

UNIVERSITÀ DEGLI STUDI DI NAPOLI
“FEDERICO II”



Scuola Politecnica e delle Scienze di Base
Area Didattica di Scienze Matematiche Fisiche e Naturali

Dipartimento di Fisica “Ettore Pancini”

Laurea Triennale in Fisica

**Entanglement, la fisica dell’informazione e i fondamenti
della meccanica quantistica**

Relatori:
Prof. Massimo Taronna

Candidato:
Francesco Nappi
Matr. N85001221

Anno Accademico 2019/2020

Indice

1	Introduzione	3
2	Matrice Densità e sistemi aperti	9
2.1	Sfera di Bloch	12
2.2	Decomposizione di Schmidt	13
2.3	Ambiguità delle preparazioni e purificazione	15
3	Misure ortogonali e generalizzate	17
3.1	Misure ortogonali	17
3.2	Misure generalizzate	18
4	Canali quantistici	22
4.1	Dualismo stato-canale e i nuovi postulati della MQ	24
4.2	Master equation dei sistemi aperti	25
5	Decoerenza	28
5.1	Toy Model	28
5.2	La decoerenza e il problema della misura	30
6	Entanglement, realismo e località	32
6.1	Articolo EPR	32
6.2	Il teorema di Bell	34
7	Informazione quantistica	37
7.1	Entropia di Von Neumann	37
7.2	Teoremi dell'informazione quantistica	39
7.3	Teletrasporto Quantistico	41
8	Conclusioni	44

Sommario

In questo elaborato sono state trattate le fondamenta della Meccanica Quantistica, con particolare attenzione alla Meccanica Quantistica dei sistemi aperti e al loro studio e sono state poste le basi dell'Informazione Quantistica. Sono stati affrontati anche interessanti problemi tutt'ora aperti, come il problema della misura e i concetti di realismo e località. All'inizio viene introdotta la Teoria dell'Informazione classica. Successivamente vengono introdotti e studiati i postulati della Meccanica Quantistica dei sistemi aperti e il fenomeno dell'entanglement. In particolare vedremo a cosa viene associato un sistema aperto, come viene descritto un sistema aperto e come il concetto di misura si generalizzi in questo caso. Questo studio ci conduce a studiare ed analizzare i problemi di realismo, completezza e località della Meccanica Quantistica e il problema della misura. Infine utilizzeremo tutte le conoscenze acquisite per introdurre la teoria dell'Informazione Quantistica e tracciare un parallelismo tra elementi della teoria dell'informazione classica e quella quantistica. Quest'opera mira ad essere chiara e precisa in ogni suo punto, a costo di omettere, in alcuni passaggi, delle dimostrazioni, facilmente reperibili nella letterature scientifica, e di utilizzare esempi e toy model, proprio per facilitare la comprensione del lettore.

1 Introduzione

Fino al 1948, anno in cui Claude Shannon pubblica il suo celebre articolo, *A Mathematical Theory of Communication*[26], l'informazione non era mai stata trattata come una vera quantità fisica. Shannon ha il merito di aver fondato la teoria dell'informazione, ponendo veri e propri limiti alla quantità e alla precisione con cui è possibile comunicare un'informazione. In particolare ha fornito una risposta alle seguenti fondamentali domande:

1. Quanto un messaggio può essere compresso?
2. Qual è il massimo tasso di trasmissione di dati in un canale con un determinato rumore? Ovvero quanto deve essere ridondante il nostro messaggio per poterlo proteggere dagli errori?

In questa introduzione vedremo brevemente i teoremi dell'informazione classica e il concetto di entropia nella teoria dell'informazione. Successivamente porremo le basi della formulazione matematica dei sistemi aperti e dei canali nella meccanica quantistica. Infine tratteremo il concetto di entropia di Von Neumann e di informazione quantistica, proponendo anche un'applicazione di quest'ultima.

Prima di enunciare i teoremi di Shannon introduciamo i concetti di bit, entropia e di capacità di un canale.

Iniziamo con il definire la *quantità di informazione*. Questo concetto è strettamente legato al concetto di probabilità. Infatti l'avvenimento di un evento certo, ovvero con probabilità che si verifichi uguale ad uno, non aggiunge alcuna informazione, mentre l'avvenimento di un evento altamente improbabile apporterà molta più informazione. Quindi è naturale che la quantità di informazione sia una funzione decrescente della probabilità e in particolare valga 0 quando la probabilità è pari a 1. Si può quindi comprendere perché Shannon definisce la quantità di informazione, anche detta *autoinformazione*, di un evento E :

$$I(E) = -\log(p(E)). \quad (1)$$

La base del logaritmo non è specificata, poiché per ogni scelta della base corrisponde una diversa unità di misura dell'informazione. Solitamente la base scelta è 2 e l'unità di misura corrispondente è il bit. Un bit permette di trasmettere l'informazione relativa a quale alternativa scegliere tra due eventi equiprobabili. In generale possiamo rappresentare un bit tramite due possibili valori, 0 e 1. Immaginiamo di assegnare al valore 0 il significato di sinistra e al valore 1 il significato di destra e di dovere percorrere una strada composta da una serie di bivi. Se abbiamo n possibili bivi, avremo allora 2^n possibili destinazioni e avremo bisogno di n bit di informazione per arrivare alla giusta destinazione.

Supponiamo di avere un insieme di possibili eventi le cui probabilità che si verifichino siano rispettivamente $X = \{p_1, p_2, \dots, p_n\}$. È possibile stabilire in maniera quantitativa una misura della *quantità di scelta* presente nella selezione di quale evento avverrà e in generale dell'*incertezza del risultato* in un insieme di questo tipo? Se esistesse una tale funzione $H(X)$ dovrebbe avere 3 requisiti fondamentali:

1. H dovrebbe essere una funzione continua della variabile i -esima p_i .
2. Se tutti gli eventi fossero equiprobabili, $p_i = \frac{1}{n}$, allora H dovrebbe essere una funzione monotona crescente di n . Ciò traduce il fatto che l'incertezza è massima se tutti gli eventi sono equiprobabili e questa aumenta con n .

3. Se un singolo viene diviso in due processi consecutivi, allora la funzione H originaria dovrebbe essere uguale alla somma pesata delle due funzioni H dei singoli processi.

È possibile dimostrare che l'unica funzione H che rispetti le 3 condizioni precedenti è del tipo:

$$H(X) = -K \sum_i^n p_i \log p_i. \quad (2)$$

Dove K è una costante positiva. Ponendo $K = 1$, la funzione che si ottiene è chiamata *entropia di Shannon* dell'insieme X , proprio per la sua forma simile all'entropia di Boltzmann definita in Termodinamica. Notiamo come l'entropia di Shannon è il valore atteso dell'autoinformazione. In generale, data una variabile aleatoria discreta x e la sua distribuzione di probabilità $\{p(x_1), \dots, p(x_n)\}$ definiamo entropia di Shannon della variabile x la quantità:

$$H(x) = - \sum_i^n p(x_i) \log(p(x_i)). \quad (3)$$

Enunciamo, senza dimostrare, una serie di proprietà di $H(X)$:

1. $H = 0$ se una probabilità è uguale ad 1 e il resto è uguale a 0 (data la condizione di normalizzazione della probabilità). Ciò si intuisce perché se un evento ha probabilità uguale a 1, allora è certo mentre i restanti eventi sono impossibili, quindi non c'è incertezza nel risultato.
2. H è massimo se gli eventi sono equiprobabili, perché è appunto la situazione in cui l'incertezza è massima.
3. Supponiamo due possibili eventi, x e y . Il primo può avvenire in n modi diversi, mentre il secondo può avvenire in m modi diversi. Sia $p(i, j)$ la probabilità che il primo evento avvenga nell' i -esimo modo e il secondo evento nell' j -esimo modo. Allora si verifica:

$$H(x, y) \leq H(x) + H(y). \quad (4)$$

Dove:

$$H(x, y) = - \sum_{i,j}^{n,m} p(i, j) \log p(i, j),$$

$$H(x) = - \sum_{i,j}^{n,m} p(i, j) \log \sum_j^m p(i, j),$$

$$H(y) = - \sum_{i,j}^{n,m} p(i, j) \log \sum_i^n p(i, j).$$

L'uguaglianza è valida solo nel caso in cui $p(i, j) = p(i)p(j)$, ovvero solo nel caso in cui in cui i due eventi siano indipendenti. Quindi l'entropia congiunta di due eventi è sempre minore o uguale dell'entropia dei due singoli eventi.

Proviamo a chiarire il significato di autoinformazione ed entropia attraverso un piccolo esempio. Supponiamo di avere una regolare moneta. Allora la possibilità di ottenere testa o croce dopo

un lancio è la stessa ed è pari a 0,5. Quindi l'autoinformazione riguardante l'esito di un lancio sarà:

$$I(t) = I(c) = \log_2\left(\frac{1}{p(x)}\right) = \log_2\left(\frac{1}{0.5}\right) = 1bit. \quad (5)$$

Conseguentemente l'entropia di Shannon sarà uguale a:

$$H(x) = \frac{1}{2}\log_2\left(\frac{1}{0.5}\right) + \frac{1}{2}\log_2\left(\frac{1}{0.5}\right) = 1bit. \quad (6)$$

Come si vede essa è uguale alla media dell'autoinformazione dei due possibili eventi. Ora consideriamo il caso di una moneta non regolare, in cui la probabilità che il risultato di un lancio sia testa è pari a 0,9 mentre la probabilità che sia croce è pari a 0,1. Allora l'autoinformazione relativa ai due eventi è uguale a:

$$I(c) = \log_2\left(\frac{1}{0.1}\right) = 3,32bit, \quad I(t) = \log_2\left(\frac{1}{0.9}\right) = 0.15bit. \quad (7)$$

Quindi vediamo come l'evento meno probabile, l'esito croce, apporti molti più informazione rispetto all'evento più probabile. Ora calcoliamo l'entropia di Shannon, il valore atteso:

$$H(x) = \frac{9}{10}\log_2\left(\frac{9}{10}\right) + \frac{1}{10}\log_2\left(\frac{1}{10}\right) = 0.50bit. \quad (8)$$

Vediamo come essa sia inferiore al caso in cui gli eventi siano equiprobabili. Nel caso di più variabili è possibile definire anche due nuove quantità, l'*entropia condizionale* e la *mutua informazione*. Abbiamo visto come l'entropia sia una misura dell'incertezza e della quantità di scelta contenuta in un evento aleatorio. Come viene influenzata questa quantità dalla conoscenza di altra informazione? Dati gli insiemi $X = \{p_x, x\}$ e $Y = \{p_y, y\}$, che possono rappresentare ad esempio il messaggio da trasmettere e il messaggio effettivamente ricevuto, quanto vale l'entropia della sorgente una volta osservato il messaggio ricevuto?

Si definisce entropia di X condizionata da Y la quantità:

$$H(X|Y) = \sum_i^N \sum_k^M p(x_i, y_k) \log\left(\frac{p(y_k)}{p(x_i, y_k)}\right). \quad (9)$$

Quindi se $H(X)$ rappresenta la quantità di incertezza sulla sorgente e $H(X|Y)$ rappresenta l'incertezza residua dopo aver osservato l'informazione trasmessa, la quantità $H(X) - H(X|Y)$ assume il significato di quanta incertezza sulla sorgente è stata eliminata dall'osservazione dell'informazione trasmessa, e quindi quantifica anche la correlazione tra X e Y e viene chiamata *mutua informazione*:

$$I(X; Y) = H(X) - H(X|Y) = \sum_i^N \sum_k^M p(x_i, y_k) \log\left(\frac{p(x_i, y_k)}{p(x_i)p(y_k)}\right). \quad (10)$$

Si vede che tale quantità è simmetrica, $I(X; Y) = I(Y; X)$ ed è uguale a 0 se le due variabili sono indipendenti.

Date n variabili discrete identiche e uniformemente distribuite x , l'entropia totale, ovvero il valore atteso dell'autoinformazione, ricordando che l'entropia gode delle proprietà del logaritmo, è pari a:

$$H(x_1, x_2, \dots, x_n) = \sum_i^n H_i(x) = nH(x). \quad (11)$$

Allora è facile vedere che l'entropia non dipende dal numero di simboli presenti. Ma se le variabili non fossero uniformemente distribuite? Allora si può definire il tasso di entropia, ovvero l'entropia media per simbolo come:

$$\lim_{n \rightarrow \infty} = \frac{H(x_1, x_2, \dots, x_n)}{n}. \quad (12)$$

Ora passiamo a definire *sorgente*, *canale* e la sua capacità. Una sorgente è la fonte che emette l'informazione da trasmettere. Una sorgente discreta è un tipo di sorgente che emette simboli che appartengono ad un insieme finito ciascuno caratterizzato da una probabilità e autoinformazione. In generale un canale è il tramite attraverso cui la comunicazione avviene. Esso può essere costituito da cavi, frequenze radio, raggi di luce etc. Definiamo la capacità C di un canale discreto non affetto da rumore come:

$$C = \lim_{T \rightarrow \infty} \frac{N(T)}{T}, \quad (13)$$

dove $N(T)$ è il numero di segnali permessi in un periodo T . Ora siamo pronti per enunciare il primo teorema di Shannon. [17]

Primo Teorema di Shannon 1 *Date n variabili uguali e identicamente distribuite ognuna con entropia $H(X)$ allora esse possono essere compresse in $nH(x)$ bit con probabilità di errore nulla quando $n \rightarrow \infty$. Viceversa, se compresse in meno di $nH(x)$ bit, allora dell'informazione verrà sicuramente persa.*

Quindi l'entropia pone un limite alla compressibilità dei dati. Forniamo un piccolo esempio per capire appieno il significato di tale teorema. Consideriamo un piccolo alfabeto di appena tre lettere $\{a, b, c\}$ e associamo ad ogni lettera una probabilità $\{p(a) = 0,5, p(b) = 0,35, p(c) = 0,15\}$. Come abbiamo visto, per distinguere tra 3 possibili alternative, avremo bisogno di 2 bit di informazione. Pertanto, dato il seguente messaggio:

$$a \ a \ b \ b \ c \ a \quad (14)$$

Per trasmetterlo avremo bisogno di 12 bit. Esiste una codifica migliore di questo messaggio, in modo da poterlo comprimere in un numero minore di bit? Una codifica efficace in questo caso è

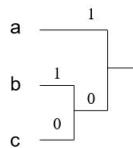


Figura 1: Schema della codifica di Huffman.

la *codifica di Huffman*. Tale codifica si basa sul partire dai simboli con la minore probabilità per poi unirli in un nuovo simbolo la cui probabilità è la somma delle probabilità dei due simboli uniti. Successivamente bisogna iterare tale processo fino a quando non sarà rimasto un solo simbolo. In questo caso specifico, assegniamo il valore 0 al simbolo meno probabile da unire e 1 al secondo simbolo. Come è evidente dallo schema, avremo che i codici di Huffman sono rispettivamente $a \ 1, b \ 01, c \ 00$. Comprimeando il messaggio precedente avremo che la codifica

di Huffman è: 1 1 01 01 00 1 . Quindi abbiamo utilizzato 9 bit, invece dei 12 bit apparentemente necessari. È questa la miglior codifica possibile? L'entropia dell'insieme è pari a 1,44, che moltiplicato per 6 è uguale a 8,64. Quindi vediamo come 9 sia effettivamente il numero di bit minimo per rappresentare un tale messaggio.

Prima di enunciare il secondo teorema di Shannon, che pone un limite alla trasmissione di un canale, formalizziamo alcuni dei concetti introdotti precedentemente. Abbiamo già visto la definizione di capacità per un canale privo di rumore. Ora invece descriveremo la trasmissione attraverso un canale affetto da rumore. Il messaggio inviato potrebbe non essere uguale al messaggio ricevuto, a causa del rumore. Il nostro obiettivo è capire quanta informazione è possibile trasmettere con una probabilità di errore trascurabile.

Dato un messaggio W , composto da n elementi presi da un alfabeto $\{1, 2, \dots, M\}$ il codificatore lo riceve e lo comprime in una sequenza X^n , schematizziamo questo passaggio con una funzione generica f del tipo $f: \{1, 2, \dots, M\} \rightarrow X^n$. Successivamente tale messaggio è trasmesso tramite un canale. In generale il messaggio ricevuto W' sarà in un alfabeto Y e quindi una sequenza Y^n , la quale viene decodificata da un decodificatore, rappresentabile da una funzione generica $g: Y^n \rightarrow \{1, 2, \dots, M\}$. Il messaggio ricevuto, W' , dovrebbe essere quanto più vicino al messaggio inviato W . Quindi per descrivere un canale, usiamo la terna $(X^n, Y^n, p(y|x))$, dove $p(y|x)$ indica la probabilità condizionata, ovvero la probabilità che il canale restituisca y dopo aver ricevuto x . Definiamo un canale *memoryless* se $p(y_k|x_k) = p(y_k|x_k, y_{k-1})$, ovvero la probabilità di decodificare correttamente un simbolo non dipende dai simboli precedenti.

Come abbiamo già evidenziato, in presenza di rumore il segnale inviato non può essere ricostruito esattamente dal ricevitore, quindi dobbiamo considerare una correzione alla quantità di informazione effettivamente trasmessa. In particolare tale correzione dovrebbe essere uguale all'incertezza sul segnale ricevuto, quindi possiamo affermare che il tasso effettivo di informazione trasmessa è uguale al tasso di produzione di informazione della sorgente, ovvero alla sua entropia, meno tale incertezza, ovvero:

$$R = H(x) - H(X|Y). \quad (15)$$

Quindi vediamo come la presenza del rumore fa diminuire la quantità di informazione. Allora è necessario ridefinire la capacità di un canale tenendo conto di questa correzione, infatti la capacità di un canale dovrebbe essere uguale al massimo tasso di trasmissione possibile, quindi la *capacità di un canale affetto da rumore* sarà uguale a:

$$C = \max_{p(x)} (H(x) - H(X|Y)) = \max(I(X; Y)). \quad (16)$$

La distribuzione di probabilità $p(x)$ che massimizza la mutua informazione è chiamata *optimal input distribution*. Ora possiamo enunciare la prima parte del secondo teorema di Shannon[25].

Secondo Teorema di Shannon 1 *Per ogni $\varepsilon > 0$ e tasso di trasmissione $R < C$ e N sufficientemente grande, esiste un codice di lunghezza N , un tasso di trasmissione $\leq R$ e un algoritmo di decodifica tali che la probabilità di errore è minore di ε*

Dove la probabilità di errore, dato un insieme di simboli $X = \{x_1, x_2, \dots, x_n\}$ e una distribuzione di probabilità $p_X(x_i)$, è definita come: $p_b = \sum_{x_i \in X} p_X(x_i) p(x'_i \neq x_i | x_i)$.

Questo teorema è un risultato teorico fondamentale, ponendo un limite al tasso di trasmissione di un canale, però allo stesso tempo nasconde delle insidie. Infatti dal punto di vista pratico non è sempre semplice trovare la capacità di un canale, inoltre il teorema non fornisce alcuna indicazione su come trovare la optimal input distribution. Però nel caso si riuscisse a

trovare il valore della capacità, potrebbe essere un parametro importante con cui confrontare il valore del tasso della trasmissione di una determinata codifica per capire quanto quest'ultima sia efficace.

Questi teoremi sono le pietre miliari della teoria dell'informazione, su cui è stata fondata successivamente questa importante teoria. Vedremo più avanti come questi risultati possono essere generalizzati in Meccanica Quantistica, descrivendo quella che è l'attuale Teoria dell'Informazione Quantistica.

2 Matrici Densità e sistemi aperti

Analogamente alla Termodinamica, anche *in Meccanica Quantistica esistono sistemi aperti*. Un sistema aperto in Meccanica Quantistica è un sistema che può scambiare energia e informazione con l'ambiente. I sistemi aperti sono estremamente importanti, in quanto sono un modello molto più vicino alla realtà rispetto ai sistemi chiusi. Infatti, è impossibile isolare perfettamente un sistema e pertanto è necessario studiare il sistema fisico dato dal sistema e da tutto ciò che lo circonda e con cui è libero di scambiare informazione: *l'ambiente*. Quindi impareremo a vedere il nostro sistema come parte di qualcosa di più grande. Ovviamente i postulati che governano la fisica dei sistemi aperti non sono gli stessi che governano la fisica dei sistemi chiusi e in particolare:

1. Uno stato fisico *non* è un vettore;
2. La misura *non* è una proiezione ortogonale;
3. L'evoluzione *non* è unitaria.

Per descrivere i sistemi aperti partiremo dalla descrizione della correlazione tra due sistemi quantistici chiusi, in particolare di due qubit.

Un qubit è un vettore in uno spazio di Hilbert di dimensione 2 e rappresenta l'unità di informazione quantistica. Quindi un qubit può essere rappresentato nella forma:

$$|\psi\rangle = a|0\rangle + b|1\rangle, \quad (17)$$

dove $a, b \in \mathbf{C}$ e sono tali che $|a|^2 + |b|^2 = 1$ e $|0\rangle, |1\rangle$ sono i vettori di base dello spazio di Hilbert. Come sappiamo dai postulati della Meccanica Quantistica, lo spazio di Hilbert di un stato nato dalla correlazione di più stati è dato dal prodotto tensoriale degli spazi di partenza. Quindi dati due qubit A e B i cui vettori di base siano rispettivamente $|0\rangle_A, |1\rangle_A$ e $|0\rangle_B, |1\rangle_B$, un possibile vettore appartenente allo spazio di Hilbert $H = H_A \otimes H_B$ che descrive lo stato dei due qubit è:

$$|\psi\rangle_{AB} = a|0\rangle_A \otimes |0\rangle_B + b|1\rangle_A \otimes |1\rangle_B. \quad (18)$$

Da questo momento useremo la notazione:

$$|0\rangle_A \otimes |0\rangle_B \equiv |0\rangle|0\rangle \equiv |00\rangle. \quad (19)$$

Dato lo stato $|\psi\rangle_{AB}$, se misuriamo il qubit A abbiamo probabilità $|a|^2$ di osservare lo stato $|0\rangle_A \otimes |0\rangle_B$ e probabilità $|b|^2$ di osservare lo stato $|1\rangle_A \otimes |1\rangle_B$. Inoltre, in base al postulato sulla riduzione del pacchetto d'onda, se misuriamo il qubit B subito dopo questa prima misura, saremo certi di osservare lo stato $|0\rangle_B$ se il risultato della prima misura è $|0\rangle_A$ e di osservare lo stato $|1\rangle_B$ se il risultato della prima misura è $|1\rangle_A$. Pertanto possiamo affermare che i possibili risultati delle misure di A e B sono completamente correlati.

Generalizziamo il nostro discorso a un qualsiasi osservabile M_A agente su A . Dato che M_A agisce solo su A possiamo descrivere la sua azione complessiva nello spazio di Hilbert H come $M_A \otimes I_B$ dove I_B è l'operatore identità dello spazio di Hilbert H_B . Troviamo il valore medio di M_A :

$$\begin{aligned} \langle M_A \rangle &= \langle \psi | M_A \otimes I_B | \psi \rangle \\ &= (a^* \langle 00 | + b^* \langle 11 |) | M_A \otimes I_B | (a | 00 \rangle + b | 11 \rangle) \\ &= |a|^2 \langle 0 | M_A | 0 \rangle + |b|^2 \langle 1 | M_A | 1 \rangle. \end{aligned} \quad (20)$$

Notiamo che tale espressione può essere scritta in una forma più compatta come:

$$\langle M_A \rangle = \text{Tr}\{M_A \rho_A\}, \quad (21)$$

dove ρ_A è la *matrice densità* del qubit A ed è definita come:

$$\rho_A = |a|^2 |0\rangle\langle 0| + |b|^2 |1\rangle\langle 1|. \quad (22)$$

Possiamo attribuire all'operatore densità il significato fisico di insieme ponderato dei possibili stati quantistici, ognuno con una propria probabilità. Infatti notiamo che il risultato di $\langle M_A \rangle$ è lo stesso che avremmo ottenuto ipotizzando che il qubit A fosse nello stato $|0\rangle$ con probabilità $|a|^2$ e nello stato $|1\rangle$ con probabilità $|b|^2$. È importante osservare che l'operatore densità non tiene conto della coerenza del qubit A. Per spiegarci meglio usiamo un esempio per ricordare il concetto di coerenza. Una combinazione lineare di stati quantistici si dice coerente se la fase relativa dei diversi stati ha conseguenze osservabili. Consideriamo a titolo d'esempio il qubit:

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle \pm |\downarrow\rangle), \quad (23)$$

dove $|\uparrow\rangle, |\downarrow\rangle$ sono gli autostati dell'operatore σ_z , l'osservabile di spin sull'asse z. Gli stati $|\pm\rangle$ sono gli autostati dell'operatore σ_x , l'osservabile di spin sull'asse x e hanno autovalori $\pm\frac{1}{2}$, pertanto la loro fase relativa è fondamentale in quanto ha conseguenze osservabili. Ora invece consideriamo l'operatore densità di un qubit A, dove A è un insieme equiprobabile di stati $|\uparrow\rangle$ e $|\downarrow\rangle$, proprio come lo stato $|+\rangle$. Però come abbiamo detto un operatore densità è un insieme ponderato di stati quantistici e non una sovrapposizione coerente come un vettore di stato. Pertanto allo stato $|+\rangle$ e $|-\rangle$, essendo miscele equiprobabili di $|\uparrow\rangle$ e $|\downarrow\rangle$, corrisponde lo stesso operatore densità:

$$\rho = \frac{1}{2}(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|) = \frac{1}{2}I. \quad (24)$$

Quindi, in base all'espressione 21, avremo che:

$$\langle \sigma_x \rangle = \text{Tr}\{|+\rangle\langle+|\rho\} = \langle+|\rho|+\rangle = \frac{1}{2}, \quad (25)$$

ovvero il valore medio dell'osservabile spin up lungo l'asse x applicato alla matrice densità di una miscela equiprobabile di $|\uparrow\rangle$ e $|\downarrow\rangle$ è pari $1/2$, mentre per una sovrapposizione coerente di questi due stati sarebbe stato 1 o 0, essendo le due possibili sovrapposizioni autostati dell'osservabile spin up o spin down.

Andiamo a generalizzare questi concetti ad uno spazio di Hilbert di un sistema bipartito di dimensione arbitraria, ovvero ad uno spazio di Hilbert $H = H_A \otimes H_B$, dove H_A e H_B sono gli spazi di Hilbert, rispettivamente di dimensione n e m , delle due parti del nostro sistema. Per introdurre la matrice densità, partiamo dal caso di uno stato puro di H (preciseremo le definizioni di stato puro e di miscela statistica più avanti). Uno stato puro di H è un vettore del tipo:

$$|\psi_{AB}\rangle = \sum_{i,\alpha}^{n,m} a_{i\alpha} |i\rangle |\alpha\rangle. \quad (26)$$

Consideriamo di nuovo il valore medio di un operatore M_A che agisca solo su A , pertanto il suo prolungamento su H è $M_A \otimes I_B$ dove I_B è la matrice identità dello spazio H_B :

$$\begin{aligned}
\langle M_A \rangle &= \langle \psi_{AB} | M_A \otimes I_B | \psi_{AB} \rangle \\
&= \sum_{j,\beta}^{n,m} a_{j,\beta}^* \langle j\beta | M_A \otimes I_B | \sum_{i,\alpha}^{n,m} a_{i,\alpha} | i\alpha \rangle \\
&= \sum_{i,j,\alpha}^{n,n,m} a_{j,\alpha}^* a_{j,\alpha} \langle j | M_A | i \rangle = \text{Tr}\{M_A \rho_A\}.
\end{aligned} \tag{27}$$

Dove:

$$\rho_A = \text{Tr}_B |\psi\rangle\langle\psi| = \sum_{i,j,\alpha}^{n,n,m} a_{j,\alpha}^* a_{i,\alpha} |i\rangle\langle j|. \tag{28}$$

È la matrice densità della parte A del nostro sistema. Con la notazione Tr_B abbiamo indicato la traccia parziale di un operatore sul sottospazio B . Diamo una definizione generale di tale operazione [18]. Sia O un operatore di uno spazio di Hilbert H di un sistema bipartito, ovvero $H = H_A \otimes H_B$ e indichiamo con $L(H)$ lo spazio degli operatori di H . La traccia parziale di O sul sottospazio B di dimensione n può essere definita come:

$$\text{Tr}_B O = \sum_{i=1}^n (\langle I_A | \otimes \langle i |) O (|I_A\rangle \otimes |i\rangle), \tag{29}$$

dove $|i\rangle$ rappresentano una base ortonormale del sottospazio B e I_A è la matrice identità agente nel sottospazio A . Quindi è un'applicazione del tipo:

$$L(H) \rightarrow L(H_A). \tag{30}$$

Dalla definizione 28 è possibile enunciare subito delle proprietà fondamentali dell'operatore densità:

1. ρ_A è hermitiano: $\rho_A^\dagger = \rho_A$;
2. ρ_A è semidefinita positiva : $\forall |\phi\rangle, \langle\phi| \rho_A |\phi\rangle = \sum_{\alpha}^n |\sum_i^n a_{i,\alpha} \langle i|\phi\rangle|^2 \geq 0$;
3. $\text{Tr} \rho_A = \sum_{i,\alpha}^{n,m} |a_{i,\alpha}|^2 = 1$ dato che $|\psi_{AB}\rangle$ è normalizzato.

Da queste proprietà segue che ρ_A può essere diagonalizzata e ha autovalori reali, non negativi e la cui somma è 1. Quindi abbiamo visto che quando studiamo un lo stato di un sottosistema fisico facente parte di un sistema più grande, il suo stato è in generale rappresentato dal suo operatore densità. Se lo stato del sottosistema è rappresentabile tramite un singolo vettore, diremo che è uno *stato puro*, altrimenti diremo che è una *miscela statistica*. La matrice densità di uno stato puro $|\psi_A\rangle$ è data semplicemente dal proiettore $|\psi_A\rangle\langle\psi_A|$. Invece nel caso più generale una matrice densità si presenta nella forma:

$$\rho_A = \sum_a^n p_a |a\rangle\langle a|, \tag{31}$$

dove $|a\rangle$ è la base in cui ρ_A è diagonale e i p_a sono tali che $\sum_a^n p_a = 1$. Inoltre notiamo che in generale la matrice densità non è un proiettore, ovvero $\rho_A^2 \neq \rho_A$. Osserviamo che anche in

questo caso generale vale una formula analoga a 21. Infatti dato un operatore M_A che agisce solo nel sottospazio H_A , il valore medio di M_A vale:

$$\langle M_A \rangle = \text{Tr}\{M_A \rho_A\} = \sum_a^n p_a \langle a | M_A | a \rangle. \quad (32)$$

Quindi è evidente che ρ_A rappresenti un'unione incoerente di stati puri $|a\rangle$, ognuno con probabilità a .

2.1 Sfera di Bloch

È possibile dimostrare che le 3 matrici di Pauli, che verranno indicate con la notazione $\sigma^i, i = 1, 2, 3$ insieme alla matrice identità I formano una base dello spazio delle matrici 2×2 . Andiamo quindi a vedere come scrivere la nostra matrice densità ρ usando questa base. In generale una matrice si può esprimere nella base delle matrici di Pauli come:

$$M = cI + a_i \sigma^i, \quad (33)$$

dove $c \in \mathbf{C}$ e \mathbf{a} è un vettore tridimensionale e $a_i \in \mathbf{C}$. Notare che la somma su indici ripetuti è implicita, seguendo la notazione di Einstein. Ricordando che la matrice ρ ha traccia unitaria e osservando che le matrici di Pauli hanno traccia nulla possiamo scrivere:

$$\rho = \frac{1}{2}(I + a_i \sigma^i) = \frac{1}{2} \begin{pmatrix} 1 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & 1 - a_3 \end{pmatrix}, \quad (34)$$

dove in questo caso gli a_i sono reali essendo ρ hermitiana. Il determinante della matrice ρ è uguale a $\frac{1}{4}(1 - \vec{a}^2)$ ed essendo la matrice densità semidefinita positiva abbiamo che il suo determinante è non negativo $|\vec{a}|^2 < 1$ quindi $0 < |\vec{a}| < 1$. Esiste quindi allora una corrispondenza tra le matrici densità e i punti di una sfera di raggio 1, detta *sfera di Bloch*. Per i punti sulla superficie della sfera vale $|\vec{a}| = 1$, conseguentemente gli autovalori sono 0,1 e la matrice densità corrispondente è uno stato puro. Invece per i punti interni vale $|\vec{a}| < 1$ e rappresentano le

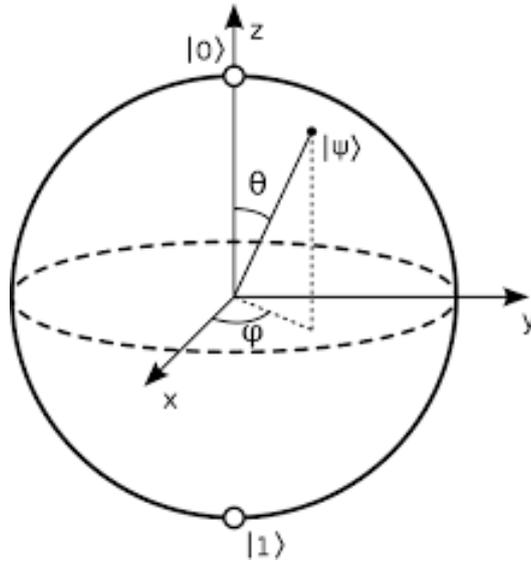


Figura 2: Esempio di Sfera di Bloch di un qubit nella base $\{|0\rangle, |1\rangle\}$.

miscela statistiche. Pertanto la sfera di Bloch ci permette di avere un'utile rappresentazione grafica delle matrici densità di dimensione 2. Questi concetti possono essere generalizzati a n-dimensioni, perdendo però la possibilità di rappresentarli graficamente.

2.2 Decomposizione di Schmidt

Abbiamo visto come ogni stato puro appartenente ad uno spazio di Hilbert H del tipo $H = H_A \otimes H_B$ è esprimibile nella forma:

$$|\psi_{AB}\rangle = \sum_{i,\alpha}^{n,m} a_{i\alpha} |i\alpha\rangle. \quad (35)$$

Ora andiamo a vedere una particolare decomposizione di $|\psi_{AB}\rangle$, nota come *Decomposizione di Schmidt* [23], in cui ρ_A e ρ_B hanno gli stessi autovalori non nulli. Riscriviamo 35 nella seguente forma:

$$|\psi_{AB}\rangle = \sum_i^n |ii'\rangle, \quad (36)$$

dove abbiamo definito i' come:

$$|i'\rangle = \sum_{\alpha}^m a_{i\alpha} |\alpha\rangle. \quad (37)$$

Notiamo che per come li abbiamo definiti, i' non sono necessariamente ortonormali. Supponiamo adesso che la base $|i_A\rangle$ sia una base tale che ρ_A è diagonale. Quindi possiamo scrivere:

$$\rho_A = \sum_i^{n'} p_i |i\rangle\langle i|. \quad (38)$$

Dove $n' \leq n$ è il numero di p_i diverse da 0. Inoltre ricordando l'espressione 28 per ρ_A vale anche:

$$\begin{aligned} \rho_A &= \text{Tr}_B |\psi_{AB}\rangle\langle\psi_{AB}| = \text{Tr}_B \sum_{i,j}^{n,n} |ii'\rangle\langle jj'| \\ &= \sum_{i,j,k}^{n,n,m} \langle k|_B |ii'\rangle\langle jj'|k\rangle_B = \sum_{i,j,k}^{n,n,m} \langle k|i'\rangle_B |i\rangle\langle j| \langle j'|k\rangle_B \\ &= \sum_{i,j,k}^{n,n,m} |i\rangle\langle j| \langle j'|k\rangle_B \langle k|_B |i'\rangle = \sum_{i,j}^{n,n} |i\rangle\langle j| \langle j'|i'\rangle. \end{aligned} \quad (39)$$

Dove si è utilizzata la relazione di completezza della base $|k\rangle$, base del sottospazio H_B . Confrontando 38 e 39 è possibile notare che:

$$p_i \delta_{ij} = \langle j'|i'\rangle. \quad (40)$$

Quindi abbiamo dimostrato che $|i'\rangle$ sono effettivamente ortogonali. Ora possiamo normalizzare:

$$|i''\rangle = p_i^{-\frac{1}{2}} |i'\rangle. \quad (41)$$

Tale espressione è corretta perché avevamo già escluso le probabilità uguali a 0 nella formula 38. Quindi ora possiamo scrivere:

$$|\psi_{AB}\rangle = \sum_i^{n'} \sqrt{p_i} |i\rangle |i''\rangle. \quad (42)$$

Abbiamo quindi scomposto il nostro stato puro di partenza in una particolare base di H . Abbiamo anche visto che questo procedimento è del tutto generale e non è stato necessario fare alcuna ipotesi sul vettore $|\psi_{AB}\rangle$, anche se è importante notare che dati due vettori distinti di H , non è possibile scomporli nella stessa base, in quanto i vettori $|i''\rangle$ dipendono dal particolare vettore considerato. Andiamo ora a vedere come è legata l'espressione generica di $|\psi_{AB}\rangle$ alla particolare decomposizione di Schmidt. Abbiamo visto come si esprime un generico stato puro di H :

$$|\psi_{AB}\rangle = \sum_{b,\beta}^{n,m} a_{b\beta} |b\rangle |\beta\rangle. \quad (43)$$

Consideriamo due trasformazioni unitarie U_A e U_B che legano rispettivamente le basi $|b\rangle$ e $|\beta\rangle$ a $|i\rangle$ e $|i''\rangle$:

$$|i\rangle_A = \sum_b^n U_{Abi} |b\rangle, \quad |i''\rangle_B = \sum_\beta^m U_{B\beta i''} |\beta\rangle. \quad (44)$$

Quindi confrontando le relazioni 43 e 42:

$$a_{b\beta} = \sum_i^n i U_{Abi} \sqrt{p_i} U_{B\beta i''}^T. \quad (45)$$

Ora andiamo invece a calcolare la matrice densità del sottosistema B, ρ_B . In maniera analoga a quanto fatto per ρ_A , scriviamo:

$$\rho_B = \text{Tr}_A |\psi_{AB}\rangle\langle\psi_{AB}| = \sum_i^n p_i |i''\rangle\langle i''|. \quad (46)$$

Notiamo quindi che ρ_A e ρ_B hanno gli stessi autovalori non nulli. In particolare se non hanno autovalori degeneri, esiste un metodo diretto per trovare la decomposizione di Schmidt. Infatti, partendo da uno stato qualsiasi, basta diagonalizzare ρ_A e ρ_B per trovare rispettivamente le basi $|i\rangle$ e $|i''\rangle$. Una volta trovate le due base, basterà effettuare il prodotto tensoriale tra gli autovettori aventi lo stesso autovalore p_i per trovare la base di ψ_{AB} e poterlo scriverlo nella forma 42. Ad ogni stato puro di un sistema bipartito è possibile attribuire un numero intero, il suo numero di Schmidt, ovvero il numero di autovalori non nulli di ρ_A e ρ_B . Se questo numero è uguale a 1, lo stato si dice separabile ed è possibile esprimerlo come prodotto tensoriale di un vettore appartenente a H_A e un vettore appartenente a H_B :

$$|\psi_{AB}\rangle = |\phi_A\rangle \otimes |\chi_B\rangle, \quad (47)$$

e i sottostati sono stati puri:

$$\rho_A = |\phi\rangle\langle\phi|, \quad \rho_B = |\chi\rangle\langle\chi|. \quad (48)$$

Se invece il numero di Schmidt è maggiore di 1, allora lo stato si dice *entangled* e i sottostati sono miscele statistiche. Se uno stato è entangled allora si dice che A e B sono correlati quantisticamente. È possibile dimostrare che non è possibile aumentare il numero di Schmidt localmente, ovvero con una trasformazione del tipo $U_A \otimes U_B$, ma per creare questa correlazione è necessario far *interagire* fra di loro le due parti del nostro sistema.

2.3 Ambiguità delle preparazioni e purificazione

Abbiamo visto che una matrice densità ρ , per poter essere definita come tale, deve avere determinate proprietà. In particolare:

1. ρ è hermitiana: $\rho^\dagger = \rho$;
2. ρ è semidefinita positiva;
3. $\text{Tr } \rho = 1$.

Date due matrici densità ρ_1 e ρ_2 e un numero $\lambda \in \mathbf{R}$ tale che $0 \leq \lambda \leq 1$ è possibile definire $\rho(\lambda)$ come :

$$\rho(\lambda) = \lambda\rho_1 + (1 - \lambda)\rho_2. \quad (49)$$

È facile verificare che $\rho(\lambda)$ verifica tutte le proprietà enunciate in precedenza, pertanto è una matrice densità. Quindi abbiamo dimostrato che, dato uno spazio di Hilbert di dimensione d le matrici densità formano un *sottospazio convesso* delle matrici hermitiane di dimensione $d \cdot d$. Un sottospazio convesso è tale se contiene tutti i segmenti che congiungono due punti qualsiasi del sottospazio, detti estremi. Dato uno stato puro $|\psi\rangle$ e la sua matrice densità $\rho = |\psi\rangle\langle\psi|$, consideriamo il vettore $|\psi_\perp\rangle$ ortogonale a $|\psi\rangle$. Ora scriviamo ρ come visto precedentemente, 49:

$$\rho = \lambda\rho_1 + (1 - \lambda)\rho_2. \quad (50)$$

Moltiplicando sia per $|\psi_\perp\rangle$ che per $\langle\psi_\perp|$ otteniamo:

$$\langle\psi_\perp|\rho|\psi_\perp\rangle = 0 = \lambda\langle\psi_\perp|\rho_1|\psi_\perp\rangle + (1 - \lambda)\langle\psi_\perp|\rho_2|\psi_\perp\rangle. \quad (51)$$

Essendo sia ρ_1 che ρ_2 semidefinite positive, abbiamo che la somma di due quantità non negativa è pari a 0, pertanto se λ è diverso sia da 0 che da 1 abbiamo che ρ_1 e ρ_2 sono ortogonali a $|\psi_\perp\rangle$. Data l'arbitrarietà di $|\psi_\perp\rangle$ abbiamo che $\rho = \rho_1 = \rho_2$. Quindi abbiamo appena dimostrato che *solo gli stati puri possono essere estremi* e che non possono essere scritti come combinazione lineare di altre due matrici. Inoltre, ricordando che la matrice densità di una miscela statistica si scrive come $\rho = \sum_i^n p_i |i\rangle\langle i|$, notiamo che solo gli stati puri sono estremi, potendo appunto sempre esprimere la matrice densità di una miscela statistica come una combinazione lineare. Andiamo ora a valutare le conseguenze fisiche di ciò. Supponiamo di avere un sistema fisico che può assumere due stati e che un fisico sperimentale decida di prepararlo con probabilità λ nello stato ρ_1 e con probabilità $1 - \lambda$ nello stato ρ_2 . Dato un osservabile M il suo valore medio vale:

$$\langle M \rangle = \lambda \text{Tr}\{M\rho_1\} + (1 - \lambda) \text{Tr}\{M\rho_2\} = \text{Tr}\{M\rho(\lambda)\}. \quad (52)$$

Quindi il valore medio M è lo stesso che avremmo ottenuto se avessimo preparato lo stato $\rho(\lambda)$. Pertanto esistono diversi modi di preparare di $\rho(\lambda)$, come una combinazione convessa di altri due stati e senza nessuna conseguenza osservabile. Invece, come abbiamo visto in precedenza, esiste un unico modo di preparare uno stato puro, mentre la preparazione di una miscela statistica è ambigua.

Data una miscela statistica descritta dalla sua matrice densità $\rho_A \in L(H_A)$:

$$\rho_A = \sum_i^n p_i |i\rangle\langle i|. \quad (53)$$

Consideriamo un secondo spazio di Hilbert, H_B , e una sua base $|\alpha_i\rangle$. Ora consideriamo lo stato puro appartenente allo spazio $H = H_A \otimes H_B$:

$$|\psi\rangle = \sum_i^n p_i |i\rangle |\alpha_i\rangle, \quad (54)$$

ritroviamo la matrice densità del sottospazio H_A eseguendo una traccia su H_B :

$$\rho_A = \text{Tr}_B |\psi\rangle\langle\psi| = \sum_i^n p_i |i\rangle\langle i|. \quad (55)$$

Quindi abbiamo appena ritrovato la matrice densità della miscela statistica iniziale come risultato della traccia parziale di uno stato puro di uno spazio di Hilbert di dimensione maggiore. In sintesi, avendo come stato iniziale una miscela statistica, e non uno stato puro di uno spazio di Hilbert bipartito, è possibile trovare uno spazio di Hilbert ausiliario tale da riottenere la miscela statistica una volta eseguita la traccia parziale sullo spazio di Hilbert ausiliario di uno stato puro del nuovo spazio di Hilbert bipartito dato dal prodotto dello spazio di partenza e dello spazio ausiliario. La procedura che abbiamo appena descritto prende il nome di *purificazione*.

3 Misure ortogonali e generalizzate

Abbiamo iniziato il nostro percorso evidenziando le differenze tra sistemi chiusi e sistemi aperti in Meccanica Quantistica e abbiamo affermato che la misura di un sistema aperto non è una proiezione ortogonale. Andiamo ora ad approfondire il tema della misura, iniziando con l'analisi della *misura ortogonale* per poi andare ad introdurre il concetto di *misura generalizzata*.

3.1 Misure ortogonali

Ora ripercorriamo *l'analisi della misura di Von Neumann*[20]. È possibile, secondo Von Neumann, legare le variabili di un sistema quantistico microscopico alle variabili macroscopiche di un secondo sistema quantistico, l'apparato di misura. Ciò significa che se vogliamo misurare un osservabile M , allora l'hamiltoniana complessiva del sistema dovrà dipendere dal prodotto tensoriale di M e di una grandezza macroscopica del nostro apparato sperimentale. Descriviamo tramite un piccolo toy model il processo di misura ortogonale. Consideriamo un osservabile M che sia diagonale nella base $|a\rangle$:

$$M = \sum_a^n p_a |a\rangle\langle a|, \quad (56)$$

e supponiamo che durante la misura tale osservabile si leghi all'operatore posizione del nostro apparato misura tramite un termine di interazione $\lambda(t)$ che tale che $\lambda(t) \neq 0$ per $0 \leq t \leq T$. Quindi l'hamiltoniana del nostro sistema è data da:

$$H = H_0 + \frac{P^2}{2m} + \lambda(t)M \otimes P, \quad (57)$$

dove H_0 è l'hamiltoniana non perturbata del sistema quantistico che vogliamo misurare e P è l'operatore momento del nostro apparato di misura. Inoltre ipotizziamo che $[H_0, M] = 0$ in modo che il nostro osservabile M non evolva nel tempo, che l'evoluzione libera del pacchetto d'onda dell'apparato di misura sia trascurabile, data la sua massa, e che la misura sia abbastanza veloce da poter trascurare l'evoluzione libera del nostro sistema quantistico. Allora l'hamiltoniana e l'operatore evoluzione assumono le seguenti forme:

$$H = \lambda(T)M \otimes P, \quad (58)$$

$$U(T) \approx e^{-i\lambda(T)tM \otimes P/\hbar}. \quad (59)$$

L'operatore momento è il generatore infinitesimale delle traslazioni, pertanto $\exp\{-ip_a P/\hbar\}\psi(x) = \psi(x - p_a)$. Quindi dato uno stato iniziale del sistema quantistico da misurare pari a, nella base delle $|a\rangle$, $|\phi\rangle = \sum_a^n \alpha_a |a\rangle$ e uno stato iniziale dell'apparato di misura $|\psi\rangle$, dopo il processo di misura avremo :

$$U(T)(|\phi\rangle \otimes |\psi(x)\rangle) = \sum_a^n \alpha_a |a\rangle \otimes |\psi(x - \lambda p_a T)\rangle. \quad (60)$$

Quindi vediamo che se l'ampiezza del pacchetto d'onda dell'apparato è piccola rispetto alla quantità $\lambda\Delta p_a T$, dove Δp_a indica la differenza minima tra due autovalori di M , è possibile misurare la traslazione dell'apparato di misura. Osservando tale traslazione, in qualsiasi modo, in base ai postulati sulla misura provocheremo il collasso della funzione d'onda del sistema composto dall'apparato di misura e dal sistema microscopico che vogliamo misurare. Con probabilità

$|\alpha_a|^2$ l'apparato di misura sarà traslato di $\lambda T p_a$ e ciò significa che abbiamo preparato l'auto-stato $|a\rangle$ del sistema quantistico. Abbiamo appena descritto il modello di misura ortogonale di Von Neumann. Ora generalizziamo il nostro discorso. Sia $E_a, a = 1, \dots, N$ un set completo di proiettori ortogonali tali che:

$$E_a E_b = \delta_{ab} E_a, \quad E_a^\dagger = E_a, \quad \sum_a^n E_a = I. \quad (61)$$

Consideriamo un apparato di misura e una sua base $|a\rangle, a = 0, 1, 2..$ e facciamo interagire con il nostro sistema quantistico. L'evoluzione del sistema dato dal stato microscopico e dall'apparato di misura è data dall'operatore :

$$U = \sum_{a,b} E_b \otimes |b+a\rangle\langle a|. \quad (62)$$

Ovvero l'apparato trasla di una quantità b . Quindi supponiamo uno stato iniziale e facciamo evolvere:

$$U(|\psi\rangle \otimes |0\rangle) = \sum_a E_a |\psi\rangle |a\rangle = |\phi\rangle. \quad (63)$$

La probabilità di misurare l'autovalore a , in base alla regola di Born, è uguale a:

$$Prob(a) = \langle \phi | (I \otimes |a\rangle\langle a |) | \phi \rangle = \langle \psi | E_a | \psi \rangle. \quad (64)$$

Dopo la misura, a causa del collasso della funzione d'onda, avremo che il sistema da misurare sarà nello stato:

$$|\Psi\rangle = \frac{E_a |\psi\rangle}{\|E_a |\psi\rangle\|}. \quad (65)$$

Se lo stato dopo la misura non è noto, allora sarà una combinazione dei possibili stati:

$$|\Psi\rangle = \sum_a^n E_a |\psi\rangle\langle\psi| E_a. \quad (66)$$

Infine, se lo stato iniziale non è uno stato puro ma una matrice densità ρ , scomponendola in stati puri è possibile dimostrare che lo stato dopo una misura sarà:

$$\rho' = \sum_a^n E_a \rho E_a. \quad (67)$$

In definitiva abbiamo visto, tramite un semplice modello, come *osservando una variabile macroscopica del nostro apparato di misura è possibile effettuare una qualsiasi misura ortogonale.*

3.2 Misure generalizzate

Per introdurre il concetto di misura generalizzata, partiamo da un esempio semplice ma istruttivo. Siano il nostro sistema microscopico e il nostro apparato macroscopico due qubit, rispettivamente A e B nello stato iniziale $|\psi\rangle_A = \alpha |0\rangle_A + \beta |1\rangle_A, |\phi\rangle_B = |1\rangle_B$ e consideriamo il prodotto della loro interazione, dovuta al processo di misura) e della successiva evoluzione, basandoci sul risultato trovato precedentemente 63:

$$U(T)(|\psi\rangle_A \otimes |\phi\rangle_B) = \alpha |0\rangle_A |0\rangle_B + \beta |1\rangle_A |1\rangle_B. \quad (68)$$

Ora se misurassimo nella base $\{|0\rangle_B, |1\rangle_B\}$ avremmo ottenuto una misura ortogonale. Infatti i possibili risultati $\alpha |0\rangle$ e $\beta |1\rangle$ sono ortogonali fra loro e una seconda misura successiva alla prima darebbe sicuramente lo stesso risultato. Ora però proiettiamo nella base $|\pm\rangle_B = \frac{1}{\sqrt{2}}(|0\rangle_B \pm |1\rangle_B)$. In questo caso i possibili risultati sono $\alpha |0\rangle_A \pm \beta |1\rangle_A$ e che a differenza essi non sono ortogonali, a meno che $|\alpha| = |\beta|$. Abbiamo appena descritto una misura generalizzata. **Notiamo come in generale è conveniente esprimere l'interazione tra l'apparato di misura e lo stato microscopico in funzione della nuova base dell'apparato di misura in cui scegliamo di misurare:**

$$U(T) |\psi\rangle_A \otimes |0\rangle_B = M_{+A} |\psi\rangle_A |+\rangle_B + M_{-A} |\psi\rangle_A |-\rangle_B. \quad (69)$$

Dove in questo caso:

$$M_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad M_- = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (70)$$

Quindi misurando il sistema nella base $|\pm\rangle_B = \frac{1}{\sqrt{2}}(|0\rangle_B \pm |1\rangle_B)$ dell'apparato di misura, abbiamo preparato il sistema A in uno dei sue seguenti stati $\{M_{+A} |\psi\rangle, M_{-A} |\psi\rangle\}$.

Andiamo ora a generalizzare il discorso. Dato uno stato iniziale del sistema quantistico $|\psi\rangle$ e uno stato iniziale dell'apparato $|0\rangle$, che supponiamo N dimensionale, esprimiamo l'evoluzione temporale in funzione della base con cui andremo effettivamente a misurare il nostro sistema, la base $|a\rangle$; $a = 1, \dots, N$:

$$U(|\psi\rangle \otimes |0\rangle) = \sum_a^N (M_a |\psi\rangle) \otimes |a\rangle = |\phi\rangle. \quad (71)$$

Dall'unitarietà di U è evidente che sussiste la seguente relazione:

$$\sum_a^N M_a^\dagger M_a = I. \quad (72)$$

Dai postulati sulla misura sappiamo che la probabilità di ottenere una misura di a è data dalla proiezione $I \otimes |a\rangle\langle a|$:

$$Prob(a) = \langle \phi | (I \otimes |a\rangle\langle a|) | \phi \rangle = \langle \psi M_a^\dagger | M_a \psi \rangle = \|M_a \psi\|^2. \quad (73)$$

Lo stato del sistema microscopico dopo la misura invece sarà dato da:

$$|\Psi\rangle = \frac{M_a |\psi\rangle}{\|M_a \psi\|}. \quad (74)$$

Come abbiamo già evidenziato, due misure successive non sono ortogonali e non restituiscono sempre lo stesso stato. Quindi la probabilità di misurare b dopo che è stato misurato a è pari a:

$$Prob(b|a) = \frac{\|M_b M_a |\psi\rangle\|^2}{\|M_a |\psi\rangle\|^2}. \quad (75)$$

Se lo stato iniziale è una miscela statistica la cui matrice densità è ρ allora la probabilità di misurare dopo l'interazione il valore a è data da:

$$Prob(a) = \text{Tr}\{M_a^\dagger \rho M_a\} = \text{Tr}\{\rho M_a^\dagger M_a\} = \text{Tr}\{\rho E_a\}, \quad (76)$$

dove abbiamo definito $E_a = M_a^\dagger M_a$. L'insieme di operatori E_a rispetta determinate proprietà :

1. E_a è hermitiano;
2. E_a è semidefinito positivo: $\langle \psi | E_a | \psi \rangle \geq 0$;
3. L'insieme dei E_a è completo: $\sum_a^N E_a = I$;
4. In generale non sono ortogonali o proiettivi: $E_a E_B \neq \delta_{ab} E_a$.

Un insieme di operatori che rispetta queste proprietà viene detto “positive operator-valued measure” o *POVM*. Ora esaminiamo il legame tra E_a e M_a . È possibile definire, per ogni operatore semidefinito positivo, una sua radice quadrata, anch'essa semidefinita positiva. Eseguendo la decomposizione polare di M_a si può esprimere nella forma:

$$M_a = U_a \sqrt{E_a}, \quad (77)$$

dove U_a è una matrice unitaria arbitraria. Osserviamo che la definizione $M_a^\dagger M_a = E_a$ è rispettata. Sostituendo questa espressione nella 74 vediamo che lo stato dopo la misura è:

$$U_a \left(\frac{\sqrt{E_a} |\psi\rangle}{\|E_a |\psi\rangle\|} \right). \quad (78)$$

Quindi lo stato dopo la misura dipende dall'operatore U_a , il quale è arbitrario (rispettando sempre la condizione di unitarietà). Diamo una piccola interpretazione fisica per concludere la trattazione.

Tramite il concetto di misure ortogonali, Von Neumann voleva descrivere un apparato di misura ideale, senza nessuna imperfezione. Tale apparato di misura doveva appunto rispettare la condizione che due misure successive dessero lo stesso risultato. Invece i POVM possono descrivere apparati di misura reali e quindi imperfetti o situazioni in cui una misura ortogonale non darebbe risultati utili (distinguere più stati non ortogonali fra loro) o ancora quando effettuiamo una misura dopo che il sistema abbia già interagito con l'apparato di misura. Vediamo un esempio. Consideriamo un apparato di misura di un qubit, il cui compito è determinare se il qubit sia nello stato $|1\rangle$ o $|2\rangle$. Questo strumento può fallire con probabilità p e nel caso in cui fallisca lo stato del qubit è inalterato. Non esiste nessuna misura proiettiva ortogonale che descriva questa situazione, però esiste un POVM i cui operatori sono:

$$M_0 = pI \quad M_1 = (1-p)|2\rangle\langle 2| \quad M_2 = (1-p)|2\rangle\langle 2|, \quad (79)$$

che descrive perfettamente questa situazione.

Infine è importante citare *il teorema di Neumark* [19], il quale collega POVM e misure ortogonali. Infatti è possibile dimostrare che si può passare da un POVM in uno spazio di Hilbert di dimensione n ad una misura ortogonale in uno spazio di Hilbert di dimensione $m > n$, tramite un'isometria. Quindi è evidente la completa equivalenza tra POVM in sistemi aperti e misure ortogonali in sistemi chiusi.

Concludiamo questa sezione esponendo *i problemi della misura* [22]. Sappiamo che l'evoluzione di un sistema chiuso, come quello formato dall'unione del nostro sistema microscopico e dall'apparato di misura è governata dall'equazione di Schrödinger. Essendo l'equazione di Schrödinger un'equazione lineare, una sovrapposizione di più soluzioni è una soluzione. Ciò è un punto fermo della Meccanica Quantistica, in quanto è stato anche confermato da esperimento di interferenza di fasci di elettroni, che sarebbero inspiegabili se non si considerasse la sovrapposizioni di funzioni d'onda. Com'è possibile che, nonostante le sovrapposizioni di

funzioni d'onda siano soluzioni dell'equazione di Schrödinger, osserviamo sempre un risultato ben definito?

Come abbiamo visto nella nostra analisi della misura, Von Neumann ha postulato l'esistenza del collasso della funzione d'onda per spiegare l'esistenza di un risultato della funzione d'onda. Eppure il collasso della funzione è un'evoluzione probabilistica e non unitaria, al contrario dell'equazione di Schrödinger che è deterministica e unitaria. Inoltre il collasso della funzione d'onda ha diversi punti critici, come ad esempio quando viene effettuato il collasso della funzione d'onda e da chi (si veda ad esempio l'esperimento l'amico di Wigner) [31]. Pertanto, nonostante il collasso della funzione d'onda sia presente nell'interpretazione di Copenaghen che ha portato molti successi ed è riuscita ad interpretare molti fenomeni, è un problema aperto come superare questo meccanismo, provando a spiegare come si passa da una sovrapposizione di stati di un sistema ad un unico stato effettivamente osservato.

Inoltre un altro problema della misura è quello della *“base preferita”*. Consideriamo lo spazio di Hilbert $H = H_1 \otimes H_2$ di due qubit 1 e 2 e interpretiamo il qubit 2 come se fosse un apparato di misura. Supponiamo uno stato iniziale:

$$|\psi\rangle_{12} = \frac{1}{\sqrt{2}}(|+\rangle_1 |-\rangle_2 - |-\rangle_1 |+\rangle_2), \quad (80)$$

dove $|\pm\rangle_{1,2}$ sono gli autostati dell'osservabile spin lungo l'asse z rispettivamente per il qubit 1 e per il qubit 2. Cambiamo base ed esprimiamo questo stesso stato usando gli autostati dell'osservabile spin lungo l'asse x.

$$|\psi\rangle_{12} = \frac{1}{\sqrt{2}}(|+x\rangle_1 |-x\rangle_2 - |-x\rangle_1 |+x\rangle_2). \quad (81)$$

Abbiamo visto nello schema di Von Neumann come un'interazione tra apparato di misura e sistema microscopico corrisponda effettivamente ad una misura, prima di essere osservata. Pertanto, potendo scegliere una base corrisponde agli autostati dello spin lungo qualsiasi direzione, potremmo misurare lo spin in una direzione arbitraria dello spazio. Si può affermare tramite un teorema generale che dato un sistema quantistico isolato composto da un apparato di misura e da uno stato da misurare, in meccanica quantistica non è possibile determinare quale osservabile sia stato misurato a posteriori. **Ciò viola ovviamente il senso comune, visto che uno strumento di misura è progettato per misurare una determinata caratteristica, ma anche la meccanica quantistica stessa, in quanto non è possibile misurare contemporaneamente due osservabili che non commutano, come ad esempio lo spin lungo due direzioni diverse.** Questi problemi sono ancora aperti e non è stata trovata una teoria completa capace di risolverli, anche se sono stati compiuti dei passi in avanti. In particolare, più avanti analizzeremo la decoerenza e le sue implicazioni sulla teoria della misura.

4 Canali quantistici

Abbiamo visto che, successivamente a una misura, uno stato puro passa da $|\psi\rangle$ a $M_a|\psi\rangle$. Pertanto immaginiamo uno stato puro del sottospazio H_A $|\psi\rangle$ che interagisce con un sistema di H_B . Ora misuriamo eseguendo la traccia parziale su H_B , ma immaginiamo di non registrare il risultato della misura ed essere pertanto costretti a considerare tutti i possibili risultati della misura. Allora l'evoluzione della matrice densità dello stato puro $\rho = |\psi\rangle\langle\psi|$ è governata da una trasformazione lineare ε che agisce, in base a quanto visto nella sezione precedente:

$$\varepsilon(\rho) = \sum_a M_a \rho M_a^\dagger, \quad (82)$$

dove, come visto, gli operatori M_a rispetto la relazione di completezza $\sum_a M_a^\dagger M_a = I$. Essendo la trasformazione lineare e potendo esprimere una qualsiasi matrice densità come combinazione di stati puri, tale ragionamento è generalizzabile ad una matrice ρ qualsiasi. Una trasformazione lineare di questo tipo, i cui operatori M_a rispettino la condizione di completezza viene detta *canale quantistico*. La forma 82 del canale quantistico è detta *operator-sum representation* e gli operatori M_a sono detti *operatori di Kraus* [15]. Come vediamo, il numero minimo di operatori di Kraus necessario per descrivere l'evoluzione di una matrice densità è pari al numero di autovalori diversi da 0 della matrice. Il formalismo dei canali quantistici sarà utilizzato, più avanti per trattare la decoerenza, ovvero l'evoluzione di stati puri in miscele statistiche. La decoerenza traduce la perdita di informazione, che passa dal sistema all'ambiente. Inoltre i canali quantistici possono anche essere utilizzati come schematizzazione del trasporto di informazione quantistica. Un esempio pratico è l'invio di un singolo fotone, il quale è un archetipo perfetto di qubit, dato che è possibile immagazzinare informazione ad esempio nella sua polarizzazione, la quale è in generale è rappresentata da un vettore di uno spazio di Hilbert di dimensione 2, attraverso una fibra ottica.

I canali quantistici, essendo trasformazioni di matrici densità in matrici densità, rispettano le seguenti proprietà:

1. Linearità: $\varepsilon(\alpha\rho_1 + \beta\rho_2) = \varepsilon(\alpha\rho_1) + \varepsilon(\beta\rho_2)$;
2. Conservano l'autoaggiuntezza: $\rho = \rho^\dagger \rightarrow \varepsilon(\rho) = \varepsilon(\rho)^\dagger$;
3. Conservano la positività: ρ semidefinito positivo $\rightarrow \varepsilon(\rho)$ semidefinito positivo;
4. Preservano la traccia: $\text{Tr}\{\rho\} = \text{Tr}\{\varepsilon(\rho)\}$

Inoltre i canali quantistici sono completamente positivi, ovvero preservano la positività anche se agiscono solo su parte del sistema. Immaginiamo di avere un canale quantistico ε che trasforma stati di H_A in stati di $H_{A'}$. Ora consideriamo un'estensione di H_A , $H_A \otimes H_B$, e un'estensione di ε , $\varepsilon \otimes I_B$. Quindi se ε aveva come operatori di Kraus M_A , allora $\varepsilon \otimes I_B$ avrà come operatori di Kraus $M_a \otimes I_B$, i quali sono ovviamente semidefiniti positivi e quindi anche $\varepsilon \otimes I_B$ sarà semidefinito positivo.

Abbiamo visto come i canali quantistici si ottengono a partire dalla misura di un sistema appartenente a $H_A \otimes H_B$ sul sottospazio H_B di base di H_B . Ciò però rende la rappresentazione di un canale quantistico non unica. Infatti immaginiamo di avere uno stato di $H_A \otimes H_B$, il quale evolve seguendo l'equazione 71. Ora immaginiamo di applicare una trasformazione unitaria ai vettori di base di H_B :

$$|a\rangle = \sum_n N_{ab} |b\rangle. \quad (83)$$

Quindi abbiamo che ora l'evoluzione dello stato sarà data da:

$$\sum_{a,b} M_a |\psi\rangle_A N_{ab} |b\rangle_B = \sum_b V_b |\psi\rangle_A |b\rangle_B. \quad (84)$$

Dove abbiamo definito i nuovi operatori di Kraus del canale come:

$$V_b = \sum_a N_{ab} M_a. \quad (85)$$

Infine illustriamo qualitativamente un'altra proprietà dei canali quantistici. Immaginiamo un canale quantistico ε_1 , composto da N operatori di Kraus V_a che descrive l'evoluzione di una matrice densità da ieri a oggi e un canale quantistico ε_2 che descrive l'evoluzione della stessa matrice da oggi a domani, il quale è composto da M operatori di Kraus L_b . Allora il canale $\varepsilon_1 \circ \varepsilon_2$, composto dai NM operatori di Kraus $V_a L_b$ descriverà l'evoluzione della miscela statistica da ieri a domani. Ciò perché, come vedremo in seguito i canali quantistici formano un semigruppato e pertanto sono legati da un'operazione binaria associativa, ma non sempre invertibile. Ora, partendo da un canale quantistico invertibile, andremo a trovare una condizione generale sugli operatori di Kraus. Partiamo da un canale quantistico ε_1 i cui operatori di Kraus siano M_a e supponiamo che esista il canale inverso ε_2 i cui operatori di Kraus siano N_b . Allora per uno stato rappresentato dalla matrice densità $|\psi\rangle\langle\psi|$ avremo :

$$\varepsilon_2 \circ \varepsilon_1(|\psi\rangle\langle\psi|) = \sum_{a,b}^{n,m} N_b M_a |\psi\rangle\langle\psi| M_a^\dagger N_b^\dagger = |\psi\rangle\langle\psi|. \quad (86)$$

Essendo la sommatoria composta da termini tutti positivi, abbiamo che ogni termine deve essere proporzionale al risultato, $|\psi\rangle\langle\psi|$, per cui possiamo scrivere:

$$N_a M_n = \lambda_{ab} I. \quad (87)$$

Usando la relazione di completezza degli operatori di Kraus troviamo:

$$M_b^\dagger M_a = M_b^\dagger \sum_{\alpha}^m N_{\alpha}^\dagger N_{\alpha} M_a = \sum_{\alpha}^m \lambda_{\alpha b}^* \lambda_{\alpha a} I = \beta_{ba} I. \quad (88)$$

Dato che il canale ε_1 trasforma il sistema A in se stesso esso è rappresentato da una matrice quadrata, si può effettuare la sua decomposizione polare:

$$M_a = U_a \sqrt{M_a^\dagger M_a} = U_a \sqrt{\beta_{aa}}, \quad (89)$$

da cui si può ricavare:

$$M_b^\dagger M_a = \sqrt{\beta_{aa} \beta_{bb}} U_b^\dagger U_a = \beta_{ba} I. \quad (90)$$

Quindi è valida la seguente espressione:

$$U_a = \frac{U_{ba}}{\sqrt{\beta_{aa} \beta_{bb} U_b}}. \quad (91)$$

In definitiva possiamo affermare che un canale quantistico è invertibile se ogni suo operatore di Kraus è proporzionale ad un singolo operatore unitario, ovvero se il canale è unitario.

4.1 Dualismo stato-canale e i nuovi postulati della MQ

Il concetto di canale quantistico $\varepsilon_{A \rightarrow A'}$ ci ha permesso di descrivere l'evoluzione di una matrice densità, la quale descriveva un sistema A fino allo stato A' . Come avevamo visto A può essere pensato come parte di uno spazio di Hilbert bipartito $H = H_A \otimes H_B$, dove per semplicità assumeremo che $\dim H_A = \dim H_B$. Consideriamo uno stato puro di H , la cui decomposizione di Schimdt è:

$$|\psi\rangle_{AB} = \sum_i^d |i\rangle_B \otimes |i\rangle_A. \quad (92)$$

Stiamo considerando stati non normalizzati per non evitare l'utilizzo costante di fattori di normalizzazione nella trattazione successiva. Come abbiamo detto in precedenza, un canale quantistico è completamente positivo, quindi $I_B \otimes \varepsilon_{A \rightarrow A'}$ trasforma la matrice densità del nostro stato puro, che come abbiamo visto è $|\psi\rangle\langle\psi|$ in un operatore semidefinito positivo, che possiamo vedere come un operatore densità e come tale può essere espresso come combinazione di stati puri:

$$(I_B \otimes \varepsilon_{A \rightarrow A'}) (|\psi\rangle_{AB} \langle\psi|_{AB}) = \sum_a^n C_a |\phi_a\rangle_{AB} \langle\phi_a|_{AB}. \quad (93)$$

Ora notiamo che dato uno stato $|\chi\rangle_A$:

$$|\chi\rangle_A = \sum_i^d \chi_i |i\rangle_A = \sum_i^d \chi_i \langle i_B | \psi_{AB} \rangle = \langle \chi_B^* | \psi_{BA} \rangle. \quad (94)$$

Allora utilizzando la linearità di ε otteniamo:

$$\varepsilon(|\chi\rangle\langle\chi|) = \sum_a^n \langle \chi_B^* | \phi_a \rangle_{BA'} \langle \phi_a |_{BA'} | \chi^* \rangle_B. \quad (95)$$

Quindi, partendo da uno stato $|\phi_a\rangle_{BA'}$, possiamo definire un operatore M_a che va da H_A a H'_A come:

$$M_a |\chi\rangle_A = \langle \chi_B^* | \phi_a \rangle_{BA'}. \quad (96)$$

Quindi abbiamo trovato una operator-sum representation del canale che descrive l'evoluzione da H_A a H'_A . Notiamo che M_a è lineare, quindi può essere applicato ad una combinazione di stati puri qualsiasi, ovvero ad una generica matrice densità

$$\varepsilon(\rho) = \sum_a^n M_a \rho M_a^\dagger. \quad (97)$$

È possibile dimostrare che se ε conserva la traccia allora l'insieme degli M_a deve rispettare la relazione di completezza.

In definitiva è possibile osservare come prima abbiamo trovato come agisce il canale quantistico su uno stato di H_A per ottenere uno stato di $H_{A'B}$, vedesi 93, e poi, dato uno stato di $H_{A'B}$, abbiamo ricavato come agisce il canale quantistico, vedesi 96. Quindi dato un canale quantistico è possibile passare ad uno stato e viceversa, ovvero esiste un isomorfismo tra lo spazio degli operatori semidefiniti positivi agenti sullo spazio $H_{A'} \otimes H_B$ e una trasformazione lineare positiva che va dallo spazio degli operatori di H_A allo spazio degli operatori di $H_{A'}$. Questa relazione è chiamata isomorfismo di Choi–Jamiołkowski [10] [13] o dualismo stato-canale.

Nei paragrafi precedenti abbiamo visto come la purificazione di una miscela statistica consiste nel ritrovare l'operatore densità che descrive la miscela statistica eseguendo la traccia parziale di uno stato puro in uno stato di uno spazio di Hilbert di dimensione maggiore rispetto a quello dello stato di partenza. Riprendiamo questo concetto e vediamo come si collega ai canali quantistici. Abbiamo visto che in generale i canali quantistici non sono invertibili. Consideriamo ora un canale quantistico $\varepsilon_{A \rightarrow A'}$ e uno spazio di Hilbert aggiuntivo H_E che chiameremo ambiente. Consideriamo la seguente isometria, una trasformazione tra due spazi metrici che conserva la distanza, che va da A a $A'E$:

$$U_{A \rightarrow A'E} |\psi\rangle = \sum_a^n M_a |\psi\rangle \otimes |a\rangle_E. \quad (98)$$

Data la relazione di completezza dell'insieme di operatori M_a possiamo vedere che U è unitario. Se eseguiamo la traccia parziale su H_E ritroviamo il canale $\varepsilon_{A \rightarrow A'}$. Questa isometria è detta Stinespring dilation [29]. Possiamo ritrovare la stessa espressione partendo dalla purificazione di uno stato di $H_{A'B}$, nella forma vista in precedenza. Usiamo lo spazio H_E per scrivere una purificazione di $|\psi\rangle_{A'B}$

$$|\phi\rangle_{A'BE} = \sum_a^n |\psi_a\rangle_{A'B} \otimes |a\rangle. \quad (99)$$

Usando l'isomorfismo visto in precedenza, possiamo ottenere gli operatori M_a come:

$$U_{A \rightarrow A'E} |\psi\rangle = \langle \phi_B | \phi_{A'BE} \rangle. \quad (100)$$

Di nuovo, effettuando la traccia su H_E ritroviamo il canale quantistico $\varepsilon_{A \rightarrow A'}$. Quindi abbiamo anche visto come l'evoluzione di uno stato sia unitaria se osservata in uno spazio di Hilbert di dimensione maggiore rispetto allo stato.

Ora, per concludere questa sezione, possiamo naturalmente riassumere tutta la trattazione precedente enunciando i postulati della Meccanica Quantistica dei sistemi aperti:

1. Uno stato è un operatore hermitiano, con traccia unitaria e semidefinito positivo chiamato operatore densità;
2. Una misura è un POVM. Dato lo stato ρ e una set di operatori E_a , se effettuiamo una misura di E_a su ρ avremo come risultato a con probabilità $\text{Tr}\{E_a \rho\}$;
3. L'evoluzione di uno stato è determinata da un trasformazione positiva che preserva la traccia.

Siamo arrivati a questi postulati sui sistemi aperti partendo dai postulati sui sistemi chiusi. Si può anche procedere in maniera opposta arrivando ai secondi partendo dai primi. Infatti abbiamo visto come un sistema aperto può essere visto come parte di un sistema chiuso di dimensione superiore.

4.2 Master equation dei sistemi aperti

Per completare questa equivalenza tra sistemi aperti e sistemi chiusi cerchiamo una master equation per i sistemi aperti, ovvero una equazione differenziale al primo ordine che descriva completamente l'evoluzione di una matrice densità, anche per istanti infinitesimi di tempo e che sia conveniente da utilizzare. Cerchiamo quindi un equivalente dell'equazione di Schrödinger,

che descrivi lo stato tramite la sua hamiltoniana. Non è detto che un'equazione di questo tipo esista, e in generale non esiste. Ciò a causa dei continui passaggi di informazioni tra il sistema e il suo ambiente, che possono avvenire in qualsiasi momento. Dobbiamo quindi imporre un'ipotesi di località nel tempo e compiere un'approssimazione, ovvero che affinché $\rho(t)$ sia descritta da un'equazione differenziale al primo ordine in t , $\rho(t + dt)$ deve essere completamente determinata da $\rho(t)$ e non dipendere da come si è arrivati a $\rho(t)$. Un'evoluzione di questo tipo è detta markoviana.

Sappiamo che per uno stato chiuso vale:

$$|\psi(t + dt)\rangle = (I - iH(t)dt) |\psi(t)\rangle. \quad (101)$$

Quindi la matrice densità corrispondente evolverà :

$$\rho(t + dt) = \rho(t) - idt[H(t), \rho(t)]. \quad (102)$$

Generalizziamo questa equazione per un qualsiasi sistema aperto. Applicando l'ipotesi di evoluzione markoviana, l'evoluzione di ρ sarà data da :

$$\rho(t + dt) = \varepsilon(\rho(t)). \quad (103)$$

Riscriviamo il canale usando i suoi operatori di Kraus:

$$\rho(t + dt) = \sum_a^n M_a \rho(t) M_a^\dagger. \quad (104)$$

Come abbiamo detto, cerchiamo un'equazione al primo ordine in t , quindi possiamo espandere ρ nel seguente modo:

$$\rho(t + dt) = \rho(t) + \dot{\rho}(t)dt. \quad (105)$$

Tenuto conto di 104 e di 105, possiamo esprimere senza ledere la generalità gli operatori di Kraus nel seguente modo¹:

$$M_0 = I + (-iH + K)dt \quad M_a = dtL_a, \quad (106)$$

dove K è un operatore hermitiano. Gli operatori L sono denominati operatori di Lindblad. Quindi sostituendo questa espressione otteniamo:

$$\begin{aligned} \rho(t + dt) &= (I + (-iH + K)dt)\rho(t)(I + (iH + K)dt) + dt \sum_a^n M_a \rho(t) M_a^\dagger \\ &= \rho(t) - idt[H, \rho] + dt(K\rho + \rho K) + dt \sum_a^n L_a \rho(t) L_a^\dagger. \end{aligned} \quad (107)$$

Dalla relazione di completezza degli operatori di Kraus ricaviamo :

$$K = -\frac{1}{2} \sum_a^n L_a^\dagger L_a. \quad (108)$$

¹In particolare, la scelta effettuata è tale che solo M_0 contiene l'identità

Quindi confrontando le ultime due espressioni otteniamo:

$$\dot{\rho}(t) = \mathcal{L}(\rho) = -idt[H, \rho] + \sum_a^n L_a \rho(t) L_a^\dagger - \frac{1}{2} L_a^\dagger L_a \rho(t) - \frac{1}{2} \rho(t) L_a^\dagger L_a. \quad (109)$$

Infine, abbiamo trovato l'espressione che cercavamo, nota come equazione di Lindblad [16] o equazione di Gorini-Kossakowski-Sudarshan-Lindblad [30]. Abbiamo anche definito l'operatore Lindbladiano \mathcal{L} , il quale è l'operatore che genera l'evoluzione di $\rho(t)$. Notiamo che il primo termine corrisponde all'evoluzione unitaria che avevamo trovato nel caso di uno stato puro, mentre tutti i termini del tipo $L_a \rho(t) L_a^\dagger$ corrispondono a dei possibili "salti", delle variazioni non unitarie di $\rho(t)$ e per ultimi i termini $-\frac{1}{2} L_a^\dagger L_a$ garantiscano la normalizzazione nel caso in cui non avviene nessun salto. Come abbiamo già visto precedentemente, gli operatori di Lindblad, essendo legati agli operatori di Kraus possono essere ridefiniti tramite una trasformazione unitaria. L'insieme delle matrici densità generate da un Lindbladiano per diversi valori di t si chiama semigruppato quantistico dinamico.

5 Decoerenza

Con gli strumenti fin'ora acquisiti, andiamo ad analizzare in dettaglio un canale quantistico, chiamato dephasing channel. Questo canale è particolarmente importante perché fornisce un modello, molto semplificato, della decoerenza in una situazione fisica. Abbiamo visto nella sezione relativa alla Stinespring dilation come l'evoluzione di un sistema aperto sia unitaria se consideriamo il sistema come parte di un sistema chiuso di dimensione maggiore che contiene il suo ambiente. Partiamo quindi dal toy model del canale per poi analizzare in dettaglio il fenomeno della decoerenza e le sue conseguenze.

5.1 Toy Model

Consideriamo un qubit, espresso nella base $|1\rangle, |2\rangle$ su cui facciamo agire questo canale, espresso nella forma isometrica 98, dove con E indichiamo l'ambiente. I vettori di base del qubit evolvono quindi come:

$$|1\rangle \rightarrow \sqrt{1-p}|1\rangle_A \otimes |0\rangle_E + \sqrt{p}|1\rangle_A \otimes |1\rangle_E, \quad (110)$$

$$|2\rangle \rightarrow \sqrt{1-p}|2\rangle_A \otimes |0\rangle_E + \sqrt{p}|2\rangle_A \otimes |2\rangle_E. \quad (111)$$

Vediamo come con probabilità $1-p$ non succeda nulla, invece con probabilità p lo stato dell'ambiente subisca uno "scattering" e passi dallo stato fondamentale e ad uno eccitato. Questo passaggio però non coinvolge anche il qubit, i cui vettori di base non cambiano mai. Quindi è come se il canale avesse scelto una base preferita del qubit.

Dall'espressione 99 possiamo ricavare gli operatori di Kraus del canale:

$$M_0 = \sqrt{1-p}I, \quad M_1 = \sqrt{p} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad M_2 = \sqrt{p} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \quad (112)$$

Ricordando però che il numero minimo di operatori di Kraus necessario per descrivere un canale è uguale al numero di autovalori diversi da 0 della matrice densità, pari a 2 in questo, quindi consideriamo solo gli operatori M_1 e M_2 e riscriviamoli in funzione della matrice di Pauli σ_3 :

$$M_1 = \frac{\sqrt{p}}{2}(I + \sigma_3), \quad M_2 = \frac{\sqrt{p}}{2}(I - \sigma_3). \quad (113)$$

Quindi troviamo che:

$$\varepsilon(\rho) = \sum_{a=1}^2 M_a \rho M_a = (I - \frac{p}{2})\rho + \frac{p}{2}\sigma_3 \rho \sigma_3. \quad (114)$$

Applicando questa evoluzione al caso di una matrice generica:

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{12} & \rho_{22} \end{pmatrix} \rightarrow \varepsilon(\rho) = \begin{pmatrix} \rho_{11} & (1-p)\rho_{12} \\ (1-p)\rho_{12} & \rho_{22} \end{pmatrix}. \quad (115)$$

Quindi vediamo come solo i termini non sulla diagonale vengono "smorzati" dal fattore $(1-p)$, mentre i termini diagonali non subiscono alcuna variazione.

Come vedremo in maniera più approfondita nell'analisi successiva, la decoerenza è data dall'interazione con l'ambiente, quindi è qualcosa di continuo. Pertanto definiamo la probabilità di "scattering" per unità di tempo λ , tale che $p = \lambda\Delta t \ll 1$ per un Δt molto piccolo. Ora usiamo l'approssimazione di tempo discreto $t' = n\Delta t$ dove Δt sono appunto intervalli molto brevi di tempi e dopo faremo tendere n a infinito, con t fisso, per considerare il tempo continuo.

Consideriamo che il canale agisce ogni Δt , quindi per considerare la sua azione su un tempo t' dobbiamo considerare ε^n . Come abbiamo visto, l'azione di ε modificava solo i termini non diagonali. Ora consideriamo l'azione di ε^n su quei termini:

$$(1-p)^n = \left(1 - \frac{\lambda t'}{n}\right)^n \rightarrow e^{-\lambda t'}. \quad (116)$$

Dove è stato considerato il limite di n che tende all'infinito, data la continuità del tempo. Consideriamo l'applicazione di questo canale su uno stato puro $|\psi\rangle = \alpha|1\rangle + \beta|2\rangle$ per un tempo $t \gg \lambda^{-1}$ ovvero per un tempo molto maggiore rispetto al tempo di interazione medio tra la il qubit e l'ambiente:

$$\varepsilon \left[\begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix} \right] = \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}. \quad (117)$$

Quindi osserviamo come solo i termini sulla diagonale siano effettivamente presenti.

Diamo un'interpretazione fisica al modello appena visto. Possiamo considerare il nostro sistema come una particella immersa in un bagno di termico di fotoni ad una certa temperatura, che costituiscono l'ambiente visto in precedenza. Allora un fotone che interagisce con la particella ha una probabilità di subire uno scattering. Nel momento in cui la particella interagisce con l'ambiente, non possiamo considerare solo la sua funzione d'onda, ma dobbiamo considerare anche tutti i fotoni che interagiscono con essa. Ovviamente si tratta di un numero di gradi di libertà troppo grande per seguire l'evoluzione di ogni singolo fotone, quindi ci interesseremo solo della particella, recuperando il suo stato dal vettore d'onda che descrive tutto il sistema eseguendo una traccia parziale sullo spazio dell'ambiente. Consideriamo uno stato iniziale della particella:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|x\rangle + |-x\rangle), \quad (118)$$

dove $|x\rangle$ e $|-x\rangle$ sono autostati dell'operatore posizione. Considereremo questa come la base "preferita" dalla decoerenza. Ciò perché l'interazione tra il fotone e la particella sono localizzate nello spazio. Abbiamo visto che la decoerenza fa evolvere questa sovrapposizione di autostati in una miscela statistica degli stati $|x\rangle$ e $|-x\rangle$. Quindi quando misureremo la particella potremo trovare come risultati x e $-x$ e non una sovrapposizione di questi due. Ciò è un importante successo di questa teoria, spiega il passaggio da una sovrapposizione coerente di stati alla miscela incoerente di stati mutuamente ortogonali. La decoerenza però non risolve il problema della misura, che resta un problema aperto, come vedremo.

Per trovare la probabilità per unità di tempo λ , possiamo far evolvere, senza farlo interagire con l'ambiente. Facendo agire l'operatore di evoluzione U , troviamo che dopo un tempo t lo stato è:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{-i\omega t} |1\rangle) \quad (119)$$

Come abbiamo visto, nel momento in cui facciamo interagire lo stato con l'ambiente otteniamo una matrice densità del tipo :

$$\rho(t) = \frac{1}{2} \begin{pmatrix} 1 & e^{i\omega t} e^{-\lambda t} \\ e^{-i\omega t} e^{-\lambda t} & 1 \end{pmatrix}. \quad (120)$$

Misurando lo stato nella base $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ otteniamo una probabilità dipendente dal tempo:

$$P(+, t) = \langle + | \rho(t) | + \rangle = \frac{1}{2}(1 + e^{-\lambda t} \cos(\omega t)). \quad (121)$$

Quindi ripetendo la misura molte volte a istanti diversi si può ritrovare la probabilità di trovare lo stato e quindi ottenere una stima di λ . Attraverso un esperimento molto più complesso di questo toy model appena descritto, nel 1996 la decoerenza è stata osservata sperimentalmente dal premio Nobel Serge Haroche [5] e dal suo team, partendo da una sovrapposizione di stati e facendoli interagire con dei fotoni in una cavità.

5.2 La decoerenza e il problema della misura

Riprendiamo il problema della misura. Abbiamo già evidenziato come il collasso della funzione d'onda vada superato a favore di una evoluzione unitaria del sistema in grado di spiegare il perché non vengano osservate sovrapposizioni di stati in natura.

Un tentativo in questa direzione è appunto la decoerenza, la quale ci ha permesso di vedere come una sovrapposizione coerente di stati quantistici evolve in uno degli stati ortogonali della “base preferita”. La novità introdotta dalla decoerenza è proprio il considerare l'ambiente. Nella nostra descrizione del processo di misura è stata considerata solo l'interazione tra l'apparato di misura e il sistema da misurare, ma è stato ignorato l'ambiente, come se l'apparato fosse isolato dal resto, cosa assolutamente non vera. Consideriamo ora i risultati ottenuti dal toy model precedente. Abbiamo trovato che gli unici stati possibili sono gli stati $|1\rangle$ e $|2\rangle$ e che per un tempo sufficientemente lungo lo scattering dell'ambiente aveva probabilità 1. Quindi dopo l'evoluzione $U_{A \rightarrow A'E} |\psi\rangle$ i due possibili stati sono $\{|1\rangle_A |1\rangle_E, |2\rangle_A |2\rangle_E\}$. Questi due stati sono però ortogonali e non possono essere risultati della stessa trasformazione unitaria agente sullo stesso stato di partenza, come è facilmente dimostrabile [1]:

$$0 = \langle 2|_E \langle 2|_A |1\rangle_A |1\rangle_E = \langle 0|_E \langle \psi|_A U^\dagger U |\psi\rangle_A |0\rangle_E = 1. \quad (122)$$

Quindi abbiamo visto come la decoerenza spiega il passaggio da una sovrapposizione coerente di stati a una miscela statistica di possibili risultati, ma non è conciliabile con una evoluzione unitaria del sistema in grado di spiegare la presenza di risultati definiti in Meccanica Quantistica.

Ora vediamo i risvolti della decoerenza per il problema della base preferita. Abbiamo visto come questo problema consista nel non poter determinare, dalla sola interazione tra apparato di misura e stato che vogliamo misurare, che osservabile stiamo effettivamente misurando. Per trattare questo problema, ritorniamo all'analisi del processo di misura. Durante la nostra analisi, avevamo compiuto una tacita ipotesi, quella di “*faithful measurement*”, ovvero avevamo ipotizzato che l'apparato di misura riproducesse fedelmente lo stato. Questa ipotesi si è tradotta nell'interazione tra apparato di misura e sistema. Infatti abbiamo considerato solo un'interazione del tipo:

$$U(|\psi\rangle \otimes |0\rangle) = \sum_a^N E_a |\psi\rangle \otimes |a\rangle, \quad (123)$$

e non una più generica interazione:

$$U(|\psi\rangle \otimes |0\rangle) = \sum_{a,b}^{N,N} E_a |\psi\rangle \otimes |b\rangle, \quad (124)$$

che corrisponderebbe a riportare una misura diversa dallo stato effettivo del sistema. Come abbiamo visto, questo modello per essere più vicino alla realtà deve tenere conto dell'ambiente in cui sono immersi l'apparato di misura e il sistema microscopico. Però la presenza dell'ambiente

non dovrebbe disturbare la misura. Quindi la “base preferita” dalla decoerenza è quella base che contiene una misura affidabile dell’ambiente. Questo criterio, enunciato da Zurek nel 1981[34], è noto come *criterio di stabilità*. Andiamo a vedere come si traduce nel formalismo studiato in precedenza. Abbiamo osservato come misurare significhi proiettare sui sottospazi generati dagli autostati dell’apparato di misura. Ora il criterio di stabilità impone che l’interazione tra l’ambiente e l’apparato di misura non influenzi la misura stessa, quindi possiamo imporre la condizione:

$$[P_n, H_{A,E}] = 0 \quad \forall n, \quad (125)$$

dove con $P_n = |a_n\rangle\langle a_n|$ abbiamo indicato la proiezione in uno specifico sottospazio generato da un vettore di una base particolare a_n e con H_{AE} abbiamo indicato l’hamiltoniana di interazione tra l’ambiente e l’apparato. Quindi l’osservabile “preferito” che andremo effettivamente a misurare deve essere del tipo:

$$A = \sum_n \lambda_n P_n. \quad (126)$$

Questa condizione è ovviamente rispettata se A è una funzione di H_{AE} . Quindi abbiamo visto come funziona il meccanismo di “scelta” della base preferita. Riassumendo, l’ambiente, tramite l’hamiltoniana H_{AB} , sceglie un osservabile “preferito” dell’apparato, ma indirettamente sceglie anche gli autostati del sistema che verrà misurato, a causa dell’ipotesi di “faithful measurement”. Infatti gli stati dell’apparato devono essere tali da interagire in maniere “fedele” con gli stati del sistema che intendiamo misurare. Quindi la funzione d’onda totale si scriverà come:

$$|\psi\rangle = \sum_n c_n |s_n\rangle_S |a_n\rangle_A |e_n\rangle_E, \quad (127)$$

dove con S, A ed E abbiamo indicato rispettivamente i sottospazi relativi al sistema, all’apparato e all’ambiente. Il teorema di unicità della “tridecomposizione” afferma che se un vettore di uno spazio di Hilbert dato dal prodotto tensoriale di tre sottospazi si può scrivere in forma diagonale, ovvero del tipo:

$$|\phi\rangle_H = \sum_n \alpha_n |a_n\rangle_A |b_n\rangle_B |c_n\rangle_C, \quad (128)$$

dove A, B e C sono i tre sottospazi di H, questa decomposizione è unica. Quindi questo teorema ci garantisce che, una volta trovata la “base preferita” essa è unica.

Abbiamo osservato come la teoria della decoerenza ci porti a considerare l’interazione dell’ambiente con l’apparato. Da queste interazioni abbiamo visto il meccanismo tramite il quale, tra tutte le possibili misure, venga effettivamente scelto un osservabile tale che la misura sia stabile. Ma non sappiamo come determinare l’hamiltoniana d’interazione tra l’apparato e l’ambiente a causa dell’elevato numero di gradi di libertà e conseguentemente non sappiamo trovare a priori questo osservabile. Quindi, anche in questo caso, la teoria della decoerenza ha permesso di capire un processo fondamentale della misura di un sistema fisico, ma una teoria generale coerente che riesca a risolvere questi problemi è ancora un problema aperto.

6 Entanglement, realismo e località

Abbiamo definito nelle sezioni precedenti uno stato entangled come uno stato il cui numero di Schmidt, ovvero il numero di autovalori non nulli delle matrici densità delle due parti del nostro sistema, è superiore ad 1. Inoltre abbiamo anche affermato che il numero di Schmidt non può essere modificato da azioni locali sulle singole parti del nostro sistema, ma è necessario agire su tutto il sistema, facendolo interagire. Quindi l'entanglement traduce l'interazione tra le varie componenti di un sistema fisico. Andiamo ora a vedere cosa implica questa interazione e come sia collegata agli aspetti fondamentali della realtà che ci circonda.

6.1 Articolo EPR

In un famoso articolo del 1935, redatto da Einstein, Podolsky e Rosen, viene messa in discussione la completezza della Meccanica Quantistica. I tre fisici sostengono che affinché una teoria fisica si possa considerare di successo, essa deve essere corretta, ovvero le previsioni della teoria devono essere in accordo con i risultati sperimentali, e deve essere completa, ovvero che ogni elemento della realtà fisica della natura deve avere una controparte nella teoria fisica.

Ovviamente è impossibile sapere a priori quali sono gli elementi della realtà fisica, quindi bisogna fare appello alle osservazioni sperimentali e alle misure. Perciò bisogna enunciare un altro criterio per determinare quali sono effettivamente gli elementi della realtà fisica, che è quello di realismo. Il criterio di realismo viene enunciato come segue “Se è possibile predire con certezza, ovvero con probabilità pari ad 1, senza in alcun modo disturbare il sistema, il valore di una quantità fisica, allora esiste un elemento della realtà fisica corrispondente a tale quantità fisica”. Notiamo come l'interpretazione di Copenaghen non rispetti questo criterio. Infatti non è sempre possibile predire il valore di un osservabile senza effettivamente misurarlo e abbiamo visto come la misura nasca dall'interazione tra apparato di misura e sistema da misurare e quindi prima di questa interazione il risultato della misura non esiste. Si può esprimere questa condizione dicendo che l'interpretazione di Copenaghen non ha la “counterfactual definiteness” ovvero non è possibile discutere la determinatezza di un risultato di una misura che non è stata effettuata. Vedremo che la “counterfactual definiteness” sarà una delle ipotesi del teorema di Bell.

In Meccanica Quantistica, un sistema fisico è completamente determinato dal suo stato, o dalla sua matrice densità nel caso di sistemi aperti, ma abbiamo visto come una matrice densità è una sovrapposizione di stati. Quindi lo stato, ovvero la funzione d'onda associata allo stato dovrebbe fornire una descrizione completa, nel senso di completezza espresso precedentemente del sistema, ma non è così. Infatti dati due operatori che non commutano, A e B , in base al principio di indeterminazione di Heisenberg, non è possibile allo stesso tempo conoscere le due quantità fisiche corrispondenti ai due osservabili. Inoltre una misura di A altererebbe in maniera irreversibile lo stato e non sarebbe più possibile conoscere il valore della misura di B dello stato inalterato. Allora o la funzione d'onda non fornisce una descrizione completa della realtà o le quantità fisiche associate a due osservabili che non commutano non “esistono” simultaneamente, ovvero non rispettano il principio di realismo enunciato precedentemente. Ciò perché se le due quantità esistessero contemporaneamente e la funzione d'onda offrisse una descrizione completa, allora in base a tale completezza la funzione d'onda dovrebbe comprendere la descrizione di queste quantità fisiche e rispettando il principio di realtà i due valori di tali grandezze dovrebbero essere prevedibili, condizione che invece non si verifica.

Continuiamo la nostra analisi e consideriamo la funzione d'onda di due sistemi fisici che hanno

interagito per un intervallo di tempo T . Sappiamo che, data la loro interazione, esiste un'unica funzione d'onda $|\Psi\rangle$ che descrive i due sistemi. Esprimiamo $|\Psi\rangle$ nella base degli autostati $u(x_1)$ di un certo osservabile A appartenente allo spazio degli operatori agenti sullo spazio di Hilbert relativo al sistema 1 e indichiamo con x_1 la variabile che determina il primo sistema. Considerata analogamente x_2 la variabile utilizzata per descrivere il secondo sistema, possiamo scrivere $|\Psi\rangle$ nella seguente forma:

$$\Psi(x_1, x_2) = \sum_n \psi_n(x_2)u_n(x_1). \quad (129)$$

Possiamo vedere questa espressione come lo sviluppo di $\Psi(x_1, x_2)$ nella base $u_n(x_1)$ e i $\psi_n(x_2)$ come semplici coefficienti. Immaginiamo di misurare A e di trovare come risultato il generico valore a_k , dove a è un autovalore di A . Allora, in base al collasso della funzione d'onda, sappiamo che il sistema 2 sarà nello stato $\psi_k(x_2)$. Ora immaginiamo di esprimere $\Psi(x_1, x_2)$ usando un insieme di autostati, $\nu(x_1)$, di un secondo operatore, B . Analogamente $\Psi(x_1, x_2)$ può essere espresso come sviluppo di $\nu(x_1)$:

$$\Psi(x_1, x_2) = \sum_n \phi_n(x_2)\nu_n(x_1). \quad (130)$$

Anche in questo caso, se misuriamo B , otteniamo il collasso della funzione d'onda. Quindi se misuriamo il valore b_r , dove b è un autovalore di B , abbiamo che il sistema 2 sarà nello stato $\phi_r(x_2)$. Supponiamo che le misurazioni di A e B siano state effettuate ad un tempo $t > T$, ovvero dopo che il sistema 1 e il sistema 2 abbiano già interagito. Allora, a meno di una teoria non locale e di azioni a distanza, la misura del sistema 1 non può influenzare il sistema 2. Quindi abbiamo visto come è possibile associare due funzioni d'onda diverse, nel nostro caso $\phi_r(x_2)$ e $\psi_k(x_2)$, alla stesso stato fisico reale. Ora osserviamo cosa succede se queste due funzioni d'onda sono rispettivamente autostati di due osservabili, Q e P , agenti sullo spazio di Hilbert associato al sistema 2, che non commutano.

Per rendere più chiara l'esposizione, utilizzeremo come esempio il caso degli osservabili momento e posizione. Consideriamo $\Psi(x_1, x_2)$ come il prodotto di due sistemi rappresentati da due onde piane:

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} e^{\frac{2\pi ip(x_1 - x_2 + x_0)}{\hbar}} dp, \quad (131)$$

dove x_0 è una costante arbitraria. Vediamo come $e^{\frac{2\pi ipx_1}{\hbar}}$ è un autostato dell'operatore momento agente sullo spazio di Hilbert relativo al sistema 1. Identificando le espressioni 131 e 129 possiamo notare che:

$$\psi_p(x_2) = e^{-\left(\frac{2\pi ip}{\hbar}\right)(x_2 - x_0)}, \quad (132)$$

ovvero è l'autostato dell'operatore momento del secondo stato relativo all'autovalore p . Analogamente a quanto fatto in precedenza, esprimiamo $\Psi(x_1, x_2)$ in funzione di un set di autostati di un secondo osservabile, in questo caso l'operatore posizione. Sappiamo che l'autostato dell'operatore posizione relativo all'autovalore x è:

$$\nu_x(x_1) = \delta(x - x_1). \quad (133)$$

Quindi abbiamo che:

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\frac{2\pi ip(x - x_2 + x_0)}{\hbar}} dp \delta(x - x_1) dx. \quad (134)$$

Identificando le espressioni 134 e 130 possiamo scrivere:

$$\phi_x(x_2) = \int_{-\infty}^{\infty} e^{\frac{2\pi ip(x-x_2+x_0)}{\hbar}} dp = \hbar \delta(x - x_2 + x_0). \quad (135)$$

Quindi in questo caso $\phi_x(x_2)$ è autostato dell'operatore posizione del sistema 2 con autovalore $x + x_0$.

Abbiamo visto pertanto come due funzioni d'onda diverse, associate allo stesso stato fisico possano essere autostati di due operatori che non commutano. Quindi, a seconda di quale operatore del sistema 1 misuriamo, A o B, prepariamo un'autostato dell'operatore P o Q, momento o posizione nel nostro esempio, del sistema 2, quindi possiamo predire con certezza la sua evoluzione, ovvero P nel primo caso e Q nel secondo caso sono elementi della stessa realtà fisica. In precedenza eravamo arrivati alla conclusione che o la descrizione della realtà della Meccanica Quantistica non è completa o che due osservabili che non commutano non possono "esistere" simultaneamente, ovvero non possono aderire al principio di realismo enunciato in precedenza. Quindi, unendo queste due conclusioni, possiamo affermare che la Meccanica Quantistica non è una teoria completa. L'articolo si chiude con l'auspicio che una teoria completa venga trovata. Nella sezione successiva vedremo che, se vengono mantenute tutte le ipotesi fatte dai 3 autori, ciò è impossibile.

6.2 Il teorema di Bell

L'incompletezza della Meccanica Quantistica scosse profondamente tutta la comunità scientifica. Dopo l'articolo di Einstein, Podolsky e Rosen [6] sono state proposte diverse teorie delle variabili nascoste, ovvero teorie che includessero più variabili rispetto alla formulazione classica della Meccanica Quantistica, nella speranza che potessero risolvere il problema dell'incompletezza. Una teoria delle variabili nascoste divenuta celebre è la teoria di De Broglie-Bohm, i quali avevano postulato l'esistenza di "un'onda pilota" che, insieme alla funzione d'onda, caratterizza completamente un sistema. A differenza della funzione d'onda, l'onda pilota esiste anche quando non è osservata. Bell [7] pose fine alla questione nel 1964 quando ha dimostrato che nessuna teoria locale delle variabili nascoste può riprodurre gli stessi risultati delle Meccanica Quantistica.

Per dimostrare questo importantissimo risultato partiamo dalla formulazione di Bohm [9] della posizione EPR. Partiamo da una coppia di spin $1/2$ nello stato di singoletto:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle), \quad (136)$$

dove $|\uparrow\rangle$ e $|\downarrow\rangle$ sono gli autostati dell'osservabile spin lungo l'asse z. Supponiamo che le due particelle viaggino in direzioni opposte e che, dopo che si siano allontanate, due sperimentatori, Alice e Bob, misurino rispettivamente lo spin della particella 1 e delle particella 2 lungo l'asse z. Se Alice misura il valore $1/2$ allora, in base al collasso della funzione d'onda, Bob avrà probabilità 1 di misurare il valore $-1/2$. Ovviamente ciò non dipende dalla scelta dell'asse, avremmo ottenuto lo stesso risultato, ovvero una perfetta anticorrelazione, anche nel scegliendo l'asse x (ricordiamo che gli autostati dell'operatore σ_x sono $|\pm\rangle = 1/\sqrt{2}(|\uparrow\rangle \pm |\downarrow\rangle)$). Quindi lo spin di Bob ha simultaneamente valori definiti lungo due assi diversi, violando il principio di indeterminazione di Heisenberg (σ_x e σ_z non commutano). Ora immaginiamo che Alice misuri lo spin lungo l'asse z e Bob lungo l'asse x. Bob avrà probabilità $1/2$ di misurare il valore sia il valore $1/2$ sia il valore $-1/2$. Quindi è come se lo spin di Bob risentisse di un'azione a

distanza dovuta alla misura dello spin di Alice. Ciò è ovviamente impossibile in quanto viola in principio di località. Prima di continuare e dimostrare il teorema di Bell, facciamo una piccola digressione sullo spin per rendere più chiaro il proseguo della dimostrazione.

Data una direzione arbitraria $\mathbf{a}(\theta) = (\sin(\theta), 0, \cos(\theta))$ e indicato con $\boldsymbol{\sigma}$ il vettore che ha per componenti le matrici di Pauli, l'osservabile corrispondente alla misura dello spin lungo la direzione \mathbf{a} è:

$$\boldsymbol{\sigma} \cdot \mathbf{a} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}. \quad (137)$$

Esprimiamo i suoi autostati in funzione degli autostati dell'osservabile spin lungo l'asse z:

$$|\theta_+\rangle = \cos\left(\frac{\theta}{2}\right) |\uparrow\rangle + \sin\left(\frac{\theta}{2}\right) |\downarrow\rangle, \quad |\theta_-\rangle = -\sin\left(\frac{\theta}{2}\right) |\uparrow\rangle + \cos\left(\frac{\theta}{2}\right) |\downarrow\rangle. \quad (138)$$

Quindi vediamo come dato uno stato $|\uparrow\rangle$ le probabilità di misurare rispettivamente i valori $1/2$ e $-1/2$ lungo la direzione $\mathbf{a}(\theta)$ siano:

$$P(\theta_+) = \cos^2\left(\frac{\theta}{2}\right), \quad P(\theta_-) = \sin^2\left(\frac{\theta}{2}\right). \quad (139)$$

Invece dato uno stato $|\downarrow\rangle$ avremo rispettivamente:

$$P(\theta_+) = \sin^2\left(\frac{\theta}{2}\right), \quad P(\theta_-) = \cos^2\left(\frac{\theta}{2}\right). \quad (140)$$

Torniamo ad Alice e Bob. Supponiamo che Alice misuri lo spin lungo l'asse z e Bob lungo un asse \mathbf{a} . Basandosi sul collasso delle funzione d'onda dovuta alla misura di Alice e alle probabilità trovate in precedenza, possiamo determinare la probabilità delle quattro possibili combinazioni (indicheremo con l'apice A e B le misure effettuate rispettivamente da Alice e Bob):

$$P(\sigma_z^A = 1/2, \sigma_a^B = 1/2) = \frac{1}{2} \sin^2\left(\frac{\theta}{2}\right), \quad P(\sigma_z^A = 1/2, \sigma_a^B = -1/2) = \frac{1}{2} \cos^2\left(\frac{\theta}{2}\right), \quad (141)$$

$$P(\sigma_z^A = -1/2, \sigma_a^B = 1/2) = \frac{1}{2} \cos^2\left(\frac{\theta}{2}\right), \quad P(\sigma_z^A = -1/2, \sigma_a^B = -1/2) = \frac{1}{2} \sin^2\left(\frac{\theta}{2}\right). \quad (142)$$

Ora supponiamo l'esistenza di una teoria delle variabili nascoste. Supponiamo esiste una funzione f_a tale che può assumere due valori 0,1 in base al valore dello spin misurato lungo una qualsiasi direzione. Quindi, ipoteticamente, conoscendo il valore di questa funzione potremmo sapere con certezza, in ogni istante, il valore di una grandezza fisica, in questo caso lo spin. Stiamo quindi assumendo che lo spin abbia un valore ben definito, e perfettamente prevedibile anche se non misurato. Quindi stiamo utilizzando il criterio di realismo di Einstein. Abbiamo visto come esista una perfetta anticorrelazione tra gli spin di Alice e Bob. Se Alice misura $1/2$ allora Bob misurerà sicuramente $-1/2$ e viceversa. Quindi possiamo scrivere $f_a^A = -f_a^B$. Notiamo come dal punto di vista logico è stato naturale definire la precedenti uguaglianza, anche se nascondono diverse insidie dal punto di vista fisico. Infatti abbiamo affermato che lo spin di Bob sarà in una configurazione opposta rispetto allo spin di Alice, ma non abbiamo effettivamente misurato lo spin di Bob. Ovvero abbiamo applicato il concetto di "counterfactual definiteness", abbiamo discusso la determinatezza di un risultato senza averlo effettivamente misurato. Inoltre così facendo abbiamo assegnato valori simultanei a spin in direzioni diversi,

ovvero ad osservabili che non commutano, in piena violazione del principio di indeterminazione di Heisenberg. Assumendo per ipotesi che esista una tale teoria è possibile riscrivere le precedenti probabilità usando solo lo spin di Alice:

$$P(\sigma_z^A = 1/2, \sigma_a^A = -1/2) = \frac{1}{2} \sin^2 \left(\frac{\theta}{2} \right), \quad P(\sigma_z^A = 1/2, \sigma_a^A = +1/2) = \frac{1}{2} \cos^2 \left(\frac{\theta}{2} \right), \quad (143)$$

$$P(\sigma_z^A = -1/2, \sigma_a^A = -1/2) = \frac{1}{2} \cos^2 \left(\frac{\theta}{2} \right), \quad P(\sigma_z^A = -1/2, \sigma_a^A = +1/2) = \frac{1}{2} \sin^2 \left(\frac{\theta}{2} \right). \quad (144)$$

Vediamo cosa succede se aggiungiamo una terza direzione. Supponiamo che Alice e Bob possano misurare il proprio spin nelle direzioni $\mathbf{a}(\theta) = (\sin(\theta), 0, \cos(\theta))$, $\mathbf{b}(\phi) = (\sin(\phi), 0, \cos(\phi))$ e lungo l'asse z . Quindi esisteranno 3 funzioni, f_z, f_a, f_b che determinano con certezza statistica la misura di questi 3 valori e garantiscano quindi la loro esistenza. Consideriamo la probabilità di misurare $1/2$ lungo \mathbf{a} e $-1/2$ lungo \mathbf{b} :

$$P(\sigma_a^A = 1/2, \sigma_b^A = -1/2) = P(\sigma_z^A = 1/2, \sigma_a^A = 1/2, \sigma_b^A = -1/2) + P(\sigma_z^A = -1/2, \sigma_a^A = 1/2, \sigma_b^A = -1/2). \quad (145)$$

Dato che la probabilità è una quantità non negativa possiamo scrivere:

$$P(\sigma_a^A = 1/2, \sigma_b^A = -1/2) \leq P(\sigma_z^A = 1/2, \sigma_a^A = 1/2, \sigma_b^A = -1/2) + P(\sigma_z^A = 1/2, \sigma_a^A = 1/2, \sigma_b^A = +1/2) + P(\sigma_z^A = -1/2, \sigma_a^A = 1/2, \sigma_b^A = -1/2) + P(\sigma_z^A = -1/2, \sigma_a^A = -1/2, \sigma_b^A = -1/2) = P(\sigma_z^A = 1/2, \sigma_a^A = 1/2) + P(\sigma_z^A = -1/2, \sigma_b^A = -1/2).$$

Applicando i risultati ottenuti in precedenza, 143, abbiamo che:

$$\sin^2 \left(\frac{\phi + \theta}{2} \right) \leq \cos^2 \left(\frac{\phi}{2} \right) + \cos^2 \left(\frac{\theta}{2} \right). \quad (146)$$

Questa disuguaglianza, nota come disuguaglianza di Bell, non è però verificata da tutti i valori di θ e ϕ . Infatti prendendo $\theta = \frac{3\pi}{2}$ e $\phi = \frac{3\pi}{4}$ otteniamo $\frac{1}{\sqrt{2}} < \frac{1}{2}$. Ovviamente ciò è assurdo e quindi abbiamo dimostrato che non può esistere una teoria delle variabili nascoste locale e che rispetti le previsioni della Meccanica Quantistica. Essendo le previsioni della Meccanica Quantistica confermate da un grandissimo numero di evidenze sperimentali, dobbiamo abbandonare o l'ipotesi di località o il criterio di realismo. Preferendo il criterio di realismo, si parla di non-località quantistica e pertanto di azioni a distanza. È possibile dimostrare che, tramite queste azioni a distanza, non è possibile comunicare informazioni, quindi vengono rispettati i postulati della relatività speciale di Einstein. Forniamo un piccolo esempio di questa incomunicabilità delle azioni a distanza.

I fisici tendono a preferire il principio di località al criterio di realismo. La teoria quantistica dei campi è basata sulla località e si è dimostrata una teoria fisica di successo. Quindi siamo costretti a abbandonare l'idea di variabili nascoste che ci permettano di determinare in un momento qualsiasi il valore di una grandezza fisica e accettare l'indeterminatezza quantistica non come una mancanza di conoscenza, ma come una caratteristica intrinseca nella natura.

7 Informazione quantistica

Nel primo capitolo abbiamo introdotto la teoria dell'informazione classica e i suoi capisaldi, ovvero i teoremi di Shannon. Successivamente abbiamo trattato ampiamente la Meccanica Quantistica dei sistemi aperti. Ora vediamo come ritrovare, grazie proprio a questa seconda trattazione, risultati analoghi alla teoria dell'informazione classica ma nel campo della Quantistica.

7.1 Entropia di Von Neumann

Data una variabile aleatoria discreta x e la sua distribuzione di probabilità p_{x_i} , abbiamo visto come l'entropia di Shannon della variabile sia uguale a:

$$H(x) = \sum_i^n p_{x_i} \log(p_{x_i}). \quad (147)$$

Inoltre abbiamo enfatizzato come tale quantità sia uguale al *valore atteso dell'autoinformazione* del messaggio contenuto nella variabile e fornisce anche una misura della *quantità di scelta* presente nel messaggio.

Ora supponiamo che tali eventi siano stati quantistici. All' i -esimo stato puro ρ_i associamo la probabilità p_i . Allora sappiamo che il nostro sistema è interamente descritto dalla matrice densità ρ :

$$\rho = \sum_i^n p_i \rho_i. \quad (148)$$

Allora, data una matrice densità qualsiasi, è possibile definire l'*entropia di Von Neumann* [20] come:

$$S(\rho) = -\text{tr}(\rho \log(\rho)). \quad (149)$$

Effettuiamo una piccola digressione per chiarire meglio il significato di questa espressione. Il logaritmo di una matrice A è definito come quella matrice tale che l'esponenziale del logaritmo di A è uguale ad A stessa. Ricordiamo inoltre che l'esponenziale di una matrice è invece definito tramite serie: $e^A = \sum_i^n \frac{A^i}{i!}$.

Nella base $|a\rangle$ in cui ρ è diagonale è possibile osservare subito che:

$$S(\rho) = H(A), \quad (150)$$

dove $H(A)$ è l'entropia di Shannon dell'insieme $\{\lambda_a, a\}$, con λ_a autovalori relativi all'autostato $|a\rangle$. In generale però l'entropia di Shannon fornisce un limite all'entropia di Von Neumann. Infatti dato un insieme generico di stati puri $X = \{|\phi_x\rangle, p_x\}$, a cui sono state associate le rispettive probabilità, si verifica:

$$H(X) \geq S(\rho). \quad (151)$$

L'uguaglianza è valida solo nel caso di stati puri ortogonali fra loro.

Enunciamo alcune proprietà dell'entropia di Von Neumann:

- L'entropia di uno stato puro $\rho = |\psi\rangle\langle\psi|$ è nulla. Ciò perché ad uno stato puro è associata una probabilità uguale ad 1;
- L'entropia è invariata per cambiamento di base. Data una matrice unitaria U :

$$S(\rho) = S(U\rho U^{-1}). \quad (152)$$

- ρ assume il suo valore massimo $\log(N)$, dove N è la dimensione dello spazio di Hilbert in cui stiamo operando, per uno stato *maximally entagled*. Uno stato si dice maximally entagled quando la sua matrice densità ridotta è un multiplo della matrice identità. Ciò corrisponde all'equivalente quantistico della proprietà secondo la quale l'entropia di Shannon è massima per un insieme di stati equiprobabili;
- Se proviamo a misurare l'osservabile A :

$$A = \sum_i^n p_i |a_i\rangle\langle a_i|. \quad (153)$$

Avremo che la probabilità di misurare a_i sarà uguale:

$$p(a_i) = \langle a_i | \rho | a_i \rangle. \quad (154)$$

Allora considerata l'entropia di Shannon $H(A)$ dell'insieme $\{a, p(a)\}$ avremo che:

$$S(\rho) \leq H(A). \quad (155)$$

In particolare avremo l'uguaglianza quando A e ρ commutano. Fisicamente ciò significa che l'entropia, e quindi la casualità della misura, è minima se scegliamo una determinata classe di osservabili, ovvero quelli che commutano con la matrice densità.

- Dato un sistema bipartito AB si verifica che:

$$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B). \quad (156)$$

L'uguaglianza è valida nel caso di un sistema separabile ovvero del tipo $H_{AB} = H_A \otimes H_B$. Anche questa è l'analogo quantistico di una proprietà dell'entropia di Shannon. Fisicamente significa che per sistema non correlati l'entropia è additiva, mentre per sistema correlati parte dell'informazione è contenuta nella correlazione tra le due parti del sistema, quindi l'entropia del sistema è minore della somma delle sue parti.

Sfruttiamo quest'ultima proprietà per ricavare un risultato molto interessante, la formulazione matematica del *secondo principio della termodinamica*, usando l'entropia di Von Neumann. Come già visto precedentemente, supponiamo di avere un sistema aperto A immerso nel suo ambiente E e supponiamo che all'istante iniziale non siano correlati. Data quest'ipotesi, allora vale:

$$S(\rho_{AE}) = S(\rho_A) + S(\rho_E). \quad (157)$$

L'evoluzione temporale di ρ_{AE} è governata dall'operatore unitario di evoluzione temporale U . Come abbiamo visto applicare un operatore unitario non varia il valore dell'entropia quindi ad istante t' l'entropia sarà uguale a quella iniziale, ma il sistema e il suo ambiente saranno correlati, quindi:

$$S(\rho_A) + S(\rho_B) = S(\rho_{AB}) = S(\rho'_{AB}) \leq S'_A + S'_B. \quad (158)$$

Con l'uguaglianza valida nel caso in cui non ci sia nessuna correlazione. Quindi abbiamo visto che l'entropia globale, ovvero del mondo e del suo ambiente può unicamente aumentare. Ciò è risultato analogo al secondo principio della Termodinamica classica.

7.2 Teoremi dell'informazione quantistica

Una volta acquisita la nozione di entropia di Von Neumann, possiamo illustrare i risultati di Schumacher [24][8], che ha avuto il merito di trovare la codifica che massimizza la compressione di un messaggio quantistico con un errore arbitrariamente piccolo. La dimostrazione di questo risultato, data la sua complessità, esula dagli scopi di questa tesi. La codifica di Schumacher può essere vista come un *analogo quantistico del primo teorema di Shannon*. Infatti ricordiamo che il primo teorema di Shannon fornisce il massimo tasso di trasmissione di un messaggio classico.

Consideriamo un messaggio di n lettere, dove ogni lettera è presa da un insieme di N stati puri $\{|\phi_x\rangle, p_x\}$. Quindi ogni lettera è descritta dalla matrice densità:

$$\rho = \sum_i^N p_i |\phi_i\rangle\langle\phi_i|. \quad (159)$$

Allora l'intero messaggio è descritto dalla matrice:

$$\rho^n = \rho \otimes \rho \otimes \dots \rho. \quad (160)$$

Ora vorremmo sapere se è possibile trovare una codifica di questo messaggio tale che possa essere descritto da uno spazio di Hilbert di dimensione minore. Schumacher ha dimostrato che è possibile comprimere un messaggio quantistico, con probabilità di perdita di informazione tendente a 0 quando $n \rightarrow \infty$, in uno spazio di Hilbert H tale che:

$$\log(\dim(H)) = nS(\rho). \quad (161)$$

Quindi si vede che l'entropia di Von Neumann è legata alla dimensione dello spazio di Hilbert usato per descrivere il messaggio, ovvero al numero di qubit.

Cosa succede se l'insieme di partenza è composto da miscele statistiche e non da stati puri? È facile vedere che il risultato precedente non è valido. Infatti consideriamo un insieme di stati e di probabilità associate, come visto in precedenza, in cui sia presente una miscela statistica ρ_0 a cui è associata una probabilità pari ad 1. Allora il messaggio di n lettere sarà sempre uguale a $\rho_0 \otimes \rho_0 \otimes \dots \rho_0$ e non ci sarà informazione trasferita, quindi il messaggio può essere compresso in 0 qubit, che è minore dell'entropia di Von Neumann della miscela statistica considerata, infatti ricordiamo $S(\rho_0) > 0$, in contraddizione con quanto visto precedentemente. Quindi dobbiamo ricavare un'espressione che si riduca all'entropia di Von Neumann per gli stati puri e che sia uguale all'entropia di Shannon nel caso di stati puri ed ortogonali fra loro. Queste due imposizioni portano a definire l'*informazione di Holevo* [11] dell'insieme ε , $\chi(\varepsilon)$:

$$\chi(\varepsilon) = S(\rho) - \sum_i^n p_x S(\rho_x). \quad (162)$$

Questa espressione si riduce all'entropia di Von Neumann per un insieme di stati puri, in quanto $S(\rho_x) = 0$. Si nota che $\chi(\varepsilon)$ dipende non solo da ρ ma anche in che modo ρ è stato preparato come insieme di miscele statistiche (si noti come l'espressione dipenda dalla singola miscela statistica e non solo dall'insieme). Quindi l'informazione di Holevo è il numero ottimale di qubit per lettera che si può ottenere comprimendo un messaggio di n lettere con un errore arbitrariamente piccolo per grandi n . Notiamo anche che l'informazione di Holevo ricorda la

mutua informazione nella teoria dell'informazione classica. Ricordiamo la sua espressione:

$$I(X; Y) = H(X) - H(X|Y) = \sum_i^N \sum_k^M p(x_i, y_k) \log \left(\frac{p(x_i, y_k)}{p(x_i)p(y_k)} \right). \quad (163)$$

Essa rappresenta la riduzione dell'entropia di X dovuta alla conoscenza di Y . Allo stesso modo, l'informazione di Holevo rappresenta quanto l'entropia di Von Neumann viene ridotta dalla conoscenza della particolare preparazione di ρ . Concludiamo vedendo come l'informazione di Holevo è legata all'informazione classica. Supponiamo di avere un insieme di stati, ε , ognuno associato ad una certa probabilità, e di preparare uno di questi possibili stati. Un osservatore esterno conosce l'insieme di stati, ma non lo stato particolare che è stato preparato e vuole ricavare la maggior quantità di informazione possibile dalla misura di quest'ultimo. In particolare, consideriamo che sia stato preparato lo stato i -esimo allora la probabilità di osservare lo stato j -esimo tramite la misura generalizzata $\{F_j\}$ è:

$$p(j|i) = \langle j | F_y | i \rangle. \quad (164)$$

Data la relazione $p(x, y) = p(x|y)p(y)$ e data l'ipotesi seconda la quale l'osservatore conosce l'insieme, e quindi le singole probabilità associate agli stati, allora si vede come tramite questa misura è possibile trovare la mutua informazione $I(X; Y)$ della preparazione X , noto il risultato della misura Y . Quindi è possibile definire come *informazione accessibile* la mutua informazione massima ottenuta al variare della misura generalizzata:

$$Acc(\varepsilon) = Max_{F_y} I(X; Y). \quad (165)$$

Se gli stati sono stati puri mutuamente ortogonali, allora è facile vedere che la misura che massimizza $I(X; Y)$ è la misura ortogonale $F_j = |j\rangle\langle j|$. Infatti abbiamo che:

$$p(j|i) = \langle i | j \rangle = \delta_{ij}. \quad (166)$$

Essendo misure perfettamente distinguibili si vede che l'entropia condizionale è nulla e l'informazione accessibile assume il valore massimo pari a $H(X)$. Se l'insieme di partenza non è composto da stati puri mutuamente ortogonali allora non è possibile trovare un'espressione sempre valida per $Acc(\varepsilon)$, ma è possibile dimostrare che questa quantità è limitata ed è valida la relazione:

$$Acc(\varepsilon) \leq \chi(\varepsilon). \quad (167)$$

Questa relazione è nota come *limite di Holevo*. Vediamo un semplice esempio. Immaginiamo di inviare un set di fotoni non correlati e supponiamo che il rumore agisca indipendentemente su ognuno di essi. L'informazione di Holevo allora ci aiuta a determinare quanta informazione possiamo ricavare dalla misura, ad esempio della polarizzazione, di un singolo fotone. In particolare:

$$\chi(\varepsilon) \leq S(\rho) \leq 1. \quad (168)$$

Allora possiamo affermare che un singolo fotone può trasportare al massimo un bit di informazione classica.

7.3 Teletrasporto Quantistico

Concludiamo questo elaborato proponendo un'applicazione pratica di informazione quantistica: *il teletrasporto quantistico* [2]. Esso è una procedura tramite la quale è possibile trasmettere informazione quantistica da un luogo all'altro. Considerando il *Teorema No Cloning* [33], il quale afferma che è impossibile effettuare una copia esatta di uno stato quantistico, si evince l'importanza di tale risultato. Nel dettaglio una semplice formulazione di tale teorema afferma che dati tre vettori normalizzati, $|\psi\rangle, |\phi\rangle, |0\rangle$ appartenenti ad uno spazio di H , allora se esiste un operatore unitario U_c agente su $H \otimes H$ tale che $U_c(|\psi\rangle |0\rangle) = |\psi\rangle |\psi\rangle$ e $U_c(|\phi\rangle |0\rangle) = |\phi\rangle |\phi\rangle$ allora $\langle\psi|\phi\rangle = 0, 1$. Ecco una breve dimostrazione:

$$\langle\psi|\phi\rangle = \langle\psi|\langle 0|\phi\rangle|0\rangle = \langle\psi|\langle 0|U^\dagger U|\phi\rangle|0\rangle = \langle\psi|\langle\psi|\psi\rangle|\psi\rangle = \langle\psi|\phi\rangle^2 = \langle\psi|\phi\rangle \longrightarrow \langle\psi|\phi\rangle = 0, 1. \quad (169)$$

Questo teorema ha conseguenze importanti anche per il problema della misura in Meccanica Quantistica, problema già trattato in precedenza nel capitolo 3. Infatti abbiamo visto come una misura ortogonale può distinguere solo stati ortogonali fra loro. Dati due stati non ortogonali $|\psi\rangle$ e $|\phi\rangle$ esistono probabilità non nulle di ottenere ognuno dei due stati. Se fosse possibile effettuare una copia di questi stati potremmo ripetere la misura un numero arbitrario di volte, facendo tendere l'errore a 0.

Il nome, però, non deve trarre in inganno. Infatti, come già affermato, il teletrasporto quantistico è una forma di comunicazione e non di effettivo trasporto di materia o energia. Inoltre esso si basa sia su comunicazione classica che quantistica.

Supponiamo che una scienziata, Alice, sia in possesso di un sistema quantistico ignoto a due livelli, come ad esempio una particella di spin $\frac{1}{2}$, che indicheremo con $|\phi\rangle$, o alternativamente con il numero 1. Tale scienziata desidera comunicare lo stato di questo sistema ad un secondo sperimentatore, Bob. Supponiamo inoltre che sia Alice che Bob siano rispettivamente in possesso di una particella di un singoletto di spin, generato in precedenza, che indicheremo con i numeri 2 e 3:

$$|\psi_{23}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_2\rangle |\downarrow_3\rangle - |\downarrow_2\rangle |\uparrow_3\rangle). \quad (170)$$

Un sistema di questo tipo viene anche chiamato coppia EPR, per ragioni storiche. Ora vediamo come utilizzando questo sistema ausiliario, anche chiamato ancilla, sia possibile trasmettere lo stato. Per avere un quadro completo, scriviamo il vettore di stato del sistema incognito in una base arbitraria $\{|\uparrow_1\rangle, |\downarrow_1\rangle\}$:

$$|\phi\rangle = a |\uparrow_1\rangle + b |\downarrow_1\rangle, \quad |a|^2 + |b|^2 = 1. \quad (171)$$

e il vettore di stato complessivo delle tre particelle, il sistema ignoto e la coppia EPR, che in questo momento non sono correlati:

$$|\psi_{123}\rangle = |\phi\rangle \otimes |\psi_{23}\rangle = \frac{a}{\sqrt{2}}(|\uparrow_1\rangle |\uparrow_2\rangle |\downarrow_3\rangle - |\uparrow_1\rangle |\downarrow_2\rangle |\uparrow_3\rangle) + \frac{b}{\sqrt{2}}(|\downarrow_1\rangle |\uparrow_2\rangle |\downarrow_3\rangle - |\downarrow_1\rangle |\downarrow_2\rangle |\uparrow_3\rangle). \quad (172)$$

Riscriviamo tale espressione in funzione di una nuova base:

$$|\psi_{12}^\pm\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\rangle |\downarrow_2\rangle \pm |\downarrow_1\rangle |\uparrow_2\rangle), \quad |\phi_{12}^\pm\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\rangle |\uparrow_2\rangle \pm |\uparrow_1\rangle |\downarrow_2\rangle), \quad (173)$$

$$|\psi_{123}\rangle = \frac{1}{2}(|\psi_{12}^-\rangle (-a |\uparrow_3\rangle - b |\downarrow_3\rangle) + |\psi_{12}^+\rangle (-a |\uparrow_3\rangle + b |\downarrow_3\rangle) + |\phi_{12}^-\rangle (a |\uparrow_3\rangle + b |\downarrow_3\rangle) + |\phi_{12}^+\rangle (a |\uparrow_3\rangle - b |\downarrow_3\rangle)). \quad (174)$$

Ora è necessario far interagire la particella ignota 1 con la particella della coppia EPR in possesso di Alice per creare una correlazione da poter sfruttare successivamente. Per fare ciò Alice effettua una misura sui due sistemi in suo possesso, proiettandoli nella base $\{|\psi_{12}^{\pm}\rangle, |\phi_{12}^{\pm}\rangle\}$. Come si evince dalla 174 i possibili risultati di questa misura, indipendentemente dal sistema ignoto $|\phi\rangle$ sono equiprobabili e avranno come risultato il collasso della particella 3 in uno dei seguenti 4 possibili stati:

$$|1\rangle = -a |\uparrow_3\rangle - b |\downarrow_3\rangle = -|\phi_3\rangle, \quad |2\rangle = -a |\uparrow_3\rangle + b |\downarrow_3\rangle = Z |\phi_3\rangle, \quad (175)$$

$$|3\rangle = a |\uparrow_3\rangle + b |\downarrow_3\rangle = X |\phi_3\rangle, \quad |4\rangle = a |\uparrow_3\rangle - b |\downarrow_3\rangle = Y |\phi_3\rangle. \quad (176)$$

Dove con Z, X e Y sono stati indicati gli operatori di rotazione intorno agli assi, rispettivamente, x, y, z per un angolo di 180° . Quindi Alice, una volta effettuata la misure sulle particelle 1 e 2, comunica, in maniera classica, il risultato a Bob. Essendo 4 i possibili risultati, sono necessari 2 bit di informazione classica per trasportare questo messaggio. Nel caso 1, la particella 3, in possesso di Bob, si troverà, a meno di un fattore di fase, nello stesso stato della particella incognita. Nei casi 2, 3, 4 invece Bob dovrà effettuare una rotazione rispetto agli assi per affinché la sua particella sia nello stato della particella incognita, la quale nel frattempo si troverà in uno stato diverso da quello di partenza a causa dell'interazione con la particella 2. Quindi l'effetto complessivo di questo schema è stato quello di trasferire lo stato della particella incognita 1, in possesso di Alice, alla particella 3, in possesso di Bob.

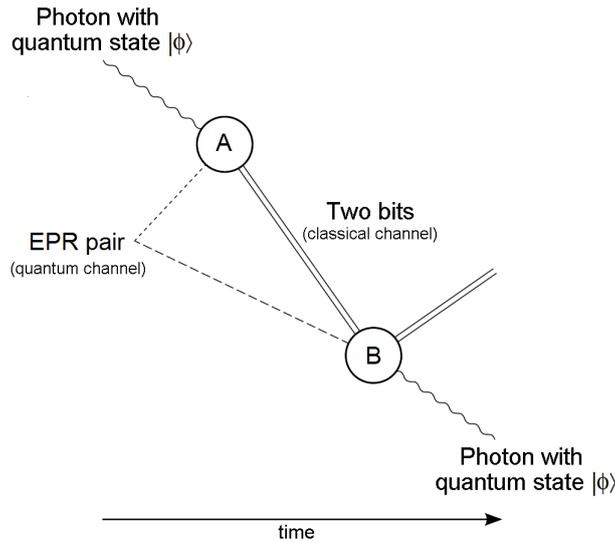


Figura 3: Rappresentazione schematica del protocollo.

Inoltre la procedura che abbiamo descritto è totalmente lineare, quindi se il sistema incognito di partenza fosse un sistema aperto, che possiamo vedere come una combinazione di stati puri, sarebbe possibile applicare la stessa procedura.

Esiste anche una procedura inversa, che permette di comunicare due bit di informazione classica attraverso un canale quantistico. Tale procedura è chiamata *dense coding* [28]. Il nome è chiaro proprio perché attraverso un singolo qubit possiamo trasmettere due bit di informazione classica ed è esattamente l'opposto del teletrasporto quantistico, in cui tramite due bit di informazione classica trasportavamo un qubit. Anche in questo caso, Alice e Bob possiedono rispettivamente una particella di un coppia EPR (utilizzeremo i pedici 1 e 2 per riferirci alla particelle della

coppia dato che non è presente più il sistema incognito 1):

$$|\psi_{12}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\rangle |\downarrow_2\rangle - |\downarrow_1\rangle |\uparrow_2\rangle). \quad (177)$$

Ora supponiamo che Alice applichi un operatore unitario alla particella in proprio possesso e consideriamo per semplicità il set di operatori $\{I, \sigma_x, \sigma_y, \sigma_z\}$, dove abbiamo indicato con I la matrice identità e con σ_i la i -esima matrice di Pauli. Quindi Alice effettuerà una delle seguenti operazioni:

$$I_1 \otimes I_2 |\psi_{12}\rangle = |\psi_{12}\rangle, \quad (178)$$

$$\sigma_{x_1} \otimes I_2 |\psi_{12}\rangle = \frac{1}{\sqrt{2}}(|\downarrow_1\rangle |\downarrow_2\rangle - |\uparrow_1\rangle |\uparrow_2\rangle), \quad (179)$$

$$\sigma_{y_1} \otimes I_2 |\psi_{12}\rangle = \frac{i}{\sqrt{2}}(|\downarrow_1\rangle |\downarrow_2\rangle + |\uparrow_1\rangle |\uparrow_2\rangle), \quad (180)$$

$$\sigma_{z_1} \otimes I_2 |\psi_{12}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\rangle |\downarrow_2\rangle + |\downarrow_1\rangle |\uparrow_2\rangle). \quad (181)$$

Una volta applicato l'operatore, Alice invia il sistema quantistico a Bob tramite un canale quantistico, praticamente realizzabile inviando un singolo fotone attraverso una fibra ottica. Una volta che Bob è in possesso della totalità della coppia EPR, che si trova in uno dei quattro stati sopra indicati, può effettuare una misura ortogonale e distinguere senza errore ognuno dei 4 stati, data la loro ortogonalità. Una volta appreso il risultato della misura, Bob ha acquisito due bit di informazione classica, date le quattro possibilità. Quindi il risultato complessivo è quello di aver acquisito due bit di informazione tramite l'invio di un qubit, ovvero l'esatto opposto del teletrasporto quantistico.

Il teletrasporto quantistico è stato introdotto teoricamente per la prima volta in un articolo del 1993 mentre i risultati della prima realizzazione pratica sono stati pubblicati nel 1997 [3]. Al momento la massima distanza alla quale sia stato teletrasportato uno stato quantistico è di circa 1400 km, come pubblicato nel 2017 [4], sfruttando un sistema di comunicazione tra la Terra e un satellite.

Il dense coding è stato introdotto in un articolo del 1992 e realizzato praticamente per la prima volta nel 1996 [14], come pubblicato da K. Mattle, H. Weinfurter, P. G. Kwiat e A. Zeilinger.

8 Conclusioni

Per concludere quest'opera, rivediamo il percorso che ci ha portato fin qui. Abbiamo iniziato introducendo la teoria dell'informazione classica di Shannon. In particolare ci siamo soffermati sul concetto di quantità di informazione e di entropia di Shannon. Successivamente abbiamo enunciato i teoremi di Shannon, che possono essere visti come i teoremi cardini della teoria dell'informazione. Poi abbiamo lasciato il mondo della fisica classica per passare allo studio dei sistemi aperti e dei canali in Meccanica Quantistica. Abbiamo introdotto e analizzato gli strumenti necessari per lo studio di questi ultimi. Entrando più nel dettaglio, è stata introdotta la matrice densità con tutte le sue caratteristiche. In seguito le conoscenze acquisite sono state utilizzate per affrontare problemi ancora aperti in Meccanica Quantistica. Nel dettaglio è stato trattato il problema della misura, il quale si articola a sua volta nel problema del collasso della funzione d'onda e della base preferita. Un passo verso la risoluzione del problema del collasso della funzione d'onda è stato fatto introducendo la teoria della decoerenza, la quale ha il merito di considerare anche l'ambiente nel quale è immerso il sistema fisico studiato. Successivamente è stata trattato il problema dell'incompletezza della Meccanica Quantistica. Tale problema, sollevato da Einstein, Podolsky e Rosen nel 1935 è stato parzialmente risolto da Bell nel 1964 il quale ha dimostrato che una teoria locale non può riprodurre gli stessi risultati della Meccanica Quantistica. Infine abbiamo introdotto la teoria dell'informazione quantistica, collegandoci a quanto visto all'inizio ma utilizzando le conoscenze acquisite. Però a cosa è dovuto questo interesse per l'informazione quantistica?

Questo campo è nato negli anni '70 quando un gruppo di scienziati, Helstrom, Gordon, Strattonovich e Holevo, hanno provato a trasmettere informazione classica attraverso sistemi ottici quantistici. Ma l'informazione trattata è sempre informazione classica, non ancora quantistica. Però solo nel 1982 Feynman discute della possibilità di creare calcolatori sfruttando sistemi quantistici allo scopo di creare simulazioni di altri sistemi quantistici. Le vere potenzialità di questa nuova area della fisica vengono intraviste nel 1994 quando Shor pubblica un algoritmo quantistico, per risolvere il problema della decomposizione in prodotto di numeri primi [27]. Un eventuale computer quantistico ideale potrebbe scomporre un numero n in un tempo polinomiale, ovvero in un tempo dell'ordine di $O(\ln(n)^3)$, quando il tempo necessario utilizzando algoritmi e calcolatori classici è di tipo superpolinomiale, ovvero maggiore di qualsiasi potenza di $\ln(n)$. Ciò avrebbe risvolti importantissimi a livello mondiale, in quanto la crittografia odierna si basa sull'incapacità di un computer di non riuscire a trovare tali dividendi in tempi rapidi. Anche il campo delle simulazioni delle evoluzioni di sistemi complessi potrebbe essere stravolto. In particolare si potrebbe studiare nel dettaglio il comportamento di grande molecole, con risvolti fondamentali in chimica e biologia.

Questi sono solo dei piccoli ma fondamentali esempi di come lo sviluppo della teoria dell'informazione quantistica e l'eventuale avvento di un computer quantistico possano essere una rivoluzione epocale.

Riferimenti bibliografici

- [1] Stephen Adler. “Why Decoherence has not Solved the Measurement Problem: A Response to P. W. Anderson”. In: *Studies in History and Philosophy of Modern Physics* (1948).
- [2] Charles Henry Bennett et al. “Teleporting an Unknown Quantum State via Dual Classical and Einstein–Podolsky–Rosen Channels”. In: *Physical Review Letters* (1993).
- [3] Dik Bouwmeester et al. “Experimental quantum teleportation ”. In: *Nature* (1997).
- [4] Pan Jianwei et al. “Ground-to-satellite quantum teleportation ”. In: *Nature* (2017).
- [5] Serge Haroche et al. “Observing the Progressive Decoherence of the “Meter” in a Quantum Measurement”. In: *Physical Review Letters* (1996).
- [6] Boris Podolsky e Nathan Rosen Albert Einstein. “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?” In: *Physical Review* (1935).
- [7] John Stewart Bell. “On the Einstein Podolsky Rosen paradox”. In: *Physics Physique Fizika* (1964).
- [8] Richard Jozsa e Benjamin Schumacher. “A New Proof of the Quantum Noiseless Coding Theorem”. In: *Journal of Modern Optics* (1993).
- [9] David Boh. *Quantum Theory*. Prentice-Hall, 1951. ISBN: 9780486659695.
- [10] Man-Duen Choi. “Completely Positive Linear Maps on Complex Matrices ”. In: *Linear Algebra and its Applications* (1974).
- [11] Alexander Semenovich Holevo. “Bounds for the quantity of information transmitted by a quantum communication channel”. In: *Problems of Information Transmission* (1973).
- [12] Michael A. Nielsen e Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. ISBN: 978-1107002173.
- [13] Andrzej Jamiolkowski. “Linear transformations which preserve trace and positive semi-definiteness of operators”. In: *Reports on Mathematical Physics* (1972).
- [14] Paul Kwiat e Anton Zeilinger Klaus Mattle Harald Weinfurter. “Dense Coding in Experimental Quantum Communication”. In: *Physical Review Letters* (1996).
- [15] Karl Kraus. *States, Effects, and Operations Fundamental Notions of Quantum Theory*. Springer, 1983. ISBN: 978-3-540-12732-1.
- [16] Göran Lindblad. “On the generators of quantum dynamical semigroups”. In: *Communications in Mathematical Physics* (1976).
- [17] David J. C. MacKay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003. ISBN: 978-0521642989.
- [18] Jonas Maziero. “Computing partial traces and reduced density matrices”. In: *International Journal of Modern Physics* (2017).
- [19] Mark Naimark. “Izv. Akad. Nauk SSSR”. In: *C. R. (Doklady) Acad. Sci. URSS (N.S.)* (1943).
- [20] John von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, 1932. ISBN: 9780691178561.
- [21] John Preskill. *Quantum information: Lecture Notes*. URL: <http://theory.caltech.edu/~preskill/ph219/index.html>.

- [22] Maximilian Schlosshauer. “Decoherence, the measurement problem, and interpretations of quantum mechanics”. In: *Reviews of Modern Physics* (2004).
- [23] Erhard Schmidt. “Zur Theorie der linearen und nichtlinearen Integralgleichungen”. In: *Mathematische Annalen* (1907).
- [24] Benjamin Schumacher. “Quantum coding”. In: *Physical Review* (1995).
- [25] Lucas Slot e Sebastian Zur. *Shannon’s Noisy-Channel Coding Theorem*. URL: https://homepages.cwi.nl/~schaffne/courses/infcom/2014/reports/Report_Lucas_Slot_Sebastian_Zur.pdf.
- [26] Claude Shannon. “A Mathematical Theory of Communication”. In: *The Bell System Technical Journal* (1948).
- [27] Peter Williston Shor. “Algorithms for quantum computation: discrete logarithms and factoring”. In: *IEEE* (1994).
- [28] Charles Henry Bennett e Stephen Wiesner. “Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states”. In: *Physical Review Letters* (1992).
- [29] William Forrest Stinespring. “Positive Functions on C*-Algebras”. In: *American Mathematical Society* (1955).
- [30] Andrzej Kossakowski e E. C. G. Sudarshan Vittorio Gorini. “Completely positive dynamical semigroups of N-level systems”. In: *Journal of Mathematical Physics* (1976).
- [31] Eugene Paul Wigner. “Remarks on the mind-body question”. In: *The Scientist Speculates* (1961).
- [32] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013. ISBN: 978-1107034259.
- [33] Wojciech Hubert Zurek. “Decoherence, Einselection, and the quantum origins of the classical”. In: *Physical Review* (2003).
- [34] Wojciech Hubert Zurek. “Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse?” In: *Physical Review* (1981).