_{l=0}^{m-1} R_{m+1-l}^{[j_l]} \right) H |j_m\rangle$$

Mettendo tutto insieme si ottiene il circuito rappresentato in figura 5.6. Si noti che l'ordine dei qubit dopo la QFT è invertito. Ciò può essere corretto con un'operazione di SWAP (di ordine $O(N)$), oppure semplicemente *rinominando* i qubit.

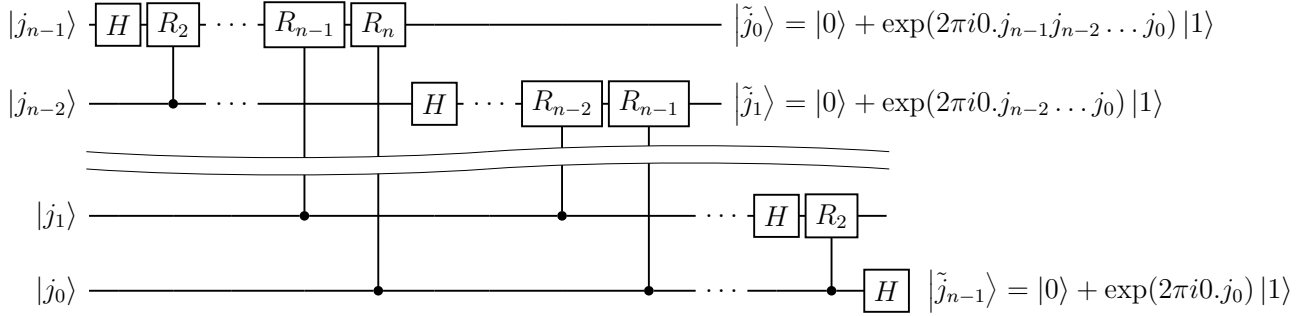


Figura (5.6) – Schema a porte logiche per la QFT

In totale stiamo usando n^2 porte logiche quantistiche, e perciò l'algoritmo è di ordine $O(n^2)$. Classicamente, l'ordine di una trasformata discreta è $O(N^2)$, e quello dell'implementazione più efficiente (FFT, Fast Fourier Transform) è $O(Nn)$. A prima vista, perciò, il guadagno quantistico sembrerebbe esponenziale. Va però ricordato che per estrarre l'informazione necessaria dallo stato finale sarà necessario fare misure di tomografia quantistica (per determinare le fasi relative), il cui numero cresce esponenzialmente con n . Ne risulta che, in sé, la QFT non offre particolari vantaggi.

5.3 Computer quantistici e crittografia classica

La trasformata di Fourier quantistica può però essere utilizzata all'interno di algoritmi quantistici più complessi - e in questo caso permette di ottenere dei *guadagni computazionali*.

In particolare, l'algoritmo di Shor permette di fattorizzare in maniera efficiente numeri primi. Ciò è estremamente utile, dato che permette di *violare* con maggiore facilità l'algoritmo (classico) di crittografia a chiave pubblica.

5.3.1 La crittografia RSA

Introduciamo una nuova tipologia di crittografia, detta **asimmetrica**. Creiamo due **chiavi**, una pubblica π e una privata k . La prima è necessaria per *criptare* il

messaggio, la seconda per *decriptarlo*:

$$E_{\pi}(P) = C; \quad D_k(C) = P$$

L'idea è che solo sapendo la chiave privata si possa decrittare il messaggio. Perciò la chiave pubblica può essere distribuita pubblicamente senza compromettere la sicurezza dei messaggi cifrati.

Ciò permette di risolvere (classicamente) il problema dello *scambio delle chiavi*. Consideriamo il seguente protocollo:

1. Bob crea π e k , e trasmette la chiave pubblica π ad Alice
2. Alice usa π per crittografare un messaggio: $E_{\pi}(P) = C$. Alice passa ora C a Bob.
3. Bob è l'unico che può decrittare il messaggio C , dato che è l'unico in possesso della chiave privata k , che non è mai stata inviata a nessuno:

$$D_k(C) = P$$

L'algoritmo RSA, alla base delle tecnologie moderne di sicurezza informatica (per esempio le connessioni sicure SSH), si basa sul modello di crittografia asimmetrica.

Costruzione dell'algoritmo RSA. L'idea alla base di RSA sta nell'uso di *one-way functions*, ossia funzioni che sono *facili* da calcolare, ma molto *difficili* (in senso di complessità algoritmica) da invertire. Una funzione di questo tipo è l'**esponenziale modulare**. Un numero P viene elevato a un numero intero e , e poi diviso per un altro intero N . Il risultato è il *resto* di tale divisione intera:

$$P^e \mod N = C$$

Si ha immediatamente $0 \leq C < P$. Questo è il passaggio di *codifica*, e la coppia (N, e) costituisce la **chiave pubblica**.

La funzione inversa dell'esponenziale modulare, detta *logaritmo discreto modulare* è molto difficile da computare. Cerchiamo ora una *trapdoor*, ossia un'informazione aggiuntiva che - se conosciuta - rende immediata l'inversione. Tale *trapdoor* costituirà la **chiave privata**, in possesso dell'unica persona che può decifrare C . Consideriamo allora un intero d che *realizzi l'inversione*, ossia che:

$$C^d \mod N = m$$

Unendo le due equazioni, ciò equivale a:

$$P^{ed} \mod N = P \tag{5.8}$$

Ci serve un modo per scegliere e , d ed N in modo che la trapdoor d sia *difficile* da ricavare conoscendo solo N ed e . Ciò è realizzato da una seconda *one-way function*, scelta per le sue proprietà in relazione all'esponenziale modulare.

L'idea è di sfruttare il problema della fattorizzazione in numeri primi. Mentre è facile calcolare $N = p_1 p_2$, con p_1 , p_2 primi, è molto più difficile trovare i fattori

conoscendo solo N .

Consideriamo ora la funzione $\varphi(n)$ (detta anche φ di Eulero, o *toziente*), che associa ad un numero n il numero di interi compresi tra 1 e n che sono coprimi con n . Per esempio, per $n = 8$, solo 1, 3, 5, 7 sono ≤ 8 e non condividono fattori con 8 (eccetto 1), e perciò $\varphi(8) = 4$. In generale, $\varphi(n)$ è difficile da calcolare - eccetto nel caso in cui n sia primo, in cui semplicemente $\varphi(n) = n - 1$. Di più, poiché φ è moltiplicativa, ossia $\varphi(ab) = \varphi(a)\varphi(b)$, per p_1, p_2 primi vale:

$$\varphi(p_1 p_2) = \varphi(p_1)\varphi(p_2) = (p_1 - 1)(p_2 - 1)$$

Possiamo mettere tutto insieme utilizzando il **teorema di Eulero** (che generalizza il piccolo teorema di Fermat):

$$P^{\varphi(n)} \equiv 1 \pmod{n}$$

Elevando entrambi i membri ad un intero k , poiché $1^k = 1$, vale:

$$P^{k\varphi(n)} \equiv 1 \pmod{n}$$

Moltiplicando per P :

$$P^{k\varphi(n)+1} \equiv P \pmod{n}$$

Confrontando con (5.8) troviamo allora:

$$k\varphi(n) + 1 = ed \Rightarrow d = \frac{k\varphi(n) + 1}{e}$$

Scegliendo $n = p_1 p_2$, con p_1, p_2 primi, solo conoscendo i fattori (difficili da ricavare sapendo solo N) possiamo calcolare facilmente $\varphi(n)$, e di conseguenza d - che funge da *trapdoor*. Del resto, k è scelto per far sì che d sia intero, mentre per e si può prendere un qualche numero $1 < e < \phi(n)$ coprimo con $\varphi(n)$.

Esaminiamo i passaggi dell'algoritmo.

1. Bob crea un numero N dato dal prodotto $p \cdot q$ di due numeri **primi** p, q sufficientemente grandi. Mentre generare numeri primi e moltiplicarli tra loro è efficiente, *fattorizzare* un numero grande in numeri primi è computazionalmente molto costoso. Perciò possiamo essere sicuri che, se uno sa solo N , non sarà in grado di ottenere p e q con facilità (il processo potrebbe richiedere con i computer odierni tempi di milioni di anni).

Bob calcola anche:

$$\Phi = (p - 1)(q - 1)$$

Inoltre sceglie un numero $1 < e < \Phi$ che sia *coprimo* con Φ (ossia che non condivide fattori $\neq 1$ con Φ) e infine calcola d tale che $d \cdot e = 1 \pmod{\Phi}$.

Tale d costituisce la **chiave privata**, mentre la coppia di numeri (e, N) è la **chiave pubblica** che viene passata ad Alice.

2. Alice, conoscendo (e, N) , può codificare messaggi nel seguente modo. Il messaggio da inviare ⁴ è codificato in una stringa di bit, ossia un numero P , e supponiamo che sia $P < N$. Allora il messaggio cifrato C si calcola come:

$$C = E_\pi(P) = P^e \mod N \quad (5.9)$$

Alice invia poi C a Bob.

3. Bob ha ora tutte le informazioni necessarie per decodificare C :

$$P = D_k(C) = C^d \mod N \quad (5.10)$$

Dimostrazione. Vogliamo dimostrare (5.10). Inserendo (5.9), ciò equivale a:

$$C^d \mod N = P^{ed} \mod N \stackrel{?}{=} P \quad (5.11)$$

e ciò deve valere $\forall P$.

Sappiamo che $p \neq q$ sono numeri primi, e che e, d sono numeri naturali tali che $ed \equiv 1 \mod \Phi$. Per definizione di modulo, $\exists k \in \mathbb{N}$ tale che:

$$ed = k\Phi + 1 \stackrel{(a)}{=} k(p-1)(q-1) + 1$$

dove in (a) si è usato $\phi = (p-1)(q-1)$. Riarrangiando:

$$ed - 1 = k(p-1)(q-1)$$

ossia $ed - 1$ è multiplo sia di $(p-1)$ che di $(q-1)$. Esplicitamente:

$$\exists h, j \in \mathbb{N} \mid ed - 1 = h(p-1) = j(q-1) \quad (5.12)$$

Poiché p e q sono primi, possiamo spezzare l'uguaglianza da verificare in due:

$$P^{ed} \mod pq \equiv P \Leftrightarrow \begin{cases} P^{ed} \mod p \equiv P \\ P^{ed} \mod q \equiv P \end{cases}$$

- Partiamo dal $\mod p$. Se $P \equiv 0 \mod p$, segue immediatamente che P^{ed} è multiplo di p , e quindi anche $P^{ed} \equiv 0 \equiv P \mod p$. Generalmente, però, $P \not\equiv 0$. In tal caso servono alcune manipolazioni:

$$P^{ed} \mod p = P^{ed-1}P \stackrel{(5.12)}{=} P^{h(p-1)}P = (P^{p-1})^h P \stackrel{(a)}{\equiv} 1^h m \equiv m \mod p$$

dove in (a) si è usato il piccolo teorema di Fermat (che non dimostriamo), per cui:

$$P^{p-1} \equiv 1 \mod p$$

⁴Generalmente RSA si usa per inviare in modo sicuro le chiavi di altri algoritmi a cifratura simmetrica, che hanno il vantaggio di essere molto più veloci.

In effetti, ciò equivale all'enunciato del teorema di Eulero, $P^{\varphi(p)} \equiv 1 \pmod{p}$, dato che p è primo e quindi $\varphi(p) = p - 1$.

- Per l'uguaglianza con \pmod{q} si ripete il procedimento. Escluso il caso banale di $P \equiv 0$, avremo:

$$P^{ed} = P^{ed-1}P \underset{(5.12)}{=} P^{j(q-1)}P = (P^{q-1})^jP \equiv 1^jP \equiv P \pmod{q}$$

Abbiamo allora mostrato che:

$$C^d \equiv P^{ed} \equiv P \pmod{N}$$

e poiché $P < N$, $P \pmod{N} = P \forall P$.

□

5.3.2 Algoritmo di Shor - parte classica

La sicurezza di RSA si basa quindi sul fatto che N non è facilmente fattorizzabile. Se riuscissimo a determinare i fattori p e q di N , infatti, avendo a disposizione anche e (che è pubblicamente distribuito), potremmo calcolare direttamente la chiave privata d .

L'algoritmo di Shor permette proprio di far ciò, con una complessità di ordine polinomiale - al contrario dell'ordine sub-esponenziale del miglior algoritmo classico per ora conosciuto.

L'idea alla base consiste nel trasformare il problema di fattorizzazione nel problema di trovare il periodo di una certa funzione, che può essere risolto in modo molto efficiente dalla QFT. Vediamo come.

1. Partiamo scegliendo un numero $N \ni x < N$, che deve essere coprimo con N , ossia tale che:

$$\text{MCD}(x, N) = 1$$

Per calcolare l'MCD si può usare l'**algoritmo di Euclide**, che ha ordine $O(\log N^2)$. Per due numeri $a > b$, si parte sottraendo b ad a il massimo numero di volte possibile (ciò equivale ad una divisione intera). Se il risultato è $c = 0$, allora a è un multiplo di b , e quindi $\text{MCD}(a, b) = b$. Altrimenti, per $c \neq 0$, si reitera l'operazione tra $b > c$, sottraendo il massimo numero di volte c a b , e così via. L'ultimo numero ottenuto prima di 0 è il risultato cercato. Per esempio⁵, per calcolare $\text{MCD}(1071, 1029)$:

$$1071 - 1 \cdot 1029 = 42$$

$$1029 - 21 \cdot 42 = 21$$

$$42 - 2 \cdot 21 = 0$$

Da cui $\text{MCD}(1071, 1029) = 21$.

⁵Una visualizzazione del funzionamento dell'algoritmo è disponibile a https://en.wikipedia.org/wiki/Euclidean_algorithm#Worked_example

2. Una volta trovato $x \in \mathbb{N}$ tale che $\text{MCD}(x, N) = 1$, cerchiamo il più piccolo $r \in \mathbb{N}$ che verifica:

$$x^r \equiv 1 \pmod{N} \quad (5.13)$$

e che denotiamo con \bar{r} . Equivalentemente, \bar{r} è il periodo della funzione:

$$f(r) = x^r \pmod{N}$$

ossia è tale che $f(0) = f(\bar{r}) = 1$. Come vedremo nella prossima sezione, questo problema è risolto in maniera efficiente da un computer quantistico. Se \bar{r} è dispari siamo di fronte ad una soluzione spuria, e risulta necessario ripetere dall'inizio l'algoritmo. Altrimenti, possiamo procedere al punto 3.

3. Poiché $r/2 \in \mathbb{N}$, possiamo definire $y = x^{r/2}$. Sostituendo in (5.13) si ha:

$$y^2 \equiv 1 \pmod{N} \Rightarrow y^2 - 1 \equiv 0 \pmod{N} \Rightarrow (y+1)(y-1) \equiv 0 \pmod{N}$$

Ciò significa che $\exists k \in \mathbb{N}$ t.c. $(y+1)(y-1) = kN$. Se $(y+1)$ o $(y-1)$ è un multiplo di N , ossia se $y \equiv -1 \pmod{N}$, è necessario ripetere l'algoritmo da capo. Altrimenti, uno tra $\text{MCD}(y+1, N)$ e $\text{MCD}(y-1, N)$ è un fattore non banale di N .

5.3.3 Algoritmo di Shor - parte quantistica

(Lezione 15 ● del 5/6/2019)

Concentriamoci sul punto 2, per cui esiste un algoritmo quantistico estremamente più efficiente di una qualsiasi alternativa classica attualmente conosciuta.

1. Consideriamo la funzione $f_a: \mathbb{N} \rightarrow [0, N] \cap \mathbb{N}$, con $a \in \mathbb{N}$, definita da:

$$f_a(x) = a^x \pmod{N}$$

Vogliamo trovare il periodo r , ossia il più piccolo numero $r \in \mathbb{N}$ per cui $f_a(x+r) = f_a(x)$. La sua implementazione quantistica (reversibile) è data dalla mappa:

$$|x\rangle_n |y\rangle_n \xrightarrow{U_f} |x\rangle_n |y \oplus f(x)\rangle_n$$

dove $|x\rangle_n$ e $|y\rangle_n$ sono due registri da n qubit ciascuno. Ciò permette di esaminare $N = 2^n$ input possibili. Fortunatamente U_f , come nel caso classico, è facilmente implementabile anche come gate quantistici.

Supponiamo, per semplicità, che il periodo r sia un sottomultiplo del *range del dominio* N , ossia che:

$$\exists m \in \mathbb{N} \mid \frac{N}{r} = m$$

In altre parole, $f(x)$ si ripete un numero intero di volte nell'insieme di input possibili.

Nota: l'algoritmo funziona anche senza tale ipotesi, ma compaiono alcune complicazioni matematiche più difficili da trattare, e che essendo di natura puramente tecnica non aggiungono nulla alla comprensione della procedura.

Inizializziamo il registro $|x\rangle_n$ con una sovrapposizione di tutti gli stati, mentre $|y\rangle_n = |0\rangle_n$:

$$|\Psi_0\rangle = (H^{\otimes n} |0\rangle_n) \otimes |0\rangle_n = \frac{1}{\sqrt{2^n}} \left(\sum_{x=0}^{2^n-1} |x\rangle_n \right) \otimes |0\rangle_n$$

Dopo l'applicazione di U_f giungiamo allo stato (figura 5.7):

$$|\Psi_1\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle_n |f(x)\rangle_n$$

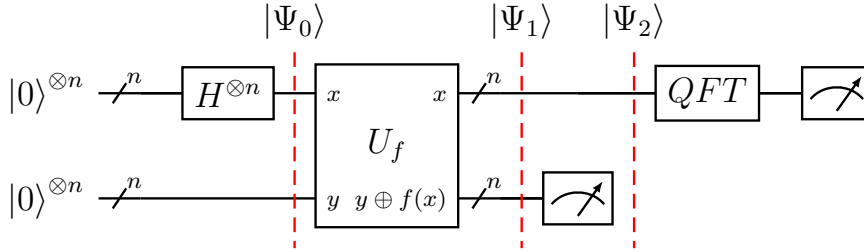


Figura (5.7) – Schema del circuito a porte logiche quantistiche per l'algoritmo di Shor

2. Notiamo che $|\Psi_1\rangle$ è uno stato **entangled** - e perciò gli stati dei registri $|x\rangle_n$ e $|y\rangle_n$ non sono separabili.

Misuriamo quindi il secondo registro, trovando un certo valore \bar{f} . Ciò proietta $|\psi\rangle$ nel prodotto tensore tra tutti i valori $|x\rangle$ tali che $f(x) = \bar{f}$ - che sono esattamente m , dato che $f(x)$ si ripete m volte nel set degli N possibili input - e il valore $|f\rangle$ stesso. Detto x_0 il primo valore per cui $f(x_0) = \bar{f}$, tutti i seguenti rispettano $x_n = x_0 + jr$ per la periodicità, con $j \in \mathbb{N}$. Perciò lo stato finale, una volta normalizzato, è dato da:

$$|\Psi_2\rangle = \frac{1}{\sqrt{m}} \sum_{j=0}^{m-1} |x_0 + jr\rangle_n |\bar{f}(x_0)\rangle_n = \left[\frac{1}{\sqrt{m}} \sum_{j=0}^{m-1} |x_0 + jr\rangle \right]_n \otimes |f(x_0)\rangle_n$$

3. Ora possiamo scartare il secondo ket, e concentrarci sul primo - che è sovrapposizione di stati tutti alla *stessa* distanza r tra loro. Per trovarla eseguiamo una Quantum Fourier Transform:

$$\text{QFT}\{|x\rangle_n\} = \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} \exp\left(\frac{2\pi i}{N} xy\right) |y\rangle$$

Per linearità, si ottiene:

$$|\Psi_3\rangle = \text{QFT}\{|\Psi_2\rangle\} = \frac{1}{\sqrt{m}} \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} \sum_{i=0}^{m-1} \exp\left(2\pi i(x_0 + jr) \frac{y}{N}\right) |y\rangle$$

Misuriamo lo stato, ottenendo un certo valore \bar{y} con probabilità:

$$\begin{aligned} P(\bar{y}) &= |\langle \Psi_3 | \bar{y} \rangle|^2 = \left| \frac{1}{\sqrt{Nm}} \sum_{j=0}^{m-1} \exp\left(\frac{2\pi i}{N} \bar{y}(x_0 + jr)\right) \right|^2 = \\ &= \underbrace{\left| \exp\left(\frac{2\pi i}{N} \bar{y}x_0\right) \right|^2}_{=1} \left| \frac{1}{\sqrt{Nm}} \sum_{j=0}^{m-1} \exp\left(\frac{2\pi i}{N} \bar{y}jr\right) \right|^2 = \\ &= \frac{1}{Nm} \frac{m^2}{m^2} \left| \sum_{j=0}^{m-1} \exp\left(\frac{2\pi i j \overbrace{r}^{N/m} \bar{y}}{N}\right) \right|^2 = \\ &= \underbrace{\frac{m}{N}}_{1/r} \left| \frac{1}{m} \sum_{j=0}^{m-1} \exp\left(\frac{2\pi i j \bar{y}}{m}\right) \right|^2 \end{aligned}$$

Si osserva che gli unici casi in cui $P(\bar{y}) \neq 0$ sono quelli in cui $\bar{y} = km$, con $k = 0, \dots, r-1$ (dato che $\bar{y} < N$, ossia $km < mr$), in cui $P(\bar{y}) = 1/r$:

$$P(km) = \frac{1}{r} \left| \frac{1}{m} \sum_{j=0}^{m-1} 1 \right| = \frac{1}{r} \quad k \in \{0, \dots, r-1\}$$

Infatti, in totale k può assumere r valori, ciascuno con probabilità $1/r$. Perciò la probabilità che lo stato finale sia uno qualsiasi in cui $\bar{y} = km$ è 1 (certezza), e quindi non vi sono altri esiti ammessi.

Ma allora una qualsiasi misura ritorna un \bar{y} , cioè:

$$\bar{y} = k \frac{N}{r}$$

Riarrangiando:

$$\frac{\bar{y}}{N} = \frac{\bar{k}}{r}$$

Se $k = 0$, cosa che succede con $p = 1/r$, non abbiamo ottenuto alcuna informazione su r , ed è quindi necessario ripetere l'algoritmo.

D'altro canto, se $\text{MCD}(k, r) = 1$, possiamo ridurre ai minimi termini la frazione \bar{y}/N , ottenendo come denominatore esattamente l' r cercato. Si dimostra che ciò avviene con $p \geq 1/(\log \log r)$.

Se così non è, risulta necessario ripetere l'algoritmo, fino a quando non si verifica $\text{MCD}(k, r) = 1$ (o fino a quando non si hanno sufficienti informazioni per ricavare r).

Da considerazioni sulla complessità di U_f , della QFT, e sulla probabilità di successo, si ottiene che l'algoritmo di Shor è di ordine $O(n^2 \log n \log \log n)$ (con $n = \log N$). Per confronto, il miglior algoritmo classico risulta in un $e^{O(n^{1/3}(\log n)^{1/3})}$.

5.3.4 Esempio

Applichiamo l'algoritmo appena esaminato ad un caso semplice. Definiamo una funzione periodica come:

$$f(x) = \frac{1}{2}(\cos \pi x + 1)$$

Poniamo $N = 2^3 = 8$, per cui lavoriamo con 3 qubit:

$$f: \{0, 1\}^3 \rightarrow \{0, 1\}$$

Si trova:

$$\begin{aligned} f(0) &= f(2) = f(4) = f(6) = 1 \\ f(1) &= f(3) = f(5) = f(7) = 0 \end{aligned}$$

E perciò $r = 2$, da cui $m = N/r = 4$.

1. Prepariamo lo stato iniziale:

$$|\psi_1\rangle = \frac{1}{\sqrt{8}} \sum_{x=0}^7 |x\rangle_1 |f(x)\rangle_2$$

2. Misuriamo la seconda componente di $|\psi_1\rangle$ e troviamo (per esempio) $|0\rangle_2$. Lo stato risultante è allora:

$$|\psi_2\rangle = \frac{1}{2}(|1\rangle + |3\rangle + |5\rangle + |7\rangle)_1 \otimes |0\rangle_2$$

3. Applichiamo la QFT alla prima componente. Esplicitamente, uno stato generico $|j\rangle$ viene così mappato in:

$$|j\rangle \mapsto \frac{1}{\sqrt{8}} \sum_{k=0}^7 \exp\left(\frac{2\pi i j k}{8}\right) |k\rangle$$

Lo stato risultante è allora dato da:

$$\begin{aligned} |\psi_3\rangle &= \frac{1}{\sqrt{8}} \frac{1}{2} \left\{ |0\rangle + \exp\left(\frac{i\pi}{4}\right) |1\rangle + \exp\left(\frac{2\pi i}{4}\right) |2\rangle + \cdots + \exp\left(\frac{7\pi i}{4}\right) |7\rangle + \right. \\ &\quad + |0\rangle + \exp\left(\frac{3\pi i}{4}\right) |1\rangle + \exp\left(\frac{6\pi i}{4}\right) |2\rangle + \cdots + \exp\left(\frac{21\pi i}{4}\right) |7\rangle + \\ &\quad + |0\rangle + \exp\left(\frac{5\pi i}{4}\right) |1\rangle + \exp\left(\frac{10\pi i}{4}\right) |2\rangle + \cdots + \exp\left(\frac{35\pi i}{4}\right) |7\rangle + \\ &\quad \left. + |0\rangle + \exp\left(\frac{7\pi i}{4}\right) |1\rangle + \cdots + \cdots \quad \cdots + \cdots + \exp\left(\frac{49\pi i}{4}\right) |7\rangle \right\} = \\ &= \frac{1}{\sqrt{2}}(|0\rangle - |4\rangle) \end{aligned}$$

4. Misuriamo $|\psi_3\rangle$. Se otteniamo $|0\rangle$ non ricaviamo nulla, mentre se otteniamo $|4\rangle$ abbiamo la soluzione, dato che:

$$\frac{y}{N} = \frac{k}{r}$$

è verificata se:

$$\frac{4}{8} = \frac{1}{2} = \frac{k}{r}$$

e quindi otteniamo $r = 2$, come desiderato.

5.4 Phase Estimation Algorithm

Consideriamo una trasformazione unitaria U che agisce “aggiungendo una fase” ϕ ad un suo autovettore $|u\rangle$:

(Lezione 16 ● del 6/6/2019)

$$U|u\rangle = e^{i\phi}|u\rangle \quad 0 \leq \phi \leq 2\pi$$

Supponiamo di poter preparare lo stato $|u\rangle$ con un registro di m qubit, e di avere a disposizione delle porte logiche *Control- U^{2^J}* , con $J \geq 0$, capaci di applicare condizionalmente (ossia a seconda del valore di certi qubit di controllo) la trasformazione U reiterata 2^J volte.

Come possiamo *stimare* la fase ϕ ?

In primo luogo, non possiamo misurare direttamente lo stato trasformato $e^{i\phi}|u\rangle$, dato che differisce da $|u\rangle$ per una sola fase globale - e quindi produce gli stessi valori attesi per ogni osservabile indipendentemente dal valore di ϕ .

Tuttavia, utilizzando i *gate* di controllo a nostra disposizione è possibile *trasformare* la fase globale in una fase relativa, e quindi creare un algoritmo per ottenere una stima arbitrariamente vicina a ϕ .

Nello specifico, possiamo determinare n bit di ϕ , ossia un numero $a = a_{n-1}a_{n-2} \dots a_1a_0$ di n bit tale che:

$$\phi = 2\pi \left(\frac{a}{2^n} + \delta \right); \quad 0 \leq |\delta| \leq \frac{1}{2^{n+1}} \quad (5.14)$$

dove δ è l'*errore* associato ad a , che diminuisce esponenzialmente al crescere del numero n di bit utilizzati. In questa notazione, denotiamo con $\tilde{\phi}$ la miglior stima della fase, e con $\delta\phi$ il relativo errore:

$$\tilde{\phi} = 2\pi \frac{a}{2^n}; \quad \delta\phi = 2\pi\delta; \quad \phi = \tilde{\phi} + \delta\phi$$

Per ottenere ciò, facciamo uso del seguente circuito quantistico:

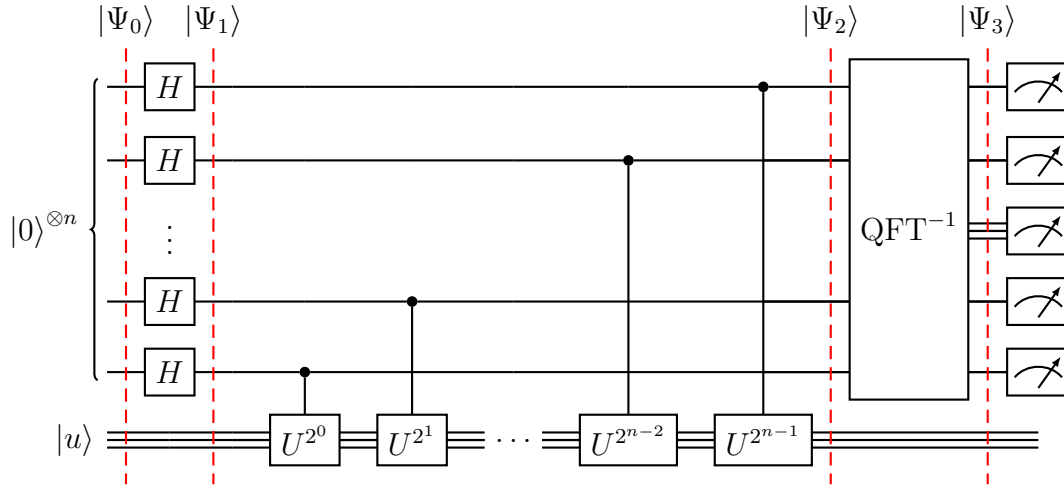


Figura (5.8) – Schema a porte logiche quantistiche dell’algoritmo di Phase Estimation

Impieghiamo due registri: uno da n qubit, inizializzato a $|0\rangle$, che conterrà la stima a della fase, e uno da m qubit, inizializzato a $|u\rangle$, su cui agiscono le $C - U^{2^j}$. Lo stato iniziale è perciò:

$$|\Psi_0\rangle = (\underbrace{|0\rangle |0\rangle \cdots |0\rangle}_n)_1 \otimes \underbrace{|u\rangle}_m_2$$

Tramite n Hadamard realizziamo una sovrapposizione di tutti gli stati nel primo registro:

$$|\Psi_1\rangle = (H^{\otimes n} \otimes \mathbb{I}^m) |\Psi_0\rangle = \left(\bigotimes_{i=1}^n \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \right)_1 \otimes |u\rangle_2$$

A questo punto, applichiamo in sequenza n $C-U^{2^j}$ sul secondo registro, ciascuna condizionata da uno degli n qubit del primo, giungendo allo stato $|\Psi_2\rangle$. Partiamo calcolando l’azione di $C-U^{2^j}$ sullo stato $(|0\rangle + |1\rangle)/\sqrt{2} \otimes |u\rangle$:

$$\begin{aligned} \underbrace{C-U^{2^j}}_W \left[\frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \otimes |u\rangle \right] &= \frac{1}{\sqrt{2}} [W |0\rangle |u\rangle + W |1\rangle |u\rangle] = \\ &= \frac{1}{\sqrt{2}} [|0\rangle |u\rangle + \exp(i2^j \phi) |1\rangle |u\rangle] = \\ &= \frac{1}{\sqrt{2}} [|0\rangle + \exp(i2^j \phi) |1\rangle] \otimes |u\rangle \end{aligned} \quad (5.15)$$

Notiamo che, poiché $|u\rangle$ è autovalore di U , U^2 , U^4 ..., la $C-U^{2^j}$ non modifica lo stato del secondo registro, mentre propaga una fase relativa (misurabile) nello stato del primo registro. Poiché allora l’input di ciascuna $C-U^{2^j}$ non cambia, basta reiterare

la (5.15) per ottenere $|\Psi_2\rangle$:

$$\begin{aligned}
|\Psi_2\rangle &= \frac{1}{\sqrt{2^n}} (|0\rangle + \exp(i\phi 2^{n-1}) |1\rangle) (|0\rangle + \exp(i\phi 2^{n-2}) |1\rangle) \otimes \dots \\
&\quad \dots \otimes (|0\rangle + e^{i2\phi} |1\rangle) (|0\rangle + e^{i\phi} |1\rangle) \otimes |u\rangle = \\
&= \frac{1}{\sqrt{2^n}} (|0\rangle |0\rangle \dots |0\rangle + e^{i\phi} |0\rangle \dots |0\rangle |1\rangle + \dots \exp(i(2^n - 1)\phi) |1\rangle |1\rangle \dots |1\rangle)_1 \otimes |u\rangle_2 = \\
&= \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} e^{i\phi y} |y\rangle_1 |u\rangle_2 \tag{5.16}
\end{aligned}$$

Concentriamoci sullo stato del solo primo registro. La (5.16) ha la forma della trasformata di Fourier dello stato $|\phi\rangle$ che codifica la fase. Per determinarlo, perciò, applichiamo l'inversa della *Quantum Fourier Transform*, che è definita da:

$$\text{QFT}^{-1}\{|y\rangle\} = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} \exp\left(-\frac{2\pi i xy}{2^n}\right) |x\rangle$$

Otteniamo allora (per il primo registro):

$$\begin{aligned}
|\Psi_3\rangle &= \text{QFT}^{-1}\{|\Psi_2\rangle\} = \frac{1}{2^n} \sum_{x=0}^{2^n-1} \sum_{y=0}^{2^n-1} \exp\left(-\frac{2\pi i xy}{2^n}\right) e^{i\phi y} |x\rangle = \\
&\stackrel{(5.14)}{=} \frac{1}{2^n} \sum_{x=0}^{2^n-1} \sum_{y=0}^{2^n-1} \exp\left(-\frac{2\pi i xy}{2^n}\right) \exp\left(\frac{2\pi i ay}{2^n}\right) \exp(2\pi i \delta y) |x\rangle = \\
&= \frac{1}{2^n} \sum_{x=0}^{2^n-1} \sum_{y=0}^{2^n-1} \exp\left(-\frac{2\pi i y}{2^n}(x - a)\right) \exp(2\pi i \delta y) |x\rangle
\end{aligned}$$

Una misura proiettiva risulta in un certo valore b con probabilità:

$$P(b) = |\langle b | \Psi_3 \rangle|^2 = \left| \frac{1}{2^n} \sum_{y=0}^{2^n-1} \exp\left(-\frac{2\pi i y}{2^n}(b - a)\right) \exp(2\pi i \delta y) \right|^2$$

dato che $\langle b | x \rangle = \delta_{xb}$.

Distinguiamo ora tra due casi:

1. Se $\delta = 0$, ossia se $|\phi\rangle$ può essere *codificata* da n bit (o meno), allora vi è un solo valore di b con probabilità non nulla, ed è quello che corrisponde alla miglior stima a . Infatti:

$$P(b = a) = \left| \frac{1}{2^n} 2^n \right| = 1$$

2. Se $\delta \neq 0$, sono in generale possibili più risultati. Quello che offre la stima corretta di ϕ , ossia $b = a$, ha probabilità data da:

$$\begin{aligned}
 P(b = a) &= \left| \frac{1}{2^n} \sum_{y=0}^{2^n-1} \exp(2\pi i \delta y) \right|^2 = \frac{1}{2^{2n}} \left| \sum_{y=0}^{2^n-1} \underbrace{\exp(2\pi i \delta)}_{\alpha}^y \right|^2 = \\
 &\stackrel{(a)}{=} \frac{1}{2^{2n}} \left| \frac{1 - \alpha^{2^n}}{1 - \alpha} \right|^2 = \frac{1}{2^{2n}} \left| \frac{1 - \exp(2\pi i \delta 2^n)}{1 - \exp(2\pi i \delta)} \right|^2 = \\
 &\stackrel{(b)}{=} \frac{1}{2^{2n}} \left| \frac{\sin(\pi \delta 2^n)}{\sin(\pi \delta)} \right|^2
 \end{aligned} \tag{5.17}$$

dove in (a) si è usata la formula per le somme parziali di una serie geometrica:

$$\sum_{k=0}^n x^k = \frac{1 - x^{n+1}}{1 - x}$$

e in (b) un'identità dell'esponenziale complesso:

$$\begin{aligned}
 |1 - \exp A| &= |1 - \cos A - i \sin A| = \sqrt{(1 - \cos A)^2 + \sin^2 A} = \\
 &= \sqrt{2} \frac{\sqrt{1 - \cos A}}{\sqrt{2}} \sqrt{2} = 2 \sin \frac{A}{2}
 \end{aligned}$$

Diamo una stima di $P(b = a)$. Poiché $2z \leq \sin(\pi z) \leq \pi z \ \forall z \in [0, 1/2]$, si ha:

$$|\sin(\pi \delta 2^n)| \geq 2|\delta|2^n; \quad |\sin(\pi \delta)| \leq \pi|\delta|$$

Sostituendo in (5.17) otteniamo:

$$P(b = a) \geq \frac{1}{2^{2n}} \frac{4|\delta|^2 2^{2n}}{\pi^2 |\delta|^2} = \frac{4}{\pi^2} \approx 0.405$$

Perciò la miglior stima a di n bit di ϕ si ottiene con una buona probabilità. Si può dimostrare che gli altri esiti probabili differiscono da a per i bit meno significativi. Nello specifico, si possono ottenere l bit della fase ϕ con una probabilità $p > 1 - \epsilon$ usando $n = l + O(\log(1/\epsilon))$ qubit.

Numericamente, con *solì* 60 qubit si raggiungono stime con 18 cifre significative - comparabili con le migliori misure tecnologicamente possibili nella fisica sperimentale.

5.5 Eigensolver

Un'applicazione importante dell'algoritmo di Phase Estimation è legata al calcolo di autovalori e autofunzioni di una certa Hamiltoniana H *indipendente dal tempo*. Lavorando in rappresentazione $\{x\}$ (con $d = 1$ per semplicità), denotiamo con $\phi_\alpha(x)$ le autofunzioni di H di autovalore \mathcal{E}_α , tali che:

$$H\phi_\alpha(x) = \mathcal{E}_\alpha \phi_\alpha(x) \quad \alpha \in \mathbb{N}$$

L'evoluzione temporale di tali autostati avviene per l'aggiunta di una fase:

$$\phi_\alpha(x, t=0) \xrightarrow{U(t)} \phi_\alpha(x, t) = \exp\left(-\frac{i}{\hbar}t\mathcal{E}_\alpha\right) \phi_\alpha(x, 0)$$

Ciò deriva direttamente dall'equazione di Schrödinger dipendente dal tempo, che per una generica funzione d'onda $\psi(x, t)$ è data da:

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = H(x) \psi(x, t)$$

Perciò gli autovettori $\phi_\alpha(x)$ sono anche autovettori dell'operatore $U(t)$ di evoluzione temporale, che ha le stesse caratteristiche dell'operatore U visto nella sezione precedente:

$$U(t) = \exp\left(-\frac{i}{\hbar}Ht\right); \quad U(t) |\phi_\alpha\rangle = \exp\left(-\frac{i}{\hbar}t\mathcal{E}_\alpha\right) |\phi_\alpha\rangle$$

5.5.1 Algoritmo classico

Classicamente, un algoritmo che permette di trovare autovalori e autovettori (e che, come vedremo, può essere adattato al caso quantistico con un certo guadagno computazionale) si basa sul calcolo di trasformate di Fourier.

Consideriamo una finestra temporale finita $0 \leq t \leq \bar{t}$, e una funzione d'onda iniziale arbitraria $\psi_0(x)$, che scriviamo nella base degli autostati di H :

$$\psi_0(x) \equiv \psi(x, t=0) = \sum_{\alpha=0}^{+\infty} a_\alpha \phi_\alpha(x)$$

La sua evoluzione temporale ad un certo istante t è quindi data da:

$$\psi_0(x, t) = \sum_{\alpha=0}^{+\infty} a_\alpha \exp\left(-\frac{i}{\hbar}\mathcal{E}_\alpha t\right) \phi_\alpha(x) = \sum_{\alpha=0}^{+\infty} a_\alpha e^{-i\omega_\alpha t} \phi_\alpha(x); \quad \omega_\alpha = \frac{\mathcal{E}_\alpha}{\hbar}$$

Fissiamo un punto $x = x_0$. La funzione $\psi(x_0, t)$ è la sovrapposizione di funzioni periodiche con pulsazioni ω_α , che possono essere evidenziate svolgendo una trasformata di Fourier⁶:

$$\begin{aligned} \tilde{\psi}(x_0, \omega) &= \mathcal{F}[\psi(x_0, t)](x_0, \omega) = \\ &= \int_0^{\bar{t}} dt e^{i\omega t} \sum_{\alpha=0}^{+\infty} a_\alpha e^{-i\omega_\alpha t} \phi_\alpha(x_0) = \sum_{\alpha=0}^{+\infty} a_\alpha \int_0^{\bar{t}} dt \exp(i(\omega - \omega_\alpha)t) \phi_\alpha(x) \end{aligned}$$

Quando $\omega = \omega_{\bar{\alpha}}$ otteniamo:

$$\tilde{\psi}(x_0, \omega_{\bar{\alpha}}) = a_{\bar{\alpha}} \phi_{\bar{\alpha}}(x_0) \bar{t} + \sum_{\alpha \neq \bar{\alpha}} a_\alpha \int_0^{\bar{t}} dt \exp(i(\omega - \omega_\alpha)t) \phi_\alpha(x)$$

⁶ Gli integrali possono essere calcolati numericamente *discretizzando* il dominio

Per \bar{t} sufficientemente grande ($\bar{t} \gg \omega_{\bar{\alpha}}/(2\pi)$) il secondo termine è la somma di integrali di funzioni oscillanti su un numero grande di oscillazioni, ed è perciò limitato (è $\rightarrow 0$), mentre il primo termine scala linearmente con \bar{t} . Del resto, per i *punti intermedi* (in cui $\omega \neq \omega_{\alpha} \forall \alpha$), la funzione è solamente limitata. Deduciamo perciò che $\tilde{\psi}(x_0, \omega)$ è piccata sulle ω_{α} (figura 5.9).

Dai massimi di $\tilde{\psi}(x_0, \omega)$ è perciò possibile stimare le ω_{α} e quindi gli autovalori⁷ \mathcal{E}_{α} di H . Per le autofunzioni, notiamo che per \bar{t} sufficientemente grande:

$$\frac{1}{\bar{t}} \tilde{\psi}(x_0, \omega_{\bar{\alpha}}) \approx a_{\bar{\alpha}} \phi_{\bar{\alpha}}(x_0)$$

Perciò:

$$\frac{\tilde{\psi}(x_1, \omega_{\bar{\alpha}})}{\tilde{\psi}(x_2, \omega_{\bar{\alpha}})} = \frac{\phi_{\bar{\alpha}}(x_1)}{\phi_{\bar{\alpha}}(x_2)}$$

Fissando arbitrariamente $\phi_{\bar{\alpha}}(x_1) \stackrel{!}{=} 1$ si può valutare l'autofunzione $\bar{\alpha}$ -esima in qualsiasi altro punto - a meno della normalizzazione (che può essere imposta successivamente).

Poiché le risorse computazionali sono finite, nella pratica ci si limita ad una certa regione, per esempio $x \in [-L, L]$ per un certo L fissato, che viene *discretizzato*, ossia suddiviso in 2^n parti, ciascuna *lunga* $\Delta x = 2L/(2^n - 1)$, dove n è il numero di bit a disposizione. Le autofunzioni (così come la funzione d'onda $\psi_0(x)$ iniziale) sono valutate solo nei punti x_i tali che:

$$x_i = -L + i\Delta x \quad i \in \{0, \dots, 2^n - 1\}$$

*Discretizzazione
del problema*

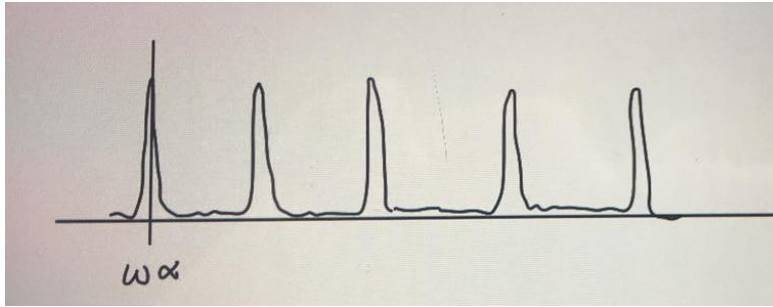


Figura (5.9) – Plot di esempio della trasformata di Fourier

5.5.2 Algoritmo quantistico

L'algoritmo appena visto si traduce in modo naturale al caso quantistico. Partiamo da uno stato $|\Psi_0\rangle$ arbitrario, che codifica una funzione d'onda $\psi_0(x)$ *valutata* in 2^n punti:

$$|\Psi_0\rangle = \sum_{i=0}^{2^n-1} \psi_0(x_i) |i\rangle$$

⁷ Più precisamente, tale metodo consente di determinare solo gli autovalori per cui $a_{\alpha} \neq 0$, ossia $\langle \psi_0 | \phi_{\alpha} \rangle \neq 0$. Per una funzione d'onda iniziale generica, senza particolari simmetrie, in genere $a_{\alpha} \neq 0$ è verificata.

Notiamo che bastano n qubit per codificare 2^n punti della funzione d'onda, con un guadagno esponenziale di memoria (ammesso che sia possibile preparare un tale stato).

Per semplicità (di calcolo e di implementazione), scegliamo $|\psi_0\rangle$ come un vettore casuale della base computazionale:

$$|\psi_0\rangle \stackrel{!}{=} |\bar{j}\rangle$$

che corrisponde ad una funzione d'onda $\psi_0(x) = \delta_{x\bar{j}}$ “localizzata in un punto”.

L'algoritmo è lo stesso usato per la Phase-Estimation, dove ora $|\psi_0\rangle$ gioca il ruolo di $|u\rangle$. Supponiamo di avere a disposizione delle gate $C-U^{2^j}$, dove U è l'operatore di evoluzione temporale per un tempo $\Delta t = \bar{t}/(2^n - 1)$:

$$U = \exp\left(-iH\frac{\Delta t}{\hbar}\right)$$

Fortunatamente tali U ammettono un'implementazione efficiente per una larga classe di Hamiltoniane.

Prepariamo allora lo stato iniziale:

$$|\Psi_0\rangle = |0\rangle_1^{\otimes n} \otimes |\psi_0\rangle_2$$

Realizziamo la sovrapposizione massima degli stati dei qubit di controllo (primo registro):

$$|\Psi_1\rangle = \left(\frac{1}{\sqrt{2^n}} \sum_{j=1}^{2^n-1} |j\rangle_1\right) \otimes |\psi_0\rangle_2$$

E applichiamo la successione delle n $C-U^{2^j}$, giungendo allo stato entangled analogo al (5.16):

$$\begin{aligned} |\Psi_2\rangle &= \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} |j\rangle_1 U^j |\psi_0\rangle_2 \stackrel{(a)}{=} \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} |j\rangle_1 |\psi(j\Delta t)\rangle = \\ &\stackrel{(b)}{=} \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} |j\rangle_1 \left(\sum_{\alpha=0}^{+\infty} a_\alpha \exp(-i\omega_\alpha j\Delta t) |\phi_\alpha\rangle \right)_2 \end{aligned}$$

dove in (a) abbiamo applicato la definizione dell'operatore di evoluzione temporale, e in (b) siamo passati nella base degli autoket $|\phi_\alpha\rangle$ di H .

Non resta allora che eseguire una QFT^{-1} sul primo registro. Misurando i primi n qubit si ottiene una stima di una $\omega_{\bar{\alpha}}$ (con probabilità $|a_{\bar{\alpha}}|^2$), e immediatamente si fa collassare il secondo registro nel relativo autostato $|\phi_{\bar{\alpha}}\rangle$ (che può essere ricostruito da ulteriori misurazioni).

Semplicemente ripetendo l'esecuzione si possono ottenere altri valori di ω_α .

Il vantaggio di tutto ciò è che la QFT è molto più efficiente dell'analogo classico. Perciò, a meno di aver scelto una $|\psi_0\rangle$ con proiezione troppo piccola lungo un $|\phi_\alpha\rangle$ desiderato, si possono calcolare tutti gli autovalori nella regione di interesse (per esempio alle energie prossime allo stato fondamentale) in tempo polinomiale.

5.6 Error correction

Ogni sistema è soggetto a *malfunzionamenti*, e ogni implementazione può - a seguito di determinati eventi più o meno rari - produrre *errori*. Risulta quindi importante studiare il modo di *ridurre quanto più possibile* tali evenienze, ed eventualmente *correggere comportamenti anomali*. Analizziamo ora alcuni metodi che permettono di raggiungere questi obiettivi.

5.6.1 Classical error correction

Sia ϵ la probabilità che un qualsiasi generico protocollo generi un *errore*. Per una buona implementazione $\epsilon \ll 1$, ma comunque $\neq 0$. Ridurre ulteriormente ϵ può però risultare molto costoso o difficile. Una strategia migliore è allora quella di inserire **ridondanza** nella comunicazione, dando così la possibilità al destinatario di rilevare e correggere alcuni tipi di errori.

Consideriamo, per esempio, una situazione in cui Alice vuole mandare un bit $a \in \{0, 1\}$ a Bob. Avremo allora una probabilità ϵ che Bob riceva un bit diverso da quello spedito da Alice, a seguito di un qualche fallimento del canale di trasmissione. Se però Alice invia 3 copie del bit a Bob, per esempio 000, un singolo errore con $p = \epsilon$ produce un messaggio tra 001, 010 e 100. In tutti e tre i casi Bob può *ricostruire* il messaggio originale applicando il principio del **voto di maggioranza**: poiché la probabilità ϵ che un bit sia corrotto è bassa, ci aspettiamo che la maggior parte dei bit siano corretti, e perciò che il messaggio originale sia quello compatibile con la maggior parte dei bit ricevuti - ossia 0.

Tale schema non funziona in ogni caso. Per esempio, se il messaggio originale è 000 e si verificano 2 errori, Bob riceve 011, 101 o 110. Applicando il *voto di maggioranza* si ricostruisce un messaggio (1) diverso da quello originale (0). Tuttavia, poiché tale situazione si verifica con probabilità decisamente inferiore, la strategia rimane valida per garantire una maggiore *stabilità* del canale di comunicazione. Lo si può vedere esaminando tutti i casi possibili:

# Errori	Messaggio ricevuto	Prob.	Fallimento?
0	000	$(1 - \epsilon)^3$	N
1	100, 010, 001	$\epsilon(1 - \epsilon)^2$	N
2	110, 011, 101	$\epsilon^2(1 - \epsilon)$	S
3	111	ϵ^3	S

Tabella (5.2) – Esiti possibili per il trasferimento di un bit con un fattore 3 di ridondanza

Mettendo tutto insieme, la probabilità di fallimento P_F^3 con 3 bit di ridondanza è data da:

$$P_F^3 = 3[\epsilon^2(1 - \epsilon)] + \epsilon^3 = O(\epsilon^2)$$

che è decisamente inferiore rispetto alla probabilità di fallimento per il protocollo a un *solo bit*: $P_F^1 = \epsilon$.

5.6.2 Quantum error correction

Proviamo ad applicare il protocollo di “ridondanza + voto” nel caso quantistico. Si rivelano subito alcuni problemi:

1. Innanzitutto non possiamo **copiare** n volte uno stato generico $|\psi\rangle$ sconosciuto (no cloning theorem) per generare ridondanza
2. Non è possibile **misurare** direttamente i qubit ricevuti per confrontarli tra loro, dato che una misura proiettiva modifica irrimediabilmente lo stato a cui si applica
3. Non è detto che gli errori consistano solamente nel *bit-flip*, ossia nella trasformazione di $|0\rangle \leftrightarrow |1\rangle$. Potrebbero esserci effetti, per esempio, che *modificano* la fase relativa tra le due componenti del qubit inviato.

Quantum bit-flip code

Supponiamo che Alice voglia inviare un qubit nello stato $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ a Bob tramite un canale quantistico affetto da **rumore**. Supponiamo inoltre che il rumore agisca **indipendentemente** su ciascun qubit, lasciando un qubit invariato con probabilità $1 - \epsilon$, o invertendolo (come un gate NOT, σ_x) con probabilità ϵ . In tale situazione, introduciamo un protocollo di correzione detto **3-qubit bit-flip code**, che si basa sul codificare gli stati della base computazionale inserendo *ridondanza*:

$$|0\rangle \mapsto |\tilde{0}\rangle = |000\rangle; \quad |1\rangle \mapsto |\tilde{1}\rangle = |111\rangle$$

Alice, perciò, applica tale codifica al proprio qubit $|\psi\rangle$:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \mapsto |\tilde{\psi}\rangle = \alpha|\tilde{0}\rangle + \beta|\tilde{1}\rangle$$

Notiamo che:

$$|\tilde{\psi}\rangle = \alpha|000\rangle + \beta|111\rangle \neq |\psi\rangle|\psi\rangle|\psi\rangle = (\alpha|0\rangle + \beta|1\rangle) \otimes (\alpha|0\rangle + \beta|1\rangle) \otimes (\alpha|0\rangle + \beta|1\rangle)$$

e perciò la codifica **non viola** il teorema del no-cloning, e infatti può essere implementata utilizzando due CNOT (figura 5.10).

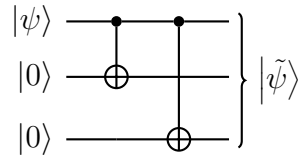


Figura (5.10) – Schema a porte logiche quantistiche per la generazione di ridondanza

Consideriamo tutte le possibili combinazioni di errore (date le ipotesi):

Messaggio Ricevuto	Prob.	Stato finale	$ x_0\rangle x_1\rangle$
$\alpha 000\rangle + \beta 111\rangle$	$(1 - \epsilon)^3$	$\alpha 000\rangle + \beta 111\rangle$	$ 00\rangle$
$\alpha 100\rangle + \beta 011\rangle$	$\epsilon(1 - \epsilon)^2$	$\alpha 000\rangle + \beta 111\rangle$	$ 11\rangle$
$\alpha 010\rangle + \beta 101\rangle$	$\epsilon(1 - \epsilon)^2$	$\alpha 000\rangle + \beta 111\rangle$	$ 10\rangle$
$\alpha 001\rangle + \beta 110\rangle$	$\epsilon(1 - \epsilon)^2$	$\alpha 000\rangle + \beta 111\rangle$	$ 01\rangle$
$\alpha 110\rangle + \beta 001\rangle$	$\epsilon^2(1 - \epsilon)$	$\alpha 111\rangle + \beta 000\rangle$	$ 01\rangle$
$\alpha 101\rangle + \beta 100\rangle$	$\epsilon^2(1 - \epsilon)$	$\alpha 111\rangle + \beta 000\rangle$	$ 10\rangle$
$\alpha 011\rangle + \beta 100\rangle$	$\epsilon^2(1 - \epsilon)$	$\alpha 111\rangle + \beta 000\rangle$	$ 11\rangle$
$\alpha 111\rangle + \beta 000\rangle$	ϵ^3	$\alpha 111\rangle + \beta 000\rangle$	$ 00\rangle$

Tabella (5.3) – Possibili messaggi ricevuti a seguito di errori

Per capire in quale caso ci si trovi non è possibile misurare direttamente i 3 qubit - ciò distruggerebbe la sovrapposizione coerente dello stato $|\psi\rangle$ che si vuole ricevere. Piuttosto, è possibile *correlare* 2 qubit ausiliari e misurarli lasciando invariati gli altri 3, che possono poi essere corretti applicando una certa operazione U determinata dall'informazione ricavata dalle misure.

L'idea è la seguente. Ipotizziamo, per esempio, che Bob abbia ricevuto lo stato:

$$|\Psi_1\rangle = \alpha |100\rangle + \beta |011\rangle$$

Bob può usare due qubit ausiliari per misurare la *correlazione* tra 2 coppie dei 3 qubit. Nel dettaglio, trova che nello stato ricevuto, primo e secondo qubit differiscono, così come il primo e il terzo. Nell'ipotesi (probabile) che vi sia stato un solo errore, una coppia di qubit differenti indica che uno dei due è quello errato. In questo caso, perciò, le informazioni ricavate puntano sul primo qubit - che può quindi essere corretto mediante un NOT (σ_x).

L'operazione appena discussa è realizzata dal circuito di figura 5.11. L'idea di base è che una sequenza di due CNOT che usano come controllo qubit nello **stesso** stato equivale ad un'identità, dato che non è altro che l'applicazione di una stessa *trasformazione unitaria* ripetuta due volte.

D'altro canto, due CNOT che partono da stati $|\psi\rangle$ e $\sigma_x |\psi\rangle$ equivalgono a un'unica NOT.

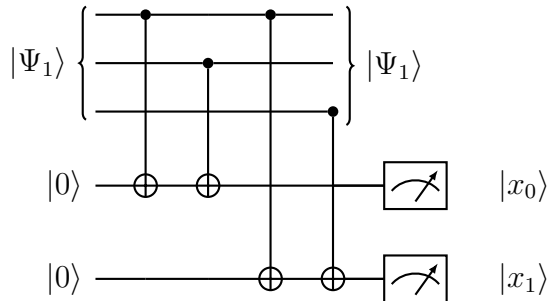


Figura (5.11) – Schema a porte logiche quantistiche per la misura di correlazioni tra i qubit ridondanti

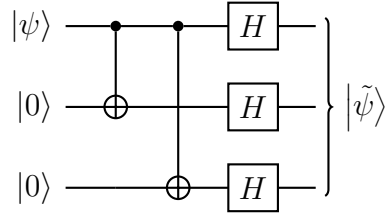


Figura (5.12) – Circuito a gate quantistici per realizzare la codifica del *3-qubit phase-flip code*

A seconda degli stati $|x_0\rangle |x_1\rangle$ misurati possiamo correggere opportunamente l'errore rilevato, come mostrato in tabella 5.3. Notiamo che a volte la correzione non consente di ricavare lo stato originale (esattamente come nel caso classico). Ciò è dovuto al fatto che, potendo misurare solo 2 correlazioni, possiamo distinguere solo 4 degli 8 casi possibili.

Uno schema molto simile può essere adottato per correggere errori che consistono in un'*inversione di fase*, ossia tali da mappare:

$$|+\rangle \mapsto |-\rangle; \quad |-\rangle \mapsto |+\rangle$$

dove:

$$|+\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle); \quad |-\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

L'idea è di usare lo stesso circuito, ma partendo da una codifica in una differente base:

$$\begin{aligned} |0\rangle &\mapsto |+++\rangle \\ |1\rangle &\mapsto |---\rangle \end{aligned}$$

Un'implementazione per tale codifica è mostrata nella figura 5.12

Correzioni avanzate

Generalizzando, possiamo correggere entrambe le tipologie di errori con uno stesso circuito, codificando la ridondanza in 9 qubit (**9-qubit Shor code**):

$$\begin{aligned} |0\rangle &\mapsto |\tilde{0}\rangle = \frac{1}{\sqrt{8}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \\ |1\rangle &\mapsto |\tilde{1}\rangle = \frac{1}{\sqrt{8}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) \end{aligned}$$

Le idee sono le stesse di prima, ma la scala del circuito rende il tutto estremamente più complicato.

Allo stato attuale si reputa che raggiungendo $\epsilon \approx 10^{-3} \div 10^{-4}$, mediante algoritmi di questo tipo si possa realizzare il *Fault tolerant quantum computing*, ossia un'architettura di computazione quantistica *resistente* a tipologie generali di errore.

Un'altra strada per la correzione degli errori è data dal codificare qubit nei sottospazi *invarianti* per le *simmetrie* degli errori possibili.

Per esempio, consideriamo un errore che aggiunga una fase:

$$\begin{cases} |0\rangle \mapsto |0\rangle \\ |1\rangle \mapsto e^{i\phi} |1\rangle \end{cases}$$

L'idea è di codificare $|0\rangle$ e $|1\rangle$ nella sovrapposizione di stati a più qubit che si trovano nello stesso autospazio dell'*operatore errore*. Nel nostro caso, se esaminiamo l'azione dell'errore sui vettori della base computazionale per 2 qubit, notiamo due stati su cui l'errore agisce “nello stesso modo”:

$$\begin{cases} |00\rangle \mapsto |00\rangle \\ |01\rangle \mapsto e^{i\phi} |01\rangle \equiv |0\rangle \\ |10\rangle \mapsto e^{i\phi} |10\rangle \equiv |1\rangle \\ |11\rangle \mapsto e^{2i\phi} |11\rangle \end{cases}$$

Scegliendo allora $|0\rangle \mapsto |01\rangle$ e $|1\rangle \mapsto |10\rangle$ come codifica, l'azione dell'errore aggiunge una fase globale e quindi non cambia lo stato.

6.1 Teoria delle perturbazioni dipendenti dal tempo

Fonti utili: [13],
[14]

(Lezione 17 del
12/6/2019)

Consideriamo un sistema quantistico - che può essere costituito da atomi, o più in generale da qubit - e che vogliamo poter regolare dall'esterno, per esempio mediante impulsi di radiazione elettromagnetica. Vogliamo poter descrivere cosa accada agli stati quantistici a seguito di un potenziale che *varia nel tempo*. Dal punto di vista energetico, alcune delle possibili conseguenze di una perturbazione sono date da:

- **Assorbimento** dell'energia della perturbazione, con passaggio tra stati $|i\rangle \rightarrow |f\rangle$, con $\mathcal{E}_i < \mathcal{E}_f$
- **Emissione** di parte dell'energia contenuta nello stato, con passaggio tra stati $|i\rangle \rightarrow |f\rangle$ con $\mathcal{E}_i > \mathcal{E}_f$
- Passaggio da un autovalore dell'energia nello spettro *discreto* a uno dello spettro *continuo*.

Nello specifico, supponiamo che il sistema sia descritto da un'Hamiltoniana $H(t)$ scomponibile in un termine H_0 indipendente da t , e un potenziale $W(t)$ che invece dipende da t :

$$H(t) = H_0 + \lambda W(t)$$

dove λ è un parametro che quantifica *il rapporto* tra le due componenti H_0 e $W(t)$. Denotiamo con $|\varphi_n\rangle$ gli autostati di H_0 (che supponiamo avere solo spettro discreto) di autovalore \mathcal{E}_n , che consideriamo conosciuti:

$$H_0 |\varphi_n\rangle = \mathcal{E}_n |\varphi_n\rangle$$

Per $\lambda = 0$ l'evoluzione temporale dipende solo da H_0 , ossia da un potenziale costante. In tal caso, se il sistema si trova in un autostato $|\varphi_n\rangle$ di H_0 a $t = 0$, allora è nello stesso autostato a qualsiasi altro istante (e infatti $|\varphi_n\rangle$ sono detti **stati stazionari**).

In particolare, la probabilità che il sistema nello stato $|\psi(t=0)\rangle = |\varphi_i\rangle$ si trovi in $|\varphi_f\rangle$ ad un istante t è esattamente nulla:

$$P_{if}^0(t) = |\langle \varphi_f | \psi(t) \rangle|^2 = \left| \langle \varphi_f | \varphi_i \rangle \exp\left(-\frac{i}{\hbar} t \mathcal{E}_f\right) \right|^2 \equiv 0$$

Tale $P_{if}(t)$ non è altro che la **probabilità di transizione** tra l'evoluzione temporale di $|\varphi_i\rangle$ al tempo t e $|\varphi_f\rangle$.

Se $\lambda \neq 0$, in generale, $P_{if}(t) \neq 0$. In altre parole, la presenza di un potenziale che varia nel tempo fa sì che quelli che prima erano stati stazionari ora non lo siano più.

Cerchiamo quindi un modo per calcolare tale $P_{if}(t)$ nel caso generale. Nella visuale di Schrödinger, l'evoluzione temporale di uno stato è data dall'equazione di Schrödinger dipendente dal tempo:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = [H_0 + \lambda W(t)] |\psi(t)\rangle \quad (6.1)$$

Basta allora trovare la $|\psi(t)\rangle$ che risolve tale equazione con la condizione iniziale $|\psi(t=0)\rangle = |\varphi_i\rangle$, e si può calcolare la probabilità desiderata:

$$P_{if}(t) = |\langle \varphi_f | \psi(t) \rangle|^2$$

Per farlo, riscriviamo la (6.1) in coordinate, scegliendo come base quella formata dagli autostati di H_0 . Partiamo trasformando la $|\psi(t)\rangle$:

$$|\psi(t)\rangle = \mathbb{I} |\psi(t)\rangle \underset{(a)}{=} \sum_{k=0}^{+\infty} \underbrace{\langle \varphi_n | \psi(t) \rangle}_{c_n(t)} |\varphi_n\rangle = \sum_{k=0}^{+\infty} c_n(t) |\varphi_n\rangle \quad (6.2)$$

dove in (a) si è usata la completezza di Dirac. Sostituendo in (6.1) si ottiene:

$$i\hbar \frac{d}{dt} \sum_{k=0}^{+\infty} c_k(t) |\varphi_k\rangle = H_0 \sum_{k=0}^{+\infty} c_k(t) |\varphi_n\rangle + \lambda W(t) \sum_{k=0}^{+\infty} c_k(t) |\varphi_k\rangle$$

Prendiamo il prodotto scalare con $|\varphi_n\rangle$:

$$i\hbar \frac{d}{dt} \sum_{k=0}^{+\infty} c_k(t) \underbrace{\langle \varphi_n | \varphi_k \rangle}_{\delta_{nk}} = \sum_{k=0}^{+\infty} c_k(t) \underbrace{\langle \varphi_n | H_0 | \varphi_k \rangle}_{\delta_{nk} \mathcal{E}_k} + \lambda \sum_{k=0}^{+\infty} c_k(t) \underbrace{\langle \varphi_n | W(t) | \varphi_k \rangle}_{W_{nk}(t)}$$

Giungiamo allora al sistema di equazioni:

$$i\hbar \dot{c}_n(t) = \mathcal{E}_n c_n(t) + \lambda \sum_{k=0}^{+\infty} W_{nk}(t) c_k(t) \quad (6.3)$$

Poiché in generale $W(t)$ non è diagonale in questa base, le equazioni differenziali sono *accoppiate* tra loro, e perciò risultano di difficile soluzione.

Cerchiamo allora di semplificare il problema. Partiamo passando in visuale di interazione, ponendo:

$$|\psi(t)\rangle_I = \exp\left(\frac{i}{\hbar}tH_0\right) |\psi(t)\rangle_S \quad (6.4)$$

dove $|\psi(t)\rangle_S$ è la funzione d'onda nella visuale di Schrödinger (finora utilizzata). Nel caso di (6.2), otteniamo:

$$\sum_{k=0}^{+\infty} b_n(t) |\varphi_n\rangle \equiv |\psi(t)\rangle_I = \sum_{k=0}^{+\infty} \exp\left(\frac{i}{\hbar}t\mathcal{E}_n\right) c_n(t) |\varphi_n\rangle$$

Proiettando in coordinate (ossia prendendo il prodotto scalare per $|\varphi_n\rangle$), si giunge a:

$$b_n(t) = c_n(t) \exp\left(\frac{i}{\hbar}\mathcal{E}_n t\right) \Rightarrow c_n(t) = b_n(t) \exp\left(-\frac{i}{\hbar}\mathcal{E}_n t\right) \quad (6.5)$$

Tale manipolazione permette di semplificare notevolmente la notazione. Infatti, sostituendo (6.5) in (6.3) otteniamo:

$$\begin{aligned} i\hbar \left(\cancel{b_n(t) \left(\frac{i}{\hbar}\mathcal{E}_n \right)} + \dot{b}_n(t) \right) \exp\left(-\frac{i}{\hbar}\mathcal{E}_n t\right) &= \cancel{\mathcal{E}_n b_n(t) \exp\left(-\frac{i}{\hbar}\mathcal{E}_n t\right)} + \\ &+ \lambda \sum_{k=0}^{+\infty} W_{nk}(t) b_k(t) \exp\left(-\frac{i}{\hbar}\mathcal{E}_k t\right) \end{aligned}$$

e dividendo per l'exp:

$$i\hbar \dot{b}_n(t) = \lambda \sum_{k=0}^{+\infty} W_{nk}(t) b_k(t) \exp(i\omega_{nk}t); \quad \omega_{nk} = \frac{\mathcal{E}_n - \mathcal{E}_k}{\hbar} \quad (6.6)$$

Abbiamo perciò rimosso l'evoluzione data dal potenziale costante H_0 , focalizzandoci solo sull'azione della componente dipendente dal tempo $W(t)$.

Geometricamente, possiamo immaginare l'evoluzione data da $H(t)$ come la sovrapposizione di due *rotazioni* a diverse velocità - una costante (data da H_0) e una variabile (da $W(t)$). In quest'ottica, le $b_n(t)$ non sono altro che le coordinate di $|\psi(t)\rangle$ rispetto ad una base che “evolve come prescritto da H_0 ”, e che funge da una sorta di “sistema di riferimento rotante solidale alla rotazione data da H_0 ”.

In altre parole, nella base data da:

$$\left\{ \exp\left(\frac{i}{\hbar}t\mathcal{E}_n\right) |\varphi_n\rangle \right\}_{n \in \mathbb{N}}$$

l'evoluzione unitaria data dall'Hamiltoniana H_0 coincide con l'**identità**.

Nonostante la notazione appena più semplice, le (6.6) sono ancora molto difficili da risolvere. Per procedere assumiamo perciò l'**ipotesi perturbativa**, per cui le $b_n(t)$ ammettono uno sviluppo in serie di potenze attorno a $\lambda = 0$:

$$b_n(t) = b_n^{(0)}(t) + \lambda b_n^{(1)}(t) + \lambda^2 b_n^{(2)}(t) + \dots \quad (6.7)$$

Equivalentemente, per il relativo ket vale:

$$\begin{aligned} |\psi(t)\rangle_I &= (b_n^{(0)}(t) + \lambda b_n^{(1)}(t) + \dots) |\varphi_n\rangle = \\ &\equiv |\psi^{(0)}(t)\rangle + \lambda |\psi^{(1)}(t)\rangle + \lambda^2 |\psi^{(2)}(t)\rangle + \dots \end{aligned}$$

Sostituendo (6.7) in (6.6) giungiamo a:

$$i\hbar \left(\dot{b}_n^{(0)}(t) + \lambda \dot{b}_n^{(1)}(t) + \dots \right) = \sum_{k=0}^{+\infty} W_{nk}(t) \left[\lambda b_k^{(0)}(t) + \lambda^2 b_k^{(1)}(t) + \dots \right] \exp(i\omega_{nk}t)$$

Poiché tale equazione deve valere $\forall \lambda \in \mathbb{R}$, possiamo uguagliare i coefficienti delle λ di ciascun ordine. Per λ^0 otteniamo:

$$i\hbar \dot{b}_n^{(0)}(t) = 0 \Rightarrow b_n^{(0)}(t) = b_n^{(0)}(0)$$

Imponendo la condizione iniziale $(|\psi(t=0)\rangle_I = |\varphi_i\rangle)$ otteniamo:

$$\begin{cases} b_n^{(0)}(t=0) = \delta_{ni} \\ b_n^{(r)}(t=0) = 0 \quad \forall r > 0 \end{cases} \quad (6.8)$$

E quindi:

$$b_n^{(0)}(t) = b_n^{(0)}(0) = \delta_{ni}$$

Per un generico ordine $r > 0$, uguagliando i coefficienti si giunge a:

$$i\hbar \dot{b}_n^{(r)}(t) = \sum_{k=0}^{+\infty} \exp(i\omega_{nk}t) W_{nk}(t) b_k^{(r-1)}(t)$$

Supponendo che $b_n^{(r-1)}(t)$ sia conosciuta, si può integrare tale equazione differenziale grazie alle condizioni iniziali specificate in (6.8). Avendo fornito la soluzione $b_n^{(0)}(t)$ al passo 0, per induzione si possono perciò calcolare le soluzioni per qualsiasi ordine.

Concentriamoci sul caso $r = 1$:

$$i\hbar \dot{b}_n^{(1)}(t) = \sum_{k=0}^{+\infty} W_{nk} \underbrace{b_k^{(0)}(t)}_{\delta_{ni}} \exp(i\omega_{nk}t) = W_{ni}(t) \exp(i\omega_{ni}t)$$

Integrando otteniamo:

$$b_n^{(1)}(t) = \frac{1}{i\hbar} \int_0^t \exp(i\omega_{ni}\tau) W_{ni}(\tau) d\tau \quad (6.9)$$

Non resta che sostituire nel calcolo della probabilità di transizione:

$$\begin{aligned} P_{if}(t) &= \langle \varphi_f | \psi(t) \rangle = |c_f(t)|^2 \stackrel{(6.5)}{=} |b_f(t)|^2 \stackrel{(a)}{\approx} |\cancel{b_f^{(0)}(t)} + \lambda b_f^{(1)}(t)|^2 = \\ &= \frac{\lambda^2}{\hbar^2} \left| \int_0^t \exp(i\omega_{fi}\tau) W_{fi}(\tau) d\tau \right|^2 \end{aligned} \quad (6.10)$$

dove in (a) abbiamo troncato la serie al primo ordine.

Perciò la probabilità di una transizione dipende dall'elemento di matrice che accoppia lo stato iniziale i allo stato finale f . Se tale elemento di matrice è nullo, allora la transizione (almeno *al primo ordine*) non può avvenire.

6.1.1 Perturbazione sinusoidale

Consideriamo una perturbazione sinusoidale di pulsazione ω (es. radiazione di un laser):

$$W(t) = -W \sin(\omega t)$$

dove W è una matrice, le cui entrate W_{ij} caratterizzano gli *accoppiamenti* tra gli autostati i e j di H_0 .

Inserendo in (6.9):

$$b_f^{(1)}(t) = -\frac{W_{fi}}{i\hbar} \int_0^t \exp(i\omega_{fi}\tau) \sin(\omega\tau) d\tau$$

Poniamo $\omega_{fi} \equiv \omega_0$. Espandendo il sin otteniamo:

$$\begin{aligned} b_f^{(1)}(t) &= -\frac{W_{fi}}{i\hbar} \int_0^t d\tau \exp(i\omega_0\tau) \frac{1}{2i} [e^{i\omega\tau} - e^{-i\omega\tau}] = \\ &= \frac{W_{fi}}{2\hbar} \left[\frac{\exp[i(\omega_0 + \omega)\tau]}{\omega_0 + \omega} \Big|_0^t - \frac{\exp[i(\omega_0 - \omega)\tau]}{\omega_0 - \omega} \Big|_0^t \right] = \\ &= \frac{W_{fi}}{2\hbar} \left[\frac{\exp[i(\omega_0 + \omega)t] - 1}{\omega_0 + \omega} - \frac{\exp[i(\omega_0 - \omega)t] - 1}{\omega_0 - \omega} \right] \end{aligned} \quad (6.11)$$

Supponiamo ora che $\omega_0 + \omega \gg |\omega_0 - \omega|$, ossia che la pulsazione ω del potenziale sia simile a quella ω_0 di transizione tra gli stati che stiamo considerando. Infatti, ci aspettiamo che P_{if} sia significativamente non nulla proprio in questi casi (almeno in questa trattazione al primo ordine).

Poiché i numeratori in (6.11) sono funzioni limitate, possiamo così trascurare il primo termine:

$$\begin{aligned} b_f^{(1)}(t) &\underset{\omega \approx \omega_0}{\approx} -\frac{W_{fi}}{2\hbar} \frac{\exp[i(\omega_0 - \omega)t] - 1}{\omega_0 - \omega} = \\ &= -\frac{W_{fi}}{2\hbar} \frac{\exp\left(\frac{i(\omega_0 - \omega)t}{2}\right) \exp\left(\frac{i(\omega_0 - \omega)t}{2}\right) - \exp\left(-\frac{i(\omega_0 - \omega)t}{2}\right)}{\omega_0 - \omega} \textcolor{red}{2i} = \\ &= -i \frac{W_{fi}}{\hbar} \frac{\exp\left(\frac{i(\omega_0 - \omega)t}{2}\right)}{\omega_0 - \omega} \sin\left(\frac{(\omega_0 - \omega)t}{2}\right) \end{aligned}$$

La probabilità di transizione (per $i \neq f$) è allora data da:

$$P_{if}(t, \omega) \approx |b_f^{(1)}(t)|^2 = \frac{|W_{if}|^2}{4\hbar^2} F(t, \omega - \omega_0); \quad F(t, \omega) = \left(\frac{\sin\left(\frac{\omega t}{2}\right)}{\frac{\omega}{2}} \right)^2$$

Il grafico di $P_{if}(t, \omega)$ per un tempo fissato e in funzione di ω (figura 6.1) presenta un picco (di risonanza) attorno a $\omega_{fi} = \omega_0$ (e a $-\omega_{fi}$), di altezza $P_{\max} = |W_{fi}|^2 t^2 / (4\hbar^2)$ e “larghezza” $\Delta\omega = 4\pi/t$. Notiamo che la P_{\max} può potenzialmente divenire > 1 : ciò è indice che la validità della teoria (al primo ordine) è per soli λt sufficientemente piccoli.

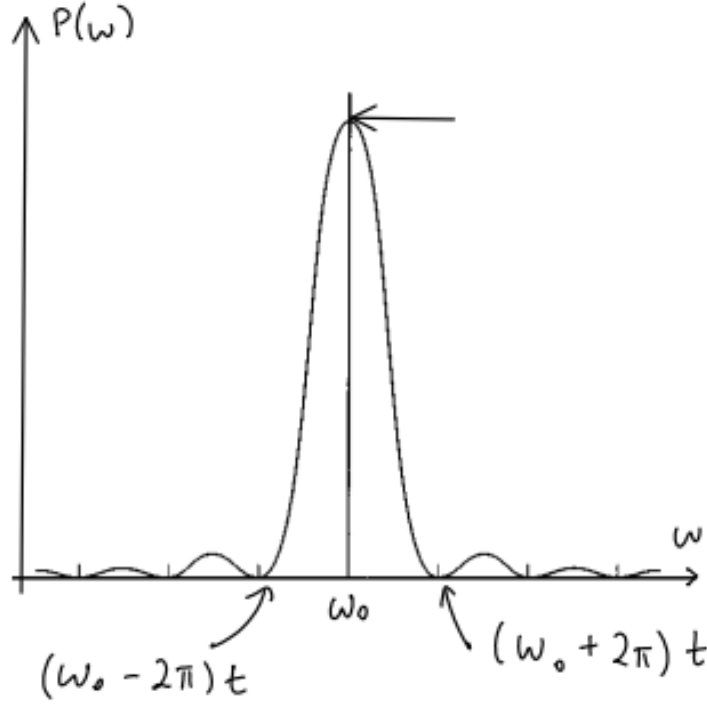


Figura (6.1) – Grafico di $P_{if}(\omega)$

6.1.2 Fermi Golden Rule

Estendiamo quanto appena visto al caso di uno spettro generale, sia discreto che continuo. Nello specifico, partiamo da $|\varphi_i\rangle$ autostato di H_0 nello spettro discreto, e vogliamo calcolare le P_{if} con uno stato $|f\rangle = |\alpha_f\rangle$ nello spettro continuo. Poiché per gli autostati del continuo vale la relazione di normalizzazione:

$$\langle \alpha | \alpha' \rangle = \delta(\alpha - \alpha')$$

si ha che la probabilità di transizione ad un *singolo specifico stato* è sempre nulla, visto che uno stato nel continuo ha *misura nulla*.

Bisogna perciò considerare un **intervallo di stati**, che prendiamo *vicini* allo stato α_F che ci interessa, ossia in un suo intorno D_f . Supponendo che tali possibili transizioni siano indipendenti tra loro, possiamo ottenere la P_{if} integrando sul range dell'indice $\alpha \in D_f$ che identifica gli autoket:

$$P(\alpha_F, t) = \int_{\alpha \in D_f} |\langle \alpha | \psi(t) \rangle|^2 d\alpha \quad (6.12)$$

Poiché stiamo parlando di autoket di H_0 , α è una funzione dell'autovalore \mathcal{E} , che possiamo usare perciò come indice per gli autoket. Differenziando otteniamo:

$$\alpha = \alpha(\mathcal{E}) \Rightarrow d\alpha = \underbrace{\frac{d\alpha}{d\mathcal{E}}}_{\rho(\mathcal{E})} d\mathcal{E}$$

dove $\rho(\mathcal{E})$ è la **densità di stati** (numero di stati per unità di energia) ad energia \mathcal{E} .

Più precisamente, ciò vale solo in assenza di degenerazione. Nel caso generale, per identificare un autoket di H_0 sono necessari due indici: α e l'*indice di degenerazione* β . Ripetendo il passaggio precedente, giungeremo perciò a:

$$d\alpha = \rho(\beta, \mathcal{E}) d\mathcal{E} d\beta$$

Sostituendo in (6.12) otteniamo:

$$P(\mathcal{E}_f, t) = \int_{\mathcal{E} \in \Delta\mathcal{E}} |\langle \mathcal{E}, \beta | \psi(t) \rangle|^2 \rho(\mathcal{E}, \beta) d\mathcal{E} \quad (6.13)$$

Al primo ordine, analogamente a quanto ricavato in (6.9):

$$\langle \mathcal{E}, \beta | \psi(t) \rangle = \frac{1}{i\hbar} \int_0^t \exp(i\omega_{fi}\tau) \langle \mathcal{E}, \beta | W(\tau) | \varphi_i \rangle d\tau; \quad \omega_{fi} = \frac{\mathcal{E} - \mathcal{E}_i}{\hbar}$$

Calcoliamo esplicitamente la probabilità di transizione in due casi semplici. Per semplicità di notazione, lavoreremo nel caso nondegenere.

1. **Potenziale costante** $W(t) \equiv W$. Questo è per esempio il caso di un potenziale che “si accende improvvisamente” a $t = 0$. Si ha allora:

$$\begin{aligned} \langle \mathcal{E} | \psi(t) \rangle &= \frac{1}{i\hbar} \langle \mathcal{E} | W | \varphi_i \rangle \int_0^t \exp(i\omega_{fi}\tau) d\tau = \\ &= \frac{1}{i\hbar} \langle \mathcal{E} | W | \varphi_i \rangle \frac{\exp(i\omega_{fi}t) - 1}{i\omega_{fi}} = \\ &= -\frac{1}{\hbar} \frac{\langle \mathcal{E} | W | \varphi_i \rangle}{\omega_{fi}} \frac{\exp\left(\frac{i\omega_{fi}t}{2}\right) - \exp\left(-\frac{i\omega_{fi}t}{2}\right)}{2i} \exp\left(\frac{i\omega_{fi}t}{2}\right) = \\ &= -\frac{2i}{\hbar} \frac{\langle \mathcal{E} | W | \varphi_i \rangle}{\omega_{fi}} \sin\left(\frac{\omega_{fi}t}{2}\right) \exp\left(\frac{i\omega_{fi}t}{2}\right) \end{aligned}$$

Prendendo il modulo quadro:

$$|\langle \mathcal{E} | \psi(t) \rangle|^2 = \frac{1}{\hbar^2} |\langle \mathcal{E} | W | \varphi_i \rangle|^2 F\left(t, \frac{\mathcal{E} - \mathcal{E}_i}{\hbar}\right); \quad F(t, \omega) = \left(\frac{\sin\left(\frac{\omega t}{2}\right)}{\frac{\omega}{2}}\right)^2$$

Possiamo ora sostituire in (6.13), ottenendo:

$$P(\mathcal{E}_f, t) = \int_{\mathcal{E} \in \Delta\mathcal{E}} \frac{1}{\hbar^2} |\langle \mathcal{E} | W | \varphi_i \rangle|^2 F\left(t, \frac{\mathcal{E} - \mathcal{E}_i}{\hbar}\right) \rho(\mathcal{E}) d\mathcal{E}$$

Sappiamo che F è piccata attorno a \mathcal{E}_i , ed è tanto più stretta quanto più t è grande. Supponendo allora che t sia sufficientemente grande, si ha che $P(\mathcal{E}_f, t)$ è significativamente $\neq 0$ solo per $\mathcal{E}_f = \mathcal{E}_i$. Supponendo poi che W e $\rho(\mathcal{E})$ siano *approssimativamente costanti* in $\Delta\mathcal{E}$ (che pensiamo centrato in \mathcal{E}_i ,

per poter ottenere un risultato non ≈ 0). Possiamo allora valutarli nel punto medio \mathcal{E}_i e portarli fuori dall'integrale:

$$P(\mathcal{E}_f = \mathcal{E}_i, t) = \frac{1}{\hbar^2} |\langle \mathcal{E}_i | W | \varphi_i \rangle|^2 \rho(\mathcal{E}_i) \int_{\mathcal{E} \in \Delta \mathcal{E}} F\left(t, \frac{\mathcal{E} - \mathcal{E}_i}{\hbar}\right)$$

Con tali ipotesi, possiamo estendere il dominio di integrazione a tutto \mathbb{R} , dato che $F \approx 0$ per valori di poco diversi da \mathcal{E}_i . Così facendo possiamo calcolare l'integrale:

$$\int_{\mathbb{R}} \sin^2\left(\frac{\omega_{fi}t}{2}\right) \frac{4}{\omega_{fi}^2} d\mathcal{E} = 4\hbar \int_{\mathbb{R}} \frac{1}{\omega_{fi}^2} \sin^2\left(\frac{\omega_{fi}t}{2}\right) d\omega_{fi}$$

dove abbiamo usato il cambio di variabile:

$$\omega_{fi} = \frac{\mathcal{E} - \mathcal{E}_i}{\hbar} \Rightarrow d\omega_{fi} = \frac{d\mathcal{E}}{\hbar}$$

Con un ulteriore cambio di variabile:

$$u = \frac{\omega_{fi}t}{2} \Rightarrow du = \frac{t}{2} d\omega_{fi}$$

giungiamo infine a:

$$4\hbar \int_{\mathbb{R}} \frac{\sin^2(u)}{\frac{4u^2}{t^2}} \frac{2}{t} du = 2\hbar t \underbrace{\int_{\mathbb{R}} \frac{\sin^2(u)}{u^2} du}_{=\pi} = 2\pi\hbar t$$

E perciò la probabilità di transizione tra lo stato $|\varphi_i\rangle$ e lo stato $|\alpha_F\rangle$ di energia $\mathcal{E}(\alpha_F) = \mathcal{E}_i$ è data da:

$$P(|\varphi_i\rangle \rightarrow |\alpha_f\rangle, t) = \frac{2\pi t}{\hbar} |\langle \alpha_f | W | \varphi_i \rangle|^2$$

Tenendo conto anche della degenerazione si ottiene un'espressione leggermente più complessa:

$$dP(|\varphi_i\rangle \rightarrow |\alpha_f\rangle, t) = d\beta \frac{2\pi t}{\hbar} |\langle \alpha_f, \beta_f | W | \varphi_i \rangle|^2$$

Definiamo infine il **rate** di transizione come la probabilità di transizione per unità di tempo:

$$w \equiv \frac{1}{t} P_{fi}(t)$$

E giungiamo così alla regola d'oro di Fermi per perturbazioni costanti:

$$w = \frac{2\pi}{\hbar} |\langle \alpha_f | W | \varphi_i \rangle|^2 \rho(\mathcal{E}_f); \quad \mathcal{E}_f = \mathcal{E}_i$$

2. **Perturbazione sinusoidale** $W(t) = W \sin(\omega t)$. Ripetendo tutti i passaggi del caso precedente, si giunge all'espressione (**regola d'oro di Fermi**):

$$w(|\varphi_i\rangle \rightarrow |\alpha_f\rangle) = \frac{2\pi}{\hbar} \rho(\beta_f, \mathcal{E}_f) |\langle \beta_f, \mathcal{E}_f | W | \varphi_i \rangle|^2 \quad \mathcal{E}_f = \mathcal{E}_i + \hbar\omega$$

che descrive il *rate* di transizione per *assorbimento* di radiazione. Nel caso di *emissione* basta sostituire $\mathcal{E}_f = \mathcal{E}_i - \hbar\omega$.

Hardware quantistico

Per realizzare nella pratica un computer quantistico è necessario soddisfare alcuni criteri, che furono formalizzati per la prima volta da Di Vincenzo nel 2000:

1. **Scalabilità:** qubit con una realizzazione ben definita e replicabile
2. **Reset:** possibilità di creare (stabilmente) stati $|00000 \dots 0\rangle$
3. **Tempi di coerenza lunghi** rispetto alla durata di esecuzione - e perciò ben più lunghi della durata di un gate.
Detto cioè τ_d il *rate* di decoerenza, e τ_g la durata media necessaria per effettuare un'operazione (*gate time*, equivalente al *clock* di un computer classico), vorremmo che $\tau_d/\tau_g \gg 1$. Nello specifico, sarebbe un buon risultato avere $\tau_d/\tau_g > 10^4$ - dato che ciò permetterebbe di realizzare anche gli algoritmi di *error-correction*.
4. **Set universale di gate**
5. **Read-out efficiente**, ossia possibilità di misurare gli stati prodotti dal computer quantistico in maniera affidabile

Negli ultimi decenni si sono esaminati diversi sistemi hardware per cercare di raggiungere tali obiettivi. I principali sono:

- Cavity QED
- Solid-state devices (quantum dot, circuiti in grado di *intrappolare* singoli elettroni)
- Atomi freddi
- Ioni intrappolati
- Circuiti superconduttivi

I più promettenti sono gli ultimi due, che ora introdurremo.

7.0.1 Ioni intrappolati

In questa implementazione fisica, i qubit sono dati da *ioni*, che vengono confinati grazie ad una *Paul trap*, mediante un campo elettrico.

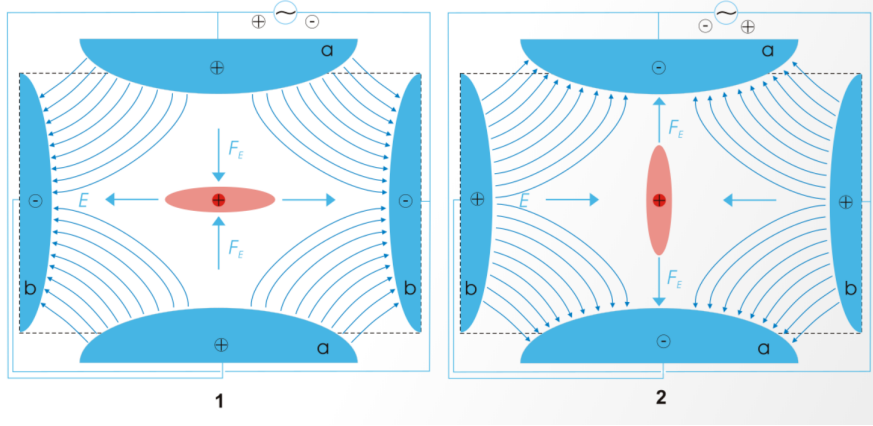


Figura (7.1) – Schema di una Paul trap. Uno ione positivo è sospeso tra coppie di elettrodi, che esercitano forze repulsive o attrattive in maniera periodica.

Non è possibile realizzare una trappola per soli effetti elettrostatici (teorema di Earnshaw), e perciò è necessario introdurre *correnti* e *scambiare periodicamente le cariche* sugli elettrodi, ad una frequenza sufficientemente alta (figura 7.1).

A livello fisico, perciò, lo ione può essere considerato come un oscillatore armonico in $d = 3$, con pulsazioni ω_z , ω_x e ω_y . Due di queste sono fissate “in modo stretto” dalla trappola, e l’altra viene utilizzata come grado di libertà quantistico. In altre parole: è richiesta molta energia per eccitare i modi lungo x o y , e molta meno per quelli lungo z - che quindi sono quelli effettivamente raggiunti.

Consideriamo perciò N ioni nella stessa trappola, per cui si ha l’Hamiltoniana:

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i=1}^N \frac{1}{2} \omega_z^2 z_i^2 + \sum_{i=1}^N \sum_{j>i}^N \frac{q^2}{4\pi\epsilon_0 |r_i - r_j|}$$

Tale hamiltoniana si risolve con i procedimenti utilizzati per gli oscillatori armonici accoppiati. In particolare ne consideriamo i *modi normali di oscillazione* - e in particolare i due in cui N particelle che oscillano “in fase” o “in opposizione di fase”.

Lo stato del sistema è allora dato da:

$$|\alpha_1\rangle |\alpha_2\rangle \dots |\alpha_N\rangle |n\rangle$$

dove $|\alpha_i\rangle$ sono stati interni agli ioni (es. due livelli energetici ben conosciuti) e $|n\rangle$ quelli dell’oscillatore armonico lungo z . In ordine di energia avremo allora:

$$|g, 0\rangle, |g, 1\rangle, |g, 2\rangle \dots |e, 0\rangle, |e, 1\rangle, |e, 2\rangle, \dots$$

(dove si sottintende il prodotto tensore con $|0\rangle_x |0\rangle_y$ che indica lo stato fondamentale degli oscillatori lungo x e y) dove con g indichiamo lo stato fondamentale e con e

il primo eccitato. Utilizzando *laser* possiamo stimolare delle transizioni tra i vari livelli.

Il reset avviene eccitando le transizioni $|g, n\rangle \rightarrow |e, n-1\rangle$. A tal punto uno ione eccitato *tende* a decadere nello stato fondamentale, e quindi $|e, n-1\rangle \rightarrow |g, n-1\rangle$. Poiché tali transizioni hanno tutte la stessa pulsazione ω (per effetto degli autovalori di un oscillatore armonico) si può usare un unico laser per *portare i qubit tutti allo stato fondamentale*. Tale procedimento è detto **side-band cooling**. Ciò ha un'efficienza $P > 99.9\%$.

Utilizzando poi *due laser* indipendenti è possibile realizzare il *Cirac-Zoller Gate*.

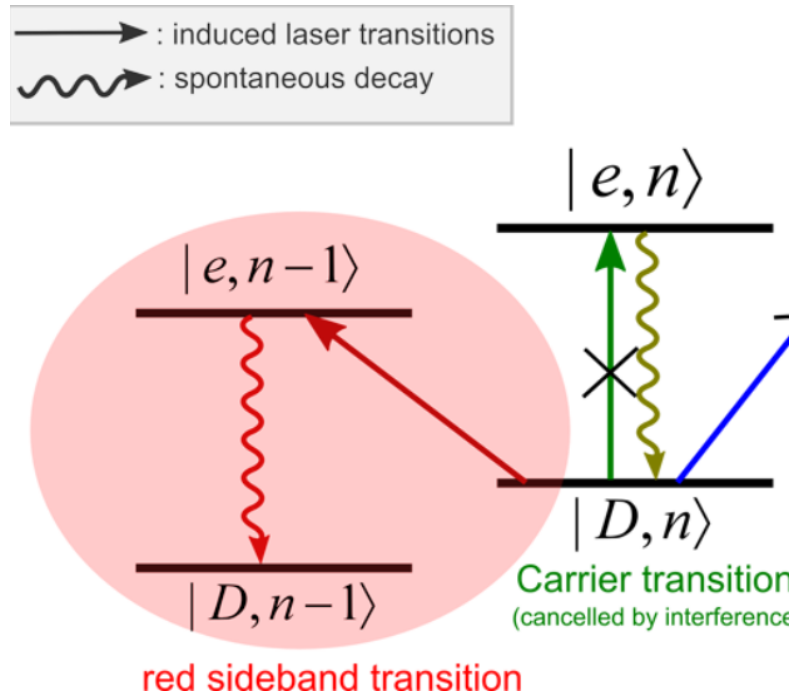



Figura (7.2) – Schema del procedimento di side-band cooling

7.0.2 Superconducting qubit

Raffreddando alcuni materiali si creano naturalmente *coppie di Cooper*, ossia coppie di elettroni “entangled”, che si comportano “come un bosone” e possono condensare¹. Tali coppie possono muoversi “senza resistenza”, e quindi produrre “correnti che non decadono” - nel fenomeno della **superconduttività**.

 Parte da ricontrollare

Per realizzare i qubit si usano delle **giunzioni Josephson**, realizzate inserendo un sottile strato isolante tra due strati di superconduttore - che si comportano come un condensatore con capacità C_j . Le coppie di Cooper possono *passare attraverso la barriera* per effetto tunnel. Connettendo una giunzione Josephson a un condensatore C_γ si crea un **charge qubit**. Le coppie di Cooper non possono

¹ L'entanglement in realtà avviene nello spazio reciproco, non in quello diretto - e quindi si ha una *trasformata di Fourier*. Tutto ciò si tratta nel framework della *seconda quantizzazione*, che è oltre agli obiettivi di questo corso.

attraversare le armature del condensatore (la distanza è troppo ampia) - e perciò lo stato superconduttore che si affaccia al condensatore è “un’isola”, nel senso che è completamente scollegato dal resto del circuito - e per depositarvi una coppia è necessario spendere una certa energia.

L’Hamiltoniana di tale sistema si dimostra essere:

$$H = \mathcal{E}_c(n - n_g)^2 - \mathcal{E}_j \cos \phi \quad [n, \phi] = i$$

dove n è il numero di coppie che occupano l’isola, e $n_g = C_\gamma V/(2e)$ è un numero “base” di coppie (corrispondente allo “zero dell’energia”), fissato dal potenziale V che alimenta il circuito. Si trova poi $\mathcal{E}_c = (2e)^2/(2(C_j + C_\gamma))$. $\phi = i\partial_n$ è il secondo numero quantico importante per il sistema, ed è detto *fase del superconduttore*.

Detto $|n\rangle$ lo stato in cui vi sono n coppie nell’isola si ha:

$$H = \mathcal{E}_c \sum (n - n_g)^2 |n\rangle \langle n| - \frac{1}{2} \mathcal{E}_j \sum |n+1\rangle \langle n| + |n\rangle \langle n+1|$$

Graficando i livelli energetici \mathcal{E} in funzione di n_g si ottiene:

[Missing]

Figura (7.3) – Grafico dei livelli

La presenza dei termini di interazione ($\cos \phi$) fa sì che i tratti del grafico di \mathcal{E} non si intersechino tra loro (*avoided crossing*). Si realizzano allora livelli separati che possono essere usati come qubit.

Elenco delle figure

1.1	Schema del funzionamento del motore di Szilard. In (a) si ha la configurazione iniziale, a temperatura T fissata. Si pone un pistone a dividere A e B , e dopo aver misurato la posizione della particella, in (b) si collega una massa m al pistone, dalla parte in cui si trova la particella. In tal modo, in (c) è possibile sfruttare il moto di quest'ultima per fare lavoro. Ipotizzando che il pistone (e carrucola e massa) si muovano senza attrito, il motore ha come unico effetto la produzione di <i>lavoro</i> senza alcuna spesa.	9
1.2	Schema a blocchi di un computer classico o quantistico	11
1.3	Schema di una comunicazione classica o quantistica	11
1.4	Schemi delle principali porte logiche	12
1.5	Schema del funzionamento delle porte logiche copy e swap	13
1.6	Come nel caso classico, rappresentiamo i <i>qubit</i> come linee, che saranno opportunamente collegate agli input delle porte logiche . . .	13
1.7	Rappresentazione grafica della sfera di Bloch	14
1.8	Schema di un computer quantistico, dove l'input consiste in 4 qubit inizializzati a $ 0\rangle$	15
1.9	Un qualsiasi operatore O (a) che esegue una certa operazione su n qubit può essere scomposto in una serie di porte logiche fondamentali (b) opportunamente collegate che svolgono la medesima operazione. . .	16
1.10	Si può estendere una qualsiasi funzione binaria $f : \{0, 1\}^n \rightarrow \{0, 1\}^m$ a una funzione reversibile $f' : \{0, 1\}^n \times \{0, 1\}^m \rightarrow \{0, 1\}^n \times \{0, 1\}^m$ dove i primi n qubit dell'output sono esattamente pari ai primi n dell'input.	22
1.11	Seppure una computazione quantistica possa avvenire in "parallelo" su tutte le singole configurazioni di cui lo stato iniziale è combinazione lineare, per poter "usare" l'output è necessario effettuare una misura, che collassa lo stato finale $ \psi_o\rangle = \sum_{x=0}^{2^n-1} c_x x\rangle y + f(x)\rangle$ in uno solo dei 2^n esiti possibili, con probabilità $ c_x ^2$	22
1.12	Schema di una porta logica <i>copy</i> quantistica	24

1.13	Schema della macchina di Turing	26
1.14	Attualmente sappiamo che $\text{NPC} \subset \text{NP}$, e $\text{P} \subset \text{NP}$, ma non è chiaro se valgano (eventualmente) inclusioni inverse.	27
1.15	Realizzazione dell'operatore δ tramite la composizione di 5 porte logiche quantistiche	29
2.1	Setup sperimentale per il teletrasporto quantistico	33
2.2	Schema circuitale del protocollo per il teletrasporto quantistico	33
2.3	Schema a porte logiche quantistiche del teletrasporto quantistico - con anche la generazione dello stato EPR	35
2.4	Schema del funzionamento del Beam Splitter.	37
2.5	Schema dell'interferometro di Mach-Zehnder.	38
2.6	Schema con porte logiche quantistiche dell'interferometro di Mach- Zehnder.	39
2.7	Interferometro di Mach-Zehnder utilizzato per misure senza intera- zione.	40
2.8	Nel caso in cui vi sia effettivamente una bomba, solo uno dei due percorsi è possibile, e perciò non si ha alcuna interferenza nel secondo beam-splitter.	40
2.9	Evoluzione temporale di un sistema quantistico instabile. Sull'asse y è riportata la <i>survival probability</i> , ossia la probabilità che il sistema rimanga nello stato iniziale (instabile)	43
2.10	Effetto di Zeno quantistico per $N = 5$ misure "alla Von Neumann". La linea blu continua mostra la <i>survival probability</i> nel caso di misure ripetute, mentre quella azzurra tratteggiata nel caso <i>senza misure</i> . La linea in verde puntinata è l'esponenziale che <i>interpola</i> l'effetto Zeno definito in (2.11).	45
2.11	L'evoluzione unitaria (a) fa sì che un vettore non abbandoni mai la superficie della sfera di Bloch, cosa che invece succede nel caso non unitario (b).	49
2.12	<i>Survival probability</i> $p(t)$ al variare del parametro di coupling V . Coupling maggiori equivalgono ad un decadimento <i>più lento</i> dallo stato iniziale.	53
2.13	Probabilità di transizione a $ 0\rangle$ per uno stato di partenza $ 0\rangle$ ($P_g(t)$) o $ 1\rangle$ ($P_e(t)$). Si nota che dopo un tempo π/Ω lo stato finale è <i>invertito</i> rispetto a quello iniziale.	62
3.1	La sfera di Bloch comprende tutti gli stati possibili per un sistema a 1 qubit. Gli stati <i>interni</i> sono stati misti , mentre quelli <i>sulla</i> <i>superficie</i> sono stati puri	71
3.2	Schema dell'azione CNOT che "distrugge" le coerenze di ρ_S	100
3.3	Rappresentazione geometrica della trasformazione $\rho \mapsto \rho'$ dal punto di vista dei vettori nella sfera di Bloch. Generalmente il processo non è invertibile, e il volume può solo diminuire.	102

3.4	L'operazione di bit-flip deforma la sfera di Bloch contraendola lungo le direzioni \hat{y} e \hat{z} di una quantità $1 - 2 \alpha ^2$, e lasciandola invariata lungo \hat{x}	103
3.5	Schema del circuito a gate quantistici equivalente all'operazione di bit-flip	103
3.6	L'operazione di phase-flip deforma la sfera di Bloch contraendola lungo le direzioni \hat{x} e \hat{y} di una quantità $1 - 2 \alpha ^2$, e lasciandola invariata lungo \hat{z}	104
3.7	Schema a gate quantistici dell'operazione di phase-flip (= schema del bit-flip, ma con σ_z al posto di σ_x)	105
3.8	Canale di Bit-Phase Flip	105
3.9	Azione del depolarizing channel sulla Sfera di Bloch	106
3.10	Schema grafico del canale di amplitude damping: la sfera di Bloch viene contratta verso il punto corrispondente allo stato $ 0\rangle\langle 0 $. . .	107
3.11	Azione asintotica del canale di Amplitude Damping	108
3.12	Esempio di distribuzione quantistica di una chiave tramite il protocollo BB84. Alice parte da una sequenza di bit iniziale e la converte in una successione di qubit in 4 stati possibili. Solo quando Bob misura il qubit nella stessa base usata da Alice per generare gli stati allora il bit classico iniziale è stato correttamente ricevuto.	117
3.13	Schema del protocollo di Dense Coding	119
3.14	Rappresentazione del canale come azione di gate quantistici	119
4.1	In una teoria locale, l'informazione completa contenuta nella regione 3 del diagramma di Minkowski è sufficiente a determinare le probabilità degli eventi nella regione 1, indipendentemente da quanto accade nella regione 2.	124
4.2	Schema dell'esperimento di Bell per esaminare la non-località della MQ. Una sorgente S di coppie di particelle entangled invia una particella ad Alice e una a Bob, che eseguono rispettivamente una misura X e Y sulla loro particella, ottenendo come risultati a e b . .	125
4.3	Diagramma di Minkowski per lo scenario di Bell. Gli apparati di Alice e Bob si trovano nelle regioni 1 e 2, e lo stato del sistema è definito nella regione 3.	125
4.4	Schema delle relazioni tra i possibili insiemi di scenari di Bell, delimitati da vincoli nello spazio \mathcal{P} (qui proiettati in $d = 2$). Si trova infatti che NS ha la struttura di un <i>politopo</i> . I piani che separano le varie classi sono detti, in generale, disuguaglianze di Bell	134
4.5	Relazioni tra Q , \mathcal{L} e NS per $\Delta = 2$, $m = 2$	134
4.6	A sinistra: schema del processo di <i>spontaneous parametric downconversion</i> . Come visibile a destra, i due fotoni risultanti sono emessi in un cono di angolo specifico.	135

4.7	I cristalli #1 e #2 hanno assi ottici perpendicolari: il primo verticale, il secondo orizzontale. Perciò #1 converte fotoni $ V\rangle \rightarrow H\rangle \otimes H\rangle$, mentre il secondo $ H\rangle \rightarrow V\rangle \otimes V\rangle$. Utilizzando come input fotoni polarizzati a 45° (sovrapposizione a pari coefficienti di $ H\rangle$ e $ V\rangle$) allora ogni fotone ha pari probabilità di essere splittato da #1 o da #2. Fino a quando non viene fatta una misura, entrambi i processi <i>sono esplorati</i> , e perciò lo stato finale è entangled.	136
4.8	Setup sperimentale per la violazione delle disuguaglianze CHSH . . .	136
4.9	Schema del funzionamento di una lamina $\lambda/2$. Poiché la lamina “riflette” la polarizzazione incidente, orientandola a $\theta = 45^\circ$ è possibile convertire $ H\rangle$ in $ V\rangle$ (e viceversa). Analogamente, per $\theta = 22.5^\circ$, si converte una polarizzazione a 45° ($ +\rangle$) in $ H\rangle$	137
4.10	Grafico dell’entropia di Shannon $H(p)$ per un messaggio binario. . .	139
4.11	Grafici relativi ad una mistura statistica di stati quantistici senza analogo classico (ossia non ortogonali)	144
5.1	Scatola nera: versione <i>classica</i> e <i>quantistica</i>	157
5.2	Schema a porte logiche quantistiche dell’algoritmo di Deutsch ($n = 1$)	157
5.3	Schema in porte logiche quantistiche dell’algoritmo di Grover per indici a 2 qubit	162
5.4	Interpretazione geometrica dell’azione di U_f su un generico vettore $ v\rangle$	165
5.5	Interpretazione grafica dell’algoritmo di Grover.	166
5.6	Schema a porte logiche per la QFT	171
5.7	Schema del circuito a porte logiche quantistiche per l’algoritmo di Shor	177
5.8	Schema a porte logiche quantistiche dell’algoritmo di Phase Estimation	181
5.9	Plot di esempio della trasformata di Fourier	185
5.10	Schema a porte logiche quantistiche per la generazione di ridondanza	188
5.11	Schema a porte logiche quantistiche per la misura di correlazioni tra i qubit ridondanti	189
5.12	Circuito a gate quantistici per realizzare la codifica del <i>3-qubit phase-flip code</i>	190
6.1	Grafico di $P_{if}(\omega)$	197
7.1	Schema di una Paul trap. Uno ione positivo è sospeso tra coppie di elettrodi, che esercitano forze repulsive o attrattive in maniera periodica.	201
7.2	Schema del procedimento di side-band cooling	202
7.3	Grafico dei livelli	203

Elenco delle tabelle

1	Cronologia di modifiche/aggiornamenti agli appunti	6
1.1	Tabella di verità per <i>not</i> (\bar{a}) e <i>identità</i> (buffer)	12
1.2	Tabella di verità per <i>and</i> $a \wedge b = ab$, <i>or</i> $a \vee b = a + b$ e <i>xor</i> $a \otimes b$. .	13
1.3	Tavola di verità per il gate CNOT	19
1.4	Tavola di verità per la porta CPHASE	21
2.1	Le prime due colonne riportano i possibili risultati della misura di Alice e il conseguente stato di B . Per ricondurre B allo stato del qubit C che si voleva teletrasportare, Bob deve eseguire l'operazione V indicata, dove indichiamo con \hat{X} e \hat{Z} gli operatori dati dalle rispettive matrici di Pauli σ_x e σ_z	34
3.1	Operazioni U svolte da Alice a seconda della combinazione dei 2 bit classici che vuole inviare a Bob	120
5.1	Vi sono solo 4 funzioni a 1-bit, 2 bilanciate (f_1 e f_2) e 2 costanti (f_0 e f_3)	157
5.2	Esiti possibili per il trasferimento di un bit con un fattore 3 di ridondanza	187
5.3	Possibili messaggi ricevuti a seguito di errori	189

Bibliografia

- [1] Kwiat et al. *Realistic Interaction-Free Detection of Objects in a Resonator* Foundation of Physics, 1998
- [2] Manzali Francesco, Mattia Morgavi *Trascrizione degli appunti del corso di MQ tenuto dal prof. Pieralberto Marchetti* https://drive.google.com/open?id=1q0yrQNs-9_2EH3r6irQlpMr8MoiWNe-C
- [3] Eric Lutz, Sergio Ciliberto. *Information: From Maxwell's demon to Landauer's eraser* Physics Today 68, 9, 30 (2015) <https://doi.org/10.1063/PT.3.2912>
- [4] Saverio Pascazio *All you ever wanted to know about the quantum Zeno effect in 70 minutes* 2014, <https://arxiv.org/abs/1311.6645>
- [5] <http://www.pas.rochester.edu/~howell/mysite2/Tutorials/Beamsplitter2.pdf> François Hénault *Quantum physics and the beam splitter mystery* <https://arxiv.org/ftp/arxiv/papers/1509/1509.00393.pdf>
- [6] Noam Nisan, Shimon Schocken *From Nand To Tetris Course* Coursera, <https://www.coursera.org/learn/build-a-computer/home/welcome>
- [7] Giuliano Benenti, Giulio Casati, Giuliano Strini *Principles of Quantum Computation and Information*
- [8] Nicolas Brunner, Daniel Cavalcanti et. al., *Bell nonlocality*, Rev. Mod. Phys. 86, 419 (2014), <https://arxiv.org/abs/1303.2849>
- [9] Travis Norsen, *John S. Bell's concept of local causality*, Am. J. Phys., Vol. 79, No. 12, December 2011
- [10] Sheldon Goldstein et al. (2011), *Bell's theorem*, Scholarpedia, 6(10):8378.
- [11] Reinhard F. Werner, *Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model*, Phys. Rev. A, Vol. 40, n. 8, 1989.

- [12] V. Vedral, *The role of relative entropy in quantum information theory*, Reviews of Modern Physics, Vol. 74, January 2002
- [13] Emmanuel N. Koukaras, *Fermi's Golden Rule*, A.M. 198. <http://staff.ustc.edu.cn/~yuanzs/teaching/Fermi-Golden-Rule-No-II.pdf>
- [14] B. Zwiebach, *Time Dependent Perturbation Theory*, MIT 8.06 Quantum Physics III - Spring 2018 lecture notes, https://ocw.mit.edu/courses/physics/8-06-quantum-physics-iii-spring-2018/lecture-notes/MIT8_06S18ch4.pdf
- [15] Philip Pearle, *Simple Derivation of the Lindblad Equation*, 11/4/2012, <https://arxiv.org/pdf/1204.2016.pdf>

Indice analitico

A			
Ancilla	91	Elitzur-Vaidman bomb tester	39
B		Entanglement of formation	148
Base computazionale		Equazione	
1 qubit	16	Liouville-von Neumann	66
2 qubit	18	Esempio	
Beable	123	Applicazione del noiseless coding	
C		140	
Classi di complessità	25	Calcolo di entropia di Von	
Coefficiente di Pearson	76	Neumann	143
Computer		Circuito per stato di massima	
classico	10	sovrapposizione	23
quantistico	10, 15	Correlazioni classiche vs	
Comunicazione		quantistiche	148
classica	11	Effetto Zenone per oscillazioni di	
quantistica	11	Rabi	46
Concurrence	151	Matrice densità di un qubit	69
Correlazioni	131	POVM Measurement	98
Crittografia	113	Purezza di matrici ridotte	73
D		Purificazione di 1 qubit	84
Decomposizione di Schmidt	78	Stati con correlazioni classiche	146
E		Weak measurement di 2 qubit	95
Effetto Zenone		Esercizio	
Hamiltoniana di interazione	48	Circuito per stato generico	28
Oscillazioni di Rabi	46	Correlazioni ed entanglement	153
Survival probability	43	CPHASE in termini di CNOT e	
Tempo di Zenone	44	phase-shift	28
Effetto Zenone:Hamiltoniane non		Proprietà di fidelity	30
hermitiane	49	F	
I		Fidelity	30
		I	
		Interferometro di Mach-Zehnder	38

L			
Località	122	Hadamard	16, 23
Fattorizzabilità	126	Phase-shift	17
LOCC	146	Postulato	
M		Proiezione di Von Neumann	36
Macchina di turing	26	Potenziale ottico	49
Matrice densità	65	Principio	
ridotta	72	Landauer	9
Maxwell's demon	8	Purificazione	83
Mistura statistica	64	Purità	68
Misura		R	
di Von Neumann	90	Rappresentazione di Kraus	87
Generalizzata	90	S	
N		Scenario di Bell	131
Notazione		Sfera di Bloch	14
Matrici e braket	54	Szilard's engine	8
O		T	
Operatore		Teletrasporto quantistico	32
Kraus	87	Teorema	
P		Bell	127
Parallelismo quantistico	22	Neumark	91
Porte logiche	12	No cloning	24
1-bit classiche	12	Noiseless Coding (Shannon)	140
2-bit classiche	12	Quantum Noiseless Coding	145
Beam-splitter	37	Termini di coerenza	69
CNOT	18	Termini di popolazione	69
CPHASE	20	Traccia	65
		parziale	72