
TRASCRIZIONE DEGLI APPUNTI DELLE LEZIONI DI

Istituzioni di Fisica teorica

TENUTE DAL PROF. *Pieralberto Marchetti*

PRESSO L'UNIVERSITÀ DI PADOVA

A CURA DI: *Francesco Manzali, Mattia Morgavi*

ANNO ACCADEMICO 2018-2019

Sorgente disponibile a <https://github.com/Einlar/FisicaTeorica>

Compilato il 18 febbraio 2020

Indice

1	Motivazione del formalismo	7
1.1	Interferenza degli elettroni	7
1.2	Polarizzazione dei fotoni	9
2	Descrizione matematica di un sistema fisico	16
2.1	Osservabili	16
2.1.1	L'interpretazione di Heisenberg	17
2.1.2	L'approccio di Schrödinger	20
2.2	Stati	21
2.2.1	Stati in Meccanica Classica	21
2.2.2	Stati in Meccanica Quantistica	22
2.3	Valor medio	23
2.3.1	Relazioni tra osservabili e stati	24
2.3.2	Complemento matematico: Teoria della misura	24
2.3.3	La struttura probabilistica	27
2.3.4	Autostati	31
2.4	Evoluzione temporale per sistemi isolati	33
3	Formulazione assiomatica Hilbertiana	35
3.1	Stati puri	35
3.1.1	Spazio degli stati puri e sue proprietà	35
3.1.2	Isomorfismi tra spazi di Hilbert	37
3.1.3	Equivalenza tra gli spazi della MQ	39
3.1.4	Funzionali lineari e spazio duale	40
3.1.5	Osservazioni sullo spazio \mathcal{S} degli stati puri in MQ	41
3.2	Osservabili	42
3.2.1	Prime proprietà: linearità e densità del dominio	42
3.2.2	Definizione mediante gli elementi di matrice	44
3.2.3	Valori medi reali: aggiunto e operatori simmetrici	45
3.2.4	Proiettori e operatori unitari	47
3.2.5	Autoaggiuntezza: uguaglianza dei domini di A e A^\dagger	48
3.2.6	Autoaggiuntezza in dimensione N : decomposizione spettrale	50
3.2.7	Autoaggiuntezza in dimensione ∞ : famiglie spettrali	55
3.2.8	Osservazioni sulle proprietà matematiche delle osservabili	62
3.3	Spettro di un operatore	67
3.3.1	Osservazioni sullo spettro	69

3.3.2	Spettro matematico	71
3.4	Il Formalismo di Dirac	73
3.4.1	Notazione e formalismo per lo spettro discreto	73
3.4.2	Estensione del formalismo allo spettro continuo	74
3.4.3	Formalismo per uno spettro generale	76
3.4.4	Spazio degli autovettori dello spettro continuo	77
3.4.5	Generalizzazione ad uno spazio equipaggiato	80
3.4.6	Vantaggi e svantaggi del formalismo di Dirac	82
3.5	L'operatore momento	84
3.5.1	Momento in \mathbb{R}	84
3.5.2	Momento in \mathbb{R}_+	85
3.5.3	Momento in $[0, 2\pi]$	86
3.6	Stati misti in MQ	91
3.7	Concetti di teoria dei gruppi	94
3.8	Evoluzione temporale	104
3.8.1	Formulazione di Heisenberg	104
3.8.2	Formulazione di Schrödinger	105
3.8.3	Formulazione di Dirac	105
3.8.4	Visuali di Schrödinger e Heisenberg	108
3.8.5	Buca infinitamente profonda in 1D	110
3.8.6	Calcolo di probabilità	117
3.8.7	Note sull'equazione di Schrödinger (Esercizio 1)	119
3.8.8	Teoria della misura	121
3.8.9	Differenze tra MC e MQ	126
3.9	Esempi di evoluzione di sistemi quantistici isolati	128
3.9.1	Particella libera in 1D	128
3.9.2	Buca infinitamente profonda in \mathbb{R}^3	131
3.9.3	Evoluzione di una particella in 1D con potenziale costante a tratti: caso generale	133
3.9.4	Il potenziale a gradino	134
3.9.5	Barriera di potenziale	148
3.9.6	Buca finitamente profonda	162
3.9.7	Regole generali per $\sigma(H)$ per potenziali in 1D (con al più salti)	168
3.9.8	Esercizio 2	170
4	Conseguenze degli assiomi	173
4.1	Principio di indeterminazione	173
4.1.1	Osservazioni sul principio di indeterminazione	176
4.2	Osservabili compatibili	178
4.3	Esercizio 3	187
4.3.1	Soluzione	187
4.4	Sistemi quantistici composti	197
4.5	Simmetrie e Algebre di Lie	205
4.5.1	Teoria dei gruppi e delle algebre di Lie	205
4.6	Le traslazioni spaziali	212

4.7	Momento angolare e rotazioni	213
4.8	Momento angolare generale	218
4.8.1	Gli autostati del momento angolare	226
4.8.2	Lo spin	235
4.8.3	Composizione dei momenti angolari	240
4.8.4	Esempio di calcolo dei coefficienti di Clebsh-Gordan	246
4.9	Esercizio 4	249
4.9.1	Soluzione	250
4.10	Oscillatore armonico	256
4.10.1	Esercizio 5	263
4.11	Potenziale centrale	269
4.11.1	Particella libera in coordinate sferiche	277
4.12	Stati legati per l'atomo di idrogeno	281
4.13	Esercizio 6	287
4.13.1	Soluzione	288
4.14	Scattering	293
4.14.1	Soluzioni stazionarie	295
4.14.2	Pacchetto d'onda incidente	296
4.14.3	Pacchetto d'onda diffuso	298
4.14.4	Approssimazione di Born per la sezione d'urto	304
4.14.5	Applicazione a potenziali noti	306
4.15	Particelle identiche	307
4.15.1	Il principio di esclusione di Pauli	313
4.15.2	Effetti in dimensioni inferiori	314
4.15.3	L'effetto di Hall quantistico	317
4.16	Potenziali periodici	322
4.16.1	Esercizio 7	324
4.16.2	Soluzione	326
4.17	Teoria delle perturbazioni	331
4.17.1	Conseguenze delle perturbazioni	337
4.18	Esercizio 8	343
4.18.1	Soluzione	344
4.19	Paradosso Einstein-Podolski-Rosen (1935)	349
4.19.1	Realismo e controfattualità	349
4.19.2	I principi del paradosso EPR	350
4.19.3	Versione classica di EPR	352
4.19.4	Versione quantistica di EPR	353
4.19.5	L'argomento di EPR	356
4.19.6	Risolvere il paradosso	357
4.19.7	Le disuguaglianze di Bell	359
4.19.8	Il gatto di Schrödinger	364
4.19.9	Riepilogo finale	366

Introduzione

Buonsalve!

In questo documento ho cercato di riordinare gli appunti del corso di Istituzioni di Fisica Teorica tenuto dal professor Pieralberto Marchetti presso il Dipartimento di Fisica dell'Università di Padova nel corso del primo semestre del 2018-19.

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

Prima di iniziare, ultimo disclaimer (che dovrebbe essere scontato dato che non ho una laurea): questi appunti non sono da intendere come sostituzione delle lezioni, o di altre dispense già presenti.

Buon viaggio! :)

Francesco Manzali, 23/10/2018

Watashi kininarimasu!
-Eru Chitanda, *Hyoka*

Aggiornamenti

Data	Aggiunte	Errata corrige	Commenti
23/10/2018	Prima pubblicazione		
29/10/2018	Lezione 25/10	Corretti errori da diverse segnalazioni	
29/10/2018	Lezione 26/10	Corretti altri errori/aggiunte immagini e plot	
6/11/2018	Lezione 5/11		
13/11/2018	Lezione 7/11	Corretta derivazione della buca infinita	
17/11/2018	Lezione 8/11		
23/11/2018	Lezioni 12-15-16/11	Svariate correzioni di refusi	Manca la lezione 9/11
22/12/2018	Lezioni 9,19-28/11	Taaanti errori corretti	
28/12/2018	Lezioni 29-30/11	Varie correzioni	
8/2/2019	Tutte lezioni tranne scattering	Apportate tutte le modifiche segnalate (200+ errori corretti)	
22/2/2019	Manca ancora lo scattering	Corretti conti buca finita, chiariti gli ordini dei coniugati per simmetrie ed evoluzione temporale. Corretti errori di battitura, parentesi, forma, aggiunti passaggi di chiarificazione	
18/2/2020	Inserito lo scattering		Mancano figure da tradurre in tikz

Tabella 1: Cronologia di modifiche/aggiornamenti agli appunti

Capitolo 1

Motivazione del formalismo

(Lezione 1 del
01/10/2018)

Lo scopo del corso è comprendere il formalismo della meccanica quantistica basandosi sulle conoscenze acquisite in fisica moderna di come i concetti si siano sviluppati storicamente, ed imparare ad operare su sistemi quantistici elementari. La meccanica quantistica mette in crisi concetti come l'onda e la particella, le osservabili come la posizione ed il momento, la misura e lo stesso concetto di "realtà fisica". Anche se risulta una teoria controintuitiva (se guardata in chiave classica) senza di essa il mondo risulterebbe inspiegabile (si pensi a fenomeni come l'impenetrabilità dei corpi, stabilità della materia ecc...). I fenomeni quantistici, meno evidenti di quelli classici ma altrettanto reali, sono conseguenze delle proprietà del mondo macroscopico come descritto dalla meccanica quantistica. Inoltre l'elettronica moderna senza meccanica quantistica risulterebbe impossibile, dunque la meccanica quantistica ha diversi risvolti anche sulla tecnologia.

Per motivare il **formalismo** che sarà utilizzato nelle prossime lezioni, esaminiamo qualitativamente due fenomeni peculiari che manifestano l'incompatibilità tra mondo classico e quantistico: l'interferenza degli elettroni e la misura di polarizzazione dei fotoni.

1.1 Interferenza degli elettroni

Sparando degli elettroni **uno alla volta** contro due fenditure, dietro le quali viene posto uno schermo rivelatore, osserviamo i seguenti fenomeni:

- I. Sullo schermo compare sempre un numero intero di e^- (come se fossero particelle).

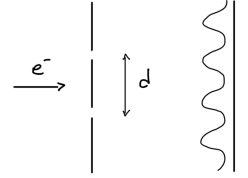
*Comportamento
particellare*

Con due fenditure aperte otteniamo una figura interferenza per il numero di elettroni rilevati sullo schermo alla posizione x , dato da $N_{12}(x)$. Tale figura scompare se chiudiamo una delle due fenditure, e in tal caso si osserva un picco in prossimità della fenditura rimasta aperta. Chiamiamo $N_1(x)$ e $N_2(x)$ le misurazioni eseguite in questi due casi.

- II. Si ha che $N_{12}(x) \neq N_1(x) + N_2(x)$, o, equivalentemente, dividendo per N (n. degli e^-) per ottenere delle probabilità, $p_{12}(x) \neq p_1(x) + p_2(x)$. In

*Comportamento
ondulatorio*

altre parole, la probabilità di trovare in x un elettrone quando entrambe le fenditure sono aperte *non* è la somma delle probabilità che si hanno nei due casi in cui solo una delle due fenditure è aperta.



Possiamo interpretare I e II utilizzando una funzione d'onda $\psi(x)$ complessa, introdotta da Schrödinger e interpretata in maniera probabilistica da Born. Il modulo quadro $|\psi(x)|^2$ dà infatti la **densità di probabilità** di trovare l'elettrone in posizione x .

In questo modo è soddisfatta I, e anche II, in quanto, analogamente alle onde:

$$\psi_{12}(x) = \psi_1(x) + \psi_2(x) \Rightarrow p_{12}(x) = |\psi_{12}(x)|^2 \neq |\psi_1(x)|^2 + |\psi_2(x)|^2 = p_1(x) + p_2(x)$$

Perciò gli elettroni sono “particelle” la cui probabilità di essere trovate con una misura in una data posizione x è data dal modulo quadrato della funzione d'onda $\psi(x)$.

Tale descrizione, seppur funzionale, è di difficile “intuizione fisica”. Per esempio, potremmo stimare la “lunghezza” di un elettrone e^- tramite la sua lunghezza d'onda Compton $\lambda = \hbar/(mc) \approx 10^{-12}$ m. Ci aspettiamo che l'elettrone “conosca” il mondo esterno a questa scala. Ma una figura di interferenza compare anche quando la distanza tra le fenditure è molto maggiore, sulla scala di 10^{-5} m. Come può l'elettrone, giunto in prossimità di una fenditura, “sapere” se l'altra è aperta o meno ed eventualmente “produrre” la figura di interferenza?

Per ogni singolo elettrone è possibile determinare da *quale* delle due fenditure è passato, per esempio predisponendo un sistema di illuminazione al di là della fenditura, che segnali il suo eventuale passaggio. Modificando così l'apparato sperimentale, tuttavia, la figura di interferenza scompare, e si ha che $N_{12}^{\text{ill}} = N_1(x) + N_2(x)$, e quindi $p_{12}^{\text{ill}} = p_1(x) + p_2(x)$.

Il formalismo deve quindi anche spiegare che:

III. La possibilità di conoscere la posizione *intermedia* di una particella ne modifica il comportamento

L'atto di misura modifica il comportamento delle particelle

La spiegazione “euristica” di questa situazione è che la luce è composta da particelle quantistiche, i fotoni, e quando questi urtano gli elettroni liberi (effetto Compton) li disturbano¹ modificandone l'impulso.

Se la luce fosse fatta di onde potremmo ridurre il disturbo riducendo l'intensità dell'onda, ma se osservo l'elettrone vuol dire che un fotone l'ha colpito (e quindi si sta comportando come particella).

Possiamo allora minimizzare l'effetto riducendo il momento del γ , ma così facendo, per de Broglie, $p = h/\lambda$, e quindi ridurre p significa aumentare la lunghezza d'onda λ , e così si perde l'informazione su quale fenditura è stata utilizzata per passare. Sperimentalmente, quando λ del fotone usato per l'illuminazione è dell'ordine della distanza tra le fenditure, gli elettroni “non sono disturbati” e producono la figura di interferenza.

Come vedremo, questa “spiegazione” non è completamente soddisfacente, ma almeno salva la meccanica quantistica da un problema di autoconsistenza.

¹ Rimane aperta la questione: come è possibile che questo “disturbo” sia esattamente tale da ricostruire la situazione classica, e cioè da rendere $p_{12}(x) = p_1(x) + p_2(x)$

Infatti, senza la doppia natura ondulatorio-corpuscolare delle particelle coinvolte i fenomeni descritti non sarebbero possibili.

Notiamo che per gli elettroni con funzione d'onda $\psi_{12}(x) = \psi_1(x) + \psi_2(x)$ nessuna delle seguenti affermazioni è corretta:

1. L'elettrone passa dalle fenditure 1 o 2.
Esclusa perché $p_{12}(x) \neq p_1(x) + p_2(x)$.
2. L'elettrone passa da entrambe le fenditure.
Esclusa perché se cerchiamo di verificare da quale fenditura è passato la troviamo in una sola (e non appena lo sappiamo l'interferenza scompare).
3. L'elettrone arriva allo schermo senza passare dalle fenditure.
Falsa, poiché se chiudiamo le fenditure non arrivano elettroni allo schermo.

Da queste considerazioni vediamo che non è affatto ovvio poter assumere che l'elettrone “abbia” una posizione (intesa come un valore numerico) se non lo osserviamo, ma solo che lo troviamo in una posizione (come valore) se eseguiamo una misura, con una probabilità determinata da $\psi_{12}(x)$.

Possiamo rendere più preciso il concetto di “disturbo”. Se conosciamo la posizione ad un istante introduciamo un'indeterminazione nel momento (allo stesso istante), e di conseguenza, con l'evoluzione temporale, anche un'indeterminazione nelle posizioni agli istanti successivi.

Es. camera a nebbia: ogni “punto” visibile è dato dall'interazione tra una particella e una delle goccioline nella camera, e costituisce una misurazione (affetta da incertezza) della sua posizione, che introduce un'indeterminazione in quelle successive. Perciò alla fine non si misura la traiettoria effettiva della particella, ma una successione di punti non completamente “allineati lungo una curva”. In ogni caso non è considerabile una traiettoria, in quanto per le indeterminazioni non è possibile capire cosa succede tra un punto e l'altro.

Possiamo riassumere ciò in un altro punto che deve essere soddisfatto dal formalismo utilizzato:

- IV. È impossibile determinare simultaneamente posizione e momento di una particella quantistica con precisione maggiore di per il prodotto delle indeterminazioni.

Principio di indeterminazione di Heisenberg

Scopo della prima parte del corso sarà costruire un formalismo matematico consistente che sia in grado di spiegare le proprietà da I a IV o l'analogo per la polarizzazione dei fotoni che ora discutiamo.

1.2 Polarizzazione dei fotoni

Lo stato dell'elettrone nell'esempio precedente è descritto da una funzione d'onda $\psi(x)$, elemento di uno spazio vettoriale infinito-dimensionale (tipico delle onde). Esistono tuttavia sistemi quantistici con spazi degli stati finito-dimensionali.

Un esempio istruttivo e familiare nella sua versione classica è quello che descrive la **polarizzazione** di un'onda elettromagnetica, classicamente, e dei fotoni che la compongono, quantisticamente.

Seguendo Dirac, partendo dal comportamento classico delle onde polarizzate con $N \gg 1$ fotoni, cercheremo di estrapolare il comportamento quantistico del singolo fotone.

Un'onda elettromagnetica classica *piana* ideale (estensione infinita), di pulsazione $\omega = 2\pi/T$ e vettore d'onda di modulo $k = \omega/c$ che punta lungo la direzione di propagazione \hat{z} , è data da un campo elettromagnetico oscillante:

$$\begin{cases} E_x(\vec{x}, t) = \text{Re}(E_{0x}e^{i(kz - \omega t + \alpha_x)}) = E_{0x} \cos(kz - \omega t + \alpha_x) \\ E_y(\vec{x}, t) = \text{Re}(E_{0y}e^{i(kz - \omega t + \alpha_y)}) = E_{0y} \cos(kz - \omega t + \alpha_y) \\ E_z(\vec{x}, t) = 0 \\ c\vec{B}(\vec{x}, t) = \hat{z} \times \vec{E}(\vec{x}, t) \end{cases}$$

con E_{0x} e E_{0y} positivi, costanti e indipendenti da \vec{x} e t , e α_x, α_y sono parametri che regolano la *fase* dei rispettivi campi.

Unendo l'informazione su intensità e fase lungo gli assi x e y (che è l'unica necessaria per definire la polarizzazione dell'onda) definiamo $E_j = E_{0j}e^{i\alpha_j} \in \mathbb{C}$. Perciò E_x e E_y così definiti (complessi) sono le due componenti del vettore polarizzazione ψ :

$$\begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = \begin{pmatrix} E_x \\ E_y \end{pmatrix}$$

che d'ora in poi indicheremo seguendo la notazione di Dirac come $|\psi\rangle$. La polarizzazione di un'onda gioca un ruolo analogo alla posizione (data tramite la funzione d'onda) di una particella.

Analizziamo meglio $|\psi\rangle$. Sappiamo che un'onda che ha una *polarizzazione lineare* di angolo α rispetto a \hat{x} può essere vista come la sovrapposizione di due onde di *stessa fase* che hanno polarizzazioni lungo \hat{x} e \hat{y} . Se l'onda iniziale ha ampiezza E_0 , le due onde “proiettate” avranno ampiezza $E_{0x} = E_0 \cos \alpha$ e $E_{0y} = E_0 \sin \alpha$.

Perciò le polarizzazioni “possono essere sovrapposte”, cioè sommate, e quindi $|\psi\rangle$ è un elemento di uno spazio vettoriale complesso, che sarà isomorfo a \mathbb{C}^2 in quanto $|\psi\rangle$ è definito da due coordinate complesse.

Nota: Abbiamo scelto di utilizzare le coordinate in \mathbb{C} per includere in E_x e E_y anche l'informazione sullo sfasamento. Per esempio, la sovrapposizione di due onde di polarizzazione ortogonale ma *sfasate* di $\pi/2$ produce la *polarizzazione ellittica*. Data questa premessa, possiamo utilizzare per $|\psi\rangle$ gli “strumenti” disponibili su \mathbb{C}^2 . Per esempio su \mathbb{C}^2 è definito un prodotto scalare, che denotiamo, seguendo la notazione di Dirac, come $\langle \cdot | \cdot \rangle$, cioè: $\langle \phi | \psi \rangle = \phi_x^* \psi_x + \phi_y^* \psi_y$, con $\langle \phi |$ il funzionale lineare su \mathbb{C}^2 dato da $\langle \phi | : |\psi\rangle \rightarrow \langle \phi | \psi \rangle \in \mathbb{C}$.

Notiamo quindi che il vettore $|\psi\rangle$ contiene informazioni ridondanti. In particolare non ci serve sapere il modulo del campo elettrico - ci interessa solo la direzione di oscillazione. Perciò potremmo fissarlo a 1, imponendo:

(Lezione 2 del
03/10/2018)

Vettore di
polarizzazione

Polarizzazione
lineare

$|\psi\rangle \in \mathbb{C}^2$

Polarizzazione
ellittica

I riduzione di \mathbb{C}^2

$$\|\vec{E}\| = \sqrt{|E_x|^2 + |E_y|^2} = 1$$

Scegliamo quindi $|\psi\rangle$ tale che $\langle\psi|\psi\rangle = 1$ (norma unitaria).

Agendo nella base canonica di \mathbb{C}^2 , alcuni esempi di diverse polarizzazioni (e relativi vettori) sono:

Esempi di polarizzazione

- **Polarizzazione lineare**, che può essere lungo \hat{x} se $|\psi\rangle = |x\rangle \equiv (1, 0)^t$, oppure lungo \hat{y} se $|\psi\rangle = |y\rangle \equiv (0, 1)^t$. Oppure può avere, ad esempio, una polarizzazione a 45° rispetto a \hat{x} se $|\psi\rangle = |45^\circ\rangle = \frac{1}{\sqrt{2}}(|x\rangle + |y\rangle)^t$;
- **Polarizzazione circolare**, che può essere destrorsa se $|\psi\rangle = |R\rangle = \frac{1}{\sqrt{2}}(1, i)$, oppure sinistrorsa se $|\psi\rangle = |L\rangle = \frac{1}{\sqrt{2}}(1, -i)^t$.

In realtà imporre $\|\psi\| = 1$ per rimuovere informazioni ridondanti restringe lo spazio considerato da \mathbb{C}^2 a $\frac{\mathbb{C}^2 \setminus \{0\}}{\mathbb{R}_+} \cong S^3$.

Matematicamente, questo è uno *spazio vettoriale quoziente*, ottenuto “identificando” tutti i punti che soddisfano una certa condizione, che in questo caso è “il differire solamente per un fattore di scala”. In maniera pittoresca lo spazio quoziente si può pensare (in questo caso) come il prodotto di “incollare” tra loro tutti i vettori che “puntano nella stessa direzione”, ma hanno diverse “lunghezze”. In tal modo si passa da una situazione (in \mathbb{C}^2) in cui ho infiniti modi di scrivere la stessa polarizzazione, a una (nello spazio quoziente) in cui ad ogni punto è associata una sola “direzione” di polarizzazione (non ancora una sola polarizzazione - c’è un’altra ambiguità da rimuovere che tratteremo tra poco).

Spazio vettoriale quoziente

Formalizziamo quanto appena detto in termini matematici. La “condizione” è data da una *relazione di equivalenza*. Detti $x, y \in I$, con I un insieme qualsiasi, una relazione di equivalenza, indicata con $x \sim y$, è una qualsiasi condizione che sia *riflessiva* (se $x \sim x$ sempre), *simmetrica* (se $x \sim y$ allora $y \sim x$) e *transitiva* (se $x \sim y$ e $y \sim z$ allora $x \sim z$).

Relazione di equivalenza

Data una relazione di equivalenza su un insieme, si possono costruire le *classi di equivalenza*, ossia i sottoinsiemi $[x] = \{y \in I | y \sim x\}$. L’insieme di tutte le classi di equivalenza è lo spazio quoziente stesso, indicato con I / \sim .

Nel nostro caso la relazione di equivalenza è data da \mathbb{R}_+ così definita:

$$|\phi\rangle \sim |\psi\rangle \quad \Leftrightarrow \quad \frac{|\psi\rangle}{\|\psi\|} = \frac{|\phi\rangle}{\|\phi\|}$$

Nota: dobbiamo rimuovere 0 da \mathbb{C}^2 , in quanto il vettore nullo non definisce alcuna polarizzazione. Se non lo facessimo, inoltre, la relazione di equivalenza non sarebbe ben definita, dato che dividiamo per la norma.

Possiamo poi specificare ogni classe di equivalenza prendendo un suo rappresentativo (ossia un suo elemento qualsiasi). Per convenienza prenderemo, per ogni classe “direzione dei vettori di $\mathbb{C}^2 \setminus \{0\}$ ”, l’unico vettore che “punta in quella direzione” che ha anche norma unitaria.

Fissando la norma a 1 lo spazio è quindi ristretto dal seguente vincolo:

$$|\psi_x|^2 + |\psi_y|^2 = 1 \quad \Rightarrow \quad (\operatorname{Re} \psi_x)^2 + (\operatorname{Im} \psi_x)^2 + (\operatorname{Re} \psi_y)^2 + (\operatorname{Im} \psi_y)^2 = 1$$

Lo spazio risultante è perciò isomorfo a S^3 , ossia a una 3-sfera². In realtà, anche fissando il modulo del campo elettrico, non c'è ancora una corrispondenza biunivoca tra le polarizzazioni “fisiche” e i vettori $|\psi\rangle \in (\mathbb{C}^2 \setminus \{0\})/\mathbb{R}_+$. O meglio, c'è nel caso ci limitassimo solo alle polarizzazioni lineari, mentre per quelle “che ruotano” c'è ancora un'informazione ridondante: abbiamo specificato due fasi (α_x e α_y), ma quella che conta è la loro differenza. Potremmo quindi fissare una delle due e specificare solamente l'altra.

II riduzione

Definendo un'opportuna relazione di equivalenza S^1 (dati due vettori (E_x, E_y) e (E'_x, E'_y) , se E'_x e E'_y sono semplicemente la rotazione - sul piano complesso \mathbb{C} - di E_x e E_y , allora i due vettori sono equivalenti, poiché la “differenza di fase” è la stessa in entrambi³), e considerando il rispettivo spazio quoziente che la incorpori (oltre a quella sulla norma): $\mathcal{S}_{\text{pol}} \cong (\mathbb{C}^2 \setminus \{0\})/(\mathbb{R}_+ \times S^1) \cong S^2$ otteniamo finalmente lo spazio delle polarizzazioni, detto **sfera di Bloch**.

Sfera di Bloch

L'isomorfismo con S^2 si nota scrivendo il vincolo precedente:

$$(\text{Re } \psi_x)^2 + (\text{Im } \psi_x)^2 + (\text{Re } \psi_y)^2 + (\text{Im } \psi_y)^2 = 1$$

e fissando una fase, ponendo, per esempio, $\text{Im } \psi_x = 0$. Si ottiene così una forma del tipo $x^2 + y^2 + z^2 = 1$, corrispondente alla comune sfera in \mathbb{R}^3 (detta S^2). Un punto nella sfera di Bloch corrisponde ad una e una sola polarizzazione “fisica” di un fotone.

Nota: la sfera di Bloch **non** è uno spazio vettoriale! Per questo, nei conti che faremo, continueremo ad usare $|\psi\rangle \in \mathbb{C}^2$, nonostante essi contengano un sacco di informazione ridondante.

La coppia $\{|x\rangle, |y\rangle\}$ costituisce una base ortonormale in \mathbb{C}^2 , cioè un insieme di stati ortogonali ($\langle x | y \rangle = 0$) tra loro e normalizzati ($\langle x | x \rangle = \langle y | y \rangle = 1$) tali che ogni vettore di \mathbb{C}^2 può essere scritto come un'unica combinazione lineare dei due elementi:

Basi ON

$$(\psi_x, \psi_y) = |\psi\rangle = \psi_x |x\rangle + \psi_y |y\rangle \quad (1.1)$$

I coefficienti della rappresentazione nella base ortonormale sono dati da opportuni funzionali $\langle x |$ e $\langle y |$:

$$\psi_x = \langle x | \psi \rangle; \quad \psi_y = \langle y | \psi \rangle$$

La stessa decomposizione si può effettuare, naturalmente, rispetto ad una qualsiasi altra base, per esempio rispetto a $\{|R\rangle, |L\rangle\}$

$$|\psi\rangle = \psi_R |R\rangle + \psi_L |L\rangle; \quad \psi_R = \langle R | \psi \rangle; \quad \psi_L = \langle L | \psi \rangle$$

Tutto ciò corrisponde, fisicamente, alla possibilità di generare un'onda con una data polarizzazione (qualsiasi) sovrapponendone due di polarizzazioni opportune. Ciò è possibile perché il campo elettrico verifica il **principio di sovrapposizione** (il campo totale in un punto è la somma vettoriale dei campi generati da ciascuna sorgente).

Principio di sovrapposizione

Unendo le scritture (1.1) e (3) si ottiene:

Completezza di Dirac

²Una n -sfera è definita come $S^n = \{x \in \mathbb{R}^{n+1} : \|x\| = r\}$

³Matematicamente, $E_x = ae^{i\varphi}$, $E_y = be^{i\varphi}$ sono equivalenti a $E'_x = ae^{i\varphi+\alpha}$, $E'_y = be^{i\varphi+\alpha}$, per $\alpha \in S^1$, ossia α che va da 0 a 2π per poi ripartire da 0 e “continuare a girare”

$$|\psi\rangle = \langle x | \psi \rangle |x\rangle + \langle y | \psi \rangle |y\rangle = |x\rangle \langle x | \psi \rangle + |y\rangle \langle y | \psi \rangle$$

Possiamo interpretare ora $|x\rangle \langle x| + |y\rangle \langle y|$ come un operatore che “non cambia il risultato”: la sua azione è quindi l'**identità**. Tale risultato è detto **completezza di Dirac** relativa ad una base ortonormale. Vale cioè, per ogni base ON, per esempio quella data da $\{|x\rangle, |y\rangle\}$, o da $\{|R\rangle, |L\rangle\}$:

$$\mathbb{I} = |x\rangle \langle x| + |y\rangle \langle y|; \quad \mathbb{I} = |R\rangle \langle R| + |L\rangle \langle L|$$

Ciò è utile per passare da una base ad un'altra:

Cambio di base

$$\begin{aligned} \psi_x &= \langle x | \psi \rangle = \langle x | \mathbb{I} | \psi \rangle = \langle x | (|R\rangle \langle R| + |L\rangle \langle L|) | \psi \rangle \\ &= \langle x | R \rangle \langle R | \psi \rangle \psi_R + \langle x | L \rangle \langle L | \psi \rangle \psi_L \end{aligned}$$

Tentiamo ora di derivare il comportamento quantistico dei fotoni in presenza di polarizzazione come limite “ad un fotone” del comportamento delle onde elettromagnetiche.

Supponiamo di avere un filtro polarizzatore lungo \hat{x} , la teoria elettromagnetica classica ci dice che, se facciamo incidere su di esso un'onda con intensità $\|\vec{E}\|^2$, dopo il passaggio dal polarizzatore l'intensità si sarà ridotta a una frazione $|E_x|^2 / \|\vec{E}\|^2$ e presenterà esclusivamente la componente polarizzata lungo \hat{x} .

Filtro polarizzatore: interpretazione quantistica

In meccanica quantistica possiamo ridurre l'intensità fino a che non passi un singolo fotone alla volta attraverso il polarizzatore. Non è ovviamente possibile definire una “frazione di un fotone”, quindi il rapporto $|E_x|^2 / \|\vec{E}\|^2 = |\psi_x|^2$ (se $\langle \psi | \psi \rangle = 1$ non può essere interpretato come “frazione di fotone”.

L'unico modo di interpretare il risultato classico in termini di singoli fotoni è perciò che $|\psi_x|^2 = |\langle x | \psi \rangle|^2$ dia la probabilità che un fotone nella polarizzazione iniziale $|\psi\rangle$ passi attraverso il polarizzatore e sia poi quindi in polarizzazione $|x\rangle$ (e ciò costituisce una *misura* di polarizzazione).

Con questa interpretazione si giustifica la terminologia $|\langle x | \psi \rangle|^2 =$ probabilità di transizione da $|\psi\rangle$ a $|x\rangle$ e $\langle x | \psi \rangle =$ ampiezza di probabilità di transizione da $|\psi\rangle$ a $|x\rangle$.

$|\langle x | \psi \rangle|^2$ è l'analogo (per la polarizzazione) di $|\psi(x)|^2$ (per la posizione di una particella) di Schrödinger.

Attenzione! $|\psi_x|^2$ non dà la probabilità che il fotone sia, cioè indipendentemente dalla misura, in polarizzazione $|x\rangle$ prima della misura, ma solo la probabilità di essere trovato in polarizzazione $|x\rangle$ se eseguiamo una misura di polarizzazione.

Le probabilità hanno senso solo dopo una misura

Questa differenza è evidenziata sperimentalmente, per esempio utilizzando dei cristalli birifrangenti.

Il fenomeno della birifrangenza, che si verifica in cristalli come la calcite, consiste nella divisione di un raggio incidente in due raggi sfalsati spazialmente, la cui polarizzazione è ortogonale. Per semplicità consideriamo la polarizzazione del primo raggio uscente pari a $|x\rangle$ e l'altra pari a $|y\rangle$.

Birifrangenza

Ciò avviene perché la struttura anisotropica del cristallo fa sì che onde polarizzate **parallelamente** all'asse ottico proprio del cristallo producano “**raggi straordinari**”, che *non obbediscono* a Snell. Nell'esempio indichiamo la polarizzazione di

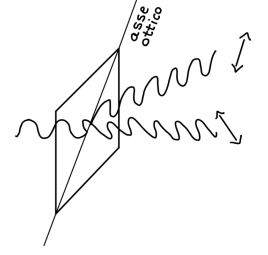
tali raggi con $|y\rangle$.

D'altro canto, onde polarizzate **perpendicolarmente** all'*asse ottico* producono “**raggi ordinari**” che *obbediscono* alle leggi di Snell.

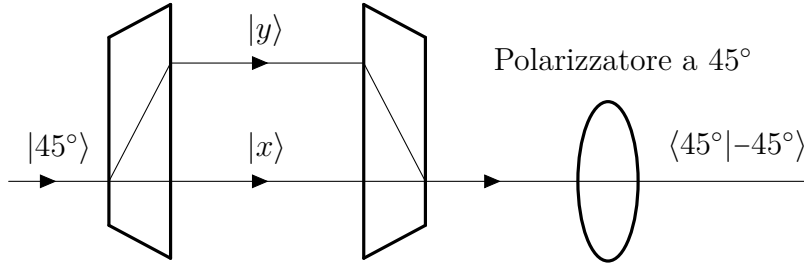
È poi possibile utilizzare un cristallo di forma speculare per “ricomporre” i due raggi prodotti da un primo cristallo e ricostruire un'onda unica, identica - nei limiti sperimentali - a quella originaria.

Setup dell'esperimento Lungo il cammino ottico di un fascio luminoso in luce polarizzata a 45° rispetto a \hat{x} si pongono due cristalli di calcite (birifrangenti) speculari. Il primo scompone il fascio in due raggi, di polarizzazioni $|x\rangle$ e $|y\rangle$, che vengono poi ricomposti dal secondo in un unico fascio, che infine incide contro un filtro polarizzatore tale da lasciar passare solo le onde con polarizzazione a $|-45^\circ\rangle$ rispetto a \hat{x} .

Ci aspettiamo che dal filtro non passi nulla, poiché $\langle 45^\circ | -45^\circ \rangle = 0$ (la polarizzazione della luce incidente è esattamente ortogonale a quella che lascerebbe passare il filtro, quindi non passa niente), e infatti ciò si verifica sperimentalmente.



Esperimento sulla polarizzazione



Proviamo a verificarlo analiticamente. Supponiamo che ogni fotone abbia sempre una polarizzazione ben definita, che venga eventualmente modificata dalle interazioni.

Allora, se un raggio è in polarizzazione $|y\rangle$ il polarizzatore a -45° ne lascia passare una frazione:

$$|\langle y | -45^\circ \rangle|^2 = \left| \left\langle y \left| \frac{1}{\sqrt{2}}(|x\rangle - |y\rangle) \right. \right\rangle \right|^2 = \frac{1}{2}$$

e analogamente $|\langle x | -45^\circ \rangle|^2 = \frac{1}{2}$.

Esaminiamo ora l'incidenza del fascio iniziale contro il primo cristallo. Partendo da $|45^\circ\rangle$ la probabilità di passare alla polarizzazione $|x\rangle$ è $|\langle 45^\circ | x \rangle|^2 = \frac{1}{2}$, e analogamente nel caso del passaggio a $|y\rangle$: $|\langle 45^\circ | y \rangle|^2 = \frac{1}{2}$.

Ma se $|\langle 45^\circ | x \rangle|^2$ desse effettivamente la probabilità che il fotone sia nello stato $|x\rangle$ dopo il primo cristallo, allora la probabilità che passi il polarizzatore è:

$$|\langle 45^\circ | x \rangle|^2 |\langle x | -45^\circ \rangle|^2 = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

e analogamente per $|y\rangle$.

Se il fotone ha una polarizzazione sempre ben definita, allora tra i due cristalli essa sarà o $|x\rangle$ o $|y\rangle$, e le due possibilità sono mutualmente esclusive. Perciò la probabilità che un fotone superi il filtro polarizzatore sarà data dalla somma delle

probabilità dei possibili modi (mutualmente esclusivi) che lo portano a superarlo. Otteniamo quindi:

$$\mathcal{P} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \neq 0$$

che contrasta con il risultato sperimentale.

Proviamo a svolgere lo stesso conto utilizzando la completezza di Dirac:

Risultati con il formalismo di Dirac

$$\begin{aligned} 0 &= |\langle 45^\circ | \mathbb{I} | -45^\circ \rangle|^2 = |\langle 45^\circ | x \rangle \langle x | -45^\circ \rangle + \langle 45^\circ | y \rangle \langle y | -45^\circ \rangle|^2 = \\ &= |\langle 45^\circ | x \rangle \langle x | -45^\circ \rangle|^2 + |\langle 45^\circ | y \rangle \langle y | -45^\circ \rangle|^2 + (\text{termini misti}) \end{aligned}$$

dove i termini misti sono dati dal doppio prodotto (termine di interferenza). In questo caso, svolgendo i conti, si scopre che l'interferenza è completamente distruttiva, e ciò spiega il risultato finale.

Ciò significa che se non osservo il risultato intermedio tra i cristalli le alternative $|x\rangle$ e $|y\rangle$ classicamente mutualmente esclusive (sono percorsi spazialmente separati!) interferiscono tra di loro e quindi non possiamo dire che il fotone tra i cristalli abbia polarizzazione $|x\rangle$ o $|y\rangle$. Ma allora la probabilità quantistica di eventi intermedi che sono mutualmente esclusivi classicamente, se non osservati, non soddisfa la regola della probabilità classica.

Più precisamente, se partiamo da un insieme I e andando in un insieme $\{m\}$ di alternative mutualmente esclusive che portano tutte ad uno stato f , classicamente la probabilità di andare da I a F è

$$\mathcal{P}_{I \rightarrow F} = \sum_m \mathcal{P}_{I \rightarrow m} \mathcal{P}_{m \rightarrow F}$$

Quantisticamente, tuttavia, non si sommano le probabilità, ma le ampiezze (complesse) di probabilità, per poi calcolare il modulo quadro per ottenere la probabilità vera e propria.

La ragione per cui è possibile violare la legge di composizione delle probabilità classiche è nuovamente che **se non osserviamo la posizione del fotone, tale grandezza non assume valori.**

Capitolo 2

Descrizione matematica di un sistema fisico

(**DISCLAIMER:** Questo appunto contiene ricerca personale, e quindi potrebbero esserci scritte cavolate. Attenzione!)

Ci occupiamo ora di introdurre il **formalismo matematico** che utilizzeremo per descrivere i **sistemi fisici** in Meccanica Quantistica. In particolar modo, ci servirà definire gli **enti fisici** (osservabili e stati) che caratterizzano il sistema, e le **regole** a cui obbediscono (valor medio ed evoluzione temporale). Ci concentreremo sul perché la MQ renda necessario ridefinire cose apparentemente *ovvie*, lasciando in sospeso alcune importanti questioni che saranno opportunamente riprese e concluse nel prossimo capitolo.

(Lezione 3 del
04/10/2018)

2.1 Osservabili

Si definiscono **osservabili**¹ le quantità fisiche O che possiamo misurare in linea di principio (ad un certo istante), come ad esempio *posizione, impulso, energia*.

Osservabile

Lo **spettro** $\sigma(O)$ dell'osservabile O è l'insieme dei valori che possiamo ottenere con una misura di O . Ovviamente deve essere $\sigma(O) \subseteq \mathbb{R}$, in quanto lo spettro di un osservabile deve essere confrontabile con l'esperimento, e le misure sono numeri reali.



In **Meccanica Classica** (MC), nel formalismo hamiltoniano, ad un sistema sono associate le coordinate $q = \langle q_1, \dots, q_n \rangle$ (posizioni) e i momenti coniugati $p = \langle p_1, \dots, p_n \rangle$, dove n è il numero di gradi di libertà. Lo spazio delle coppie (q, p) è detto **spazio delle fasi** e lo denoteremo con Ω .

Le osservabili sono *funzioni reali* su Ω e formano un'algebra abeliana su \mathbb{R} , per addizione e moltiplicazioni puntuali (ossia, se ho un'osservabile $f(q, p)$ e una $g(q, p)$ esiste sempre l'osservabile-prodotto $(fg)(q, p) = f(q, p)g(q, p)$ e, essendo un'algebra abeliana, vale la proprietà commutativa $f(q, p)g(q, p) = g(q, p)f(q, p)$).

¹ \wedge Sostantivo femminile

Tale struttura algebrica appena definita è necessaria perché osservabili che contengono prodotti per scalari, somme e prodotti di altre osservabili siano ben definiti. Un esempio è l'energia: $O = \mathcal{H}(q, p) = \frac{1}{2m}p \cdot p + V(q)$.

In MC, inoltre, lo spettro $\sigma(f(q, p))$ è semplicemente il codominio della funzione f , ossia i valori che f può assumere. Questa proprietà, intuitiva e tipica della MC, non segue però dalla definizione di osservabile che abbiamo dato (non è detto che l'insieme dei valori ottenuti dalla misura siano gli stessi che può assumere l'osservabile).

Riepilogando, in MC:

1. Le osservabili sono funzioni **reali**
2. Lo spettro di un'osservabile è il **codominio** della funzione ad esso associata
3. Le osservabili si possono sommare tra loro, dando luogo ad altre osservabili, e anche moltiplicare con **commutatività**

Tutte queste proprietà, che sembrano estremamente ovvie, non necessiterebbero nemmeno di essere enunciate se non fosse che in Meccanica Quantistica (MQ) vengono sistematicamente violate, come è stato verificato sperimentalmente.

2.1.1 L'interpretazione di Heisenberg

La prima intuizione in tal senso fu di Heisenberg (1925), partendo dalle regole della **spettroscopia** dell'atomo di idrogeno, per cui la frequenza delle righe di emissione è data dal principio combinatorio di Ritz-Rydberg:

$$\frac{1}{\lambda} = R \left(\frac{1}{n^2} - \frac{1}{m^2} \right); \quad n, m \in \mathbb{N}$$

Ogni frequenza ω può essere quindi indicizzata da due numeri m, n come ω_{mn} .

Ci si accorse di questa possibilità osservando che non tutte le frequenze facevano parte dello spettro dell'idrogeno (come ci si aspetterebbe in fisica classica), e quelle permesse avevano la proprietà di poter essere scritte sempre come differenza di numeri che non erano frequenze: $\omega_{mk} = \omega_m - \omega_k$.

Questo fatto è il prodotto della quantizzazione dell'energia degli elettroni: quando questi effettuano transizioni tra stati alti e stati bassi con energie fissate emettono un fotone di una determinata frequenza.

In particolare, notiamo che la somma di due frequenze (osservabili) non sempre corrisponde ad un'altra frequenza dello spettro. Più precisamente, la somma di due frequenze ω_{mk} e ω_{rs} porta ad una nuova frequenza permessa solo se $k = r$ o $m = s$, poiché in tal caso si effettua una "cancellazione", visto che è possibile scomporre le frequenze come:

$$\omega_{mk} + \omega_{rs} = \omega_m - \omega_k + \omega_r - \omega_s$$

e sappiamo che le frequenze "accettate" sono differenze del tipo $\omega_i - \omega_j$. Se $k = r$, per esempio, avremo che $\omega_{mk} + \omega_{ks} = \omega_m - \omega_s = \omega_{ms}$.

Un qualsiasi formalismo adottato per la MQ deve poter giustificare in maniera

naturale tale fenomeno.

Heisenberg partì considerando una grandezza osservabile $q(t)$ periodica con frequenza fondamentale $\omega = 2\pi\nu$ (che potrebbe per esempio essere la posizione di un oscillatore armonico²), e la sua rappresentazione in serie di Fourier³:

$$q(t) = \dots + q_{-2}e^{-i2\omega t} + q_{-1}e^{-i\omega t} + q_0 + q_1e^{i\omega t} + q_2e^{i2\omega t} + \dots$$

Per prima cosa dobbiamo assicurarci che, essendo $q(t)$ un'osservabile, sia reale. Per permettere questo una condizione sufficiente è $q_n = q_{-n}^* \forall n \in \mathbb{Z} \setminus \{0\}$, cioè che se un coefficiente di Fourier $q_n = u_n + iv_n$, il suo “corrispondente” sia il suo complesso coniugato: $q_{-n} = q_n^* = u_n - iv_n$. Così facendo i termini immaginari di ogni coppia si cancellano:

$$\begin{aligned} q_n e^{in\omega t} + q_{-n} e^{-in\omega t} &= (u_n + iv_n) e^{in\omega t} + (u_n - iv_n) e^{-in\omega t} = \\ &= (u_n + iv_n)(\cos(n\omega t) + i \sin(n\omega t)) + (u_n - iv_n)(\cos(n\omega t) - i \sin(n\omega t)) = \\ &= 2u_n \cos(n\omega t) - 2v_n \sin(n\omega t) \end{aligned}$$

e il risultato $q(t)$ sarà reale.

È facile verificare che, se $a(t)$ e $b(t)$ rispettano questa proprietà, allora anche le loro somme la rispettano, e anche le loro derivate (come ci aspettiamo: la somma di reali è reale, la derivata di reali è reale).

Esaminiamone il prodotto, raccogliendo i termini con “la stessa periodicità”:

$$\begin{aligned} c(t) = a(t)b(t) &= \dots + (\dots + a_{-2}b_1 + a_{-1}b_0 + a_0b_{-1} + a_1b_{-2} + \dots)e^{-i\omega t} + \\ &+ (\dots + a_{-2}b_2 + a_{-1}b_1 + a_0b_0 + a_1b_{-1} + \dots) + \\ &+ (\dots + a_{-1}b_2 + a_0b_1 + a_1b_0 + a_2b_{-1} + \dots)e^{i\omega t} + \dots \end{aligned}$$

(i termini con lo stesso esponenziale $k i \omega$ hanno come coefficiente la somma di prodotti del tipo $a_m b_n$ con $m + n = k$, poiché nella moltiplicazione gli esponenti vengono sommati).

Notiamo che anche il prodotto è reale: $c_{-k} = c_k^*$ in quanto l'uno è la somma dei coniugati dei termini dell'altro (ogni termine del tipo $a_m b_n$ nel coefficiente di k avrà un analogo $a_{-m} b_{-n}$ nel coefficiente di $-k$, e il complesso coniugato di un prodotto è il prodotto dei complessi coniugati).

L'intuizione di Heisenberg fu nel rappresentare un'osservabile con una matrice le cui righe sono ripetizioni “shiftate” dei coefficienti della serie di Fourier, in modo

²^Nel caso di una carica oscillante, potrebbe essere il “meccanismo” alla base dell'emissione delle righe dello spettro dell'idrogeno

³^Riferimento: <http://www.mathpages.com/home/kmath698/kmath698.htm>

che i coefficienti di ordine 0 riempissero la diagonale principale:

$$q(t) = \begin{pmatrix} \ddots & & & & & & \\ & q_0 & q_1 e^{i\omega t} & q_2 e^{2i\omega t} & q_3 e^{3i\omega t} & q_4 e^{4i\omega t} & \dots \\ & q_{-1} e^{-i\omega t} & q_0 & q_1 e^{i\omega t} & q_2 e^{2i\omega t} & q_3 e^{3i\omega t} & \dots \\ & q_{-2} e^{-2i\omega t} & q_{-1} e^{-i\omega t} & q_0 & q_1 e^{i\omega t} & q_2 e^{2i\omega t} & \dots \\ & q_{-3} e^{-3i\omega t} & q_{-2} e^{-2i\omega t} & q_{-1} e^{-i\omega t} & q_0 & q_1 e^{i\omega t} & \dots \\ & q_{-4} e^{-4i\omega t} & q_{-3} e^{-3i\omega t} & q_{-2} e^{-2i\omega t} & q_{-1} e^{-i\omega t} & q_0 & \dots \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Tale rappresentazione molto strana ha un suo perché, che ora esamineremo.

Sia q_{mn} l'elemento di riga m -esima e colonna n -esima (qui valgono anche indici negativi). Per esempio $q_{00} = q_0$, e $q_{10} = q_1 e^{i\omega t}$. In generale:

$$a_{mn} = a_{n-m} e^{i(n-m)\omega t}; \quad b_{mn} = b_{n-m} e^{i(n-m)\omega t}$$

Vale poi $a_{mn} = a_{nm}^*$ (perché siamo partiti da una serie a somma reale).

È definita poi la somma di osservabili (sommando le matrici). Per il prodotto, invece, notiamo che, per $c(t) = a(t)b(t)$, il termine c_{mn} è dato dal prodotto scalare tra la riga m -esima della matrice di a e la colonna n -esima della matrice di b (che sono uguali se non per il fatto che la seconda è “shiftata” di $n-m$ in avanti rispetto alla prima):

$$c_{mn} = \sum_{\mu=-\infty}^{\infty} a_{m\mu} b_{\mu n}$$

EsPLICITIAMO uno dei termini della somma:

$$a_{m\mu} b_{\mu n} = a_{m-\mu} b_{\mu-n} e^{i(m-\mu)\omega t} e^{i(\mu-n)\omega t} = a_{m-\mu} b_{\mu-n} e^{i(m-n)\omega t}$$

Perciò c_{mn} avrà la stessa frequenza $(m-n)\omega$ di a_{mn} e b_{mn} . Ma allora, concentrandoci solo sulle frequenze, dal calcolo di sopra notiamo che:

$$\omega_{m\mu} + \omega_{\mu n} = \omega_{mn}$$

Ora, nel caso che abbiamo costruito artificialmente (matrice della serie di Fourier), ciò è ovvio (deriva dal fatto che $\omega_{mn} = (m-n)\omega_0$, per costruzione). Tuttavia, Heisenberg considerò il caso più generale in cui tale espressione è verificata, che è quello in cui esistono determinate costanti ω_n tali che:

$$\omega(m, n) = \omega_m - \omega_n$$

Possiamo quindi modificare la matrice di prima cambiandone le frequenze seguendo questa regola appena trovata. Si otterranno colonne che non sono più serie di Fourier, ma un qualcosa di più generale - capace di spiegare gli spettri atomici trovati⁴.

⁴ Bisognerebbe anche eliminare tutti i termini con almeno un indice negativo, poiché si considerano transizioni solo tra “stati energetici positivi”.

Heisenberg congetturò che tutte le osservabili nell'atomo dovessero assumere la forma di termini di matrici, della forma $q_{mn}(t) = q_{mn}e^{i\omega_{mn}t}$ (dove q_{mn} potrebbe essere, per esempio, l'intensità di una riga di emissione e ω_{mn} la sua pulsazione). Ciò ha profonde conseguenze, poiché viola tutte le considerazioni intuitive fatte in MC:

- Le osservabili sono descritte da **matrici** con termini **complessi**;
- Il prodotto tra osservabili è un prodotto matriciale (che non è detto neppure essere commutativo, dunque l'algebra non è abeliana).

In effetti, confrontando questa intuizione con la quantizzazione di Planck, Heisenberg dimostrò che per un sistema unidimensionale, la posizione q , e il momento p verificano:

$$[q, p]_{mn} \equiv [qp - pq]_{mn} = i\hbar\delta_{mn}$$

E cioè la commutatività viene meno (l'espressione tra parentesi è detta *commutatore*).

Inoltre le matrici devono essere infinito-dimensionali. Infatti, se per assurdo le matrici fossero finito-dimensionali allora sarebbe definita una traccia, e dovrebbe valere:

$$\text{Tr}(qp) = \text{Tr}(pq) \Rightarrow \text{Tr}(qp - pq) = 0 \neq \text{Tr}(i\hbar\mathbb{I}) = i\hbar$$

(Tr è la traccia di una matrice), ovvero un assurdo.

L'unico modo perché le relazioni di commutazione siano consistenti è allora che la traccia non sia definita, e cioè che le matrici in questione siano infinito-dimensionali.

2.1.2 L'approccio di Schrödinger

Alle stesse conclusioni arrivò Schrödinger (1926) in modo più indiretto.

Scrivendo la sua equazione d'onda stazionaria per una particella con potenziale V :

$$\left(\frac{-\hbar^2}{2m} \nabla^2 \psi + V\psi \right) = \mathcal{E}\psi \quad (2.1)$$

si accorse che si poteva ricondurre il membro di sinistra all'Hamiltoniana classica:

$$H_{\text{cl}}(\vec{x}, \vec{p}) = \frac{\vec{p}^2}{2m} + V(\vec{x})$$

sostituendo \vec{x} con un operatore che agisce su $\psi(\vec{x})$ come $\vec{x}\psi(\vec{x})$, e \vec{p} con un operatore che agisce su $\psi(\vec{x})$ come $-i\hbar\vec{\nabla}\psi(\vec{x})$.

Generalizzando, sostituì allora alle \vec{x} e \vec{p} classiche questi operatori:

$$H_{\text{quant}}(\vec{x}, -i\hbar\vec{\nabla})$$

riscrivendo l'equazione (2.1) come:

$$H(\vec{x}, -i\hbar\vec{\nabla})\psi(\vec{x}) = \mathcal{E}\psi(\vec{x}) \quad (2.2)$$

In particolare, in una dimensione, l'algebra tra la posizione X ($X\psi(x) = x\psi(x)$) e il momento P ($P\psi(x) = -i\hbar \frac{d}{dx}\psi(x)$) quantistici è data da:

$$(XP - PX)\psi(x) = i\hbar\psi(x)$$

ossia un'algebra su \mathbb{C} non commutativa di operatori che agiscono sullo spazio vettoriale complesso delle funzioni d'onda, che risulta essere isomorfa a quella già trovata da Heisenberg per le matrici.

Dall'equazione (2.2), si ha che le \mathcal{E} che risolvono l'equazione di Schrödinger stazionaria nel caso dell'atomo dell'idrogeno sono proprio quelle trovate da Bohr per spiegare l'emissione spettrale dell'idrogeno, e devono perciò rappresentare lo spettro dell'osservabile-energia H quantistico. Poiché l'equazione stazionaria si presenta come un'equazione agli autovalori:

$$\hat{H}\psi = \mathcal{E}\psi$$

si nota che in questo caso **i valori ottenibili dalle misure non sono dati dal codominio delle osservabili**, ma (come verrà formalmente dimostrato più avanti) risolvendo equazioni agli autovalori. Cade perciò anche l'ultima proprietà intuitiva che avevamo osservato nella MC.

Restano comunque aperte le seguenti questioni:

- Quali operatori corrispondono a osservabili?
- Come mai sono possibili rappresentazioni delle stesse osservabili come “oggetti matematici diversi” (per esempio negli approcci di Heisenberg e Schrödinger, q_{mn} vs \hat{x} , p_{mn} vs $-i\hbar \frac{d}{dx}$)

Le risposte a queste domande costituiscono alcuni dei concetti di base della MQ.

2.2 Stati

Uno **stato** Σ caratterizza l'informazione su un sistema fisico. Dato un sistema S si assume che tutti i suoi stati Σ si possano ottenere l'uno dall'altro con operazioni fisicamente eseguibili sul sistema.

Stato

Gli stati con **informazione massimale** sono detti **puri**, quelli con informazione minore sono detti **misti** (questa distinzione sarà più chiara in seguito). Nel contesto della Meccanica *non* statistica, (dunque dove si ha sempre tutta l'informazione del sistema), spesso si omette l'aggettivo “puri”.

Stati puri vs. stati misti

2.2.1 Stati in Meccanica Classica

In MC gli **stati puri** sono descritti da punti nello spazio delle fasi, come $(q_0, p_0) \in \Omega$ e descrivono le “*condizioni iniziali*”. Gli **stati misti** in MC sono distribuzioni di probabilità $\rho(q, p)$ su Ω , che perciò soddisfano:

$$\rho(q, p) \geq 0; \quad \int_{\Omega} \rho(q, p) dq dp = 1$$

Ad esempio se il sistema è una molecola di un gas ideale monoatomico di massa m e sappiamo che il gas è in equilibrio a temperatura T e in un volume V , lo stato misto corrispondente è la distribuzione di Maxwell-Boltzmann:

$$\rho_{\text{MB}}(\vec{x}, \vec{p}) = \frac{\chi_V(\vec{x}) e^{-\frac{\vec{p}^2}{mkT}}}{\int d^3x d^3p \chi_V(\vec{x}) e^{-\frac{\vec{p}^2}{mkT}}} \quad \chi_V(\vec{x}) = \begin{cases} 1 & \text{se } \vec{x} \in V \\ 0 & \text{se } \vec{x} \notin V \end{cases}$$

Gli stati puri classici possono essere identificati anch'essi con distribuzioni di probabilità “concentrate in un solo punto” su Ω , utilizzando le distribuzioni delta di Dirac multi-dimensionali:

$$q_0, p_0 \in \Omega \rightarrow \rho_{q_0, p_0}(q, p) = \delta^{(N)}(q - q_0) \delta^{(N)}(p - p_0)$$

dove N è il numero di gradi di libertà del sistema. Notiamo che non essendo in generale Ω uno spazio vettoriale, non è possibile (e non avrebbe nemmeno senso fisico) “sommare” stati puri classici.

2.2.2 Stati in Meccanica Quantistica

Il primo indizio che non esiste uno “spazio delle fasi” in MQ viene dalla soluzione di Planck del problema del corpo nero. Per riprodurre la formula dell'energia media per i suoi oscillatori armonici non poteva pesare con una distribuzione di probabilità tutti i punti dello spazio delle fasi dell'oscillatore armonico che emetteva la radiazione, ma doveva quantizzare lo spazio delle fasi, suddividendolo in cellette di area h , detta appunto *costante di Planck*.

Questo diviene più preciso con il principio di indeterminazione di Heisenberg (1927) per cui $\Delta x \Delta p \geq \frac{\hbar}{2}$ ci indica l'impossibilità di conoscere con precisione arbitraria simultaneamente posizione e momento, cioè i punti di Ω . Ma se sperimentalmente non posso definire “un punto” nello spazio delle fasi, allora pragmaticamente quel punto “non esiste”. Dobbiamo allora reinventare la nozione di spazio delle fasi.

Se non utilizzando punti dello spazio delle fasi, come possiamo caratterizzare l'informazione massimale? Schrödinger fu il primo a dare una risposta: per eseguire i suoi conti per l'atomo di idrogeno doveva usare le funzioni d'onda $\psi(\vec{x})$, che contengono intrinsecamente le informazioni dello stato. Tuttavia, conoscendo la distribuzione sulla posizione $\psi(\vec{x})$ di una particella, non si conosce la distribuzione sul momento \vec{p} . D'altro canto, le ψ sono onde, e perciò a differenza della MC si possono sommare e formano uno spazio vettoriale. Pertanto si inizia a comprendere che lo spazio naturale degli stati in MQ non è lo spazio delle fasi, costituito da punti, bensì uno spazio di funzioni, dato che ogni stato non può essere rappresentato da un punto, ma da una funzione.

Restano in ogni caso aperte le questioni:

- Qual è il corretto spazio vettoriale delle ψ di Schrödinger?
- Qual è lo spazio vettoriale degli stati per le q_{nm} , p_{nm} di Heisenberg?

- Questi due spazi vettoriali sono gli stessi?
- Come sono definiti gli stati misti?

Riepilogando:

- In **Meccanica Classica**: **stato puro** \mapsto distribuzione “ δ di Dirac” puntuale nello spazio delle fasi, **stato misto** \mapsto distribuzione di probabilità $\rho(q, p)$ sullo spazio delle fasi.
- In **Meccanica Quantistica**: **stato puro** \mapsto funzione d’onda $\psi(\vec{x})$

2.3 Valor medio

Data un’osservabile O e uno stato Σ la teoria deve definire come si calcola il valor medio di O nello stato Σ , che indichiamo con $\langle O \rangle_\Sigma$ e che deve essere confrontabile con l’esperimento.

Questo numero reale rappresenta la stima teorica del valor medio sperimentale, che possiamo definire come la media dei valori ottenuti misurando N volte l’osservabile O nello stato Σ (in identiche condizioni) ottenendo $o_1, \dots, o_N \in \sigma(O)$ nel limite ideale $N \rightarrow \infty$, cioè:

*Valor medio
(sperimentale)*

$$\langle O \rangle_\Sigma = \lim_{N \rightarrow \infty} \frac{o_1 + o_2 + \dots + o_N}{N} \quad (2.3)$$

Ove l’esistenza del limite è assunta come fatto sperimentale.

In **MC**, in generale per gli **stati misti** il valor medio dell’osservabile $O = f(q, p)$ in uno stato $\Sigma = \rho(q, p)$ è dato da:

*(Lezione 4 del
05/10/2018)*

$$\langle O \rangle_\Sigma = \langle f \rangle_\rho = \int_\Omega dp dq \rho(q, p) f(q, p) \quad (2.4)$$

*Valor medio in
MC*

In **MQ**, la regola per gli **stati puri** ($\Sigma = \psi$) trovata da Born enuncia che $|\psi(x)|^2$ è proporzionale alla densità di probabilità di trovare la particella in x se lo stato è descritto da ψ . Per renderla una densità di probabilità dividiamo per il fattore di normalizzazione $\|\psi\|^2$:

*Valor medio in
MQ*

$$\langle \hat{x} \rangle_\psi = \int dx x \frac{|\psi(x)|^2}{\|\psi\|^2}; \quad \|\psi\|^2 = \int dx |\psi(x)|^2$$

Dove x sono i possibili valori, e la frazione costituisce la densità di probabilità. Più avanti si vedrà come non vi sia alcuna analogia con la (2.4). Infatti il concetto di funzione d’onda quantistica non ha nulla a che vedere con quello di distribuzione di uno stato misto classico.

La generalizzazione di questa regola all’osservabile momento $\hat{p} = -i\hbar \frac{d}{dx}$ ha mostrato che la corretta definizione è:

$$\langle \hat{p} \rangle_\psi = \int dx \frac{\psi^*(x) \hat{p} \psi(x)}{\|\psi\|^2} = \int dx \frac{\psi^*(x)}{\|\psi\|^2} \left(-i\hbar \frac{d}{dx} \psi(x) \right)$$

Calcolando la trasformata di Fourier⁵ si ottiene:

$$\langle \hat{p} \rangle_\psi = \int dp \frac{\tilde{\psi}^*(p)p\tilde{\psi}(p)}{\|\tilde{\psi}\|^2} = \int dp \frac{|\tilde{\psi}(p)|^2}{\|\tilde{\psi}\|^2} p$$

È perciò possibile generalizzare tale regola ad una qualsiasi osservabile O :

$$\langle \hat{O} \rangle_\psi = \int \frac{\psi^*(x)\hat{O}\psi(x)}{\|\psi\|^2} dx \quad (2.5)$$

Quest'ultima è dunque una regola per tutte le osservabili in MQ.

Domanda 1. Come si giustifica matematicamente il passaggio da (2.3) a (2.4)-(2.5)?

Nota tecnica: Se $\sigma(O)$ è illimitato, $\langle O \rangle_\Sigma$ potrebbe non esistere per qualche stato Σ . Assumiamo quindi che noti $\langle O \rangle_\Sigma$ per tutti i Σ per i quali è definito sia sufficiente a caratterizzare O (**principio di minimalità**). (Cioè se conosco tutta l'informazione sperimentalmente ottenibile da O , conosco O . In parole povere mi limito solo ai casi in cui $\langle O \rangle_\Sigma$ ha per ogni stato un valore definito).

2.3.1 Relazioni tra osservabili e stati

Poiché i **valori medi** sono tutto ciò che possiamo sperare di misurare in laboratorio, le relazioni tra **osservabili** e **stati** necessariamente fanno uso di questi. Facciamo quindi le seguenti definizioni:

- Due osservabili O_1 e O_2 tali che $\forall \Sigma$ in cui sono definiti si abbia $\langle O_1 \rangle_\Sigma = \langle O_2 \rangle_\Sigma$ sono lo stesso osservabile: $O_1 = O_2$. In altre parole due osservabili sono uguali se, a parità di condizioni iniziali (stesso stato), forniscono lo stesso valore (medio) della misura. *Identità tra osservabili*
- Analogamente, se $\langle O \rangle_{\Sigma_1} = \langle O \rangle_{\Sigma_2}$ per ogni osservabile O , si ha che $\Sigma_1 = \Sigma_2$ (sono lo stesso stato). In altre parole, due stati sono uguali se, considerando la stessa osservabile, si ottiene lo stesso valore (medio) della misura. *Identità tra stati*
- La relazione tra funzioni d'onda e stati è biunivoca solo a meno di un fattore moltiplicativo. Infatti, come si può verificare facilmente dalla (2.5), moltiplicando la funzione d'onda ψ associata allo stato Σ per un qualsiasi numero complesso α (costante) si ottengono gli stessi valori medi (per qualsiasi osservabile). Dalle definizioni di identità date sopra segue che $\alpha\psi$ corrisponde allo stesso stato Σ . *Stati e funzioni d'onda*

2.3.2 Complemento matematico: Teoria della misura

Cerchiamo ora di dare una risposta alla domanda (Q:1). Nello specifico, ci chiediamo quale sia la relazione tra $\langle O \rangle_\Sigma$ (definito come integrale sullo spazio delle

Relazione tra $\langle O \rangle_\Sigma$ e la media

⁵ΛCFR pag. 49 degli appunti di Fisica Moderna

fasi in MC (2.4) o di un operatore applicato ad una funzione d'onda in un certo modo per la MQ (2.5)) e $\lim_{N \rightarrow \infty} \frac{o_1 + \dots + o_N}{N}$? Sappiamo sperimentalmente che sono identificabili, ma come giustificarlo teoricamente?

Supponiamo che le misure dell'osservabile O (che quindi fanno parte di $\sigma(O)$) siano **indipendenti** e con la stessa **distribuzione di probabilità** P_Σ^O , e che il valor medio di P_Σ^O esista per il principio di minimalità.

Allora sotto un'assunzione di **regolarità**, la successione $\frac{1}{N} \sum_{i=1}^N o^{(i)}$ con $o^{(i)}$, $i = 1 \dots N$ nello stato Σ , converge (quasi ovunque rispetto a P_Σ^O) al valor medio $\langle O \rangle_\Sigma$ per $N \rightarrow \infty$ (legge forte dei grandi numeri).

Convergenza della media al valor medio

Nel parlare di *distribuzioni di probabilità* stiamo però facendo riferimento ad una *struttura probabilistica* propria delle osservabili, e che dobbiamo giustificare.

Per far ciò, diamo prima qualche cenno alla teoria della misura⁶ su \mathbb{R} .

Trascureremo dettagli matematici come il fatto che si considerano solo insiemi “misurabili”, cioè per cui la nozione di misura è non contraddittoria. Del resto “tutti” gli insiemi riscontrati in esempi fisici sono misurabili.

Definizione 1. Una misura μ su \mathbb{R} è una mappa dagli insiemi (misurabili) di \mathbb{R} , detti boreliani, $\mathcal{B}(\mathbb{R})$ in $[0, +\infty]$ che soddisfa:

1. $\mu(\emptyset) = 0$ (la misura di un insieme vuoto è nulla);
2. μ è **numerabilmente additiva**: se $\{\Delta_i\}_{i \in I}$ è una famiglia numerabile di boreliani disgiunti ($\Delta_i \cap \Delta_j = \emptyset, i \neq j$), allora $\mu(\cup_{i \in I} \Delta_i) = \sum_{i \in I} \mu(\Delta_i)$.

Le misure “regolari” su \mathbb{R} sono tutte generate da funzioni **monotone crescenti** $\alpha(\lambda)$, $\lambda \in \mathbb{R}$ e definite nel seguente modo.

Misura regolare su \mathbb{R}

Sia μ_α la misura. Allora la misura di un aperto $]a, b[\subset \mathbb{R}$ è data da:

$$\mu_\alpha(]a, b[) = \lim_{\epsilon \rightarrow 0} [\alpha(b - \epsilon) - \alpha(a + \epsilon)]$$

In questo modo è definita la misura di ogni $\Delta \in \mathcal{B}(\mathbb{R})$, ricordando che ogni aperto di \mathbb{R} è un'unione numerabile di aperti disgiunti, come:

$$\mu_\alpha(\Delta) = \inf_{\Delta \subset A} \mu_\alpha(A), \quad A \text{ aperto di } \mathbb{R}$$

Le misure μ_α si dicono misure di Lebesgue-Stieltjes.

Esempi:

1. $\alpha(\lambda) = \lambda$ è la **misura di Lebesgue** $\mu_{\alpha=\lambda}(]a, b[) = b - a$
Si dimostra immediatamente che la “misura di un punto” è nulla:

$$0 \leq \mu_{\alpha=\lambda}(\{x\}) = \inf_{\{x\} \subset A} \mu_{\alpha=\lambda}(A) \leq \inf_n \mu_{\alpha=\lambda}\left(\left[x - \frac{1}{n}, x + \frac{1}{n}\right]\right) = \inf_n \frac{2}{n} = 0$$

⁶Qui per “misura” si intende un'operazione in senso matematico, che a priori non ha nulla a che fare con una “misura fisica”

2. $\alpha(\lambda) = H(\lambda)$ è la **misura di Dirac**, con $H(\lambda)$ la funzione di Heaviside (che vale 0 per $\lambda < 0$, e 1 altrimenti).

$$\mu_{\alpha=H}([a, b[) = 1 \text{ se } 0 \in]a, b[\text{ e } 0 \text{ altrimenti}^7$$

In effetti, la misura di Dirac bada solo allo 0: $\mu_H(\{0\}) = 1$

Si può utilizzare la misura μ_α per definire un integrale che generalizza quello di Lebesgue (per cui $\alpha(\lambda) = \lambda$), detto **integrale di Lebesgue-Stieltjes**.

Supponiamo $f \geq 0$ (e misurabile, cioè le antimmagini di intervalli aperti del codominio sono insiemi misurabili del dominio).

Dividiamo il codominio di f in intervallini di ampiezza 2^{-n} e consideriamo l'antimmagine per f , $f^{-1}([\frac{m-1}{2^n}, \frac{m}{2^n} [)$. Questi insiemi del dominio ammettono misura di Lebesgue-Stieltjes

$$\mu_\alpha \left(f^{-1} \left(\left[\frac{m-1}{2^n}, \frac{m}{2^n} \right[\right) \right)$$

Allora:

$$\int f(x) d\alpha(x) \equiv \lim_{n \rightarrow \infty} \sum_{m=1}^{\infty} \frac{m-1}{2^n} \mu_\alpha \left(f^{-1} \left(\left[\frac{m-1}{2^n}, \frac{m}{2^n} \right[\right) \right)$$

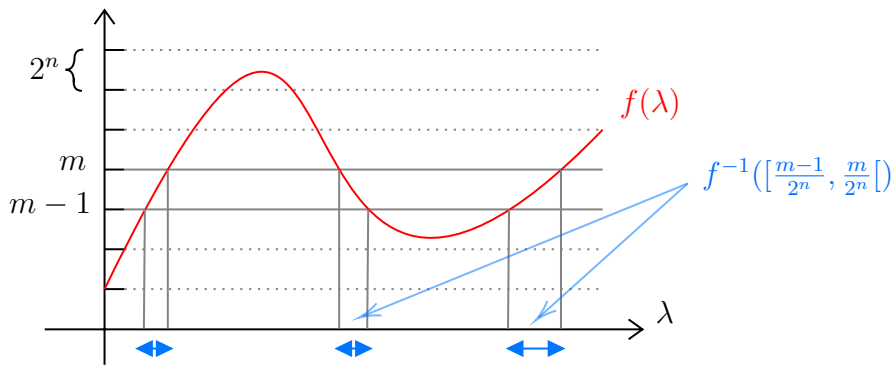


Figura 2.1: Interpretazione grafica dell'integrale di Lebesgue

Qualche **esempio**:

⁷Calcolando con la definizione, per un intervallo a destra di 0, che non include 0, il computo dà $1 - 1 = 0$, e a sinistra si ha $0 - 0 = 0$. Un intervallo con un estremo che coincide con 0 ha analogamente misura nulla, essendo un intervallo per definizione aperto, per cui 0 è escluso.

1. $\alpha = \lambda$, $\int f(\lambda) d\lambda$ è l'integrale di Lebesgue.
2. Se esiste la derivata $\frac{d\alpha(\lambda)}{d\lambda}$ ed è localmente Lebesgue-integrabile (ossia è $\in L^1$) allora:

$$\int f(\lambda) d\alpha(\lambda) = \int f(\lambda) \frac{d\alpha(\lambda)}{d\lambda} d\lambda \quad (2.6)$$

E $d\alpha$ è detta assolutamente continua rispetto alla misura di Lebesgue.

3. $\alpha(\lambda) = H(\lambda)$ Allora:

$$\int f(\lambda) dH(\lambda) = f(0) \equiv \int f(x) \delta(\lambda) d\lambda$$

Quindi H è la misura di Dirac⁸.

4. Per completezza si fa presente che esistono misure di L-S che “non pesano i punti” (come invece succede caso della misura di Dirac), ma per cui comunque non esiste $\alpha'(\lambda)$ tale che $d\alpha = \alpha'(\lambda) d\lambda$. Tali misure sono dette *singularmente continue* rispetto alla misura di Lebesgue.

2.3.3 La struttura probabilistica

Fatte le opportune premesse matematiche, il seguente teorema ci consentirà di collegare osservabili alla loro struttura probabilistica.

Teorema 2.3.1. *Sia $C_0(\mathbb{R})$ l'insieme delle funzioni continue definite su \mathbb{R} che svaniscono all'infinito, e sia $F : C_0(\mathbb{R}) \rightarrow \mathbb{R}$ un **funzionale lineare positivo**, cioè tale che*

*Teorema di
Riesz-Markov*

1. $\forall f, g \in C_0(\mathbb{R}), \forall a, b \in \mathbb{R}$ valga $F(af + bg) = aF(f) + bF(g)$ (linearità);
2. se $f \geq 0$ allora $F(f) \geq 0$ (positività: F “non cambia il segno” di f).

Allora esiste una **funzione monotona crescente** $\alpha_F(\lambda)$ su \mathbb{R} tale che:

$$F(f) = \int f(\lambda) d\alpha_F(\lambda)$$

In altre parole, il teorema di Riesz-Markov fa sì che ogni funzionale lineare e positivo F definisca una misura α_F (la si può pensare proprio come un “metro”, una “scala di misura”) tale che integrare una funzione f utilizzando quella misura equivale ad applicare il funzionale F alla funzione f stessa.

Ora mostriamo che, con opportune assunzioni, il valore medio delle osservabili in uno stato definisce un funzionale lineare positivo a cui applicare Riesz-Markov, in modo da associare ad ogni valor medio una *misura di probabilità*, giustificando così

⁸Si noti che, al punto precedente, abbiamo chiesto che la derivata della misura $d\alpha(\lambda)/d\lambda$ sia localmente integrabile. Perciò, visto che per la funzione di Heaviside H la derivata nel senso normale non è definita - c'è un “gradino” - possiamo usare quella distribuzionale, e si ha $dH(\lambda)/d\lambda = \delta(\lambda)$. CFR appunti del corso di metodi matematici per la dimostrazione.

la “struttura probabilistica”.

Se O è un’osservabile con spettro $\sigma(O)$ è naturale richiedere che, se g è una funzione reale continua su \mathbb{R} che si annulla all’infinito, esista un’osservabile $g(O)$ con spettro $\sigma(g(O)) = g(\sigma(O))$. Ciò corrisponde ad usare, processo di misura⁹ di O , invece della scala λ , la scala $g(\lambda)$ (come si fa, per esempio, nell’uso di una scala logaritmica).

Il valor medio di $g(O)$ è dato da:

$$\langle g(O) \rangle_\Sigma = \lim_{N \rightarrow \infty} \frac{g_1(o_1) + \cdots + g(o_N)}{N} \quad (2.7)$$

Se definiamo $g(O)$ positivo, ossia tale che verifichi $g(O) \geq 0$, se tutti i risultati delle misure di $g(O)$ sono positivi, allora anche il valor medio sarà positivo $\langle g(O) \rangle_\Sigma \geq 0$ (*intuitivamente: la media di numeri positivi è a sua volta positiva*, di certo “non inverte il segno”. Se alcune misure sono negative, la media potrà essere negativa, ma mai se tutte le misure sono positive).

Dall’espressione in (2.7) segue poi che, per due funzioni degli osservabili $g(O)$ e $f(O)$ (reali continue positive che si annullano all’infinito):

$$\langle af(O) + bg(O) \rangle_\Sigma = a\langle f(O) \rangle_\Sigma + b\langle g(O) \rangle_\Sigma$$

(Questa proprietà di linearità si può vedere utilizzando la definizione, o ricordandosi le proprietà della *speranza matematica* viste in Sperimentazioni 1).

Perciò il valor medio sperimentale della funzione di un’osservabile è lineare e positivo, e dunque se fosse associabile ad un funzionale rispetterebbe le ipotesi del teorema 2.3.1.

Consideriamo quindi $F : f \in C_0(\mathbb{R}) \mapsto \langle f(O) \rangle_\Sigma$, ossia il funzionale che associa ad ogni funzione continua che si annulla all’infinito il valor medio sperimentale dell’osservabile $f(O)$ (che è ben definito per le funzioni su $C_0(\mathbb{R})$). Per quanto abbiamo appena visto, F è un funzionale lineare positivo, e quindi per Riesz-Markov esiste una funzione monotona crescente che denotiamo $P_\Sigma^O(\lambda)$ tale che $\forall g \in C_0(\mathbb{R})$ si possa scrivere:

$$\langle g(O) \rangle_\Sigma = \int_{\mathbb{R}} g(\lambda) dP_\Sigma^O(\lambda) \quad (2.8)$$

con integrale inteso in senso di Lebesgue-Stieltjes. Sempre rifacendosi a quanto riepilogato nei cenni di teoria della misura, formalmente si può scrivere:

$$dP_\Sigma^O(\lambda) = \frac{dP_\Sigma^O(\lambda)}{d\lambda} d\lambda \quad (2.9)$$

Se inoltre si considera g come funzione costante $g \equiv 1$, si ottiene l’osservabile identità $g(O) = \mathbb{I}$ che dà sempre 1 come risultato della misura, per cui $\langle \mathbb{I} \rangle_\Sigma = 1$, allora, dato un qualsiasi stato Σ e una qualsiasi osservabile O , estendendo¹⁰ (2.8) a \mathbb{I} dà:

$$\langle \mathbb{I} \rangle_\Sigma = 1 = \int_{\mathbb{R}} dP_\Sigma^O(\lambda)$$

⁹La misura va intesa qui come l’operazione fisica di misura, non la misura matematica

¹⁰Non è una funzione che si annulla ad infinito, quindi non appartiene a $C_0(\mathbb{R})$ e non potremmo “applicare” il funzionale valor medio ad essa. Allo stesso tempo il suo valor medio è ben definito, in quanto è una funzione che assume *un solo valore*, e ciò ci libera dalla restrizione che avevamo adottato. Inoltre la scelta dell’operatore identità è arbitraria: avremmo potuto usare

*Funzionale F
“valor medio”*

Ovvero $dP_\Sigma^O(\lambda)$ è una misura di probabilità, in quanto è normalizzata a 1.

Dato f misurabile reale possiamo utilizzare la misura P_Σ^O appena trovata per definire l'integrale $\int f(\lambda) dP_\Sigma^O(\lambda)$ per tutti i Σ per i quali l'integrale esiste, che nel caso di f limitata sono tutti i Σ .

È allora naturale richiedere che esista un'osservabile $f(O)$ tale che:

$$\langle f(O) \rangle_\Sigma = \int_{\mathbb{R}} f(\lambda) dP_\Sigma^O(\lambda)$$

(visto che siamo partiti da un g generico per costruire la misura di probabilità, ci aspettiamo che tale risultato si applichi a qualsiasi altra osservabile¹¹).

Con questa ulteriore assunzione, dato un insieme (misurabile) Δ di \mathbb{R} , definiamo:

$$P_\Sigma^O(\Delta) \equiv P_\Sigma^O(b) - P_\Sigma^O(a) \stackrel{(a)}{=} \int_{\Delta} dP_\Sigma^O(\lambda) = \int_{\mathbb{R}} \chi_\Delta(\lambda) dP_\Sigma^O(\lambda) = \langle \chi_\Delta(O) \rangle_\Sigma \quad (2.10)$$

dove il passaggio (a) è giustificato dal teorema fondamentale del calcolo, e $\chi_\Delta(O)$ è la funzione caratteristica di Δ , che vale 1 per $O \in \Delta$ e 0 altrimenti¹². Il suo effetto è quello di “restringere” l'integrale da \mathbb{R} a Δ , visto che azzerà la funzione integranda per tutti i $\lambda \notin \Delta$. Confrontando l'espressione con (2.8) è giustificata la scrittura di $\langle \chi_\Delta(O) \rangle_\Sigma$, interpretando quindi χ_Δ come una funzione dell'osservabile. *Esemplificando, intuitivamente: si immagini di eseguire un esperimento molte volte, ottenendo una sequenza di o_i (valori dell'osservabile). Allora il valor medio di $\chi_\Delta(O)$ è dato da:*

$$\langle \chi_\Delta(O) \rangle_\Sigma = \lim_{N \rightarrow \infty} \frac{\chi_\Delta(o_1) + \cdots + \chi_\Delta(o_N)}{N}$$

Poiché χ_Δ è 1 solo quando $o_i \in \Delta$, il numeratore è il numero di volte per cui l'osservabile “è comparso” dentro all'intervallo Δ , mentre il denominatore è il numero totale di volte per cui si è misurato l'osservabile. Ripetendo il discorso dei funzionali lineari positivi e del teorema di Riesz è possibile riscrivere ciò con un opportuno integrale, esattamente come abbiamo fatto sopra.

In particolare se $\Delta =] - \infty, \lambda]$ avremo:

$$P_\Sigma^O(] - \infty, \lambda]) = \int_{-\infty}^{+\infty} H(\lambda - \lambda') dP_\Sigma^O(\lambda') = \int_{-\infty}^{\lambda} dP_\Sigma^O(\lambda') \stackrel{(a)}{=} P_\Sigma^O(\lambda) = \langle H(\lambda - O) \rangle_\Sigma \quad (2.11)$$

ove in (a) si è usato il teorema fondamentale del calcolo integrale, tenendo conto che $P_\Sigma^O(-\infty) = 0$. Il risultato appena ottenuto è compatibile con (2.10) con

una qualsiasi funzione costante, per cui il valor medio è “ben definito a priori” e ottenere lo stesso risultato. Alternativamente possiamo vedere la funzione $f \equiv 1$ come limite, immaginando di considerare una funzione C^0 che vale 1 entro un intervallo e poi decresce a 0 e “allargare” all'infinito tale intervallo.

¹¹ \wedge Il punto è che prima siamo partiti da una generica funzione di un'osservabile, mentre ora chiediamo che $f(O)$ sia proprio un'altra osservabile

¹² \wedge Qui O è un'osservabile, la cui misura deve essere reale - e quindi può appartenere ad un intervallo reale

$\Delta =] - \infty, \lambda]$.

Ma allora, ripartendo sempre da (2.8):

$$\langle g(O) \rangle_\Sigma = \int_{\mathbb{R}} g(\lambda) dP_\Sigma^O(\lambda) \stackrel{(2.11)}{=} \int_{\mathbb{R}} g(\lambda) d\langle H(\lambda - O) \rangle_\Sigma$$

Prendendo per g l'identità arriviamo a:

$$\langle O \rangle_\Sigma = \int_{\mathbb{R}} \lambda d\langle H(\lambda - O) \rangle_\Sigma \quad (2.12)$$

Possiamo ora “estrarre” il valor medio dall'integrale. Matematicamente andrebbe giustificato: è semplice farlo nel caso classico, un po' meno nel caso quantistico (ma qui non lo faremo). Si giunge quindi a:

$$\langle O \rangle_\Sigma = \langle \int \lambda dH(\lambda - O) \rangle_\Sigma \quad \forall \Sigma$$

Poiché la relazione vale per tutti i Σ , posso omettere l'indicazione e definire una nuova misura $dP^O = dH(\lambda - O)$:

$$\langle O \rangle = \langle \int \lambda \underbrace{dH(\lambda - O)}_{dP^O} \rangle$$

dove dP^O , per come è definita la derivata distribuzionale della funzione di Heaviside, è una δ di Dirac centrata sul valore dell'osservabile. Ma allora, rimuovendo il valor medio, si ottiene infine, formalmente:

$$O = \int_{\mathbb{R}} \lambda dP^O \quad \Rightarrow \quad f(O) = \int_{\mathbb{R}} f(\lambda) dP^O(\lambda) \quad (2.13)$$

Analogamente a quanto abbiamo fatto per passare da $dP_\Sigma^O(\lambda)$ a $P_\Sigma^O(\lambda)$, potremmo passare da $dP^O(\lambda) = dH(\lambda - O)$ a $P^O(\lambda)$, che costituisce un nuovo osservabile dato da:

$$P^O(\lambda) = H(\lambda - O) \quad (2.14)$$

Per le proprietà della Heaviside, tale osservabile vale 1 se $O < \lambda$, e 0 altrimenti. Inoltre, per come è definita $dP^O(\lambda)$, possiamo anche scrivere:

$$P_\Sigma^O(\lambda) \stackrel{(2.11)}{=} \langle H(\lambda - O) \rangle_\Sigma = \langle P^O(\lambda) \rangle_\Sigma$$

L'interpretazione probabilistica di $\langle O \rangle_\Sigma = \int \lambda d\langle P^O(x) \rangle_\Sigma$ è perciò:

- λ descrive i possibili risultati a priori dalla misura di O , cioè i punti di \mathbb{R} (essendo O un'osservabile, cioè una grandezza che possiamo *misurare sperimentalmente*, e le misure sperimentali danno solo valori reali)
- $P_\Sigma^O(\Delta)$ = probabilità che una misura di O eseguita su un sistema nello stato Σ dia un valore contenuto nell'intervallo Δ
- $P_\Sigma^O(\lambda)$ = probabilità che una misura di O (eseguita su un sistema nello stato Σ) restituisca un valore $\leq \lambda$

- $dP_{\Sigma}^O(\lambda)$ = probabilità che una misura di O (eseguita su un sistema nello stato Σ) restituisca un valore tra λ e $\lambda + d\lambda$

Quindi note le $P_{\Sigma}^O(\lambda)$ e astrattamente $P^O(\lambda)$ sono noti i valori medi di tutte le funzioni di O (in quanto sappiamo l'esatta distribuzione dell'osservabile da cui tale funzioni dipendono).

Le $dP_{\Sigma}^O(\lambda)$ sono precisamente le misure di probabilità rispetto alle quali è definito il teorema del limite dei grandi numeri per il valor medio. Applicando tale teorema si ottiene che, al limite di misure infinite, la media dei risultati sperimentali tende ai valori medi calcolati teoricamente tramite integrazione.

(Lezione 5 del
08/10/2018)

Riepilogando

Sia O osservabile in uno stato Σ , esiste una **misura di probabilità** su \mathbb{R} data da $dP_{\Sigma}^O(\lambda)$ tale che:

$$\langle O \rangle_{\Sigma} = \int \lambda dP_{\Sigma}^O(\lambda); \quad P_{\Sigma}^O(\lambda) = \langle H(\lambda - O) \rangle_{\Sigma} \quad (2.15)$$

In **MC** $O = f(q, p)$, $\Sigma = \rho(q, p)$, da cui:

$$\begin{aligned} P_{\rho}^f(\lambda) &= \langle H(\lambda - f(q, p)) \rangle_{\rho} = \int_{\Omega} dq dp \rho(q, p) H(\lambda - f(q, p)) \\ \int_{\mathbb{R}} \lambda dP_{\rho}^f(\lambda) &= \int_{\mathbb{R}} \lambda d\lambda \int_{\Omega} dq dp \rho(q, p) H(\lambda - f(q, p)) = \\ &= \int_{\mathbb{R}} \lambda \int_{\Omega} dq dp \rho(q, p) \delta(\lambda - f(q, p)) d\lambda = \\ &= \int_{\Omega} dq dp \rho(q, p) f(q, p) \end{aligned}$$

Domanda 2. In **MQ** le osservabili sono operatori, perciò come si scrive la formula (2.15)?

2.3.4 Autostati

Un caso particolare di funzione di una osservabile O è il quadrato $(O - a)^2$, $a \in \mathbb{R}$. Definiamo la **fluttuazione quadratica media** di O attorno ad a nello stato Σ come:

$$\left[\langle (O - a)^2 \rangle_{\Sigma} \right]^{\frac{1}{2}} \equiv \Delta O_{\Sigma, a}$$

La controparte sperimentale è la fluttuazione delle misure di O nello stato Σ attorno ad a .

Usando (2.15):

$$(\langle \Delta O \rangle_{\Sigma, a})^2 = \int (\lambda - a)^2 dP_{\Sigma}^O(\lambda) \quad (2.16)$$

che come funzione di a ha un minimo per $a = \langle O \rangle_{\Sigma}$, infatti:

$$\begin{aligned} \frac{d}{da} \int (\lambda - a)^2 dP_{\Sigma}^O(\lambda) &= -2 \int (\lambda - a) dP_{\Sigma}^O(\lambda) = -2(\langle O \rangle_{\Sigma} - a) = 0 \text{ per } a = \langle O \rangle_{\Sigma} \\ \frac{d^2}{da^2} \int (\lambda - a)^2 dP_{\Sigma}^O(\lambda) &= 2 > 0 \end{aligned}$$

Definizione 2. Uno stato Σ per cui $(\Delta O)_{\Sigma, \langle O \rangle_\Sigma} \equiv \langle \Delta O \rangle_\Sigma = 0$ è detto **autostato** di O relativo all'autovalore $a = \langle O \rangle_\Sigma$.

Autostato

In altre parole, se Σ è autostato di O allora qualsiasi misura di O darà come risultato a , cioè il valor medio di O .

Nota: Se Σ è un autostato di O relativo ad a , una misura di O su Σ dà con probabilità 1 il risultato a , e quindi $a \in \sigma(O)$.

In **MC** tutti gli stati puri sono autostati di tutte le osservabili. Infatti: $O = f(q, p)$, e uno stato puro

In MC tutti gli stati puri sono autostati

$$\Sigma_{puro} = \delta^{(N)}(q - q_0) \delta^{(N)}(p - p_0); \quad (q_0, p_0) \in \Omega.$$

Ma allora, facendo tutti i passaggi a costo di essere pedanti:

$$\begin{aligned} \langle f \rangle_{\Sigma_{q_0, p_0}} &= f(q_0, p_0) \\ (\Delta f)_{\Sigma_{q_0, p_0}}^2 &\stackrel{(2.16)}{=} \int_{\mathbb{R}} (\lambda - \langle f \rangle_{\Sigma_{q_0, p_0}})^2 dP_{\Sigma_{q_0, p_0}}^f(\lambda) = \\ &\stackrel{(2.11)}{=} \int_{\mathbb{R}} (\lambda - f(q_0, p_0))^2 dH(\lambda - f(q_0, p_0)) = \\ &\stackrel{(2.9)}{=} \int_{\mathbb{R}} (\lambda - f(q_0, p_0))^2 \frac{dH(\lambda - f(q_0, p_0))}{d\lambda} d\lambda = \\ &= \int (\lambda - f(q_0, p_0))^2 \delta(\lambda - f(q_0, p_0)) d\lambda = 0 \end{aligned}$$

(Nell'ultimo passaggio si è applicata la definizione di derivata distribuzionale della funzione di Heaviside).

Quindi la conoscenza dello stato puro porta alla conoscenza con **certezza** del valore di **tutte le osservabili** (la conoscenza massimale che posso avere è tutta, non ci sono incertezze di fondo).

Chiaramente in MQ questo è impossibile, ad esempio se siamo certi che una misura di energia su un sistema ci dà il valore \mathcal{E} , allora lo stato puro è descritto da una funzione d'onda $\psi_{\mathcal{E}}$ che è soluzione dell'equazione stazionaria $H\psi_{\mathcal{E}}(x) = E\psi_{\mathcal{E}}(x)$. Ma allora non può avere un valore certo in una misura di posizione! Si ha infatti una $\psi_{\mathcal{E}}(x)$, che al modulo quadro mi indica la probabilità che la particella sia ad una certa x , che perciò non ha un unico valore definito.

Ciò si vede ancora più esplicitamente dal principio di indeterminazione di Heisenberg:

$$(\Delta X)_{\psi} (\Delta P)_{\psi} \geq \frac{\hbar}{2}$$

Quindi non possiamo conoscere con certezza posizione e momento per nessuno stato!

In MQ nessuno stato è autostato per tutte le osservabili

Quindi non esiste nessuno stato in MQ che sia autostato di tutte le osservabili. In pratica, in MC per gli stati puri, per cui ho informazione massima, possiamo ignorare completamente la struttura probabilistica - poiché in tal caso abbiamo una conoscenza completa del sistema. La probabilità nasceva da una conoscenza incompleta (stati misti). In MQ, tuttavia, anche la massima informazione (stati puri) racchiude una incertezza (Heisenberg).

Questioni aperte:

- Come si possono caratterizzare gli autostati in MQ?
- Quando è possibile misurare 2 osservabili con arbitraria precisione?

2.4 Evoluzione temporale per sistemi isolati

In un sistema isolato c'è invarianza per traslazioni temporali, come conseguenza dell'omogeneità del tempo (fare misure ad istanti diversi genera le stesse distribuzioni di probabilità).

È allora naturale richiedere che l'evoluzione temporale sia descritta in termini di un gruppo corrispondente alle traslazioni, quindi isomorfo a $(\mathbb{R}, +)$ e che l'evoluto al tempo t dell'osservabile O sia data da un'osservabile O_t , che soddisfa $O_0 = O$ e $(O_{t_1})_{t_2} = O_{t_1+t_2}$.

In altre parole: l'evoluzione temporale non dipende dal valore assoluto della variabile tempo, ma solo da “quanto tempo è passato” da un istante iniziale definito arbitrariamente. Ci aspettiamo che tale proprietà intuitiva sia rispettata anche in MQ.

Poiché sperimentalmente posso verificare solo i valori medi (e non le osservabili stesse), è possibile definire un'evoluzione temporale invece che tramite le osservabili, tramite gli stati: invece di considerare $O \rightarrow O_t$, consideriamo $\Sigma \rightarrow \Sigma_t$, purché i valori medi restino invariati. Poniamo allora:

$$\langle O_t \rangle_{\Sigma} = \langle O \rangle_{\Sigma_t}$$

(il valor medio dell'osservabile O al tempo t nello stato Σ lo interpreto come il valor medio di O nello “stato evoluto” Σ_t).

In **MC** l'evoluzione delle osservabili nel formalismo Hamiltoniano una volta data l'Hamiltoniana H è data dalle equazioni di Hamilton:

$$\frac{dq_i(t)}{dt} = \frac{\partial H}{\partial p_i}; \quad \frac{dp_i(t)}{dt} = -\frac{\partial H}{\partial q_i}$$

Da cui un'osservabile $f(q, p)$ si evolve come:

$$\frac{df}{dt}(q(t), p(t)) = \sum_{i=1}^N \frac{\partial f}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial f}{\partial p_i} \frac{dp_i}{dt} = \sum_{i=1}^N \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} = \{f, H\}$$

(Nel penultimo passaggio si è effettuata la sostituzione con le equazioni di Hamilton, mentre $\{f, H\}$ è detta **parentesi di Poisson**).

Lo stato $\rho(q, p)$ invece evolve come

$$\rho_t(q, p) = \rho(q(-t), p(-t)) \quad (2.17)$$

Giustificiamolo matematicamente. Per il **teorema di Liouville**, in un sistema Hamiltoniano, presa una regione Ω nello spazio delle fasi (q, p) , facendo “evolvere” tutti i suoi punti tramite le equazioni di Hamilton per un certo tempo t , si giunge

Omogeneità del tempo

Evoluzione degli stati

Evoluzione in MC

ad una nuova regione $\Omega(t)$ che ha lo stesso volume della prima. In altre parole “gli stati *si evolvono* come un fluido incompressibile” e vale $dq dp = dq(t) dp(t)$. Perciò, il valor medio di un’osservabile f nello stato ρ al tempo t è dato da:

$$\langle f(t) \rangle_\rho = \int dq dp \rho(q, p) f(q(t), p(t))$$

Applicando Liouville si passa a:

$$= \int dq(t) dp(t) \rho(q, p) f(q(t), p(t))$$

Ed effettuando il cambio di variabili (“traslazione temporale”) $q(t), p(t) \mapsto q, p$ e $q, p \mapsto q(-t), p(-t)$:

$$= \int dq dp \rho(q(-t), p(-t)) f(q, p) = \langle f \rangle_{\rho_t}$$

In **MQ**, invece, l’evoluzione temporale è data dalle formule di Born-Heisenberg-Jordan (1925):

$$\frac{dq_{mn}(t)}{dt} = \frac{[q, H]_{mn}}{i\hbar} \quad (2.18)$$

O dalla versione equivalente elaborata da Schrödinger nel 1926:

$$i\hbar \frac{\partial \psi}{\partial t}(t) = H\psi(t) \quad (2.19)$$

Restano aperte le seguenti questioni:

- Qual è la descrizione matematica dell’evoluzione temporale?
- Come si passa da (2.18) a (2.19)?

Capitolo 3

Formulazione assiomatica Hilbertiana

La meccanica quantistica è uno degli ambiti scientifici nei quali la fisica e la matematica più astratta manifestano reciproche influenze. Nella prima metà del novecento diversi matematici e fisici tentarono di costruire una solida base formale capace di assiomatizzare tutto ciò che fino ad allora era solo un fatto sperimentale. La prima formulazione fu quella Hilbertiana (o Standard), elaborata da Hilbert, Dirac (1930) e von Neumann (1932).

Anche se la formulazione di Dirac e quella di von Neumann si possono considerare equivalenti, tuttavia la matematica utilizzata da Dirac (un “autentico” fisico) è diversa da quella utilizzata da von Neumann (un fisico-matematico).

Il formalismo di Dirac è il più seguito dai fisici per la sua maggiore flessibilità e semplicità¹, seppur a costo di molta difficoltà di formalizzazione rigorosa. Solo nel dopoguerra, con la teoria delle distribuzioni, fu possibile giustificare esaurientemente la matematica di Dirac.

Il formalismo di von Neumann è invece quello più seguito dai fisici-matematici, essendo caratterizzato da un maggior rigore matematico e una solida base formale, anche se formalmente è meno flessibile.

3.1 Stati puri

Gli stati puri sono codificati dalle funzioni d'onda $\psi(x)$ di Schrödinger, analogamente agli stati $|\psi\rangle$ stati di polarizzazione, e sono gli stati che comprendono la **massima informazione**.

3.1.1 Spazio degli stati puri e sue proprietà

Esaminiamo dunque le proprietà che vogliamo imporre sull'insieme (lo spazio) degli stati puri.

¹ΛCiò rispecchia il minimalismo di Dirac, il quale riteneva che le teorie potessero essere corrette solo se semplici.

1. Tutti i vettori di stato sono elementi di uno **spazio vettoriale**.

Infatti gli stati possono essere sommati e moltiplicati per un numero complesso, e da questo segue immediatamente il **principio di sovrapposizione**: ogni vettore che rappresenta uno stato puro (detto **vettore di stato** o, con abuso di linguaggio, stato) può essere scritto come sovrapposizione, cioè **combinazione lineare** di altri vettori di stato dello spazio vettoriale a cui appartiene.

Principio di sovrapposizione

2. Lo spazio vettoriale degli stati è uno spazio vettoriale **complesso**.

Per la prima volta in fisica l'algebra naturale degli stati è in \mathbb{C} e non in \mathbb{R} .

3. In realtà abbiamo visto che $\langle X \rangle_\psi$ e $\langle P \rangle_\psi$ non cambiano se sostituiamo ψ con $\alpha\psi$, con $\alpha \in \mathbb{C} \setminus \{0\}$. Infatti, sia $|\psi\rangle$ uno stato, e $|\psi'\rangle = e^{i\varphi} |\psi\rangle$ la sua versione “sfasata” di una fase φ . Prendendo il coniugato di entrambi i membri si ha la relazione per i bra: $\langle\psi'| = e^{-i\varphi} \langle\psi|$.

Vettori definiti a meno di proporzionalità

Si verifica che $|\psi\rangle$ e $|\psi'\rangle$ producono gli stessi valor medi per ogni osservabile O , e quindi devono descrivere lo *stesso* stato:

$$\langle O \rangle_{\psi'} = \langle\psi'| O |\psi'\rangle = \langle\psi| e^{-i\varphi} O e^{i\varphi} |\psi\rangle = \quad (3.1)$$

$$\stackrel{(a)}{=} \langle\psi| e^{-i\varphi} e^{i\varphi} O |\psi\rangle = \langle\psi| O |\psi\rangle = \langle O \rangle_\psi \quad \forall O \quad (3.2)$$

dove in (a) abbiamo usato la linearità dell'osservabile O (che, come vedremo nella prossima sezione, è una delle proprietà necessarie perché un operatore descriva un'osservabile).

Inoltre nella definizione dei valori medi (o delle probabilità) abbiamo dovuto normalizzare le funzioni dividendo per $\|\psi\| = (\int |\psi(x)|^2 dx)^{1/2}$, e quindi necessariamente $0 < \|\psi\| < \infty$, in pratica escludiamo vettori di norma nulla o infinita.

Per queste ragioni è più corretto definire lo spazio vettoriale a cui appartengono le ψ come segue.

Definizione 3 (Raggio vettore). Dato V spazio vettoriale normato, $\psi \in V$, l'insieme $\{\alpha\psi, \alpha \in \mathbb{C} \setminus \{0\}\}$ o equivalentemente $\{e^{i\alpha}\psi, e^{i\alpha} \in S^1, \|\psi\| = 1\}$, è detto **raggio vettore** di V . Quindi ad ogni **stato puro** corrisponde **uno e uno solo** raggio vettore².

Raggio vettore

4. Lo spazio vettoriale V dei vettori di stato ha bisogno di un **prodotto scalare** $\langle \cdot | \cdot \rangle$, che si usa per esempio per calcolare il valor medio della posizione:

Prodotto scalare

$$\langle X \rangle_\psi = \frac{\langle\psi| X \psi\rangle}{\langle\psi| \psi\rangle} = \int \frac{\psi^*(x) x \psi(x)}{\|\psi\|^2} dx$$

Nel caso di un vettore 2-dimensionale il prodotto scalare è definito da $\langle\psi| \phi\rangle = \psi_x^* \phi_x + \psi_y^* \phi_y$. Vogliamo poi che per il momento valga:

$$\langle P \rangle_\psi = \int \frac{\tilde{\psi}^*(p) p \tilde{\psi}(p)}{\|\tilde{\psi}\|^2} dp$$

²Questa in realtà non è una semplice sottigliezza matematica. Tale biiezione “a meno di un fattore moltiplicativo” è alla base dell'esistenza dei fermioni - e perciò della “materia”.

per cui per ogni funzione di stato $\psi(x)$ deve esistere la sua trasformata di Fourier.

Esistenza della trasformata

Uno spazio su \mathbb{R}^n che ha tutte queste proprietà è $L^2(\mathbb{R}^n, d^n x)$. Per definire L^2 si considerano le funzioni “modulo quadro integrabili” secondo Lebesgue, ovvero funzioni f che soddisfano:

$$\int |f(x)|^2 d^n x < \infty$$

(dove l'integrale è quello di Lebesgue). Due funzioni di questo tipo si considerano **equivalenti** se sono uguali quasi ovunque (q.o.) rispetto alla misura di Lebesgue.

Classi di Equivalenza in L^2

Definizione 4 (Spazio vettoriale L^2). Si denota con $L^2(\mathbb{R}^n, d^n x)$ lo spazio delle classi di equivalenza delle funzioni complesse f su \mathbb{R}^n modulo quadro integrabili secondo Lebesgue.

L^2 è dotato del **prodotto scalare** definito come:

$$\langle f, g \rangle = \int d^n x f^*(x) g(x) \quad \forall f, g \in L^2(\mathbb{R}, d^n x)$$

Da qui si nota il perché abbiamo dovuto considerare le classi di equivalenza. Se non lo avessimo fatto $\langle f, f \rangle = 0$ non implicherebbe che $f = 0$, poiché anche funzioni quasi ovunque nulle avrebbero la stessa “norma” pari a 0!

Definizione 5. Uno spazio vettoriale dotato di un **prodotto scalare** è detto **pre-Hilbertiano**. Se è anche **completo** è allora detto di **Hilbert**.

Spazi pre-Hilbertiani, e di Hilbert

3.1.2 Isomorfismi tra spazi di Hilbert

Definizione 6. Due spazi di Hilbert \mathcal{H}_1 e \mathcal{H}_2 sono **isomorfi** se esiste una mappa biiettiva U , detta **trasformazione unitaria**, con $U : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ che preserva il prodotto scalare:

Spazi di Hilbert isomorfi, trasformazioni unitarie

$$\forall \psi, \varphi \in \mathcal{H} \quad \langle U\psi, U\varphi \rangle_{\mathcal{H}_2} = \langle \psi, \varphi \rangle_{\mathcal{H}_1}$$

In altre parole, per ogni coppia di elementi di \mathcal{H}_1 , il loro prodotto scalare (seguendo la “regola” di \mathcal{H}_1) è lo stesso dei loro corrispondenti $U\psi, U\varphi$ in \mathcal{H}_2 (seguendo la “regola” di \mathcal{H}_2).

Problema: perché occorre la completezza per lo spazio delle funzioni d'onda, cioè perché ci serve prendere L^2 ? Potremmo considerare ad esempio per le $\psi \in \mathbb{R}^3$, lo spazio $V = L^2(\mathbb{R}^3) \cap C^2(\mathbb{R}^3)$ come suggerito dall'equazione di Schrödinger:

$$-\frac{\hbar^2}{2m} \nabla^2 \psi + V\psi = \mathcal{E}\psi$$

Infatti siccome nell'equazione si deriva la ψ due volte, è naturale richiedere che ciò sia possibile, mantenendo lo stesso prodotto scalare di L^2 . La risposta a questa domanda è il fatto che lo spazio così definito risulta pre-Hilbertiano, ma non di Hilbert, e dunque manca la completezza. Ma tale proprietà è *necessaria* per alcuni importanti teoremi, e quindi non possiamo rinunciarvi.

Per uno spazio di Hilbert (che è quindi uno spazio **completo**) valgono infatti i due seguenti teoremi:

Teorema 3.1.1. Per ogni base ortonormale di \mathcal{H} è possibile scrivere ogni vettore di \mathcal{H} in modo **univoco** come combinazione lineare (eventualmente infinita) di elementi della base (Teorema di Riesz-Fisher)

Basi ON in spazio di Hilbert

Come vedremo il teorema precedente è cruciale per l'interpretazione statistica della meccanica quantistica.

Teorema 3.1.2. A meno di isomorfismi esiste un **unico spazio di Hilbert complesso** per ogni **cardinalità** della base di vettori di \mathcal{H} .

Cardinalità della base \leftrightarrow unico spazio di Hilbert

Quindi ci sono due casi:

- La cardinalità (di una base) dello spazio di Hilbert \mathcal{H} è di dimensione N finita, per cui \mathcal{H} è isomorfo a \mathbb{C}^N ($\mathcal{H} \cong \mathbb{C}^N$). In questo caso, fissata una base ON di \mathcal{H} , ogni vettore in \mathcal{H} è identificato univocamente da N coordinate complesse, ossia da un vettore in \mathbb{C}^N .
- Se la cardinalità (di una base) di \mathcal{H} è infinita ma numerabile, allora \mathcal{H} è isomorfo a l_2 , ossia lo spazio delle successioni (numerabili) in \mathbb{C} che convergono in modulo-quadro:

Spazio l_2

$$l_2 = \left\{ \{a_n\}_{n \in \mathbb{N}}, a_n \in \mathbb{C} \text{ t.c. } \sum_n |a_n|^2 < \infty \right\}$$

In questo spazio vettoriale è definito il prodotto scalare come estensione di quello in \mathbb{C}^N :

$$(\{a_n\}, \{b_n\})_{l_2} = \sum_n a_n^* b_n$$

Rispetto a questo prodotto scalare l_2 è completo, e quindi è di Hilbert.

Definizione 7 (Spazio separabile). Uno spazio si dice **separabile** se in esso esiste un sottoinsieme **denso** e **numerabile**. Nel caso degli spazi di Hilbert, tale condizione è equivalente ad avere una **base ortonormale numerabile**.

Spazio di Hilbert separabile

Un esempio di spazio separabile è \mathbb{R} , in quanto contiene \mathbb{Q} , che è numerabile e denso in \mathbb{R} (ogni intorno non vuoto di un numero reale contiene almeno un numero razionale^a).

^a^Ciò significa che si può approssimare con precisione arbitraria i reali (non numerabili) partendo solamente dai razionali, che sono numerabili. Analogamente, per uno spazio di Hilbert è possibile approssimare bene qualsiasi suo elemento usando una combinazione lineare a coefficienti razionali di vettori di una base *numerabile*.

Se \mathcal{H} è separabile, per quanto detto prima sugli spazi con basi ON numerabili, è allora isomorfo a \mathbb{C}^N o a l_2 . Data una base ON numerabile $\{\phi_n\}$ di \mathcal{H} , tale isomorfismo si ottiene con:

\mathcal{H} separabile isomorfo a l_2

$$\psi = \sum_n a_n \phi_n \mapsto \{a_n\}_{\mathbb{C}^N, l_2} \quad \forall \psi \in \mathcal{H}$$

Analogamente al caso finito-dimensionale, identifichiamo un vettore $\psi \in \mathcal{H}$ con la successione - che sarà numerabile, essendo \mathcal{H} separabile - dei suoi coefficienti rispetto ad una base ON.

(Lezione 6 del
11/10/2018)

A prima vista sembrerebbe che il numero di funzioni di L^2 sia molto più grande di quante siano le successioni in \mathbb{C}^N . Ma questo è falso perché quando si costruisce L^2 si creano le classi di equivalenza, e il risultato è “ridurre” di molto il numero di elementi. Intuitivamente si può pensare di “ruotare” la base delle funzioni di L^2 (si pensi al caso di \mathbb{C}^N) per passare da una rappresentazione all'altra.

3.1.3 Equivalenza tra gli spazi della MQ

Tramite questi isomorfismi possiamo garantire l'**equivalenza** tra la **meccanica ondulatoria** di **Schrödinger**, dove le funzioni d'onda $\psi(x)$ vivono in $\mathcal{H} = L^2(\mathbb{R}, dx)$ e la **meccanica matriciale** di **Heisenberg** in cui q_{mn}, p_{mn} agiscono come matrici ∞ -dimensionali su l_2 .

Equivalenza tra
Heisenberg e
Schrödinger

Ad ogni elemento $\psi(x)$ dello spazio delle funzioni di Schrödinger è associata una successione $\{a_n\}$ (un vettore ∞ -dimensionale), cioè uno stato dello spazio di Heisenberg. Ad ogni operatore nello spazio di Schrödinger (il quale data una funzione restituisce un'altra funzione) è associata una matrice ∞ -dimensionale che agisce sui vettori ∞ -dimensionali di Heisenberg secondo le regole del prodotto matriciale. Il teorema non fornisce un modo pratico per passare da una all'altra rappresentazione, ma dimostra che questo modo esiste.

Più precisamente data una funzione ψ di Schrödinger, siano ϕ_n le soluzioni dell'equazione stazionaria di Schrödinger per l'oscillatore armonico, che sono anche una base ortonormale di $L^2(\mathbb{R}, dx)$. Allora la mappa di isomorfismo è:

$$L^2(\mathbb{R}, dx) \ni \psi \mapsto \{a_n = (\phi_n, \psi)\}_{i=1}^{\infty} \in l_2$$

Lo stesso teorema garantisce inoltre l'equivalenza tra lo spazio delle **funzioni d'onda** $\psi(x) \in L^2(\mathbb{R}, dx)$ e delle funzioni d'onda del **momento** $\tilde{\psi}(p) \in L^2(\mathbb{R}, dp)$ (sono entrambi spazi della stessa cardinalità, e di conseguenza sono isomorfi), con l'isomorfismo che è definito dalle **trasformate di Fourier** in L^2 .

Equivalenza tra
rappresentazione
in posizione e in
momento

Perché tutta l'informazione fisica sullo stato è contenuta sia in $\psi(x)$ che in $\tilde{\psi}(p)$ che in $\{a_n\}$ evidentemente l'informazione sullo stato non è contenuta in uno spazio di Hilbert “concreto” come $L^2(\mathbb{R}, dx)$ o $L^2(\mathbb{R}, dp)$ o l_2 , ma è contenuta in uno spazio di Hilbert “astratto”³, cioè **definito a meno di isomorfismi**, in cui sono definite

Spazio di Hilbert
“astratto”

³La distinzione tra i due può essere ben compresa facendo riferimento alla nozione di *vettore* in un normale spazio finito-dimensionale, come per esempio \mathbb{R}^3 . Potremmo definire un vettore come una *terna di numeri* (coordinate rispetto ad una base), oppure come le classi di equivalenza della relazione di equipollenza, che associa tra loro tutte le “coppie di punti” di tipo AB (in uno spazio affine) che “hanno la stessa distanza” e puntano nella stessa direzione (che è quella che va da A a B). Chiaramente in questo secondo modo stiamo definendo i vettori *indipendentemente* da una base, come “entità matematica astratta” che semplicemente soddisfa determinate proprietà. Nell'introdurre la notazione dei “ket” Dirac compie un processo simile per la MQ.

le rappresentazioni delle osservabili del sistema quantistico (questa era la visione di **Dirac**).

In altre parole, un qualsiasi spazio di Hilbert “concreto” ha delle informazioni ridondanti, che non ci servono per la rappresentazione fisica. Ma come facciamo a distinguerle?

Con una notazione dovuta a Dirac, un **vettore di stato** dello spazio di Hilbert “**astratto**” è denotato con il simbolo $|\psi\rangle$ e chiamato “**Ket**”.

Ket

Warning: Talora la distinzione tra “astratto” e “concreto” può non essere rispettata (per esempio per ragioni di brevità), e viene data per scontata.

Quando consideriamo spazi concreti $L^2(\mathbb{R}^3, d^3x)$, o $L^2(\mathbb{R}^3, d^3p)$, diremo che i vettori (di stato) e le osservabili che agiscono su di essi sono considerati “in **rappresentazione**”, rispettivamente, \vec{x} e \vec{p} . (Troveremo poi un metodo per “costruire” questi spazi di Hilbert concreti)

*Rappresentazione
 \vec{x} e \vec{p}*

Nella notazione di Dirac, $|\psi\rangle$ in rappresentazione \vec{x} sarà denotato come:

$$\langle \vec{x} | \psi \rangle = \psi(\vec{x}) \in L^2(\mathbb{R}^3, d^3x)$$

In rappresentazione \vec{p} invece si avrà:

$$\langle \vec{p} | \psi \rangle = \tilde{\psi}(\vec{p}) \in L^2(\mathbb{R}^3, d^3p)$$

3.1.4 Funzionali lineari e spazio duale

I **funzionali lineari continui** (se e solo se limitati in quanto sono definiti su uno spazio di Hilbert⁴) F su \mathcal{H} che associano a un vettore un numero complesso $\psi \in \mathcal{H} \mapsto F(\psi) \in \mathbb{C}$ e soddisfano (limitatezza):

$$\|F\| = \sup_{\psi \in \mathcal{H} \setminus \{0\}} \frac{|F(\psi)|}{\|\psi\|} < \infty \quad (3.3)$$

costituiscono uno spazio di Hilbert \mathcal{H}^* , detto duale di \mathcal{H} , e (anti-)isomorfo ad \mathcal{H} con mappa di isomorfismo data da:

*Spazio duale \mathcal{H}^**

$$F \in \mathcal{H}^* \rightarrow \psi_F \in \mathcal{H}, \quad F(\phi) = (\psi_F, \phi) \forall \phi \in \mathcal{H}, \quad \|F\|_{\mathcal{H}^*} = \|\psi_F\|_{\mathcal{H}} \quad (3.4)$$

(tale corrispondenza è garantita dal teorema di Riesz⁵).

In altre parole, applicare un funzionale $F \in \mathcal{H}^*$ a un vettore ϕ in \mathcal{H} equivale a moltiplicare scalarmente il vettore ψ_F “associato a” F tramite la relazione di dualità per ϕ stesso.

Combinando le definizioni (3.3) e (3.4), otteniamo una nuova espressione per la norma del funzionale F :

$$\|F\|_{\mathcal{H}^*} = \|\psi_F\|_{\mathcal{H}} = \sup_{\phi \in \mathcal{H} \setminus \{0\}} \frac{|(\psi_F, \phi)|}{\|\phi\|} \quad (3.5)$$

In notazione di Dirac i funzionali lineari continui su \mathcal{H} astratto sono detti “**Bra**” e denotati come $\langle \psi |$. Applicando un “bra” $\langle \psi | \in \mathcal{H}^*$ ad un “ket” $|\phi\rangle \in \mathcal{H}$:

Bra

⁴ΛCFR pag. 18 [1]

⁵ΛCFR pag. 19 [1]

$$\langle \psi | : | \phi \rangle \rightarrow \langle \psi | \phi \rangle \in \mathbb{C}$$

otteniamo una “**braket**” $\langle \psi | \phi \rangle$ (parentesi). Come vedremo, anche la notazione $\psi(\vec{x}) = \langle \vec{x} | \psi \rangle$ è “quasi” consistente.

3.1.5 Osservazioni sullo spazio \mathcal{S} degli stati puri in MQ

1. Tutte le operazioni **lineari** hanno senso e vanno eseguite in \mathcal{H} e poi eventualmente proiettate nell'**insieme dei raggi vettori** in

Gli stati sono definiti a meno di una fase

$$\mathcal{S} = \frac{\mathcal{H} \setminus \{0\}}{\mathbb{C} \setminus \{0\}}$$

In quanto ogni stato puro è rappresentabile come un elemento di \mathcal{H} *a meno di un complesso non nullo* (come avevamo visto alla sezione 3.1).

Nota: tale proiezione va fatta *per ultima*. Ad esempio, se uno stato è descritto da un $|\phi\rangle$ normalizzato a meno di una fase e $|\psi\rangle$ similmente (con $|\psi\rangle$ e $|\phi\rangle$ non paralleli), $\frac{1}{\sqrt{2}}(|\psi\rangle + |\phi\rangle)$ non descrive uno stato *a meno di due fasi*, ma sempre a meno di una! Questo perché la fase relativa tra $|\psi\rangle$ e $|\phi\rangle$ è fisicamente osservabile. Matematicamente, infatti, i seguenti due stati *non* sono equivalenti:

$$\frac{1}{\sqrt{2}}(|\phi\rangle + |\psi\rangle) \not\equiv \frac{1}{\sqrt{2}}(|\phi\rangle + e^{i\gamma}|\psi\rangle)$$

Anche se (intendendo sempre l'uguaglianza come equivalenza *di stati*):

$$\frac{1}{\sqrt{2}}(|\phi\rangle + |\psi\rangle) \equiv e^{i\alpha} \frac{1}{\sqrt{2}}(|\phi\rangle + |\psi\rangle)$$

dove è importante notare che le uguaglianze o disuguaglianze si intendono come uguaglianze o disuguaglianze *tra stati*.

Insomma, l'idea è che *il vettore di stato finale* a cui si giunge al termine dei conti è definito a meno di una fase, esattamente come tutti i vettori di stato iniziali, se normalizzati.

2. In \mathcal{S} è naturalmente definita una **distanza** (o metrica):

Distanza su \mathcal{S}

$$d(|\psi\rangle, |\phi\rangle) = \left[1 - \frac{|\langle \psi | \phi \rangle|^2}{\|\psi\|^2 \|\phi\|^2} \right]^{\frac{1}{2}}$$

Si noti che vale $d(\alpha|\psi\rangle, |\phi\rangle) = d(|\psi\rangle, |\phi\rangle)$ come dev'essere.

3. \mathcal{H} e \mathcal{S} sono comunque **spazi differenti**, e in quanto tali hanno differenti proprietà topologiche. Per esempio, benché \mathcal{H} sia contraibile a un punto, $\mathcal{S} = \frac{\mathcal{H} \setminus \{0\}}{\mathbb{C} \setminus \{0\}}$ non lo è (in gergo matematico: \mathcal{S} è uno spazio topologicamente non banale).

Ad esempio esistono curve chiuse in \mathcal{S} non riducibili a un punto con continuità (essendo uno spazio proiettivo, intuitivamente può succedere che la “proiezione” di una curva aperta sia chiusa)

Consideriamo una curva chiusa C che descrive un moto ciclico (uno stato che ritorna dopo un po’ su se stesso), e in questo ciclo si può dimostrare che lo stato acquista una fase che dipende solo dalla struttura topologica⁶ di \mathcal{S} e C ed è la cosiddetta **fase di Berry** $e^{\oint_C \langle \psi | d\psi \rangle}$, su cui torneremo in un capitolo futuro.

3.2 Osservabili

Occupiamoci ora di analizzare la descrizione matematica delle osservabili.

Abbiamo già visto che nella descrizione di Heisenberg le osservabili sono ottenute da matrici infinito-dimensionali \vec{q}_{mn} , \vec{p}_{mn} , mentre per Schrödinger le osservabili sono operatori \vec{x} , $-i\hbar\vec{\nabla}$ che agiscono sulle funzioni d’onda ψ e i loro valor medi sono dati da:

$$\langle X \rangle_\psi = \frac{(\psi, \vec{x}\psi)}{\|\psi\|^2}, \quad \langle P \rangle_\psi = \frac{(\psi, -i\hbar\vec{\nabla}\psi)}{\|\psi\|^2}$$

3.2.1 Prime proprietà: linearità e densità del dominio

Per le osservabili dunque richiediamo le seguenti proprietà:

1. **Linearità:** Le osservabili sono descritte da *operatori lineari* su \mathcal{H} (spazio dei vettori di stato). Oltre al fatto che gli operatori comuni (posizione e momento) sono evidentemente lineari, questa proprietà garantisce che i valori medi dipendano dallo stato e non dal vettore scelto per rappresentarlo. Infatti se l’osservabile O è descritta dall’operatore A , con $A : \mathcal{H} \rightarrow \mathcal{H}$, dato $\alpha \in \mathbb{C} \setminus \{0\}$ si ha:

Osservabili = operatori lineari

Valor medio e osservabile

$$\langle O \rangle_{\alpha\psi} = \langle A \rangle_{\alpha\psi} = \frac{(\alpha\psi, A(\alpha\psi))}{\|\alpha\psi\|^2} \stackrel{(a)}{=} \frac{\alpha^2 (\psi, A\psi)}{\alpha^2 \|\psi\|^2} = \langle A \rangle_\psi = \langle O \rangle_\psi$$

dove il passaggio (a) è valido se A è lineare, ovvero se $A(\alpha\psi) = \alpha A\psi$. Pertanto il **valor medio** dell’**osservabile** nello **stato** ψ non cambia se applichiamo A ad una funzione d’onda ψ moltiplicata per un $\alpha \in \mathbb{C}$, dato che $\alpha\psi$ descrive lo stesso **stato** di ψ , esattamente come vogliamo.

2. **Densità del dominio:** se la dimensione di \mathcal{H} è finita ($\dim \mathcal{H} < \infty$) allora il **dominio** di un operatore *lineare* A è certamente \mathcal{H} , ossia A è applicabile senza problemi a tutti i vettori di \mathcal{H} .

Densità del dominio

Questo fatto non vale tuttavia nel caso infinito-dimensionale, cioè per $\dim \mathcal{H} = +\infty$, come si vede dal seguente esempio:

⁶Queste proprietà che sembrano puramente matematiche e astratte hanno in realtà applicazione in fisica - e spiegano alcune proprietà di materiali. In effetti il premio Nobel per la fisica del 2016 fu assegnato proprio per la ricerca sugli effetti fisici di proprietà topologiche.

Esempio 1. Sia X l'operatore "posizione" $X \equiv \hat{x}$, che agisce su ψ come: $X\psi(x) = x\psi(x)$. In generale, se $\psi(x) \in L^2$ non è vero che $D(X) = L^2(\mathbb{R}, dx)$. Possiamo infatti trovare una funzione di L^2 a cui, applicando l'operatore, si ottiene un qualcosa di non definito. Per esempio consideriamo $\psi(x) = \frac{1-e^{-x^2}}{x} \in L^2$. Allora⁷ per $x \sim 0$, $\psi(x) \sim x$, e per $x \sim \infty$ $\psi(x) \sim \frac{1}{x}$, in modo che $\int_{\mathbb{R}} dx \frac{1}{x^2} < \infty$. Ma se applichiamo l'operatore otteniamo $x\psi(x) = 1 - e^{-x^2}$, $x \sim \infty$, e $\psi(x) \sim 1$, per cui $\int_{\mathbb{R}} dx 1 = \infty$ (l'integrale diverge).

Perciò non si può sperare di definire gli operatori su tutti i vettori di \mathcal{H} . Chiediamo allora la cosa più vicina a \mathcal{H} che possiamo chiedere: che $D(A)$ sia **denso** in \mathcal{H} , cioè:

$$\forall \psi \in \mathcal{H}, \exists \{\psi_n\} \subset D(A) \text{ t.c. } \|\psi_n - \psi\| \xrightarrow{n \rightarrow \infty} 0$$

Il significato fisico della densità è la possibilità di "approssimare qualsiasi ψ bene quanto vogliamo restando in $D(A)$ ".

Dominio denso

Partendo da un operatore A lineare e con dominio $D(A)$ denso in \mathcal{H} possiamo estenderlo con unicità a un operatore definito su tutto \mathcal{H} solo se A è *continuo*, cosa che negli spazi di Hilbert è equivalente a essere **limitato**, ossia a verificare la proprietà⁸:

$$\|A\| \equiv \sup_{\psi \in \mathcal{H} \setminus \{0\}} \frac{\|A\psi\|}{\|\psi\|} < \infty$$

Operatori lineari limitati

Lo spazio degli operatori (lineari⁹) limitati su \mathcal{H} è uno spazio vettoriale normato con la norma appena definita, ed è denotato con $\mathcal{B}(\mathcal{H})$ (*bounded*).

In particolare se A descrive un'osservabile, $\|A\|$ ha significato fisico: è il più grande valore in modulo che si possa ottenere misurando A (o meglio, l'osservabile O descritto da A).

Spazio $\mathcal{B}(\mathcal{H})$ degli operatori limitati

Notiamo ora che il dominio degli operatori è strettamente legato al significato fisico delle informazioni che possiamo ricavare da essi. Infatti, se $\psi \in D(A)$, allora il valor medio $\langle A \rangle_\psi$ è ben definito, e in particolare $A\psi \in \mathcal{H}$, $(\psi, A\psi)$ sono ben definiti, per cui:

$$\langle A \rangle_\psi = \frac{(\psi, A\psi)}{\|\psi\|^2}$$

è effettivamente il valor medio (che otterremmo da misure ripetute sperimentali) dell'osservabile descritto da A .

Se invece $\psi \notin D(A)$, in generale $\langle A \rangle_\psi = \frac{(\psi, A\psi)}{\|\psi\|^2}$ **non** è il valor medio.

Operatori al di fuori del loro dominio

Esempio 2 (Operatore energia nella buca ∞ -profonda). Consideriamo l'operatore energia:

$$H = \frac{P^2}{2m} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2}$$

⁷^Sviluppando in serie di MacLaurin: $e^{-x^2} = 1 - x^2 + \dots$

⁸^Pag. 39 di [1]

⁹^In seguito, quando scriveremo "operatori" faremo riferimento implicitamente agli operatori lineari.

e appliciamolo al caso della buca 1-dimensionale infinitamente profonda¹⁰ tra 0 e 1. Allora $D(H) = \{\psi \text{ regolare, e poi } \psi(0) = \psi(1) = 0\}$.¹¹ Prendiamo una $\psi = 1$ in $[0, 1]$, che chiaramente non appartiene a $D(H)$. Allora: $(\psi H, \psi) = 0$ (la derivata seconda di 1 è nulla). Ma le energie della buca infinitamente profonda sono: $E_n \sim n^2, n > 0$ e la media di numeri positivi non può essere un numero nullo, e quindi il risultato che abbiamo ottenuto applicando l'operatore ad una funzione non nel suo dominio porta ad un risultato assurdo!

3.2.2 Definizione mediante gli elementi di matrice

Se $D(A)$ è **denso** in \mathcal{H} (cosa che abbiamo assunto per tutte le osservabili), per determinare A possiamo dare gli “elementi di matrice di A ”, ossia

$$\{(\phi, A\psi), \psi \in D(A), \phi \in D \text{ denso in } \mathcal{H}\}$$

Mostriamo che tale identificazione è univoca. Siano A e A' due operatori, con lo stesso dominio $D(A) = D(A')$, e stessi elementi di matrice $(\phi, A\psi) = (\phi, A'\psi) \forall \psi \in D(A), \forall \phi \in D$ (con D denso in \mathcal{H}). Vogliamo che da ciò derivi che $A = A'$. Scrivendo $(\phi, A\psi) - (\phi, A'\psi) = 0$, prendendo la norma e dividendo per $\|\phi\|$ all'interno di un sup possiamo costruire l'espressione della norma di un operatore $A - A'$ (come data da Riesz in (3.5)):

$$0 = \sup_{\phi \in D \setminus \{0\}} \frac{|(\phi, A'\psi) - (\phi, A\psi)|}{\|\phi\|} = \sup_{\phi \in D \setminus \{0\}} \frac{|(\phi, (A' - A)\psi)|}{\|\phi\|}$$

che per Riesz corrisponde al modulo:

$$\|(A' - A)\psi\| = 0 \quad \forall \psi \in D(A)$$

Ma $D(A)$ è denso e A lineare, quindi $A' = A$.

Notiamo che possiamo esprimere gli elementi di matrice di A per $D = D(A)$ in termini di valori medi utilizzando l'**identità di polarizzazione**:

$$(\phi, A\psi) = \sum_{n=0}^3 \frac{(-i)^n}{4} (\phi + i^n \psi, A(\phi + i^n \psi))$$

In questo modo, riscrivendo $\phi + i^n \psi = \varphi_n \in D(A)$, abbiamo riscritto il prodotto scalare $(\phi, A\psi)$ come somma di norme del tipo $(\varphi_n, A\varphi_n)$ che sono, a meno di un fattore di normalizzazione, interpretabili fisicamente come valori medi (vi è quindi un modo per *misurare* il prodotto scalare tra due vettori).

Elementi di matrice di un operatore con dominio denso
Dimostrazione univocità

Identità di polarizzazione: da prodotto scalare a valori medi

¹⁰^Che può essere immaginata come una “scatola” con pareti impenetrabili, dove il potenziale è $+\infty$

¹¹^L'ipotesi di regolarità è necessaria poiché applicare l'operatore H significa calcolare una derivata seconda, e perciò le ψ in $D(H)$ devono essere derivabili *quasi ovunque* due volte. In generale rimarremo sul vago nella definizione precisa di regolarità, che richiederà essenzialmente che le operazioni che eseguiamo siano ben definite.

Verifichiamo l'identità di polarizzazione. Per prima cosa svolgiamo la sommatoria, tralasciando il fattore $1/4$ per adesso (lo aggiungeremo di nuovo alla fine):

$$(\phi, A\psi) = (\phi + \psi, A(\phi + \psi)) - i(\phi + i\psi, A(\phi + i\psi)) + \\ - (\phi - \psi, A(\phi - \psi)) + i(\phi - i\psi, A(\phi - i\psi))$$

Svolgendo tutti i prodotti arriviamo a:

$$= (\phi, A\phi) + (\phi, A\psi) + (\psi, A\phi) + (\psi, A\psi) \\ - i(\phi, A\phi) - i(\phi, iA\psi) - i(i\psi, -iA\phi) - i(i\psi, iA\psi) \\ - (\phi, A\phi) - (\phi, -A\psi) - (-\psi, A\phi) - (-\psi, -A\psi) \\ + i(\phi, A\phi) + i(\phi, -iA\psi) + i(-i\psi, A\phi) + i(-i\psi, -iA\psi)$$

dove i termini evidenziati si elidono.

Osserviamo ora, che per proprietà del prodotto scalare, se $a \in \mathbb{C}$, $a(A, B) = (A, aB)$, $(aA, B) = (A, a^*B)$ (dove l'asterisco indica il complesso coniugato).

Così facendo possiamo effettuare alcune semplificazioni. Per esempio

$$i(-i\psi, iA\phi) = i(\psi, -A\phi) = -i(\psi, A\phi)$$

e analogamente per gli altri termini.

Giungiamo a:

$$= (\phi, A\psi) + (\psi, A\phi) + (\psi, A\psi) \\ + (\phi, A\psi) - (\psi, A\phi) - i(\psi, A\psi) \\ + (\phi, A\psi) + (\psi, A\phi) - (\psi, A\psi) \\ + (\phi, A\psi) + (\psi, -A\phi) + i(\psi, A\psi)$$

Notiamo che tutti i termini tranne quelli della prima colonna si elidono, e il risultato è $4(\phi, A\psi)$. Riportando il fattore $1/4$ prima trascurato si giunge al prodotto scalare $(\phi, A\psi)$, e pertanto l'identità è dimostrata.

3.2.3 Valori medi reali: aggiunto e operatori simmetrici

Un'altra proprietà degli operatori che rappresentano osservabili viene dalla richiesta che i **valori medi siano reali**. Matematicamente ciò significa che il complesso coniugato del valor medio sia il valor medio stesso¹²:

$$(\psi, A\psi) = (\psi, A\psi)^* = (A\psi, \psi)$$

Specifichiamo questa scrittura con la definizione di aggiunto.

Definizione 8 (Aggiunto di un operatore). Sia A con $D(A)$ denso, allora l'**aggiunto** di A , denotato con A^\dagger , è un operatore definito sul dominio:

Valori medi reali

Aggiunto di un operatore

¹²Essendo il prodotto scalare $(\psi, \phi) = \int \psi^* \phi$, si ha $(\psi, A\psi)^* = (\int \psi^* A\psi) = \int \psi (A\psi)^* = (A\psi, \psi)$

$$D(A^\dagger) = \left\{ \phi \in \mathcal{H} \text{ t.c. } \sup_{\psi \in D(A)} \frac{|(\phi, A\psi)|}{\|\psi\|} < \infty \right\}$$

e tale che

$$(A^\dagger \phi, \psi) = (\phi, A\psi) \quad \forall \psi \in D(A), \phi \in D(A^\dagger)$$

Si noti che per A^\dagger , l'insieme denso $D(A)$ svolge il ruolo di D (denso in \mathcal{H}) nella definizione degli elementi di matrice di A^\dagger , che perciò sono ben definiti.

Giustificiamo la scelta del dominio di A^\dagger . Si ha che $\forall \psi \in D(A^\dagger) \ A^\dagger \phi \in \mathcal{H}$. Ciò significa che:

$$\infty > \|A^\dagger \phi\| \stackrel{(a)}{=} \sup_{\psi \in \mathcal{H} \setminus \{0\}} \frac{|(A^\dagger \phi, \psi)|}{\|\psi\|} \stackrel{(b)}{=} \sup_{\psi \in D(A)} \frac{|(A^\dagger \phi, \psi)|}{\|\psi\|} \stackrel{(c)}{=} \sup_{\psi \in D(A)} \frac{|(\phi, A\psi)|}{\|\psi\|}$$

dove in (a) si è applicata la definizione della norma (per teorema di Riesz), mentre (b) è giustificato dal fatto che $D(A)$ è denso in \mathcal{H} , e (c) deriva dalla definizione di operatore aggiunto A^\dagger .

Nota: se $\dim \mathcal{H} < \infty$ allora la matrice di A^\dagger (cioè la matrice composta dai suoi elementi di matrice) è la matrice trasposta e complesso-coniugata della matrice A :

$$(A\psi, \phi) = \sum_{i,j=1}^N (A_{ji}\psi_i)^* \phi_j = (\psi, A^\dagger \phi) = \sum_{i,j=1}^N \psi_i^* A_{ij}^* \phi_j \quad \Rightarrow \quad A_{ij}^\dagger = A_{ji}^*$$

Se allora consideriamo un'osservabile O , descritta da A operatore lineare, con dominio $D(A)$ denso, affinché i valori medi siano reali $\langle A \rangle_\psi \in \mathbb{R}$, dobbiamo poter definire A^\dagger aggiunto in $D(A^\dagger)$, ossia:

$$(A^\dagger \phi, \psi) = (\phi, A\psi) \quad \forall \psi \in D(A) \quad \forall \phi \in D(A^\dagger) \quad (3.6)$$

Inoltre, è necessario imporre che tra operatore A e il suo aggiunto A^\dagger esista una relazione precisa, che definiamo successivamente.

Definizione 9 (Estensione di un operatore). Dato un operatore A , un operatore B si dice un'estensione di A , e si indica con $A \subseteq B$, se $D(A) \subseteq D(B)$ e se vale

$$A\psi = B\psi \quad \forall \psi \in D(A)$$

Definizione 10 (Operatore simmetrico). A si dice **simmetrico** se $A \subseteq A^\dagger$, ovvero se il suo aggiunto è una sua estensione.

*Operatore
simmetrico*

Per verificare che A è simmetrico partendo da (3.6) è sufficiente verificare:

$$(\phi, A\psi) = (A\phi, \psi) \quad \forall \phi, \psi \in D(A)$$

Da cui segue immediatamente che se $\phi \in D(A)$ allora $\phi \in D(A^\dagger)$, e quindi $D(A) \subseteq D(A^\dagger)$, come richiesto.

Teorema 3.2.1. *Se A è simmetrico allora per ogni $\psi \in D(A)$ si ha $(\psi, A\psi) \in \mathbb{R}$. In particolare ciò significa che un operatore simmetrico produce valori medi **reali**.*

Operatori
simmetrici hanno
valori medi reali

Dimostrazione. Sia A simmetrico, e quindi $D(A) \subseteq D(A^\dagger)$ e $A^\dagger\psi = A\psi \ \forall \psi \in D(A)$. Moltiplicando scalarmente entrambi i membri di quest'ultima equazione per ψ otteniamo:

$$(\psi, A\psi) \underset{(a)}{=} (\psi, A^\dagger\psi) \underset{(b)}{=} (A\psi, \psi) = (\psi, A\psi)^*$$

Nel passo (a) si ha che $\psi \in D(A)$ è anche $\psi \in D(A^\dagger)$, essendo $D(A) \subseteq D(A^\dagger)$, e ciò rende ben definita l'uguaglianza $A^\dagger\psi = A\psi$ della simmetria, che qui applichiamo. In (b), invece, applichiamo la definizione di aggiunto. Ma allora, per come è definito il prodotto scalare, otteniamo il coniugato di $(\psi, A\psi)$.

Perciò $(\psi, A\psi)$ è uguale al suo coniugato, e quindi è reale. \square

(Lezione 7 del
12/10/2018)

3.2.4 Proiettori e operatori unitari

Utilizzando la definizione di **aggiunto** si possono definire due importanti **classi** di operatori limitati, che ci torneranno utili più avanti: i proiettori e gli operatori unitari.

Proiettori

Definizione 11 (Proiettore). Un operatore limitato $P \in \mathcal{B}(\mathcal{H})$ è un proiettore se $P = P^\dagger$ e $P = P^2$.

Un proiettore P_ψ non è altro che quell'operatore che applicato ad un vettore ne restituisce la sua proiezione lungo ψ (intesa come vettore¹³). In particolare si noti che ogni vettore $\psi \in \mathcal{H}$ con $\|\psi\| = 1$ definisce (tramite Riesz) un proiettore P_ψ

$$P_\psi\phi = (\psi, \phi)\psi \quad \forall \phi \in \mathcal{H} \quad (3.7)$$

In notazione di Dirac $P_\psi = |\psi\rangle\langle\psi|$ e $P_\psi|\phi\rangle = |\psi\rangle\langle\psi|\phi\rangle$

L'analogo dei proiettori nelle funzioni sono le funzioni caratteristiche χ_Δ . Infatti tali funzioni sono a valori reali (e quindi la simmetria è immediatamente verificata) e hanno come valori solo 0 e 1, che non cambiano dopo essere stati elevati al quadrati¹⁴ (e quindi soddisfano anche $\chi_\Delta^2 = \chi_\Delta$)¹⁵.

Operatori unitari

Definizione 12 (Operatore unitario). Un operatore **limitato** $U \in \mathcal{B}(\mathcal{H})$ è unitario se soddisfa:

$$UU^\dagger = U^\dagger U = \mathbb{I} \quad (3.8)$$

¹³^Nel senso, non come semplice “lunghezza della proiezione”, ma come un vettore che punta nella direzione su cui si proietta e che viene riscalo alla lunghezza della proiezione.

¹⁴^In particolare sono gli unici numeri reali con questa proprietà.

¹⁵^In effetti moltiplicando scalarmente una funzione caratteristica χ_Δ con una qualsiasi altra funzione si ottiene la “proiezione” di quest'ultima sull'intervallo Δ , cioè una funzione che è nulla ovunque e vale quanto la funzione originaria entro Δ . Ipoteticamente, se scomponessimo \mathbb{R} in intervalli e sommassimo le proiezioni su ciascuno di essi generate dalle funzioni caratteristiche riatterremo la funzione di partenza, così come sommando le proiezioni di un vettore sui vettori di una base si riottiene il vettore di partenza.

Nota: in $\dim \mathcal{H} = \infty$ nella definizione devono essere verificate entrambe le uguaglianze, in quanto potrebbero esserci operatori U per cui $UU^\dagger = \mathbb{I}$ ma non $U^\dagger U = \mathbb{I}$.

In generale se dato $A \in B(\mathcal{H}) \exists B \mid AB = BA = \mathbb{I}$, diciamo che B è l'**inverso** di A e lo denotiamo con A^{-1} .

Si ha perciò che un operatore è unitario se e solo se il suo aggiunto coincide con il suo inverso:

$$U \text{ unitario} \Leftrightarrow U^\dagger = U^{-1}$$

Nelle funzioni a variabile complessa l'analogo dell'aggiunto è il complesso coniugato. Perciò le funzioni "unitarie" sono della forma $e^{i\alpha(x)}$ (il loro complesso coniugato $e^{-i\alpha(x)}$ è effettivamente pari al loro inverso).

3.2.5 Autoaggiuntezza: uguaglianza dei domini di A e A^\dagger

Diversi indizi mostrano che la simmetria non basta per definire un operatore da poter associare ad un'osservabile quantistica. L'esempio seguente mostra come sia necessaria una proprietà più forte che include la simmetria: l'autoaggiuntezza. Consideriamo "un candidato" per l'operatore momento, che denominiamo P_0 , in $\mathcal{H} = L^2([0, 1], dx)^{16}$:

$$P_0 = -i\hbar \frac{d}{dx}; \quad D(P_0) = \{\psi \in \mathcal{H} \mid \psi(0) = \psi(1) = 0 \text{ e regolari}\}$$

P_0 è l'operatore del momento, e per il suo dominio abbiamo scelto tutte le funzioni sufficientemente regolari (in modo che la derivata sia definita) che si annullano in 0 e 1 (essendo qui il potenziale pari a ∞).

Dimostriamo che P_0 è simmetrico, ossia che $P_0 \subseteq P_0^\dagger$, e che $\forall \phi, \psi \in D(P_0)$ si ha $(\psi, P_0\phi) = (P_0\psi, \phi)$

$$\begin{aligned} (\psi, P_0\phi) &= \int_0^1 \psi^*(x) \left(-i\hbar \frac{d}{dx} \phi \right) (x) dx = \\ &\stackrel{(a)}{=} -i\hbar (\psi^* \phi)(x) \Big|_0^1 + i\hbar \int_0^1 \frac{d}{dx} \psi^*(x) \phi(x) dx = (P_0\psi, \phi) \end{aligned}$$

In (a) si è integrato per parti, e il termine evidenziato è nullo in quanto $\psi(0) = 0$ e $\psi(1) = 0$, per come abbiamo scelto il dominio $D(P_0)$ (in particolare, notiamo che perché si annulli non è necessario richiedere alcuna condizione su $\phi \in D(P_0^\dagger)$, se non la regolarità). Quali sarebbero i possibili risultati di misura del momento di P_0 ? Basta risolvere l'equazione agli autovalori:

$$-i\hbar \frac{d}{dx} \psi_\lambda(x) = \lambda \psi_\lambda(x) \Rightarrow \psi_\lambda(x) = c e^{\frac{i}{\hbar} \lambda x}$$

Ma imponendo la $\psi_\lambda(0) = 0$ si ottiene $c = 0$, dunque non esiste alcuna ψ_λ che non sia identicamente nulla.

¹⁶ Tale spazio corrisponde fisicamente a quello delle funzioni d'onda nella buca di potenziale 1-dimensionale di profondità infinita tra 0 e 1, ossia alla "scatola" con pareti impenetrabili in una dimensione

*Esempio:
l'operatore
momento e la
necessità di
operatori
autoaggiunti*

In realtà a noi interessa solo che il termine evidenziato sparisca, in modo che il momento sia un operatore simmetrico. Perciò, considerando sia ψ che ϕ che si annullano ai bordi abbiamo imposto una condizione molto forte, quando basterebbe imporre che $\psi(0) = \psi(1)$. Definiamo quindi una “nuova versione” dell’operatore P_0 , che chiamiamo P per distinguerla da quella vecchia, in cui il dominio è allargato considerando la nuova condizione (meno impositiva):

$$P = -i\hbar \frac{d}{dx}; \quad D(P) = \{\psi \mid \psi(0) = \psi(1), \psi \text{ regolare}\}$$

Analogamente a prima dobbiamo verificare che P sia simmetrico $(\psi, P\phi) = (P\psi, \phi)$, con $\psi, \phi \in D(P)$, e $D(P) \subseteq D(P^\dagger)$.

Ma perché il termine evidenziato si annulli ora non basta più la condizione specificata sulle ψ . Infatti (essendo $\psi(0) = \psi(1) \Rightarrow \psi(0)^* = \psi(1)^*$ per ipotesi):

$$(\psi^* \phi)_0^1 = \psi^*(1)\phi(1) - \psi^*(0)\phi(0) = 0 = \psi(0)^*(\phi(1) - \phi(0))$$

Bisogna quindi imporre una condizione anche sulle ϕ , e in particolare che $\phi(0) = \phi(1)$, perciò:

$$D(P^\dagger) = \{\phi \mid \phi(0) = \phi(1), \text{ e } \phi \text{ regolare}\}$$

Notiamo allora che $D(P) = D(P^\dagger)$, ed essendo P simmetrico possiamo scrivere $P = P^\dagger$.

Nota: Per P_0 avevamo imposto solo la regolarità per le ϕ , e quindi il dominio di P_0^\dagger era effettivamente più grande di quello di P_0 . Ora, stringendo il dominio da una parte, si è allargato quello dall’altra. Qui abbiamo ristretto P^\dagger e allargato P , e alla fine abbiamo ottenuto che i due hanno lo stesso dominio.

Se ripetiamo per P il calcolo dello spettro:

$$-i\hbar \frac{d}{dx} \psi_\lambda(x) = \lambda \psi(x)$$

la cui soluzione, imponendo che $\psi(0) = \psi(1)$ come da richiesta del dominio, porta a:

$$1 = e^{i\frac{\lambda}{\hbar}0} = e^{i\frac{\lambda}{\hbar}1}; \quad \lambda_n = 2\pi n\hbar$$

Stavolta, P non è solo simmetrico, ma soddisfa anche $P = P^\dagger$.

Definizione 13 (Operatore autoaggiunto). Un operatore A simmetrico si dice **autoaggiunto** se $A = A^\dagger$, ovvero se $D(A) = D(A^\dagger)$ e

$$A\psi = A^\dagger\psi \quad \forall \psi \in D(A)$$

*Operatore
autoaggiunto*

In effetti, nel processo di “allargare un dominio” e “accorciare quello del suo aggiunto” l’unica “situazione speciale” è quella in cui entrambi i domini coincidono. Nel definire operatori, richiedere che sia soddisfatta questa condizione (ossia che siano autoaggiunti e non semplicemente simmetrici) è *matematicamente soddisfacente* dunque ciò sembra naturale e non arbitrario. Il fatto che la stessa condizione si ritrovi esaminando “entità fisiche” - come abbiamo visto nell’esempio del momento (è l’unico modo perché sia simmetrico e dia risultati sensati) - è uno di quei

casi speciali in cui la fisica segue le nostre idee di “regole naturali”.

Tuttavia, seppur tutto questo sia matematicamente naturale e soddisfacente, è necessaria una giustificazione “fisica” del perché succede. Una spiegazione di ciò emerge da una proprietà naturale che non abbiamo ancora richiesto.

3.2.6 Autoaggiuntezza in dimensione N : decomposizione spettrale

Tra le richieste per le osservabili nella descrizione matematica di un sistema fisico c’era la possibilità di definire funzioni di osservabili $f(O)$ dell’osservabile O . Qual è la relazione tra richiedere ciò e la proprietà dell’autoaggiuntezza?

Operatori autoaggiunti e funzioni degli osservabili

Partiamo dal caso **finito dimensionale**.

Se $\dim \mathcal{H} < \infty$, allora ogni A simmetrico è anche autoaggiunto, perché $D(A) = \mathcal{H} = D(A^\dagger)$.

Verifichiamo che gli autovalori di A sono proprio gli autovalori ottenibili dall’osservabile descritto da A .

Autovalori di A autoaggiunto sono anche autovalori dell’osservabile

Ricordiamo che se il valor medio dell’osservabile O nello stato Σ è $\lambda = \langle O \rangle_\Sigma$, e la fluttuazione di O nello stato Σ è nulla:

$$(\Delta O)_\Sigma^2 = \langle (O - \langle O \rangle_\Sigma)^2 \rangle_\Sigma = 0 \quad (3.9)$$

allora diciamo che λ è **autovalore** di O e ovviamente $\lambda \in \sigma(O)$.

Prendendo $\|\psi\| = 1$ (ψ normalizzato), in MQ scriviamo che $(\psi, A\psi) = \lambda$. Allora la condizione di λ autovalore (3.9) si scrive come¹⁷:

$$0 = (\psi, [A - (\psi, A\psi)]^2 \psi) =$$

Vogliamo verificare che da essa discenda che λ sia autovalore di A .

Sappiamo che $A = A^\dagger$ (autoaggiunto) e in particolare $(\psi, A\psi) \in \mathbb{R}$.

“Trasportando indietro” uno dei due fattori del quadrato, per la definizione di *aggiunto*, scriviamo:

$$= ([A^\dagger - (\psi, A\psi)^*] \psi, [A - (\psi, A\psi)] \psi) =$$

Poiché A è autoaggiunto, e quindi $A = A^\dagger$:

$$\begin{aligned} &= ([A - (\psi, A\psi)] \psi, [A - (\psi, A\psi)] \psi) = \left\| \left[A - \underbrace{(\psi, A\psi)}_\lambda \right] \psi \right\|^2 = \|(A - \lambda) \psi\|^2 = 0 \\ &\Rightarrow A\psi = \lambda\psi \quad (3.10) \end{aligned}$$

Quindi si ha che da $A\psi = \lambda\psi$, dove ψ è un autovettore in senso “matematico”, deriva immediatamente che $(\Delta O)_\Sigma^2 = 0$, e quindi ψ è anche autovettore in senso

¹⁷∧ Nella differenza $A - (\psi, A\psi)$, in cui apparentemente stiamo sottraendo un numero da una matrice, si intende che $(\psi, A\psi)$ è opportunamente moltiplicato per la matrice identità \mathbb{I} , che ometteremo per brevità di notazione.

“fisico”, cioè costituisce autostato di un’osservabile. Segue che tutti gli autovalori *in senso matematico* sono autovalori *in senso fisico*, cioè che tra i valori $\lambda \in \sigma(O)$ si trovano gli autovalori (matematici) di A .

$$\sigma(A) \supseteq \{\lambda \text{ autovalori dell'operatore } A\}$$

Resta da dimostrare se tali autovalori “matematici” coincidono con *tutti e soli* gli autovalori (in senso fisico).

Teorema 3.2.2. *Se ψ_n e ψ_m sono autovettori (anche se $\dim \mathcal{H} = \infty$) di A simmetrico, appartenenti agli autovalori λ_n, λ_m , con $\lambda_n \neq \lambda_m$, allora $(\psi_n, \psi_m) = 0$*

*Decomposizione
spettrale*

Dimostrazione. Dalla definizione di autovettore ψ_m di A con autovalore λ_m si ha che $A\psi_m = \lambda_m\psi_m$ e dunque $(\psi_n, A\psi_m) = \lambda_m(\psi_n, \psi_m)$. Poiché A è simmetrica, vale $(\psi, A\phi) = (A\psi, \phi)$, e poiché anche ψ_n è autovettore si ha:

$$(\psi_n, A\psi_m) = (A\psi_n, \psi_m) = \lambda_n(\psi_n, \psi_m)$$

Unendo le due uguaglianze (essendo $(\lambda_n - \lambda_m) \neq 0$) si ottiene:

$$\lambda_m(\psi_n, \psi_m) = \lambda_n(\psi_n, \psi_m) \Rightarrow 0 = (\lambda_n - \lambda_m)(\psi_n, \psi_m) \Rightarrow (\psi_n, \psi_m) = 0$$

□

Si noti che in $\dim \mathcal{H} < \infty$ ogni A simmetrico è anche hermitiano (autoaggiunto).

Per le matrici hermitiane (ossia auto-aggiunte, uguali alla propria trasposta coniugata) sappiamo (dal teorema spettrale) che possiamo scegliere gli autovettori ψ_n come una base ON di \mathcal{H} . Ciò implica che:

*Proprietà di
operatori
autoaggiunti*

1. $\forall \psi \in \mathcal{H} \quad \psi = \sum_n (\psi_n, \psi) \psi_n$ cioè ogni vettore $\psi \in \mathcal{H}$ può essere espresso in maniera univoca come somma delle sue proiezioni sugli autovettori ψ_n ortogonali.

In notazione di Dirac ciò è dato dalla **completezza**:

$$|\psi\rangle = \sum_n |\psi_n\rangle \langle \psi_n | \psi \rangle$$

Con $\sum_n |\psi_n\rangle \langle \psi_n| = \mathbb{I}$

2. $A\psi = \sum_n (\psi_n, A\psi) \psi_n \stackrel{(a)}{=} \sum_n (A\psi_n, \psi) \psi_n \stackrel{(b)}{=} \sum_n \lambda_n (\psi_n, \psi) \psi_n$

dove in (a) si è usata la simmetria di A , e in (b) la definizione di ψ_n autovettore di A con autovalore λ_n ($A\psi_n = \lambda_n\psi_n$).

Applicando quanto appena visto al calcolo dei valor medi dell’osservabile descritto da A nello stato ψ :

$$\langle A \rangle_\psi = \frac{(\psi, A\psi)}{\|\psi\|^2} = \sum_n \lambda_n \frac{(\psi_n, \psi)(\psi, \psi_n)}{\|\psi\|^2} = \sum_n \lambda_n \frac{|(\psi_n, \psi)|^2}{\|\psi\|^2}$$

(che è consistente perché $(\psi_n, \psi_m) = 0$ se $\lambda_n \neq \lambda_m$), dove i λ_n sono autovalori, per semplicità qui assunti non degeneri, e il restante fattore $p_n = \frac{|\psi_n, \psi|^2}{\|\psi\|^2}$ è compreso tra 0 e 1, e lo interpretiamo come probabilità.

Ma allora:

$$\sum_n \frac{|\langle \psi_n, \psi \rangle|^2}{\|\psi\|^2} = \sum_n \frac{\langle \psi | \psi_n \rangle \langle \psi_n | \psi \rangle}{\langle \psi | \psi \rangle} = 1$$

Poiché allora per il teorema spettrale la somma delle probabilità associate agli autovalori è esattamente 1, si ha che gli autovalori esauriscono tutte le possibilità.

Quindi:

$$\begin{aligned} \sigma(A) &= \{\text{insieme dei valori che posso ottenere misurando } A\} = \\ &= \{\lambda_n : \text{insieme degli autovalori di } A\} \end{aligned}$$

Gli autovalori di A sono tutti e soli quelli “fisici” di O

3. Possiamo sfruttare la decomposizione spettrale per definire funzioni di A . Partiamo definendo P_n come il proiettore associato a ψ_n , che se applicato ad un vettore lo proietta nel sottospazio 1-dimensionale generato da ψ_n . Possiamo riscrivere gli elementi di matrice di A usando P_n :

Funzioni $f(A)$

$$(\psi, A\phi) = \sum_n \lambda_n (\psi, P_n \phi) = \left(\psi, \left(\sum_n \lambda_n P_n \right) \phi \right) \Rightarrow A = \sum_n \lambda_n P_n$$

Abbiamo perciò ottenuto A come la somma delle sue “proiezioni” $\lambda_n P_n$.

Ma allora data una funzione reale f , possiamo “calcolarla in A ” calcolandola su tutte le *proiezioni* e sommando i risultati, nel seguente modo:

$$f(A) = \sum_n f(\lambda_n) P_n \quad (3.11)$$

E lo spettro $\sigma(f(A)) = f(\sigma(A))$. In questo modo giustifichiamo anche in MQ la richiesta che avevamo fatto nei paragrafi di descrizione matematica di un sistema fisico, quando si era supposta l'esistenza di funzioni $g(O)$ di un osservabile O . Poiché qui l'osservabile è descritto dall'operatore A , se la corrispondenza tra i due è ben definita dobbiamo poter definire funzioni $f(A)$, che sono appunto descritte nel modo appena presentato.

Riepilogando:

- Le osservabili sono descritte da operatori A lineari (in modo che dipendano dallo stato e non dal vettore usato per rappresentarlo), con dominio $D(A)$ denso in \mathcal{H} (vogliamo che siano definiti per “tutti” gli stati, e gli altri li vogliamo poter approssimare)
- Gli operatori A sono **simmetrici** ($\forall \psi \in D(A)$ si ha $A^\dagger \psi = A\psi$), in modo che i valori medi siano reali
- Gli operatori A sono **autoaggiunti** ($A = A^\dagger$ con $D(A) = D(A^\dagger)$) in modo che si possano definire funzioni di osservabili.

Esempio numerico. Proviamo a fare un esempio concreto di decomposizione spettrale, in $\mathcal{H} = \mathbb{C}^3$. Un operatore lineare su \mathbb{C}^3 è quindi dato da una matrice 3×3 a valori complessi. Consideriamo A data da:

Esempio di decomposizione spettrale

$$A = \begin{pmatrix} 2 & -i & i \\ i & 2 & 1 \\ -i & 1 & 2 \end{pmatrix}$$

Si verifica immediatamente che A è uguale alla sua trasposta coniugata, e quindi definisce un operatore simmetrico, che è anche autoaggiunto (essendo \mathbb{C}^3 finito-dimensionale): $A = A^\dagger$.

1. Determiniamo gli **autovalori** λ di A . Da $\det(A - \lambda \mathbb{I}) \stackrel{!}{=} 0$, otteniamo (sviluppo di Laplace lungo la prima colonna)

$$(2-\lambda)[(2-\lambda)^2-1]-i(-2i+\lambda i-i)-i(-i-2i+\lambda i) = -\lambda^3+6\lambda^2-9\lambda = -\lambda(\lambda-3)^2 = 0$$

Perciò gli autovalori sono $\lambda_1 = 0$ (senza degenerazione) e $\lambda_2 = 3$ (con degenerazione 2).

2. Troviamo gli **autovettori** associati a λ_1, λ_2 .

- Per $\lambda_1 = 0$, avremo un autovettore $|\psi_0\rangle$

$$\begin{aligned} (A - 0\mathbb{I}) \begin{pmatrix} a \\ b \\ c \end{pmatrix} &= 0 \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 2 & -i & i \\ i & 2 & 1 \\ -i & 1 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} 2a - ib + ic = 0 & (a) \\ ia + 2b + c = 0 & (b) \\ -ia + b + 2c = 0 & (c) \end{cases} \\ &\Rightarrow \begin{cases} a = ib & (a \text{ dopo sost.}) \\ c = -b & (b + c) \end{cases} \\ &\Rightarrow |\psi_0\rangle = N \begin{pmatrix} i \\ 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{3}} \begin{pmatrix} i \\ 1 \\ -1 \end{pmatrix} \end{aligned}$$

Dove abbiamo scelto di normalizzare il vettore a 1 (cioè in modo che $\langle \psi_0 | \psi_0 \rangle = 1$)

- Per $\lambda_2 = 3$, avremo due autovettori $|\psi_{3,1}\rangle, |\psi_{3,2}\rangle$

$$(A - 3\mathbb{I}) \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0 \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -1 & -i & i \\ i & -1 & 1 \\ -i & 1 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} -a - ib + ic = 0 & (a) \\ ia - b + c = 0 & (b) \\ -ia + b - c = 0 & (c) \end{cases} \Rightarrow -ia + b - c = 0$$

(Notiamo infatti che le a-b-c sono equazioni equivalenti, infatti (c) = -(b) e (a) = i(b), quindi basta prendere una qualunque tra le tre. In questo caso otteniamo un autospazio di dimensione massima (2), pari alla degenerazione dell'autovalore λ_2).

Riscrivendo $c = -ia + b$, possiamo prendere un autovettore del tipo:

$$|\psi_3\rangle = \begin{pmatrix} a \\ b \\ -ia + b \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \\ -i \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Cerchiamo una qualsiasi combinazione di a e b per ottenere un vettore che sia normale a $|\psi_0\rangle$. Prendendo per esempio $a = 0$ e $b = 1$ e normalizzando otteniamo:

$$|\psi_{3,2}\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

E si ha immediatamente che $\langle\psi_{3,2}|\psi_0\rangle = 0$.

Cerchiamo ora un altro autovettore nello stesso autospazio che sia ortogonale a $\langle\psi_{3,1}\rangle$:

$$\langle\psi_{3,1}|\psi_{3,2}\rangle = 0 \Rightarrow \begin{pmatrix} a & b & -ia + b \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 0 \Rightarrow b = i\frac{a}{2}$$

Che, assieme alla condizione $c = -ia + b$ di prima, per $a = 2$ produce:

$$|\psi_{3,1}\rangle = N \begin{pmatrix} 2 \\ i \\ i - 2i \end{pmatrix} = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ i \\ -i \end{pmatrix}$$

3. Otteniamo ora i **proiettori** a partire dagli autovalori. Partendo dalla definizione di (3.7) in notazione di Dirac, e riconoscendo che in \mathbb{C}^3 un “bra” è

un vettore-riga 1×3 (essendo un funzionale che agisce su vettori di dim. 3) mentre un “ket” è un normale vettore-colonna 3×1 , basta svolgere dei prodotti matriciali, che daranno come risultato un operatore (ossia una matrice 3×3 , dato che moltiplichiamo $(3 \times 1) \cdot (1 \times 3)$):

$$P_0 = |\psi_0\rangle \langle \psi_0| = \frac{1}{\sqrt{3}} \begin{pmatrix} i \\ 1 \\ -1 \end{pmatrix} \frac{1}{\sqrt{3}} (-i \quad 1 \quad -1) = \frac{1}{3} \begin{pmatrix} 1 & i & -i \\ -i & 1 & -1 \\ i & -1 & 1 \end{pmatrix}$$

$$P_{3,1} = |\psi_{3,1}\rangle \langle \psi_{3,1}| = \frac{1}{6} \begin{pmatrix} 4 & -2i & 2i \\ 2i & 1 & -1 \\ -2i & -1 & 1 \end{pmatrix}$$

$$P_{3,2} = |\psi_{3,2}\rangle \langle \psi_{3,2}| = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

Nota: moltiplicando ciascun proiettore per il suo autovalore riotteniamo l'operatore iniziale (come ci si aspetta, dato che stiamo sommando le sue “proiezioni” - intese come operatori):

$$0 \cdot P_0 + 3 \cdot P_{3,1} + 3 \cdot P_{3,2} = A$$

4. Si usa infine la decomposizione spettrale per calcolare funzioni di A , tramite la formula in (3.11). Per esempio:

$$e^A = e^0 P_0 + e^3 P_{3,1} + e^3 P_{3,2}$$

$$\cos A = \cos 0 P_0 + \cos 3 P_{3,1} + \cos 3 P_{3,2}$$

(Lezione 8 del
15/10/2018)

3.2.7 Autoaggiuntezza in dimensione ∞ : famiglie spettrali

Potremmo essere tentati di estendere questa strategia per costruire una funzione di un'osservabile O ($f(O)$) anche nel caso $\dim \mathcal{H} = \infty$, per esempio in $\mathcal{H} = L^2(\mathbb{R}, dx)$. Tuttavia, ciò porta subito a dei problemi.

Per esempio, sia X l'operatore autoaggiunto ($X = X^\dagger$) con dominio $D(X) = \{\psi \in L^2(\mathbb{R}, dx) \mid x\psi \in L^2(\mathbb{R})\}$ che agisce su ψ come $X\psi = x\psi$.

Allora l'equazione agli autovalori diviene:

$$X\psi_\lambda(x) = \lambda\psi_\lambda(x)$$

Applicando la definizione di X :

$$X\psi_\lambda(x) = x\psi_\lambda(x) \Rightarrow (x - \lambda)\psi_\lambda(x) = 0$$

Scartiamo la soluzione banale $\psi_\lambda(x) \equiv 0$ che non è fisica¹⁸, ma allora l'equazione di sopra deve valere solo per $x = \lambda$.

¹⁸La posizione di una particella non può essere “da nessuna parte”! Inoltre tale $\psi = 0$ non è neanche normalizzabile.

*Decomposizione
spettrale in
 $\dim \mathcal{H} = \infty$:
problemi iniziali*

Tuttavia $\{x = \lambda\}$ ha misura di Lebesgue nulla: in altre parole stiamo cercando una soluzione (non nulla) che sia 0 per tutte le x tranne λ (ossia tranne un insieme Lebesgue-trascurabile). Ma tale funzione, in L^2 , è quella nulla, che abbiamo appena scartato!

Siamo costretti ad affermare perciò che X non possiede autovettori - e con ciò termina il primo tentativo di generalizzare la decomposizione spettrale al caso infinito-dimensionale: come possiamo trovare una base ON di autovettori se di autovettori non ce ne sono?

Proviamo a tornare al caso $\dim \mathcal{H} < \infty$ e studiare una strategia alternativa che possa essere generalizzata.

Poco fa, per dimostrare che gli autovalori “matematici” (di un operatore) sono tutti e soli quelli “fisici” (dell’osservabile rappresentato) nel caso finito-dimensionale avevamo fatto uso di un argomento probabilistico (le probabilità p_n associate agli autovalori λ_n esauriscono tutte le possibilità). Potremmo perciò cercare di partire da qui, riesaminando la struttura probabilistica nel caso finito-dimensionale - essenzialmente declinando quanto detto nei paragrafi di descrizione di un sistema fisico alla luce di quanto appena visto sulle proprietà degli operatori e sulla decomposizione spettrale.

Partiamo dallo scrivere l’equivalente di (2.14)¹⁹ per un operatore, tramite la decomposizione spettrale:

*Struttura
probabilistica per
operatori in MQ*

$$P^A(\lambda) = \sum_n H(\lambda - \lambda_n) P_n \equiv H(\lambda \mathbb{I} - A) \quad (3.12)$$

(che si intende valida per ogni stato ψ . Nota: con $P^A(\lambda)$ indichiamo un operatore che dipende da λ , non una funzione reale a valori reali!)

Ma allora da (2.12)²⁰ sappiamo che:

$$\langle O \rangle_\Sigma = \int_{\mathbb{R}} \lambda d\langle P^O(\lambda) \rangle_\Sigma$$

Nel caso degli operatori basta quindi calcolare il valor medio dell’operatore definito in (3.12), e si ottiene quindi, nel caso normalizzato (ossia con $\|\psi\| = 1$):

$$\langle A \rangle_\psi = \int_{\mathbb{R}} \lambda dP_\psi^A(\lambda) = \int_{\mathbb{R}} \lambda d \underbrace{(\psi, P^A(\lambda) \psi)}_{P_\psi^A(\lambda)} \quad (3.13)$$

Abbiamo quindi adattato la scrittura “probabilistica” dei valori medi anche al caso degli operatori.

Formalmente possiamo anche “estrarre” il valor medio dall’integrale, come avevamo fatto in (2.13)²¹, ottenendo un’uguaglianza che - essendo vera in ogni stato - deve valere anche rimuovendo il valor medio:

$$A = \int \lambda dP^A(\lambda); \quad f(A) = \int f(\lambda) dP^A(\lambda) \quad (3.14)$$

¹⁹ $\wedge P^O(\lambda) = H(\lambda - O)$

²⁰ $\wedge \langle O \rangle_\Sigma = \int_{\mathbb{R}} \lambda d\langle H(\lambda - O) \rangle_\Sigma$

²¹ $\wedge O = \int_{\mathbb{R}} \lambda dP^O \Rightarrow f(O) = \int_{\mathbb{R}} f(\lambda) dP^O(\lambda)$

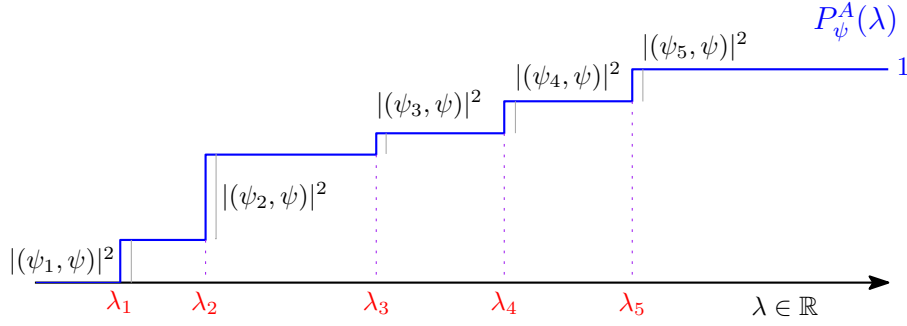


Figura 3.1: Andamento di $P_\psi^A(\lambda)$

Specifichiamo nel dettaglio cosa significa la notazione in (3.13):

$$\begin{aligned}
 d_\lambda(\psi, P^A(\lambda)\psi) &\stackrel{(3.12)}{=} d_\lambda(\psi, \sum_n H(\lambda - \lambda_n) P_n \psi) = \\
 &\stackrel{(a)}{=} \sum_n d_\lambda H(\lambda - \lambda_n)(\psi, \psi_n)(\psi_n, \psi) = \\
 &\stackrel{(b)}{=} \sum_n \delta(\lambda - \lambda_n) |(\psi_n, \psi)|^2 d\lambda
 \end{aligned} \tag{3.15}$$

In (a) si è usata la definizione di proiettore, (in notazione di Dirac) $P_n = |\psi_n\rangle \langle \psi_n|$, da cui $(\psi, P_n \psi) = \langle \psi | |\psi_n\rangle \langle \psi_n| | \psi \rangle$, e si è portato fuori dal prodotto scalare il termine reale $H(\lambda - \lambda_n)$.

In (b) si è calcolata la derivata della somma, ricordando che la derivata distribuzionale della Heaviside è la δ di Dirac, e notando che $(\psi, \psi_n) = (\psi_n, \psi)^*$, e il prodotto dei complessi coniugati dà il modulo quadro.

Possiamo ora calcolare l'integrale completo di (3.13):

$$\int_{\mathbb{R}} \lambda d(\psi, P^A(\lambda)\psi) = \int_{\mathbb{R}} \lambda \sum_n \delta(\lambda - \lambda_n) |(\psi_n, \psi)|^2 d\lambda \stackrel{(a)}{=} \sum_n \lambda_n |(\psi_n, \psi)|^2$$

Come ci si aspetta, nel passaggio (a) la δ di Dirac annulla l'integrale per tutti i $\lambda \neq \lambda_n$, e il risultato sarà quindi una sommatoria di quei valori.

Poniamo, seguendo la notazione già stabilita alla sezione precedente:

$$P_\psi^A(\lambda) = (\psi, P^A(\lambda)\psi)$$

Si tratta di una funzione reale, che interpretiamo come la probabilità che effettuata una misura di A nello stato ψ essa abbia un valore $\leq \lambda$. Osserviamone l'andamento nel dettaglio. Supponiamo di porre sulla retta reale gli autovalori, in ordine crescente. In particolare, vi sarà un autovalore minimo λ_1 , prima del quale la $P_\psi^A(\lambda)$ è nulla. Quando incontra il primo autovalore vi è un “gradino” (dato dalla Heaviside), e allo stesso modo per i successivi autovalori, con “il salto” tra ciascuno che è dato dal peso $|(\psi_n, \psi)|^2$, come illustrato in figura 3.1.

Alla fine (dopo tutti gli autovalori), giungerà ovviamente a 1, essendo normalizzata la ψ normalizzata: $\|\psi\| = 1$. Perciò:

$$\sum_n |(\psi_n, \psi)|^2 = 1$$

In particolare:

$P_\psi^A(\lambda)$ = probabilità in una misura di A nello stato ψ di trovare un valore $\leq \lambda$

Perciò costruendo l'operatore P^A grazie alla decomposizione spettrale possiamo ricreare la struttura probabilistica (e in particolare ottenere la “probabilità cumulativa” P_ψ^A).

Elenchiamo ora alcune **proprietà** degli operatori $\{P^A(\lambda), \lambda \in \mathbb{R}\}$ dati da $P^A(\lambda) = \sum_n H(\lambda - \lambda_n) P_n$ (siamo sempre nel caso di $\dim \mathcal{H} < \infty$). L'idea è di richiedere le stesse proprietà anche nel caso infinito-dimensionale, in modo da svincolarsi dalle equazioni agli autovalori e poter generalizzare. Proprietà di $P^A(\lambda)$

1. $P^A(\lambda)$ è un **proiettore**. Infatti è autoaggiunto, come si può dimostrare lavorando in rappresentazione spettrale, e cioè scrivendo $P^A(\lambda)$ come somma di proiettori sui singoli autovettori “pesati” da funzioni dei singoli autovalori: $P^A(\lambda)$ è un proiettore

$$\left(\sum_n H(\lambda - \lambda_n) P_n \right)^\dagger = \sum_n H(\lambda - \lambda_n)^* P_n^\dagger = \sum_n H(\lambda - \lambda_n) P_n$$

Dove si è usato il fatto che H è reale (quindi pari al suo coniugato), e P_n è autoaggiunto.

Vale anche la seconda condizione per i proiettori, ossia che $P^A(\lambda)^2 = P^A(\lambda)$:

$$P^A(\lambda)^2 = \left(\sum_n H(\lambda - \lambda_n) P_n \right) \left(\sum_m H(\lambda - \lambda_m) P_m \right) \quad (3.16)$$

Dove si ha che:

$$\begin{aligned} P_n P_m \psi &= \underset{(a)}{P_n (\psi_m, \psi) \psi_m} = \underset{(b)}{(\psi_m, \psi) (\psi_n, \psi_m) \psi_n} = \underset{(c)}{(\psi_m, \psi) (\psi_n, \psi_m) \psi_n} \\ &= \delta_{nm} (\psi_m, \psi) \psi_n = \delta_{nm} \underbrace{\langle \psi_m | \psi \rangle \langle \psi_n | \psi \rangle}_{P_n \psi} \underset{(d)}{=} P_n \psi \delta_{nm} \end{aligned}$$

In (a) si applica P_m a ψ , che genera $|\psi_m\rangle \langle \psi_m | \psi \rangle$, mentre in (b) si applica P_n al $|\psi_m\rangle$ appena generato, che porta a $\langle \psi_n, \psi_m \rangle |\psi_n\rangle \langle \psi_m | \psi \rangle$. Ma per un operatore simmetrico come A (che è autoaggiunto) possiamo scegliere una base ortonormale di autovettori ψ_n , per cui in (c) $\langle \psi_n | \psi_m \rangle = \delta_{mn}$ (il prodotto scalare restituisce 1 per la normalizzazione solo quando un vettore è moltiplicato per se stesso, e altrimenti 0 perché gli autovettori sono ortogonali tra loro). Infine in (d) usiamo la δ per identificare $n = m$, e riconoscere in $\langle \psi_m | \psi \rangle |\psi_n\rangle$ il proiettore P_n applicato a ψ .

Se consideriamo solo l'operatore “puro” avremo perciò $P_n P_m = P_n \delta_{nm}$

Sostituendo questo risultato in (3.16):

$$\begin{aligned} P^A(\lambda)^2 &= \sum_{n,m} H(\lambda - \lambda_n) H(\lambda - \lambda_m) P_n P_m = \\ &= \sum_{n,m} H(\lambda - \lambda_n) H(\lambda - \lambda_m) P_n \delta_{nm} = \\ &= \sum_n H(\lambda - \lambda_n) H(\lambda - \lambda_n) P_n \\ &\underset{(a)}{=} \sum_n H(\lambda - \lambda_n) H(\lambda - \lambda_n) P_n \\ &\underset{(b)}{=} \sum_n H(\lambda - \lambda_n) P_n = P^A(\lambda) \end{aligned}$$

In (a) abbiamo usato la δ per identificare $n = m$ e collassare la doppia sommatoria a un'unica variabile.

In (b), invece, si è usato il fatto che la Heaviside assume come valori solo 0 e 1, e in particolare è una funzione caratteristica, e il suo quadrato è ovviamente se stessa - in effetti l'avevamo vista come analogo nelle funzioni dei proiettori.

$$2. \lim_{\lambda \rightarrow -\infty} P^A(\lambda) = 0; \quad \lim_{\lambda \rightarrow +\infty} P^A(\lambda) = \sum_n P_n = \sum_n |\psi_n\rangle \langle \psi_n| = \mathbb{I}$$

Seguendo un po' il discorso già fatto per $P_\psi^A(\lambda)$, per $\lambda \rightarrow -\infty$ non vi sarà alcun autovalore $\lambda_n \leq \lambda$, e perciò l'operatore darà 0 (nella definizione, l'Heaviside è costantemente 0 per $\lambda \rightarrow -\infty$). All'estremo opposto, invece, avremo che l'operatore è semplicemente la somma di tutti i proiettori P_n - e la somma di tutte le proiezioni (vettoriali) di un vettore su una base ON è il vettore stesso - da cui l'operatore agisce come l'identità.

*Comportamento
asintotico di
 $P^A(\lambda)$*

$$3. P^A(\lambda) P^A(\lambda') = P^A(\lambda') P^A(\lambda) = P^A(\min\{\lambda, \lambda'\})$$

Per comprenderlo applichiamo la definizione, e concentriamoci sulle Heaviside. In particolare, esaminiamo la seguente espressione nella variabile λ'' :

*Composizione di
 $P^A(\lambda)$*

$$H(\lambda - \lambda'') H(\lambda' - \lambda'')$$

Il primo termine è 1 per $\lambda'' < \lambda$, e il secondo è 1 solamente se $\lambda'' < \lambda'$.

Perciò il valore sarà 1 nell'intersezione tra i due insiemi, che è data da $\lambda'' < \min\{\lambda, \lambda'\}$, e la funzione si può riscrivere come:

$$H(\min\{\lambda, \lambda'\} - \lambda'')$$

$$4. \lim_{\lambda \rightarrow \lambda_0^+} P^A(\lambda) = P^A(\lambda_0)$$

Poiché la Heaviside è 1 in 0 (è continua da destra, ma non da sinistra), perciò il "valore da sopra" è lo stesso della funzione in quel punto. Più precisamente:

*$P^A(\lambda)$ sono
continui a destra*

$$\lim_{x \rightarrow 0^+} H(x) = 1; \quad \lim_{x \rightarrow 0^-} H(x) = 0$$

Definizione 14. Una famiglia di proiettori $\{P(\lambda), \lambda \in \mathbb{R}\}$ che soddisfa le proprietà 2-4, è detta **famiglia spettrale**²².

*Famiglia
spettrale*

Teorema 3.2.3. Esiste una corrispondenza biunivoca tra operatori autoaggiunti $A = A^\dagger$ e famiglie spettrali definita da:

*Operatore
autoaggiunto \leftrightarrow
famiglia spettrale*

$$\{P(\lambda), \lambda \in \mathbb{R}\} \rightarrow A$$

$$D(A) = \left\{ \phi \in \mathcal{H} \mid \int_{\mathbb{R}} \lambda^2 d(\phi, P(\lambda) \phi) < \infty \right\} \quad (3.17)$$

Si ha che $D(A)$ è denso, quindi per definire A basta darne gli elementi di matrice, che sono ben definiti grazie alla scelta fatta per il dominio $D(A)$:

$$(\phi, A\phi) = \int_{\mathbb{R}} \lambda d(\phi, P(\lambda) \phi)$$

²²La proprietà 1 è già soddisfatta quando richiediamo che sia una famiglia di proiettori

Più in generale, se f è una funzione misurabile su \mathbb{R} , esiste un operatore $f(A)$ con dominio $D(f(A))$ denso in \mathcal{H} , dato da:

$$D(f(A)) = \left\{ \phi \in \mathcal{H} \mid \int |f(x)|^2 d(\phi, P(\lambda)\phi) < \infty \right\}$$

L'operatore $f(A)$ è quindi definito completamente dai suoi elementi di matrice:

$$(\phi, f(A)\phi) = \int f(\lambda) d(\phi, P(\lambda)\phi)$$

(Qui negli elementi di matrice stiamo usando tutti vettori nel dominio denso dell'operatore, poiché abbiamo visto che è possibile utilizzare l'**identità di polarizzazione** per riscrivere il tutto in termini di valori medi)

Dim. (molto tecnica e lunga, omessa)

In particolare, per la speciale funzione $H(\lambda - A)$ si ha:

$$H(\lambda - A) = P(\lambda)$$

(per chiarire, prova a pensare tutto ciò in termini di spazi finito-dimensionali).
Con questa strategia non abbiamo usato autovettori/autovalori, e perciò tramite questa è possibile estendere quanto detto nel caso finito-dimensionale all'infinito-dimensionale.

Esempio In $\mathcal{H} = L^2(\mathbb{R}, dx)$, consideriamo la **famiglia spettrale**:

$$(P(\lambda)\psi)(x) = H(\lambda - x)\psi(x) \quad (3.18)$$

*Esempio (1) di
famiglia
spettrale:
operatore
posizione*

Verifichiamo che soddisfa tutte le proprietà:

1. H reale e $H^2 = H$, quindi è un proiettore.
2. $\lim_{\lambda \rightarrow -\infty} H(\lambda - x)\psi(x) = 0$; $\lim_{\lambda \rightarrow +\infty} H(\lambda - x)\psi(x) = \psi(x) \quad \forall \psi \in \mathcal{H}$
3. $H(\lambda - x)H(\lambda' - x) = H(\min\{\lambda, \lambda'\} - x)$
4. $\lim_{\lambda \rightarrow \lambda_0^+} H(\lambda - x) = H(\lambda_0 - x)$

Di quale operatore autoaggiunto è la famiglia spettrale? Calcoliamone il dominio:

$$\begin{aligned} D(A) &= \left\{ \phi \in L^2(\mathbb{R}) \mid \int_{\mathbb{R}} \lambda^2 d(\phi, P(\lambda)\phi) < \infty \right\} \\ (\phi, P(\lambda)\phi) &= \int_{\mathbb{R}} dx \phi^*(\lambda) H(\lambda - x)\phi(x) \\ d(\phi, P(\lambda)\phi) &= \left(\int_{\mathbb{R}} dx \phi^*(x) \delta(\lambda - x)\phi(x) \right) d\lambda = \phi^*(\lambda)\phi(\lambda) d\lambda \end{aligned}$$

E quindi:

$$D(A) = \left\{ \phi \in L^2(\mathbb{R}, dx) \mid \int \lambda^2 \phi^*(\lambda)\phi(\lambda) d\lambda < \infty \right\} = \{ \|x\phi\|^2 < \infty \}$$

E gli elementi di matrice sono dati da:

$$(\phi, A\phi) = \int \lambda d(\phi, P(\lambda)\phi) = \int \lambda \phi^*(\lambda) \phi(\lambda) d\lambda = \int x \phi^*(x) \phi(x) dx$$

Riconosciamo l'azione dell'operatore **posizione**:

$$X\phi(x) = x\phi(x) \\ D(X) = \{\phi \in L^2 \mid \|\phi\|^2 < \infty\}$$

Quindi $A = X$, e $dP_\phi^X(\lambda) = |\phi(\lambda)|^2 d\lambda$.

Se ϕ non è normalizzato, perché sia

$$\int_{\mathbb{R}} dP_\phi^X(\lambda) = 1$$

Bisogna scrivere:

$$dP_\phi^X(\lambda) = \frac{|\phi(\lambda)|^2}{\|\phi\|^2} d\lambda$$

La grandezza evidenziata in azzurro è la **densità di probabilità** che misurando X nello stato ϕ si ottenga λ . In fisica si denota con:

Densità di probabilità

$$w_\phi^X(\lambda) = \frac{dP_\phi^X(\lambda)}{d\lambda} = \frac{|\phi(\lambda)|^2}{\|\phi\|^2}$$

Il valor medio è quindi:

$$\langle X \rangle_\phi = \int \lambda dP_\phi^X(\lambda) = \int \lambda w_\phi^X(\lambda) d\lambda$$

Riepilogando:

Riepilogo delle proprietà delle osservabili

1. Le osservabili sono **operatori** A e dalla indipendenza del rappresentativo del raggio vettore nel calcolo dei valor medi $\langle A \rangle_\psi$ si ha che A sono **lineari** (e richiediamo abbiano un dominio denso in \mathcal{H} , ossia il più largo possibile).
2. Vogliamo poi che $\langle A \rangle_\psi \in \mathbb{R}$, da cui A deve essere **simmetrico**.
3. La richiesta che esista $f(A)$ implica che A è **autoaggiunto**. Ciò consente anche di costruire una interpretazione probabilistica, potendo definire la funzione $P_\psi^A(\lambda)$

Enunciamo questi risultati, che abbiamo costruito nei paragrafi precedenti e che sono estensivamente corroborati dall'evidenza, come **assiomi** del formalismo quantistico.

Assioma 1. Le **osservabili** di un sistema quantistico, i cui **stati puri** sono descritti dai **raggi vettori** di \mathcal{H} , sono descritti da **operatori autoaggiunti** $A = A^\dagger$ in \mathcal{H} (lineari e densamente definiti).

Assioma delle osservabili

Assioma 2. La **probabilità** che il risultato di una misura di A nello stato ψ (ricavata ripetendo tante volte la misura sperimentale nelle stesse condizioni) sia in $]-\infty, \lambda]$ è data da:

Assioma della probabilità

$$P_{\psi}^A(\lambda) = \frac{(\psi, P^A(\lambda) \psi)}{\|\psi\|^2}$$

Dove $P^A(\lambda)$ è la **famiglia spettrale** di A e $\forall \psi \in D(A)$:

$$\langle A \rangle_{\psi} = \frac{(\psi, A\psi)}{\|\psi\|^2}$$

3.2.8 Osservazioni sulle proprietà matematiche delle osservabili

1. La condizione di autoaggiuntezza è delicata: Ad esempio, se A e B sono autoaggiunti **limitati**, $A + B$ lo è, ma AB no. Infatti $(AB)^{\dagger} = B^{\dagger}A^{\dagger} = BA \neq AB$.

Perciò gli operatori autoaggiunti non formano un'algebra sui reali.

Possiamo costruire un'algebra degli operatori autoaggiunti su \mathbb{C} , e in effetti $\frac{(AB-BA)}{i}$ è autoaggiunto.

2. Se invece A e B sono autoaggiunti, ma **illimitati** in $D(A)$ e $D(B)$ densi in \mathcal{H} . Allora $(A+B)$ è ben definito in $D(A) \cap D(B)$ ma $A+B$ in generale non è autoaggiunto in $D(A) \cap D(B)$ che potrebbe addirittura non essere denso. Questo problema appare in fisica ad esempio per l'osservabile energia:

$$H = \frac{\vec{p}^2}{2m} + V(\vec{x})$$

Con $A = \frac{\vec{p}^2}{2m}, B = V(\vec{x})$

Fortunatamente, molti casi importanti sono coperti dal seguente risultato:

Teorema 3.2.4. Se V è limitato in \mathbb{R}^n , oppure se $n = 3$ e V è Coulombiano (es. potenziale elettrostatico), allora H è autoaggiunto nel dominio di A definito da:

$$D(\vec{p}^2) = \{\varphi \in L^2(\mathbb{R}^n, d^n x) \mid \vec{p}^2 \tilde{\varphi}(\vec{p}) \in L^2(\mathbb{R}^n, d^n p)\}$$

Autoaggiuntezza dell'Hamiltoniana H

3. **Attenzione:** Conoscere $\sigma(A)$ e $\sigma(B)$ non porta alla conoscenza di $\sigma(A+B)$. Per esempio, $A = \frac{p^2}{2m}, B = V(x)$, con $\sigma(A) = \mathbb{R}_+, \sigma(V(x)) = \text{cod } V(x)$ sono entrambi operatori a spettro continuo. Tuttavia la loro somma contiene uno spettro discreto: sono per esempio i livelli energetici di una buca infinita. (In MC, d'altro canto, conoscere $\sigma(A)$ e $\sigma(B)$ porta subito alla conoscenza di $\sigma(A+B)$.)

Riepilogando: abbiamo visto che A deve essere lineare, di dominio denso, simmetrico (valor medi reali), e se vogliamo che esistano funzioni $f(A)$, si ha:

(Lezione 9 del 16/10/2018)

$$\langle A \rangle_{\psi} = \int \lambda dP_{\psi}^A(\lambda) \Rightarrow A = A^{\dagger} \quad (3.19)$$

Osservazione 1. Sia $f(\lambda) = e^{i\lambda a}$, $a \in \mathbb{R}$. Applicandola ad un operatore generico otteniamo, formalmente, $f(A) = e^{iaA}$. Dimostriamo che il nuovo operatore $f(A)$ è unitario.

e^{iaA} è un operatore unitario

Nella notazione formale di (3.14) si ha che:

$$f(A) = e^{iaA} = \int e^{iaA} dP^A(\lambda)$$

Calcolandone l'aggiunto:

$$(e^{iaA})^\dagger = \left(\int e^{ia\lambda} dP^A(\lambda) \right)^\dagger \underset{(a)}{=} \int e^{-ia\lambda} dP^A(\lambda) \underset{(b)}{=} e^{-iaA} = (e^{iaA})^{-1} \quad (3.20)$$

dove in (a) si è usato il fatto che $dP^A(\lambda)$ è autoaggiunto, e perciò l'unico cambiamento sta nel prendere il complesso coniugato dell'esponentiale. Riconosciamo allora in (b) la funzione $g(A) = e^{-iaA}$ scritta nel formalismo di (3.14), che è esattamente l'inversa di f , e ciò rende $f(A)$ un operatore unitario.

Osservazione 2. Dato un vettore in notazione di Dirac $|\phi\rangle$ il proiettore $P_\phi = |\phi\rangle\langle\phi|$ è autoaggiunto, quindi descrive un'osservabile (per la corrispondenza biunivoca tra autoaggiunti e famiglie spettrali), che corrisponde alla domanda: il sistema si trova nello stato $|\phi\rangle$?

Il valor medio di $|\phi\rangle\langle\phi|$ nello stato $|\psi\rangle$ è infatti dato da (assumendo stati normalizzati, con $\langle\phi|\phi\rangle = 1 = \langle\psi|\psi\rangle$):

$$(\psi, P_\phi \psi) = \langle\psi|\phi\rangle\langle\phi|\psi\rangle = |\langle\psi|\phi\rangle|^2$$

che viene detto **probabilità di transizione** (e $\langle\psi|\phi\rangle$ ampiezza di transizione) da $|\phi\rangle$ a $|\psi\rangle$.

Probabilità di transizione

Poiché tale probabilità è non nulla anche per $|\phi\rangle \neq |\psi\rangle$ (anche come raggi vettori), vediamo che in MQ **gli stati puri non sono mutualmente esclusivi**, come invece avviene nel caso classico, in cui o due punti (intesi nello spazio delle fasi Ω) coincidono o sono distinti.

Stati puri distinti non sono mutualmente esclusivi

In MQ, dando l'analogo delle condizioni iniziali (vettore nello spazio di Hilbert, stato puro = massima informazione), provando a fare una misura in generale possiamo ottenere il risultato di uno stato diverso. Alternativamente, se preparo il sistema ad un istante $t = 0$ nello $|\phi\rangle$, a $t = 0^+$ è possibile trovarlo nello stato $|\psi\rangle \neq |\phi\rangle$!

Matematicamente ciò è dato dal fatto che due generici vettori, se moltiplicati scalarmente, non sempre devono dare un risultato nullo - tale ovvietà produce un risultato fisico per nulla intuitivo.

Del resto, in MC gli stati puri sono costituiti da delta di Dirac, e il prodotto di due δ diverse è per forza nullo.

Esercizio. Si consideri la famiglia di operatori $\{P(\lambda), \lambda \in \mathbb{R}\}$ definiti in $L^2(\mathbb{R}, dx)$ da:

Esempio (2) di famiglia spettrale: il momento

$$(P(\lambda)\psi)(x) = \frac{1}{2\pi\hbar} \int_{\mathbb{R}} \left(\int_{-\infty}^{\lambda} e^{i\frac{p}{\hbar}(x-x')} dp \right) \psi(x') dx'$$

Si dimostri che $P(\lambda)$ è una **famiglia spettrale** (solo il fatto che $P(\lambda)$ sono proiettori) e si trovi la corrispondente osservabile.

Verifichiamo innanzitutto l'autoaggiuntezza, ossia che: $(P(\lambda))^\dagger = P(\lambda)$ Partendo dagli elementi di matrice:

$$(\phi, P(\lambda) \psi) = (P^\dagger(\lambda) \phi, \psi)$$

Svolgendo il prodotto scalare, il primo termine è pari a:

$$\begin{aligned} (\phi, P(\lambda) \psi) &= \int_{\mathbb{R}} \phi^*(x) \int_{\mathbb{R}} \frac{1}{2\pi\hbar} \left(\int_{-\infty}^{\lambda} e^{i\frac{p}{\hbar}(x-x')} dp \right) \psi(x') dx' dx = \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \left(\frac{1}{2\pi\hbar} \int_{-\infty}^{\lambda} \exp\left(i\frac{p}{\hbar}(x' - x)\right) dp \phi(x) dx \right)^* \psi(x') dx' \end{aligned}$$

dove si è “*spostata l'azione dell'operatore al primo membro del prodotto scalare*”, portando tutto ciò che dipenda da x sotto lo stesso segno di integrale ed evidenziando una coniugazione complessa (si noti il cambio di segno nell'esponenziale). Possiamo ora notare che il termine tra parentesi è l'azione di $P^\dagger(\lambda)$ su un ϕ :

$$(P^\dagger(\lambda) \phi)(x') = \frac{1}{2\pi\hbar} \int_{\mathbb{R}} \left(\int_{-\infty}^{\lambda} \exp\left(i\frac{p}{\hbar}(x' - x)\right) dp \right) \phi(x) dx = (P(\lambda) \phi)(x')$$

Dimostriamo quindi che P è un proiettore, e cioè che $P^2(\lambda) = P(\lambda)$. La cosa non è immediata - ci servirà passare in una notazione più conveniente. Notando la presenza di esponenziali e integrali nella definizione di P , un'idea è di sfruttare le trasformate di Fourier²³

Osserviamo che se definiamo trasformata e antitrasformata di Fourier come:

$$\begin{aligned} (\mathcal{F}\psi)(p) &\equiv \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \exp\left(-i\frac{p}{\hbar}x\right) \psi(x) dx = \tilde{\psi}(p) \\ (\mathcal{F}^{-1}\tilde{\psi})(x) &\equiv \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \exp\left(i\frac{p}{\hbar}x\right) \tilde{\psi}(p) dp \end{aligned}$$

Possiamo riscrivere $(P(\lambda)\psi)(x)$ tramite trasformate di Fourier, e quindi usare le proprietà di queste ultime per dimostrare $P = P^2$. Partiamo dall'azione di $P(\lambda)$ su ψ :

$$(P(\lambda)\psi)(x') = \int_{\mathbb{R}} \left[\frac{1}{2\pi\hbar} \int_{-\infty}^{\lambda} \exp\left(i\frac{p}{\hbar}(x' - x)\right) dp \right] \psi(x) dx$$

Notiamo che l'integrale “esterno” su x è un *integrale di convoluzione*, in quanto ha la struttura di:

$$(f * g) = \int_{\mathbb{R}} f(x) g(x' - x) dx$$

dove la $f(x) = \psi(x)$, e per la $g(x' - x)$ (che corrisponde al termine tra parentesi quadre, ossia all'integrale in p), riconosciamo la struttura di un'antitrasformata di Fourier, con due differenze chiave: un fattore $\sqrt{2\pi\hbar}$ in più e un dominio “troppo stretto” (non tutto \mathbb{R}).

Per quanto riguarda il fattore basta tenerne conto, mentre per correggere il dominio passiamo ad un integrale su tutto \mathbb{R} , e annulliamo il suo valore fuori dalla regione

²³^In effetti da Fisica Moderna possiamo già intuire che questo operatore sia l'operatore momento, che è ricavato a partire dall'operatore posizione tramite trasformate di Fourier

che ci interessa moltiplicando la funzione integranda per l'opportuna funzione di Heaviside $H(\lambda - p)$.

Possiamo quindi riscrivere:²⁴:

$$\begin{aligned}(P(\lambda)\psi)(x) &= \int_{\mathbb{R}} \left[\frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \frac{H(\lambda - p)}{\sqrt{2\pi\hbar}} \exp\left(+i\frac{p}{\hbar}(x - x')\right) dp \right] \psi(x') dx' = \\ &= \mathcal{F}^{-1} \left(\frac{H(\lambda - p)}{\sqrt{2\pi\hbar}} \right) * \psi(x)\end{aligned}$$

Sfruttando “al contrario” la proprietà delle trasformate²⁵:

$$\mathcal{F}^{-1}(\phi\psi) = \frac{1}{\sqrt{2\pi\hbar}} \mathcal{F}^{-1}(\phi) * \mathcal{F}^{-1}(\psi)$$

Otteniamo:

$$\begin{aligned}(P(\lambda)\psi)(x) &= \mathcal{F}^{-1} \left(\frac{H(\lambda - p)}{\sqrt{2\pi\hbar}} \right) * \psi(x) = \mathcal{F}^{-1} \left(\frac{H(\lambda - p)}{\sqrt{2\pi\hbar}} \right) * \mathcal{F}^{-1}(\mathcal{F}\psi)(x) = \\ &= \frac{1}{\sqrt{2\pi\hbar}} \mathcal{F}^{-1}(H(\lambda - p) * \mathcal{F}^{-1}(\mathcal{F}\psi))(x) = \mathcal{F}^{-1}(H(\lambda - p)\mathcal{F}\psi)\end{aligned}$$

(dove si è usata la linearità dell'antitrasformata per portar fuori il fattore di normalizzazione $1/\sqrt{2\pi\hbar}$).

Definendo:

$$\phi(x) = (P(\lambda)\psi)(x) = \mathcal{F}^{-1}(H(\lambda - p)\mathcal{F}\psi)(x)$$

Si ha che applicando $P(\lambda)$ a ϕ otteniamo $(P(\lambda)^2\psi)(x)$. Facendo allora il conto:

$$\begin{aligned}(P(\lambda)\phi)(x) &= \mathcal{F}^{-1}(H(\lambda - p)\mathcal{F}\phi)(x) = \\ &= \mathcal{F}^{-1}[H(\lambda - p)\mathcal{F}\mathcal{F}^{-1}(H(\lambda - p)\mathcal{F}\psi)](x) = \\ &\stackrel{(a)}{=} \mathcal{F}^{-1}(H(\lambda - p)H(\lambda - p)\mathcal{F}\psi)(x) = \\ &\stackrel{(b)}{=} \mathcal{F}^{-1}(H(\lambda - p)\mathcal{F}\psi)(x) = (P(\lambda)\psi)(x)\end{aligned}$$

In (a) si è sfruttata la proprietà di Fourier, per cui trasformata e antitrasformata sono una l'inversa dell'altra (e la loro composizione è l'identità). Per (b), invece, basta notare che $H(\lambda - p)$ ha come valori solo 1 e 0 (in particolare è la funzione caratteristica di $] - \infty, \lambda]$) e quindi è pari al suo quadrato (come si era visto tra gli esempi di “funzioni unitarie”).

Avendo quindi dimostrato che P è autoaggiunto e $P = P^2$ si ha che P è un proiettore.

Si può verificare (ma non lo faremo) che i $P(\lambda)$, $\lambda \in \mathbb{R}$ costituiscono una *famiglia spettrale* (cioè soddisfano anche le altre tre proprietà che abbiamo elencato). Ogni

²⁴^Rinominando la variabile x' con x per comodità

²⁵^Il fattore $1/\sqrt{2\pi\hbar}$ compare perché stiamo usando la “normalizzazione angolare” per le trasformate di Fourier

famiglia spettrale è associata ad un'osservabile: esaminiamola. Sia $|\psi\rangle$ normalizzata ($\langle\psi|\psi\rangle = 1$). Il valor medio dell'operatore A associato alla famiglia spettrale $P(\lambda)$ nello stato ψ è dato da:

$$\langle A \rangle_\psi = \int \lambda d(\psi, P(\lambda)\psi) \quad (3.21)$$

Espandendo il prodotto scalare:

$$\begin{aligned} d(\psi, P(\lambda)\psi) &= d_\lambda \int \int \psi^*(x) \left(\int_{-\infty}^{\lambda} \frac{e^{i\frac{p}{\hbar}(x-x')}}{2\pi\hbar} dp \right) \psi(x') dx dx' = \\ &\stackrel{(a)}{=} \left[\int \int \psi^*(x) \exp\left(i\frac{\lambda}{\hbar}(x-x')\right) \frac{1}{2\pi\hbar} \psi(x') dx dx' \right] d\lambda = \\ &\stackrel{(b)}{=} \underbrace{\left[\frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \left(\psi(x) \exp\left(-i\frac{\lambda}{\hbar}x\right) \right)^* dx \right]}_{(\mathcal{F}\psi)^*(\lambda)} \cdot \underbrace{\left[\frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \psi(x') \exp\left(-i\frac{\lambda}{\hbar}x'\right) dx' \right]}_{(\mathcal{F}\psi)(\lambda)} d\lambda = \\ &= (\mathcal{F}\psi)^*(\lambda) (\mathcal{F}\psi)(\lambda) d\lambda = |\tilde{\psi}(\lambda)|^2 d\lambda \end{aligned}$$

In (a) notiamo che la derivata per λ agisce solo sull'integrale, e per il teorema fondamentale del calcolo restituisce la funzione integranda calcolata in $p = \lambda$. In (b) abbiamo “spezzato” l'esponenziale in due fattori, dove si riconosce la struttura delle trasformate di Fourier. Il prodotto di una trasformata per la sua complessa coniugata dà allora il modulo quadro.

Sostituendo questo risultato in (3.21):

$$\langle A \rangle_\psi = \int \lambda |\tilde{\psi}(\lambda)|^2 d\lambda$$

Possiamo finalmente calcolare il dominio di A , completandone la definizione (facendo riferimento al teorema in (3.17)):

$$\begin{aligned} D(A) &= \left\{ \psi \in L^2 \mid \int \lambda^2 d(\psi, P(\lambda)\psi) < \infty \right\} = \\ &= \left\{ \psi \in L^2 \mid \int \lambda^2 |\tilde{\psi}(\lambda)|^2 d\lambda < \infty \right\} \end{aligned}$$

Se ora esaminiamo:

$$\mathcal{F}\left(-i\hbar \frac{d}{dx} \psi\right)(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \exp\left(-i\frac{p}{\hbar}x\right) \left(-i\hbar \frac{d}{dx} \psi(x)\right) dx \quad (3.22)$$

La condizione imposta nel determinare il dominio fa sì che $\psi(x) \xrightarrow{x \rightarrow \infty} 0$ (lo dimostreremo più avanti in un altro esempio). Ciò significa che **integrando per parti** il primo termine (che viene valutato tra $-\infty$ e ∞) si annulla, e perciò lo trascuriamo. Svolgendo quindi i conti rimanenti:

$$\begin{aligned} (3.22) &= \cancel{\frac{1}{\sqrt{2\pi\hbar}}} \int_{\mathbb{R}} \cancel{i\hbar} \psi(x) \frac{d}{dx} \left(\exp\left(-i\frac{p}{\hbar}x\right) \right) dx = \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} i\hbar \psi(x) \left(-i\frac{p}{\hbar} \exp\left(-i\frac{p}{\hbar}x\right) \right) dx = \\ &= p \underbrace{\frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \exp\left(-i\frac{p}{\hbar}x\right) \psi(x) dx}_{\tilde{\psi}(p)} = p \tilde{\psi}(p) \end{aligned}$$

Ciò significa che la rappresentazione in posizioni e quella in momenti sono collegate da una trasformata di Fourier. Infatti vale, per il teorema di Plancherel²⁶:

$$\langle A \rangle_\psi = \int \tilde{\psi}(p)^* p \tilde{\psi}(p) dp = \int \psi(x)^* \left(-i\hbar \frac{d}{dx} \psi(x) \right) dx$$

Perciò l'operatore A in $D(A)$, usando la rappresentazione in posizioni, è dato da:

$$A = -i\hbar \frac{d}{dx}$$

3.3 Spettro di un operatore

Ci poniamo ora il problema di determinare quali valori si possono ottenere con una misura di un operatore autoaggiunto $A = A^\dagger$ (o meglio, dell'osservabile da esso descritto).

Nel caso $\dim \mathcal{H} < \infty$ si ha che $\sigma(A) = \{\lambda_n \text{ autovalori}\}$, analogamente al caso classico, dove un'osservabile è data da una funzione $f(q, p)$ i cui valori, che costituiscono lo spettro, sono tutti autovalori (cod $f = \{\text{autovalori di } f\}$).

Tuttavia, per $\dim \mathcal{H} = \infty$, sorgono fin da subito dei problemi. Infatti, un autovalore è definito come (3.10):

$$(\Delta O)_{\lambda, \psi_\lambda} = 0 = \|(A - \lambda)\psi_\lambda\|$$

Con ψ_λ autovettore di autovalore λ (l'equivalenza tra le due scritture è stata ricavata in (3.10)).

Proviamo a imporre tale condizione per trovare gli autovalori dell'operatore **posizione** su \mathbb{R} $X\psi(x) = x\psi(x)$:

$$\langle (X - a)^2 \rangle_\psi = \int (x - a)^2 |\psi(x)|^2 dx = 0 \Rightarrow (x - a)^2 \psi(x) = 0 \Rightarrow \psi(x) = 0$$

(un integrale di quantità positive è nullo se e solo se la funzione integranda è nulla, a meno di insiemi Lebesgue-trascurabili, e perciò in ogni caso corrisponde alla funzione $f \equiv 0$ in L^2 , considerata come classe di equivalenza).

Poiché l'unico risultato è la funzione nulla, ciò significa che in L^2 non esiste alcun autovalore. Fisicamente vuol dire che non esiste la possibilità di localizzare una particella in un punto in MQ, e matematicamente che:

$$\nexists \psi_a \mid (\Delta X)_{a, \psi_a} = 0$$

Come già visto in precedenza (nell'esempio sull'operatore momento), il fatto che compaia la sola soluzione nulla fa pensare che forse abbiamo imposto delle condizioni troppo stringenti. In effetti potremmo accontentarci di poter determinare la

²⁶ Il teorema di Plancherel enuncia che una trasformazione di Fourier è un'isometria rispetto alla norma di L^2 - e, visto che norma e prodotto scalare sono collegati dall'identità di polarizzazione - anche rispetto al prodotto scalare di L^2 . In altre parole, il prodotto scalare tra due elementi di L^2 dà lo stesso risultato se fatto tra le trasformate di quei due elementi. Qui nello specifico si ha $(\tilde{\psi}, A\tilde{\psi}) = (\psi, A_x\psi)$, dove $A = p$ (operatore momento in rappresentazione dei momenti), e $A_x = -i\hbar \frac{d}{dx} = \mathcal{F}(A)$ è lo "stesso operatore" in rappresentazione di posizioni.

posizione in un range arbitrariamente fine (e quindi con precisione alta quanto si vuole), ma comunque di misura di Lebesgue non nulla, ossia non composto da un singolo punto (che richiederebbe una precisione infinita, cosa assurda).

Matematicamente ciò corrisponde a cercare, tra tutte le $\psi \in D(X)$, quelle che si avvicinano molto alla situazione di fluttuazione nulla, ossia tali che l'*estremo inferiore* della loro fluttuazione (che non è detto venga raggiunto, e quindi sia anche un minimo!) sia nullo:

$$\inf_{\psi \in D(X)} (\Delta X)_{a,\psi} = 0$$

Possiamo allora costruire la successione di funzioni che “si stringono” su a , ossia che valgono \sqrt{n} in un intervallino largo $1/n$ attorno ad a , per ogni n finito:

$$\psi_n(x) = \sqrt{n} \chi_{[a-\frac{1}{2n}; a+\frac{1}{2n}]}(x) \in D(X) \quad \forall n < +\infty$$

Così definite, le ψ_n sono normalizzate:

$$\|\psi_n\|^2 = \int_{a-\frac{1}{2n}}^{a+\frac{1}{2n}} n \, dx = 1$$

E l'applicazione dell'operatore X su una ψ_n è ben definita:

$$\|X\psi_n\|^2 = \int_{a-\frac{1}{2n}}^{a+\frac{1}{2n}} x^2 n \, dx = \frac{nx^3}{3} \Big|_{a-\frac{1}{2n}}^{a+\frac{1}{2n}} < \infty$$

Cerchiamo allora di calcolare gli autovalori con la richiesta “più larga” precedentemente discussa:

*Dimostrazione
che $\sigma(X) = \mathbb{R}$*

$$\begin{aligned} \inf_{\psi \in D(X)} (\Delta X)_{a,\psi}^2 &\leq \inf_{(a)} \inf_{n \in \mathbb{N}} (\Delta X)_{a,\psi_n}^2 \stackrel{(3.10)}{=} \inf_{n \in \mathbb{N}} \int_{a-\frac{1}{2n}}^{a+\frac{1}{2n}} \|X - \lambda\|^2 \|\psi_n\|^2 \, dx = \\ &= \inf_{n \in \mathbb{N}} \int_{a-\frac{1}{2n}}^{a+\frac{1}{2n}} n (x-a)^2 \, dx \\ &\stackrel{(x-a) \rightarrow x}{\Longrightarrow} \inf_{n \in \mathbb{N}} n \int_{-\frac{1}{2n}}^{\frac{1}{2n}} x^2 \, dx = \inf_{n \in \mathbb{N}} n \frac{2}{3} \left(\frac{1}{2n}\right)^3 = 0 \end{aligned}$$

In (a) abbiamo maggiorato l'inf calcolato sull'insieme di *tutte* le funzioni d'onda, con quello sull'insieme ben più ristretto delle ψ_n “che si restringono” su a . Chiaramente, poiché stiamo escludendo molti casi, otterremo un valore di inf che sarà maggiore o uguale a quello generale. Poiché dimostriamo che quest'ultimo inf è esattamente 0 (e l'estremo inferiore di un insieme di numeri positivi è al minimo 0), abbiamo dimostrato l'uguaglianza.

Ciò significa che la fluttuazione di X attorno ad un qualsiasi $a \in \mathbb{R}$ può essere resa arbitrariamente piccola tramite un'opportuna successione $\{\psi_n(x)\} \subset D(X)$, e perciò definendo gli autovalori in questo modo, troveremo che lo spettro $\sigma(X) = \mathbb{R}$ come ci aspettiamo!

Assioma 3. Assioma (dello spettro) L'insieme dei valori ottenibili da una misura dell'osservabile descritta dall'operatore autoaggiunto A è dato da

*Assioma dello
spettro*

$$\sigma(A) = \left\{ a \in \mathbb{R} \mid \inf_{\psi \in D(A) \setminus \{0\}} (\Delta A)_{a,\psi} = 0 \right\} \quad (3.23)$$

Nota: chiaramente la funzione d'onda sempre nulla viene esclusa in queste considerazioni, in quanto banalmente per essa la fluttuazione è nulla (la sua norma è 0). Se ciò non viene specificato sarà comunque da intendersi come sottinteso (come prima, nell'analisi dello spettro di X , ci siamo concentrati solo sulle ψ non nulle).
Nota: Da questa definizione si vede subito come uno spettro non contenga solamente gli autovalori (e quindi elementi “discreti”), ma tutte le misure possibili, che possono avere anche un range continuo (e che rispettano la condizione sull'inf). Analizzeremo meglio la differenza tra queste due “classi di misure” nei prossimi paragrafi.

3.3.1 Osservazioni sullo spettro

1. Per come è definito lo spettro di un operatore autoaggiunto A , vi sono due possibilità: che $\inf = 0$ sia anche un minimo (ossia viene “raggiunto” da qualche funzione d'onda), oppure non lo sia.

- Tra i punti di $\sigma(A)$ vi sono quelli attorno ai quali le fluttuazioni possono annullarsi, per cui:

$$\exists \psi_a \in D(A) \mid (\Delta A)_{a,\psi_a} = 0$$

Questi valori di a sono detti costituire lo **spettro discreto** $\sigma_d(A)$ (o **puntuale** $\sigma_P(A)$) di A . In effetti sono quelli analoghi al caso finito dimensionale, con $\dim \mathcal{H} < \infty$. Come questi ultimi, in particolare, sono le soluzioni dell'equazione agli autovalori:

Spettro discreto

$$(\Delta A)_{a,\psi_a}^2 = \|(A - a)\psi_a\|^2 = 0 \Rightarrow (A - a)\psi_a = 0$$

Perciò gli ψ_a sono detti gli **autovettori** di A appartenenti all'**autovalore** a e $(A - a)\psi_a = 0$ è l'**equazione agli autovalori**.

Nota: lo spettro discreto è costituito al più da un insieme numerabile di punti. Se così non fosse, avremmo un insieme non numerabile di autovettori ortogonali tra loro nello spazio di Hilbert, cosa che è assurda perché \mathcal{H} è separabile. Tuttavia, lo spettro discreto potrebbe essere denso²⁷ in \mathcal{H} .

Numerabilità dello spettro discreto

- D'altro canto, i punti dello spettro di $\sigma(A)$ in cui la fluttuazione ha estremo inferiore nullo in $D(A)$, ma **non minimo**, sono detti **spettro continuo** $\sigma_C(A)$ di A e per essi **non c'è** autovettore $\psi_a \in D(A)$. (Come vedremo, nel formalismo di Dirac c'è una caratterizzazione più semplice di $\sigma_C(A)$)

Spettro continuo

2. È intuitivo che lo spettro non possa dipendere dalle trasformazioni unitarie (“rotazioni” dello spazio infinito-dimensionale): essendo una caratteristica

Trasformazioni unitarie preservano lo spettro

²⁷ciò si verifica, per esempio, nel caso di elettroni in un cristallo disordinato. La scoperta di ciò valse un premio Nobel in fisica nel 1977 a Philip Warren Anderson

fisica, cioè misurabile, non deve dipendere dal modo che adottiamo per descriverlo o determinarlo.

Se U è **unitario** e A è autoaggiunto:

$$\sigma(U^\dagger A U) = \sigma(A)$$

Osserviamo innanzitutto che se $U^\dagger A U \psi$ è ben definito, ossia se risulta in un vettore $\in \mathcal{H}$, necessariamente $\phi \equiv U \psi \in D(A)$. Perciò:

$$\psi \in D(U^\dagger A U) \Rightarrow \phi \equiv U \psi \in D(A) \quad (3.24)$$

D'altro canto, se partiamo da un generico $\phi \in D(A)$ possiamo procedere a ritroso e trovare un ψ t.c. t.c. $U \psi = \phi$ (U è unitario, perciò invertibile) e allora $U^\dagger \phi = U^\dagger U \psi = \psi$, con $\psi \in D(U^\dagger A U)$.

Perciò vale la doppia implicazione:

$$\psi \in D(U^\dagger A U) \Leftrightarrow \phi = U \psi \in D(A) \quad (3.25)$$

Ciò garantisce che le operazioni che eseguiamo nei prossimi passaggi siano ben definite. Calcolando l'estremo inferiore:

$$\begin{aligned} \inf_{\psi \in D(U^\dagger A U)} (\Delta(U^\dagger A U))_{a,\psi}^2 &= \inf_{\psi \in D(U^\dagger A U)} \|(U^\dagger A U - a)\psi\|^2 = \\ &= \inf_{\psi \in D(U^\dagger A U)} \|(U^\dagger A U - U^\dagger a U)\psi\|^2 = \inf_{\psi \in D(U^\dagger A U)} \|U^\dagger (A - a)U \psi\|^2 = \\ &= \inf_{(a) \psi \in D(U^\dagger A U)} \|(A - a)U \psi\|^2 = \inf_{(b) \phi \in D(A)} \|(A - a)\phi\|^2 = \inf_{\phi \in D(A)} (\Delta A)_{a,\phi}^2 \end{aligned}$$

dove in (a) si è usato il fatto che un operatore unitario preserva la norma, mentre in (b) si è usata $U \psi = \phi$ e la relazione sui domini trovata in (3.25).

otteniamo la tesi.

In particolare ciò significa che lo spettro di un operatore autoaggiunto non dipende dalla sua rappresentazione in uno spazio di Hilbert concreto (come tutte le quantità confrontabili con l'esperimento), perché gli isomorfismi tra spazi di Hilbert concreti in cui si rappresentano le osservabili sono unitari.

Perciò: $\sigma(-i\hbar \frac{d}{dx})$ in $L^2(\mathbb{R}, dx)$ è uguale a $\sigma(p)$ in $L^2(\mathbb{R}, dp)$, ove con p si è qui rappresentato l'operatore di moltiplicazione per p .

Per esercizio, verificalo per X (posizione). Come è rappresentato in $L^2(\mathbb{R}, dp)$?

Osservazione 3. Se A è limitato e autoaggiunto:

$$\|A\| = \sup_{\lambda \in \sigma(A)} |\lambda|$$

Operatore
limitato \leftrightarrow
spettro limitato

Quindi A è limitato (in senso matematico) se e solo se l'insieme dei valori ottenibili dalla misura è limitato. Abbiamo perciò un'altra corrispondenza tra l'idea matematica di risultato e l'idea fisica di misura sperimentale.

(Lezione 10 del
17/10/2018)

3.3.2 Spettro matematico

Come già accennato in precedenza, esiste un modo più generale per definire lo spettro di un operatore. Lo **spettro matematico** di un operatore A (anche non autoaggiunto, per esempio nel caso di un operatore unitario) è dato da:

*Nuova definizione
di spettro $\sigma(A)$
di un operatore A*

$$\sigma(A)_{\text{mat}} = \{\mathbb{C} \setminus \{z \in \mathbb{C} \mid \|(A - z\mathbb{I})^{-1}\| < \infty\}\} = \{z \in \mathbb{C} \mid (A - z\mathbb{I})^{-1} \notin \mathcal{B}(\mathcal{H})\}$$

In altre parole, lo spettro di A è dato da tutti i numeri complessi z per cui $(A - z\mathbb{I})^{-1}$ non è un operatore limitato.

In effetti, nel caso matriciale, se a è un autovalore, allora $(A - a\mathbb{I})$ è singolare (è proprio la matrice di cui facciamo il ker per calcolare gli autovettori associati ad a), e quindi l'inversa $(A - a\mathbb{I})^{-1} \notin \mathcal{B}(\mathcal{H})$ proprio non esiste (e di certo non è un operatore limitato!).

Verifichiamo che tale nuova definizione è equivalente a quella precedente (3.23) nel caso di operatori autoaggiunti ($A = A^\dagger$).

Innanzitutto, se lo spettro di A è reale (come vogliamo che sia), allora ogni numero non reale $a \notin \mathbb{R}$ fa sì che $(A - a\mathbb{I})^{-1}$ sia un operatore limitato:

$$\sigma(A)_{\text{mat}} \subseteq \mathbb{R} \Rightarrow (A - a\mathbb{I})^{-1} \in \mathcal{B}(\mathcal{H}), a \notin \mathbb{R}$$

Consideriamo ora gli $a \in \mathbb{R}$ che generano un $(A - a\mathbb{I})^{-1}$ limitato. Allora, per la limitatezza, vale:

$$\|(A - a\mathbb{I})^{-1}\| < C$$

Dimostriamo che tali a non soddisfano la definizione in (3.23):

$$\begin{aligned} \inf_{\psi \in D(A)} (\Delta A)_{a,\psi}^2 &= \inf_{\psi \in D(A)} \frac{\|(A - a\mathbb{I})\psi\|^2}{\|\psi\|^2} \stackrel{(a)}{=} \inf_{\phi} \frac{\|\phi\|^2}{\|(A - a\mathbb{I})^{-1}\phi\|^2} = \\ &\stackrel{(b)}{=} \frac{1}{\sup_{\phi} \frac{\|(A - a\mathbb{I})^{-1}\phi\|^2}{\|\phi\|^2}} \stackrel{(c)}{=} \frac{1}{\|(A - a\mathbb{I})^{-1}\|^2} > \frac{1}{C^2} \end{aligned}$$

In (a) si è definito $\phi = (A - a\mathbb{I})\psi$ per $\psi \in D(A)$, da cui $(A - a\mathbb{I})^{-1}\phi = \psi$. In (b) portiamo tutto al denominatore, e l'estremo inferiore di una frazione con numeratore fissato si ottiene “massimizzando” il denominatore. Ma allora applicando Riesz (norma di un funzionale) in (c) si ha che l'inf di partenza è maggiore di un numero certamente maggiore di 0. Perciò gli $a \notin \sigma(A)_{\text{mat}}$, ossia quelli per cui $(A - a\mathbb{I})^{-1} \in \mathcal{B}(\mathcal{H})$, non sono autovalori per la (3.23). Perciò:

$$\sigma(A)_{\text{mat}}^C \subseteq \sigma(A)^C \Leftrightarrow \sigma(A) \subseteq \sigma(A)_{\text{mat}}$$

(dove B^C è l'insieme complementare di B).

Ci serve ora dimostrare l'inclusione inversa, ossia che tutti i valori nello “spettro matematico” per le osservabili siano anche autovalori per la (3.23).

Se $a \in \sigma(A)_{\text{mat}}$, cioè se $(A - a\mathbb{I})^{-1} \notin \mathcal{B}(\mathcal{H})$, abbiamo 3 casi:

1. $\exists \psi_a \mid (A - a\mathbb{I}) \psi_a = 0 \Rightarrow (\Delta A)_{a, \psi_a} = 0$

In altre parole, come nel caso matriciale, esiste un autovettore ψ_a di autovalore a dell'operatore A , per cui $(A - a\mathbb{I})^{-1}$ **non è definito**.

Inoltre si ha che il codominio (o range) di $(A - a\mathbb{I})$ non è denso in \mathcal{H} , poiché da esso manca il sottospazio generato da ψ_a (che essendo un autovettore è costituito da vettori del tipo $a\psi_a$, che vengono rimossi nella differenza).

Questo caso corrisponde allo **spettro discreto** precedentemente definito.

2. $(A - a)^{-1}$ esiste in D denso, è un **operatore illimitato**:

$$\sup_{\phi \in D(A)} \frac{\|(A - a)^{-1} \phi\|}{\|\phi\|} = \infty$$

Ciò significa che possiamo trovare una sequenza di vettori ϕ_n di norma unitaria che fanno sì che $(A - a)^{-1}$ diverga. Matematicamente:

$$\exists \{\phi_n\} \subset D((A - a)^{-1}), \|\phi_n\| = 1 \text{ t.c. } \|(A - a)^{-1} \phi_n\| \xrightarrow{n \rightarrow \infty} \infty$$

Ma allora esiste una sequenza di “autovettori approssimati” ψ_n per cui $\|(A - a)\psi_n\| \xrightarrow{n \rightarrow \infty} 0$. Infatti, definendo opportunamente:

$$\psi_n = \frac{(A - a)^{-1} \phi_n}{\|(A - a)^{-1} \phi_n\|}$$

Si ha che:

$$\|(A - a)\psi_n\| = \frac{\|(A - a)(A - a)^{-1} \phi_n\|}{\|(A - a)^{-1} \phi_n\|} = \frac{\|\phi_n\|}{\|(A - a)^{-1} \phi_n\|} \xrightarrow{n \rightarrow \infty} 0$$

Ma questi ψ_n soddisfano la definizione di autovalori che abbiamo dato in (3.23):

$$\inf_{\psi \in D(A)} (\Delta A)_{a, \psi}^2 \leq \lim_{n \rightarrow \infty} \|(A - a)\psi_n\|^2 = 0$$

(si è usata nuovamente la tecnica di calcolare l'inf maggiorandolo con l'inf di un sottoinsieme di tutte le ψ che contiene solo le ψ_n appena costruite).

Questo caso corrisponde allora allo **spettro continuo**.

3. Per esclusione, c'è anche il caso in cui non esiste ψ_a autovettore, e quindi $(A - a\mathbb{I})^{-1}$ è definito, ma non con dominio denso (perciò non è possibile estenderlo univocamente a tutto \mathcal{H} , e quindi ci saranno dei vettori di \mathcal{H} per cui non è definito).

Questo caso costituisce lo **spettro residuo**, che non ha alcuna interpretazione fisica.

Fortunatamente vale il seguente teorema:

Teorema 3.3.1. *Se A è autoaggiunto o unitario allora non ha spettro residuo.*

Avendo allora mostrato che tra le componenti dello spettro matematico $\sigma(A)_{\text{mat}}$ vi sono tutti gli autovalori che abbiamo già incontrato nell'esaminare lo spettro $\sigma(A)$, abbiamo dimostrato che $\sigma(A)_{\text{mat}} \subseteq \sigma(A)$, e ciò, unito all'inclusione inversa già discussa, fa sì che le due definizioni siano equivalenti.

Notiamo infine che basta spostarsi di poco da un operatore autoaggiunto per perdere completamente diverse proprietà fisiche di fondamentale importanza. Per esempio, quando avevamo definito l'operatore momento P_0 (che non è autoaggiunto, in quanto la definizione giusta sarebbe quella di P), ponendo $\mathcal{H} = L^2([0, 1], dx)$:

$$P_0 = -i\hbar \frac{d}{dx}$$

$$D(P_0) = \{\psi \in \mathcal{H} \mid \psi \text{ regolare e } \psi(0) = \psi(1) = 0\}$$

$$-i\hbar \frac{d}{dx} \psi(x) = \lambda \psi(x); \quad \psi(x) = \lambda e^{i\frac{\lambda}{\hbar}x}$$

Imponendo $\psi(0) = 0$ si ha che non esiste alcun λ autovalore.

Ciò significa che lo spettro di questo operatore (che non può essere l'insieme vuoto, come si può dimostrare in generale) è unicamente **residuo**, e quindi non ha interpretazione fisica.

3.4 Il Formalismo di Dirac

Osserviamo ora più nel dettaglio la notazione di Dirac²⁸, finora occasionalmente utilizzata senza definizioni precise. Dirac elaborò il suo formalismo senza l'utilizzo degli spazi di Hilbert, dato che questi ancora non erano stati ideati. Solo tramite la teoria delle distribuzioni e una matematica decisamente sofisticata è stato possibile dare un senso logico ed una giustificazione rigorosa al suo formalismo, che assume così i caratteri e la consistenza propri della matematica. L'idea alla base della notazione di Dirac è quella di partire dal formalismo per lo spettro discreto e applicarlo, in maniera naturale, anche quello continuo. Questo paragrafo espone il formalismo con dei cenni alla complessa matematica che permette a tutto questo di “reggersi” in modo consistente.

3.4.1 Notazione e formalismo per lo spettro discreto

Sia A un operatore autoaggiunto, il cui spettro è solo dato dagli autovalori “di algebra lineare”, e quindi è puramente puntuale: $\sigma(A) = \sigma_P(A) = \{\lambda_n\}$.

Per rappresentare i funzionali e gli stati in notazione di Dirac si utilizzano i bra e i ket già presentati negli scorsi paragrafi. Come già discusso in precedenza, questa notazione viene introdotta in generale per denotare i vettori “astratti”, cioè permette di rappresentare gli elementi di \mathcal{H} senza la necessità di scegliere una *base*, dunque una rappresentazione “concreta” dei vettori. Per semplicità ci limiteremo per ora ai casi in cui i λ_n non presentano degenerazione, ossia per

²⁸Un utile resoconto per il formalismo di Dirac è disponibile in [7]

cui ad ogni autovalore corrisponde un solo autovettore²⁹. In notazione di Dirac l'equazione agli autovalori diviene:

$$A|\lambda_n\rangle = \lambda_n|\lambda_n\rangle$$

Siccome i $|\lambda_n\rangle$ costituiscono una base ON, possiamo scrivere un qualsiasi altro ket $|\psi\rangle$ o bra $\langle\phi|$ come somme delle loro proiezioni su di essi:

$$|\psi\rangle = \sum_n |\lambda_n\rangle \langle\lambda_n|\psi\rangle; \quad \langle\phi| = \sum_n \langle\phi|\lambda_n\rangle \langle\lambda_n| \quad (3.26)$$

dove $\langle\lambda_n|$ sono i funzionali in \mathcal{H}^* associati a $|\lambda_n\rangle$ per Riesz. In notazione di Dirac la relazione di Parseval, che permette di scrivere il prodotto scalare in termini dei coefficienti rispetto alla base, ha la seguente forma:

$$\langle\phi|\psi\rangle = \sum_n \langle\phi|\lambda_n\rangle \langle\lambda_n|\psi\rangle \quad (3.27)$$

Si ha che (3.26) e (3.27) sono compatibili solamente se $\langle\lambda_n|\lambda_m\rangle = \delta_{mn}$, e le formule precedenti sono sintetizzate nella **relazione di completezza di Dirac**:

$$\sum_n |\lambda_n\rangle \langle\lambda_n| = \mathbb{I}$$

Tale relazione esprime il fatto che *ogni* vettore di \mathcal{H} può essere rappresentato mediante i vettori della base $|\lambda_n\rangle$. Dato che $|\lambda_n\rangle \langle\lambda_n|$ è il proiettore nell'autospazio di λ_n , mediante la relazione di completezza è definita anche la rappresentazione spettrale di A :

$$A = A\mathbb{I} = A \sum_n |\lambda_n\rangle \langle\lambda_n| = \sum_n A |\lambda_n\rangle \langle\lambda_n| = \sum_n \lambda_n |\lambda_n\rangle \langle\lambda_n|$$

3.4.2 Estensione del formalismo allo spettro continuo

Nei precedenti paragrafi abbiamo visto che l'equazione agli autovalori $A|\lambda\rangle = \lambda|\lambda\rangle$ non fornisce le soluzioni per lo spettro continuo. Infatti se $\lambda \in \sigma_C(A)$, abbiamo dimostrato che $|\lambda\rangle$ non può essere un autovettore in \mathcal{H} . L'intuizione di Dirac fu che $|\lambda\rangle$ potesse comunque risolvere l'equazione agli autovalori, dunque essere un autovettore anche se in $\sigma_C(A)$, ma solo in uno spazio diverso da \mathcal{H} , sicuramente “più grande”, ovvero uno spazio che estendesse \mathcal{H} .

Per comprendere di cosa si tratta consideriamo ad esempio l'operatore posizione $A = X$, che agisce su $|\lambda\rangle$ come $X|\lambda\rangle = x|\lambda\rangle$. Allora l'equazione agli autovalori diviene (nella rappresentazione in posizioni):

$$X|\lambda\rangle = \lambda|\lambda\rangle \quad \Rightarrow \quad x|\lambda\rangle = \lambda|\lambda\rangle \quad \Rightarrow \quad (x - \lambda)|\lambda\rangle = 0$$

Solo la funzione nulla (che non è in \mathcal{H}) soddisfa questa equazione. Tuttavia un altro “oggetto” che soddisfa questa condizione (e che quindi può fungere da “autovettore” generalizzato) è la delta di Dirac. Per definire rigorosamente la δ il formalismo

²⁹^Nel caso finito dimensionale questa è la situazione delle matrici diagonalizzabili.

matematico elementare ovviamente non basta. Tuttavia Dirac suppose che la δ fosse un “nuovo” elemento dello spazio degli stati, e che potesse comunque essere manipolata come $\delta(x - \lambda) = \langle x | \lambda \rangle$. Sostituendola al posto dell'autovettore la δ può essere considerata una soluzione dell'equazione agli autovalori:

$$(x - \lambda)\langle x | \lambda \rangle = 0 \quad \Rightarrow \quad (x - \lambda)\delta(x - \lambda) = 0$$

infatti una nota proprietà della δ (da metodi) è che $(x - \lambda)\delta(x - \lambda) = 0$ anche per $x = \lambda$ dove la δ vale (idealmente) ∞ . Un altro modo per scrivere tale proprietà è

$$x\delta(x - \lambda) = \lambda\delta(x - \lambda)$$

la quale esprime il fatto che $\delta(x - \lambda)$ sia proprio una soluzione dell'equazione agli autovalori dell'operatore posizione. La soluzione non è identicamente nulla: la proprietà più importante della δ è il fatto che

$$\int_{\mathbb{R}} \delta(x - \lambda) dx = 1$$

Dunque la δ , a meno del fatto di non essere una funzione “ordinaria”, ha le stesse proprietà di una funzione di stato “ammessa” in \mathcal{H} ! Questo fatto ha diversi vantaggi. Innanzitutto si mantiene, almeno formalmente, la struttura dell'equazione agli autovalori $A|\lambda\rangle = \lambda|\lambda\rangle$ identica a quella del caso dello spettro puramente discreto. Ma in questo modo ad ogni $\lambda \in \mathbb{R}$ corrisponde l'autostato “generalizzato” $\delta(x - \lambda)$, cosicché λ può assumere qualsiasi valore reale, e lo spettro $\sigma(X)$ è tutto \mathbb{R} , esattamente quello che ci aspettiamo dallo spettro dell'operatore posizione! In realtà gli autostati di tali autovalori λ non esistono in \mathcal{H} , per il fatto che le δ non sono funzioni rigorosamente definite nello spazio degli stati, almeno nel formalismo matematico utilizzato finora. Trattare la δ come uno stato “possibile” porta in ogni caso a conclusioni corrette sullo spettro di X , e questo, almeno a prima vista, può sembrare una coincidenza. In realtà questa formulazione è perfettamente equivalente a quella di von Neumann, ma dimostrare ciò è tutt'altro che banale. Vediamo ora le relazioni per gli stati nel caso dello spettro continuo. Consideriamo un operatore A il cui spettro $\sigma(A)$ sia puramente continuo, dunque $\lambda \in \sigma_C(A) = \sigma(A)$ limitandoci ancora al caso di autovalori senza degenerazione. Se gli autovalori “generalizzati” di $\sigma_C(A)$ sono, come per il caso dello spettro dell'operatore posizione, una quantità non numerabile in \mathbb{R} , come si può scrivere uno stato $|\psi\rangle$ come somma di un infinito non numerabile di autostati “generalizzati” di base? La notazione di Dirac lo permette, semplicemente integrando tutte le “proiezioni” di $|\psi\rangle$ su tutti gli autostati generalizzati dello spettro continuo di A :

$$|\psi\rangle = \int_{\sigma_C(A)} d\lambda |\lambda\rangle \langle \lambda | \psi \rangle \quad \langle \phi | = \int_{\sigma_C(A)} d\lambda \langle \phi | \lambda \rangle \langle \lambda | \quad (3.28)$$

Si noti l'analogia con il caso dello spettro discreto: l'unica differenza è che l'integrale sostituisce la sommatoria, non essendo possibile includere tutti gli elementi di $\sigma_C(A)$ mediante una somma di una quantità numerabile di termini. In questo modo l'estensione allo spettro continuo è molto naturale, non richiedendo di definire

oggetti complessi come le famiglie spettrali. Tutto questo rispecchia il minimalismo di Dirac, il quale antepose la semplicità a tutto il resto nella formulazione della MQ. Sempre in analogia con lo spettro discreto, la relazione di Parseval per lo spettro continuo diviene:

$$\langle \phi | \psi \rangle = \int_{\sigma_C(A)} d\lambda \langle \phi | \lambda \rangle \langle \lambda | \psi \rangle \quad (3.29)$$

E la (3.28) è compatibile con la (3.29) solamente se vale $\langle \lambda | \lambda' \rangle = \delta(\lambda - \lambda')$, la quale si sintetizza nella *relazione di completezza*, valida anche per il caso dello spettro continuo:

$$\int_{\sigma_C(A)} d\lambda |\lambda\rangle \langle \lambda| = \mathbb{I}$$

Infine l'ultima analogia con lo spettro discreto consiste nella *rappresentazione spettrale* di un operatore A con spettro continuo:

$$A \int d\lambda |\lambda\rangle \langle \lambda| = \int d\lambda \lambda |\lambda\rangle \langle \lambda| = A$$

3.4.3 Formalismo per uno spettro generale

I casi visti fin ora sono quelli con spettro puramente discreto e puramente continuo, in entrambi i casi senza degenerazioni. Fortunatamente il formalismo di Dirac si generalizza facilmente ad uno spettro qualsiasi (sia discreto che continuo) $\sigma(A) = \sigma_P(A) \cup \sigma_C(A)$ e con eventuali degenerazioni.

Definizione 15 (degenerazione). Dato un autovalore $\lambda_n \in \sigma_P(A)$ o un autovalore generalizzato ($\lambda \in \sigma_C(A)$) diremo che ha **degenerazione** $d(\lambda_n)$ o $d(\lambda)$ se le equazioni agli autovalori corrispondenti (nel senso di Dirac) hanno $d(\lambda_n)$ o $d(\lambda)$ soluzioni indipendenti (per \mathcal{H} nel senso esplicitato in seguito) che denotiamo rispettivamente con $|\lambda_n, r\rangle$ (per $r = 1, \dots, d(\lambda_n)$) e $|\lambda, r\rangle$ per $r = 1, \dots, d(\lambda)$ (con $d(\lambda) \in \mathbb{N}$ ed eventualmente ∞).

Nel caso in cui lo spettro abbia una componente discreta e una continua allora la notazione di Dirac permette semplicemente di sommare le due componenti nelle formule del paragrafo precedente. Ad esempio la relazione di completezza nel caso di uno spettro qualsiasi diviene:

$$\sum_{\lambda_n \in \sigma_P(A)} \sum_{r=1}^{d(\lambda_n)} |\lambda_n, r\rangle \langle \lambda_n, r| + \int_{\sigma_C(A)} d\lambda \sum_{r=1}^{d(\lambda)} |\lambda, r\rangle \langle \lambda, r| = \mathbb{I}_{\mathcal{H}}$$

Tramite la relazione di completezza possiamo scrivere funzioni per osservabili nel caso più generale:

$$f(A) = \sum_{\lambda_n \in \sigma_D(A)} f(\lambda_n) \sum_{r=1}^{d(\lambda_n)} |\lambda_n, r\rangle \langle \lambda_n, r| + \int_{\sigma_C(A)} d\lambda f(\lambda) \sum_{r=1}^{d(\lambda)} |\lambda, r\rangle \langle \lambda, r|$$

Notiamo che per lo spettro generale anche nei casi più “semplici” è necessaria la completezza.

Ad esempio sia H l’hamiltoniana dell’atomo di idrogeno. Negli stati in cui l’elettrone è legato l’energia è quantizzata, con degenerazione n^2 , ma quando è libero possiamo scegliere una qualsiasi energia in un range continuo, e vi è un numero di stati infinito con una determinata energia (la degenerazione è infinita):

$$\sigma(H) = \left\{ -\frac{c}{n^2}, n \in \mathbb{N} \right\} \cup \{x > 0, x \in \mathbb{R}\}$$

Abbiamo perciò una componente discreta dello spettro (con degenerazione finita) e una continua (con degenerazione addirittura ∞).

3.4.4 Spazio degli autovettori dello spettro continuo

Come visto prima, se $\lambda \in \sigma_C(A)$ gli autovettori $|\lambda\rangle$, $|\lambda, r\rangle$ non si trovano in \mathcal{H} , ma in uno spazio più grande (sono autovettori “generalizzati”). In questo paragrafo si cerca di definire questo spazio più ampio, accennando alla matematica che ha permesso di giustificare il formalismo di Dirac.

Se si definisce $F_\lambda(x) \equiv \langle x | \lambda \rangle$ e lo si applica all’operatore posizione X abbiamo visto che la soluzione dell’equazione agli autovalori risulta:

$$XF_\lambda(x) = \lambda F_\lambda(x) \quad F_\lambda(x) = \delta(x - \lambda) \notin L^2(\mathbb{R})$$

Tuttavia la δ non è una funzione reale a valori reali, pertanto non può essere una funzione di $L^2(\mathbb{R})$. Tuttavia (da metodi) sappiamo che $\delta(x - \lambda)$ appartiene allo spazio delle distribuzioni $\mathcal{S}'(\mathbb{R})$, il quale contiene $L^2(\mathbb{R})$.

Proviamo la stessa cosa con l’operatore momento. Se si definisce $F_\mu(x) \equiv \langle x | \mu \rangle$ la soluzione dell’equazione agli autovalori risulta:

$$PF_\mu(x) = \mu F_\mu(x) \quad F_\mu(x) = e^{\frac{i}{\hbar}\mu x} \notin L^2(\mathbb{R}), \in \mathcal{S}'(\mathbb{R})$$

infatti

$$PF_\mu(x) = -i\hbar \frac{\partial}{\partial x} e^{\frac{i}{\hbar}\mu x} = -i\hbar \frac{i}{\hbar} \mu e^{\frac{i}{\hbar}\mu x} = \mu F_\mu(x)$$

La situazione è analoga, anche se opposta a quella dell’operatore posizione: in questo caso l’autovettore generalizzato identifica un preciso valore del momento (una sinusoide “monocromatica”). Tuttavia si noti questa soluzione *non* è in $L^2(\mathbb{R})$, infatti tale funzione non è quadrato sommabile. Non a caso la trasformata di Fourier di $e^{\frac{i}{\hbar}\mu x}$ è proprio la δ di Dirac. La trasformata di Fourier per funzioni di questo tipo è però definita solo nello spazio delle distribuzioni \mathcal{S}' e non in L^2 .

Viste in quest’ottica le distribuzioni sono estremamente comode, ma la domanda sorge spontanea: possono essere considerate come “stati fisici”, cioè stati corrispondenti a situazioni reali? Se questo fosse possibile, allora tutta la formulazione assiomatica della MQ non sarebbe coerente con il principio di indeterminazione. Il fatto che le soluzioni dell’equazione agli autovalori non siano in $L^2(\mathbb{R})$ significa che esse non corrispondono a “stati fisici”, fisicamente possibili. In particolare, ciò

vuol dire che non è comunque possibile localizzare una particella perfettamente in un punto, né assegnarle un singolo momento definito. Questo artefatto matematico permette solo di estendere lo spazio degli autostati, ma non “aggiunge” funzioni di L^2 bensì istituisce nuove soluzioni “non fisiche” dell’equazione agli autovalori. La teoria delle distribuzioni permette di formalizzare in modo rigoroso questi concetti. Lo spazio $\mathcal{S}'(\mathbb{R})$ è definito come l’insieme dei funzionali lineari continui su $\mathcal{S}(\mathbb{R})$, ovvero lo spazio delle funzioni \mathcal{C}^∞ che decrescono all’infinito (assieme a tutte le loro derivate) più rapidamente di x^{-n} per ogni $n \in \mathbb{N}$. Questa definizione è facilmente generalizzabile in dimensione n :

- $\mathcal{S}(\mathbb{R}^n)$ è lo spazio delle funzioni \mathcal{C}^∞ su \mathbb{R}^n che decrescono più velocemente della norma $\|x\|^{-m}$ per ogni $m \in \mathbb{N}$;
- $\mathcal{S}'(\mathbb{R}^n)$ è lo spazio dei funzionali lineari continui su $\mathcal{S}(\mathbb{R}^n)$.

Osservazione 4. Lo spazio delle distribuzioni \mathcal{S}' è più esteso di L^2 e sicuramente ammette funzioni particolari come la δ . Tuttavia \mathcal{S}' ha molte limitazioni: notiamo che esistono “funzioni regolarissime” che non appartengono a $\mathcal{S}'(\mathbb{R})$. Infatti, le distribuzioni regolari (descritte da funzioni) in $\mathcal{S}'(\mathbb{R})$ non possono crescere più rapidamente di un polinomio in x , per $x \rightarrow \infty$, quindi ad esempio $e^x \notin \mathcal{S}'(\mathbb{R})$.

Definendo i nuovi spazi \mathcal{S} e \mathcal{S}' si dimostra (da metodi) che vale la seguente catena di inclusioni:

$$\mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R}) \approx [L^2(\mathbb{R})]' \subset \mathcal{S}'(\mathbb{R}) \quad (3.30)$$

Dunque l’idea, come già anticipato, è quella di risolvere l’equazione agli autovalori proprio nello “spazio più grande” $\mathcal{S}'(\mathbb{R})$. Questo spazio è quello per cui è possibile trovare soluzioni per tutti i $\lambda \in \mathbb{R}$ per gli operatori X e P , e dunque trovare lo spettro voluto per entrambi gli operatori:

$$\sigma(X) = \sigma_C(X) = \mathbb{R} \quad \sigma(P) = \sigma_C(P) = \mathbb{R}$$

Abbiamo visto che, almeno nel caso di posizione e momento, gli autovalori generalizzati sono quelli dello spettro continuo. Questo fatto vale in realtà in generale per qualsiasi osservabile: la differenza tra gli autovettori generalizzati di \mathcal{S}' e gli autovettori “algebrici” di \mathcal{H} è esattamente ciò che distingue lo *spettro discreto* dallo *spettro continuo*. Concretamente, se la soluzione dell’equazione agli autovalori è in $\mathcal{S}'(\mathbb{R})$, ma non è in \mathcal{H} l’autovalore corrispondente appartiene allo **spettro continuo**. Se invece l’autovettore è in \mathcal{H} lo spettro corrispondente appartiene allo **spettro discreto**.

(Lezione 11 del
18/10/2018)

Queste idee ovviamente non appartengono più alle intuizioni di Dirac, ma fanno parte della successiva teoria distribuzionale. Uno dei massimi esponenti di questa teoria fu Gel’fand della scuola russa. Come trovare lo spettro di un operatore autoaggiunto A nel formalismo di Dirac? L’idea di Gel’fand fu quella di trovare uno spazio Φ_A più “piccolo” di \mathcal{H} , ma tale che il suo duale Φ'_A potesse estendere lo spazio degli stati \mathcal{H} , in pratica:

*Spettro di un
operatore A
autoaggiunto per
Gel’fand*

$$\Phi_A \subset \mathcal{H} \approx \mathcal{H}' \subset \Phi'_A \quad (3.31)$$

Tale catena di inclusioni, che generalizza quella vista in (3.30), è detta **tripletta di Gel'fand**, e $(\Phi_A, \mathcal{H}, \Phi'_A)$ è detto **spazio di Hilbert equipaggiato** o **allargato** (*rigged Hilbert space*). Gel'fand comprese che creando uno spazio con una topologia più forte, dunque uno spazio Φ_A contenuto in \mathcal{H} , la topologia del suo duale Φ'_A sarebbe stata più debole, e dunque si sarebbe creata un'estensione di \mathcal{H}' . Un esempio di spazio di Hilbert equipaggiato è lo spazio dato dalle distribuzioni: $(\mathcal{S}, \mathcal{H}, \mathcal{S}')$.

La teoria che a partire da queste definizioni fornisce una spiegazione completa del formalismo di Dirac è molto sofisticata. Infatti è necessario fare le seguenti ulteriori richieste nella definizione di Φ_A :

1. Φ_A dev'essere *contenuto nel dominio* di A , dunque $\Phi_A \subseteq D(A)$. Ad esempio nel caso delle distribuzioni e dell'operatore posizione, lo spazio delle funzioni di prova $\mathcal{S}(\mathbb{R})$ è contenuto in $D(X)$;
2. Nella topologia di \mathcal{H} , Φ_A dev'essere *denso* in \mathcal{H} , ovvero $\overline{\Phi_A} = \mathcal{H}$. Nel caso particolare delle distribuzioni si dimostra che $\overline{\mathcal{S}(\mathbb{R})} = L^2(\mathbb{R})$;
3. A dev'essere *continuo* in Φ_A nella topologia di Φ_A , e questo è vero per X nella topologia di $\mathcal{S}(\mathbb{R})$;
4. Φ_A dev'essere “*nucleare*” cioè i funzionali lineari continui sullo spazio³⁰ delle coppie $\Phi_A \times \Phi_A$ sono lineari e continui anche nello spazio delle combinazioni lineari (eventualmente infinite) degli elementi dello spazio dei prodotti $\Phi_A \otimes \Phi_A$. Nel caso delle distribuzioni una coppia è ad esempio $f, g \in \mathcal{S}(\mathbb{R}) \times \mathcal{S}(\mathbb{R})$ mentre i prodotti sono $f(x, y) \in \mathcal{S}(\mathbb{R}^2) = \mathcal{S}(\mathbb{R}) \otimes \mathcal{S}(\mathbb{R})$. Dunque la nuclearità richiede che i funzionali lineari continui su $\mathcal{S}(\mathbb{R}) \times \mathcal{S}(\mathbb{R})$ siano lineari e continui su $\mathcal{S}(\mathbb{R} \otimes \mathbb{R}) = \mathcal{S}(\mathbb{R}^2)$.

Giustificazione della richiesta di nuclearità

Perché è necessario richiedere l'ultima richiesta, quella di nuclearità? Consideriamo un'osservabile A con spettro puramente discreto $\sigma(A) = \sigma_P(A)$ (e senza degenerazione). La relazione di completezza è

$$\sum_n |\lambda_n\rangle \langle \lambda_n| = \mathbb{I}_{\mathcal{H}}$$

pertanto si ha

$$\langle \phi | A | \psi \rangle = \langle \phi | \mathbb{I} A \mathbb{I} | \psi \rangle = \sum_{n,m} \langle \phi | \lambda_n \rangle \langle \lambda_n | A | \lambda_m \rangle \langle \lambda_m | \psi \rangle = \sum_{n,m} A_{nm} \phi_n^* \psi_m$$

$\langle \phi | A | \psi \rangle$ è ovviamente lineare separatamente sia in ϕ_n^* che in ψ_m . Ma da questo risultato si osserva che è lineare anche nei prodotti $\phi_n^* \psi_m$, e quindi in tutte le loro combinazioni lineari.

³⁰Il simbolo \otimes indica il prodotto tensoriale tra due spazi.

Il formalismo di Dirac deve permettere di generalizzare anche questo in modo naturale allo spettro continuo $\sigma(A) = \sigma_C(A)$. Poniamo che sia senza degenerazione (sempre per semplicità di notazione). La relazione di completezza nel caso dello spettro continuo è:

$$\int_{\sigma_C(A)} d\lambda |\lambda\rangle \langle \lambda| = \mathbb{I}_{\mathcal{H}}$$

dunque analogamente al caso dello spettro discreto si ottiene:

$$\begin{aligned} \langle \phi | A | \psi \rangle &= \langle \phi | \mathbb{I} A \mathbb{I} | \psi \rangle = \int d\lambda d\lambda' \langle \phi | \lambda \rangle \langle \lambda | A | \lambda' \rangle \langle \lambda' | \psi \rangle \\ &= \int d\lambda d\lambda' \langle \lambda | A | \lambda' \rangle \langle \lambda | \phi \rangle^* \langle \lambda' | \psi \rangle \end{aligned}$$

Ma ciò non è possibile in L^2 . Infatti si consideri ad esempio $A = \mathbb{I}$, e siano $f, g \in L^2$. Allora se $x = \lambda$ e $y = \lambda'$ deve verificarsi

$$(f, \mathbb{I}g) = \int dx dy k(x, y) f^*(x) g(y) = \int f^*(x) g(x) dx$$

ovvero il funzionale è lineare e continuo in f^* e in g separatamente e nei loro prodotti. Tuttavia non esiste alcuna funzione $k(x, y)$ che renda il funzionale lineare e continuo in $L^2(\mathbb{R}) \otimes L^2(\mathbb{R}) = L^2(\mathbb{R}^2)$. Ma una soluzione possibile è prendere $k(x, y) = \delta(x - y)$ che come sappiamo non è in $L^2(\mathbb{R}^2)$ ma in $\mathcal{S}'(\mathbb{R}^2)$. Per questo, se lo spazio equipaggiato Φ'_A è generico, è necessario richiedere come ipotesi che la condizione di nuclearità sia verificata.

(Lezione 11 del
18/10/2018)

3.4.5 Generalizzazione ad uno spazio equipaggiato

(Lezione 11 del
18/10/2018)

Vediamo più in generale il significato degli autovalori e autovettori generalizzati in uno spazio di Hilbert equipaggiato generico Φ'_A . Se A è un'osservabile, $F_\lambda \in \Phi'_A$ e $\lambda \in \sigma(A)$ l'equazione agli autovalori $AF_\lambda = \lambda F_\lambda$ si deve intendere (in analogia con il caso di \mathcal{S}'):

$$AF_\lambda(\varphi) \equiv F_\lambda(A\varphi) = \lambda F_\lambda(\varphi) \quad \forall \varphi \in \Phi_A$$

Si noti che φ deve appartenere allo spazio con la topologia più forte Φ_A . Quindi ad esempio nel caso dell'operatore posizione X si ha $\Phi_X = \mathcal{S}(\mathbb{R})$, $\Phi'_X = \mathcal{S}'(\mathbb{R})$ e $F_\lambda = \delta(\lambda - x)$, dunque la scrittura

$$x\delta(\lambda - x) = \lambda\delta(\lambda - x)$$

va in realtà interpretata come

$$\int dx x\delta(\lambda - x)\varphi(x) = \int dx \lambda\delta(\lambda - x)\varphi(x) \quad \forall \varphi \in \mathcal{S}(\mathbb{R})$$

Il teorema seguente è fondamentale per la costruzione del formalismo di Dirac.

Teorema 3.4.1. *Dato A autoaggiunto esiste sempre almeno uno spazio Φ_A che soddisfa alle condizioni volute.*

Inoltre vale il seguente risultato.

Teorema 3.4.2. *Data un'osservabile A , è possibile prendere $\Phi_A = \mathcal{S}(\mathbb{R}^n)$ se A è un polinomio di \vec{x} e $\vec{p} + V(\vec{x})$, con $V \in \mathcal{C}^\infty(\mathbb{R}^n)$ e derivate limitate. Più in generale, se V ha discontinuità in un insieme \mathcal{N} di misura nulla allora si può prendere*

$$\Phi_A = \{ \varphi \in \mathcal{S}(\mathbb{R}^n) \mid \varphi(x) = 0 \ x \in \mathcal{N} \}$$

Con il formalismo di Dirac si adottano le seguenti notazioni.

1. Si denotano con $\langle \lambda |$ gli autofunzionali in Φ'_A , cioè se \mathcal{H} è lo spazio astratto, allora l'equazione agli autovalori si scrive

$$\langle \lambda | A = \lambda \langle \lambda | \quad \lambda \in \sigma(A)$$

In pratica invece che scrivere l'equazione agli autovalori con i ket $|\lambda\rangle$ la si scrive con i bra $\langle \lambda |$.

2. Si denota con $|\varphi\rangle$ un ket nello spazio Φ_A e con

$$\langle \lambda | \varphi \rangle \equiv F_\lambda(\varphi)$$

con F_λ opportunamente normalizzato. In pratica $\langle \lambda | \varphi \rangle = F_\lambda(\varphi)$ è l'applicazione del funzionale $\langle \lambda | \in \Phi'_A$ al ket $|\varphi\rangle \in \Phi_A$. Nel caso in cui $|\lambda\rangle \in \mathcal{H}$ e $\lambda \in \sigma_P(A)$ allora $\langle \lambda | \varphi \rangle$ è semplicemente il prodotto scalare per il quale vale

$$\langle \varphi | \lambda \rangle = \langle \lambda | \varphi \rangle^* \quad (3.32)$$

3. Si noti che la scrittura $\langle \varphi | \lambda \rangle$ in generale non ha nessun significato, poiché non è stato definito un prodotto scalare. Pertanto la si estende mediante la (3.32) come una definizione, anche nel caso in cui $\langle \lambda | \in \Phi'_A$ e $\langle \lambda | \notin \mathcal{H}$.
4. Applicare il tutto alle osservabili come l'operatore posizione. Si estende dunque la notazione $\langle x | \psi \rangle = \psi(x) \in L^2(\mathbb{R}, dx)$ per denotare $|\psi\rangle \in \mathcal{H}$ in rappresentazione x . Si denota quindi con $\langle \lambda |$ l'autofunzionale in rappresentazione x , cioè $F_\lambda(x) \equiv \langle \lambda | x \rangle$ e nuovamente si definisce

$$\langle x | \lambda \rangle \equiv \langle \lambda | x \rangle^*$$

Una volta impostata questa notazione si dimostra il seguente fatto

Teorema 3.4.3. *Per ogni operatore A autoaggiunto esistono gli insiemi di autofunzionali rispettivamente per lo spettro discreto e continuo:*

$$\begin{aligned} & \{ \langle \lambda_n, r | \in \mathcal{H}' \subset \Phi'_A, \lambda_n \in \sigma_p(A), r = 1, \dots, d(\lambda_n) \} \\ & \{ \langle \lambda, r | \in \Phi'_A, \langle \lambda, r | \notin \mathcal{H}, \lambda \in \sigma_c(A), r = 1, \dots, d(\lambda) \} \end{aligned}$$

tali che per ogni $|\psi\rangle, |\varphi\rangle \in \Phi_A$ si ha

$$|\varphi\rangle = \sum_{\lambda_n \in \sigma_p(A)} \sum_{r=1}^{d(\lambda_n)} |\lambda_n, r\rangle \langle \lambda_n, r | \varphi\rangle + \int_{\sigma_c(A)} d\lambda \sum_{r=1}^{d(\lambda)} |\lambda, r\rangle \langle \lambda, r | \varphi\rangle$$

$$\langle \psi | \varphi\rangle = \sum_{\lambda_n \in \sigma_p(A)} \sum_{r=1}^{d(\lambda_n)} \langle \psi | \lambda_n, r\rangle \langle \lambda_n, r | \varphi\rangle + \int_{\sigma_c(A)} d\lambda \sum_{r=1}^{d(\lambda)} \langle \psi | \lambda, r\rangle \langle \lambda, r | \varphi\rangle$$

dove per definizione $\langle \psi | \lambda, r\rangle \equiv \langle \lambda, r | \psi\rangle^*$.

Si noti che siccome abbiamo imposto che Φ_A è denso in \mathcal{H} queste formule per densità si estendono facilmente a ogni $|\psi\rangle, |\varphi\rangle \in \mathcal{H}$. Da queste formule è poi possibile calcolare

$$\langle \psi | A | \varphi\rangle = \sum_{\lambda_n \in \sigma_p(A)} \sum_{r=1}^{d(\lambda_n)} \lambda_n \langle \psi | \lambda_n, r\rangle \langle \lambda_n, r | \varphi\rangle + \int_{\sigma_c(A)} d\lambda \lambda \sum_{r=1}^{d(\lambda)} \langle \psi | \lambda, r\rangle \langle \lambda, r | \varphi\rangle$$

e per $|\psi\rangle = |\varphi\rangle$ si ottiene il valor medio di A nello stato φ :

$$\langle A \rangle_\varphi = \int \lambda d(\varphi, P^A(\lambda)\varphi)$$

e da questa si ottengono le formule per $P^A(\lambda)$ per lo spettro discreto e continuo:

$$dP^A(\lambda) \Big|_{\sigma_p(A)} = \sum_{\lambda_n \in \sigma_p(A)} \delta(\lambda - \lambda_n) \sum_{r=1}^{d(\lambda_n)} |\lambda_n, r\rangle \langle \lambda_n, r| d\lambda$$

$$dP^A(\lambda) \Big|_{\sigma_c(A)} = \sum_{r=1}^{d(\lambda)} |\lambda, r\rangle \langle \lambda, r| d\lambda$$

Infine la relazione di completezza per un generico Φ'_A è perfettamente identica a quella già vista:

$$\sum_{\lambda_n \in \sigma_p(A)} \sum_{r=1}^{d(\lambda_n)} |\lambda_n, r\rangle \langle \lambda_n, r| + \int_{\sigma_c(A)} d\lambda \sum_{r=1}^{d(\lambda)} |\lambda, r\rangle \langle \lambda, r| = \mathbb{I}_{\mathcal{H}}$$

3.4.6 Vantaggi e svantaggi del formalismo di Dirac

Vantaggi

1. Estrema semplicità ed eleganza nella notazione;
2. potenza di calcolo, in particolare la relazione di completezza;
3. l'equazione agli autovalori funziona in modo naturale anche con lo spettro continuo $\sigma_c(A)$;
4. fornisce un modo semplice per vedere se λ appartiene allo spettro discreto o a quello continuo: $\lambda \in \sigma_p(A)$ se $|\psi\rangle \in D(A)$, mentre $\lambda \notin \sigma_p(A)$ se $|\psi\rangle \notin \mathcal{H}$.

5. La notazione di Dirac non ammette comunque soluzioni che divergono esponenzialmente, poiché non stanno in $\mathcal{S}'(\mathbb{R})$. Per esempio, se consideriamo l'hamiltoniana $H = \frac{p^2}{2m} + V(x)$, con $\Phi_H \subseteq \mathcal{S}(R)$ e $\mathcal{S}'(\mathbb{R}) \subseteq \Phi'_H$ allora non posso prendere le soluzioni che divergono esponenzialmente perché non stanno in Φ'_H . Un'onda piana invece non è in L^2 , ma in $\mathcal{S}'(\mathbb{R})$ sì, e quindi è una soluzione accettata.

Svantaggi

1. Dato A autoaggiunto, in generale Φ_A non è unico (non c'è un Φ_A “canonico”). La scelta del Φ_A ha una certa arbitrarietà, e questo può non piacere ad alcuni matematici. I matematici utilizzano spesso le famiglie spettrali perché queste invece sono uniche dato un operatore autoaggiunto A ;
2. la scrittura $\int d\lambda |\lambda\rangle \langle \lambda|$ è matematicamente imprecisa, e la sua formalizzazione rigorosa richiede una matematica sofisticata (per esempio la nuclearità);
3. è necessario non dimenticare che se $\langle \lambda| \notin \mathcal{H}'$ allora $\langle \lambda|$ non rappresenta alcuno stato fisico. Questo è analogo al caso di de Broglie, il quale scoprì che un'onda con momento ben definito corrisponde ad una particella che si muove ad una velocità superluminale:

$$p = \frac{h}{\lambda} \quad \nu = \frac{\mathcal{E}}{h} \quad \lambda\nu = \frac{\mathcal{E}}{p} = \frac{mc^2\gamma}{m\nu\gamma} = \frac{c^2}{\nu} > c$$

Dunque una soluzione semplice come un'onda piana non appartiene ad L^2 , pertanto $e^{\frac{i}{\hbar}p_0x} \notin L^2$. Tuttavia se invece si prende

$$\int f_{p_0}(p) e^{\frac{i}{\hbar}px} dp$$

questa può stare in L^2 purchè $f_{p_0} \in L^2(\mathbb{R}, dp)$. In questo modo la velocità effettiva della particella non è quella di fase ma quella di gruppo:

$$v = \frac{d\mathcal{E}}{dp} = \frac{d\omega}{dk}$$

4. quando si scrive la completezza, per esempio, per l'operatore posizione X , si scrive come

$$\int dx |x\rangle \langle x| = \mathbb{I}$$

Intuitivamente si potrebbe pensare a $\langle x|$ come ad una “base continua” analoga al caso di una base $\{\lambda_n\}$ ortonormale tale per cui

$$\sum_n |\lambda_n\rangle \langle \lambda_n| = \mathbb{I}$$

Tuttavia è necessario fare attenzione al fatto che se $\langle x| \in \Phi'_X$ e $\langle x| \notin \mathcal{H}$, allora tutti funzionali $\langle x|$ non appartengono allo spazio degli stati, e dunque non costituiscono una base di \mathcal{H} ! Questo è importante perché ad esempio nel formalismo di Dirac non è violata in ogni caso la separabilità di \mathcal{H} .

3.5 L'operatore momento

In questa sezione si analizza l'osservabile momento in meccanica quantistica. Tale analisi, oltre ad essere fondamentale in fisica, mostra come il cambiamento del dominio degli operatori in meccanica quantistica possa modificare completamente la fisica del sistema che essi descrivono.

3.5.1 Momento in \mathbb{R}

Consideriamo l'operatore momento $P = -i\hbar \frac{d}{dx}$ su $\mathcal{H} = L^2(\mathbb{R}, dx)$. Esaminiamone nel dettaglio il dominio, espandendo tutte le condizioni che avevamo prima sintetizzato con una semplice “regolarità”.

In particolar modo, è necessario chiedere che:

- Esista la derivata ψ' *quasi ovunque*.
- Poiché dovremo fare un'integrazione per parti, ci serve che $\psi' \in L^2(\mathbb{R}, dx)$ e su ogni compatto valga l'integrazione per parti:

$$\int_a^b \psi'(x) dx = \psi(b) - \psi(a) \quad \forall [a, b] \subset \mathbb{R}$$

Nota: È necessario che le condizioni siano così “larghe”. Per esempio, potremmo essere tentati di sintetizzarle chiedendo che $\psi \in \mathcal{C}^1$. Ciò, tuttavia, fa perdere l'autoaggiuntezza dell'operatore, e perciò fa sì che il suo spettro non sia più solamente reale!

Definiamo perciò il dominio di P come:

$$D(P) = \left\{ \psi \in L^2(\mathbb{R}, dx) \mid \exists \psi' \text{ q.o. } \int_a^b \psi'(x) dx = \psi(b) - \psi(a) \quad \forall [a, b] \subset \mathbb{R}, \right. \\ \left. \psi' \in L^2(\mathbb{R}, dx) \right\}$$

Nota: per *definire* un operatore è necessario indicarne in maniera esplicita il dominio. Come vedremo tra poco, infatti, lo “stesso operatore” su domini diversi può dar luogo a risultati assurdi, o essere associato a diverse osservabili.

Fissate queste condizioni, si ha subito che se $\psi, \psi' \in L^2(\mathbb{R}, dx)$ allora $\psi(x)$ si annulla all'infinito:

$$\lim_{x \rightarrow \pm\infty} \psi(x) = 0$$

A priori non è nemmeno ovvio che tale limite esista. Potremmo per esempio considerare una funzione che si annulla ovunque, se non in determinati intervalli (di misura non nulla), in cui “salta” a un certo valore positivo. Calibrando opportunamente la larghezza dei “salti” e distanziandoli tra loro si può far sì che l'integrale di tale ψ converga a un valore finito - e quindi che $\psi \in L^2$. Tuttavia, poiché tali “salti” sono sempre presenti, per quanto grande sia M , per $x > M$ la funzione assume valori che non si avvicinano a 0 più del valore del “salto”, e perciò il limite a ∞ non esiste.

Tuttavia, se ψ e $\psi' \in L^2(\mathbb{R})$, allora: $\psi, \psi' \in L^2([0, +\infty[, dx)$. Di conseguenza esiste finito:

$$\int_0^\infty dx (\psi^*(x) \psi'(x) + \psi^{*'}(x) \psi(x)) = \int_0^\infty dx \frac{d}{dx} (\psi^*(x), \psi(x)) = |\psi(\infty)|^2 - |\psi(0)|^2$$

e l'ultima uguaglianza è soddisfatta perché abbiamo richiesto sia possibile l'integrazione per parti.

Ma allora esiste $|\psi(\infty)|$, e quindi anche:

$$\psi(\infty) = \lim_{x \rightarrow \infty} \psi(x) = 0 \quad (3.33)$$

(Se fosse qualsiasi altro valore eccetto 0 si avrebbe che $\psi \notin L^2$).

Verifichiamo allora l'autoaggiuntezza $P = P^\dagger$. Essendo $D(P)$ denso in \mathcal{H} partiamo dall'uguaglianza degli elementi di matrice per definire P^\dagger e osservare quali condizioni è necessario imporre per il suo dominio:

$$(\phi, P\psi) = (P^\dagger \phi, \psi); \quad \forall \phi \in D(P^\dagger); \forall \psi \in D(P)$$

Calcolando il prodotto scalare:

$$\int_{-\infty}^{+\infty} dx \phi^*(x) \left[-i\hbar \frac{d}{dx} \psi(x) \right] \underset{(a)}{=} \left[-i\hbar \phi^*(x) \psi(x) \right]_{-\infty}^{+\infty} + \int \left[-i\hbar \frac{d}{dx} \phi(x) \right]^* \psi(x) dx$$

(dove in (a) si è integrato per parti).

Se il termine evidenziato in giallo si annullasse avremmo dimostrato che l'aggiunto P^\dagger ha la stessa forma di P , ossia che P è simmetrico (le espressioni corrispondenti a P e P^\dagger sono evidenziati in azzurro). Sappiamo che tale termine si annulla poiché, come visto in (3.33), $\psi(x)$ si annulla all'infinito. Tale conclusione è però valida solamente se l'integrazione per parti effettuata in (a) è sensata, ossia se anche per ϕ valgono le richieste che abbiamo fatto per ψ a tal proposito, e cioè che ϕ' esista quasi ovunque, $\phi, \phi' \in L^2(\mathbb{R}, dx)$ e per ϕ' valga l'integrazione per parti sui compatti.

Ma allora le condizioni che abbiamo imposto per trovare $D(P)$ sono esattamente le stesse che caratterizzano $D(P^\dagger)$, e quindi:

$$D(P) = D(P^\dagger)$$

3.5.2 Momento in \mathbb{R}_+

Paradossalmente, se cerchiamo di restringere la definizione di P ai soli reali positivi, ossia ponendo $\mathcal{H} = L^2(\mathbb{R}_+, dx)$, la costruzione di prima non porta ad alcun operatore autoaggiunto. Verifichiamolo.

Sia $P = -i\hbar \frac{d}{dx}$, e fissiamo:

$$D(P) = \{ \psi \in L^2(\mathbb{R}_+) \mid \exists \psi' \text{ q.o. e sia definita l'integrazione per parti, } \psi' \in L^2(\mathbb{R}_+) \}$$

Ripetendo gli stessi passaggi di prima, si arriva a dover definire le condizioni per cui il primo termine dell'integrazione per parti (quello evidenziato in giallo nel caso precedente) si annulli:

$$\phi^*(x)\psi(x)\Big|_0^{+\infty} = \cancel{\phi^*(+\infty)\psi(+\infty)} - \phi^*(0)\psi(0) = 0$$

Il problema è che dalle condizioni che abbiamo imposto finora sappiamo solo che ψ si annulla all'infinito, ma nulla si sa sul suo comportamento in 0. È quindi necessario imporre un'altra condizione - ma non c'è modo di farlo in maniera "simmetrica". Se infatti risolvessimo ponendo $\psi(0) = 0$ non dovremmo fare la stessa richiesta per la ϕ (non sarebbe necessario), e quindi $D(P) \subset D(P^\dagger)$. Viceversa, se imponessimo $\phi(0) = 0$ avremmo la situazione opposta, con $D(P) \supset D(P^\dagger)$.

Perciò, se richiediamo che il momento P nell'intervallo \mathbb{R}_+ sia simmetrico, allora non può essere autoaggiunto, e quindi *non è un'osservabile*.

Qual è il significato fisico di ciò?

ipotizziamo per assurdo che P su \mathbb{R}_+ sia un'osservabile. Allora avremo degli autovalori, per esempio p autovalore di P . Facendo quindi una misura del momento troveremmo allora p , e sapremmo per certo che $p > 0$.

Con un'analogia semiclassica, "fissare un momento" significa considerare un'onda piana³¹. Tuttavia, un'onda piana "confinata" a \mathbb{R}_+ vuol dire che "si è riflessa" sul piano per $x = 0$, ed è quindi in realtà la sovrapposizione di un'onda incidente e una riflessa - che hanno momenti di segno opposto³². Perciò anche in questo caso non è possibile ottenere una misura univoca del momento - e infatti l'operatore P su \mathbb{R}_+ non corrisponde a osservabili. In effetti, un'eventuale particella è costretta a rimbalzare a $x = 0$, e quindi il suo momento non può avere un *unico* valore.

(Lezione 12 del
19/10/2018)

3.5.3 Momento in $[0, 2\pi]$

È invece possibile definire l'operatore momento su $[0, 2\pi]$, cioè ponendo $\mathcal{H} = L^2([0, 2\pi], dx)$.

(Lezione 12 del
19/10/2018)

Nella costruzione del suo dominio inizialmente consideriamo un operatore "test" \tilde{P} , che poi raffineremo al P vero e proprio. Poniamo quindi:

$$D(\tilde{P}) = \{\psi \in L^2([0, 2\pi], dx) \mid \psi \text{ regolari}, \psi(0) = \psi(2\pi) = 0\}$$

Ma \tilde{P} così definito non può essere autoaggiunto, poiché $D(\tilde{P}) \subset D(\tilde{P}^\dagger)$. Infatti, partendo come prima dalla condizione di simmetria:

$$(\phi, \tilde{P}\psi) = (\tilde{P}^\dagger\phi, \psi); \quad \phi \in D(\tilde{P}^\dagger), \psi \in D(\tilde{P})$$

quando si giunge all'integrazione per parti, il primo termine il primo termine (quello evidenziato in giallo all'inizio) si annulla solamente per la condizione che

³¹Λ Pensandola con il principio di indeterminazione: l'onda piana ha estensione infinita - quindi non conosciamo la sua posizione con nessuna precisione - ma ha un *unico* momento p

³²Λ In altre parole, facendo oscillare una fune "infinitamente lunga" ma collegata ad un punto fisso - oltre il quale non c'è più nulla, e che corrisponde all'*estremo del dominio*, ossia $x = 0$, si crea un'onda stazionaria, che classicamente si ottiene come sovrapposizione di due onde con verso opposto di propagazione

abbiamo imposto sulle ψ :

$$\phi^*(x) \psi(x) \Big|_0^{2\pi} = 0$$

Ma allora sulle ϕ non serve imporre nulla oltre alla regolarità:

$$D(\tilde{P}^\dagger) = \{\phi \in L^2([0, 2\pi], dx) \mid \phi \text{ regolari}\}$$

da cui $D(\tilde{P}) \subset D(\tilde{P}^\dagger)$.

In effetti \tilde{P} non produce risultati fisici. Risolvendo infatti l'equazione agli autovalori (per separazione delle variabili):

$$\begin{aligned} \tilde{P}|\lambda\rangle &= \lambda|\lambda\rangle \Rightarrow -i\hbar \frac{d}{dx} \psi_\lambda = \lambda \psi_\lambda \Rightarrow \frac{\psi'_\lambda}{\psi_\lambda} = \frac{i\lambda}{\hbar} \\ \Rightarrow \ln \psi_\lambda &= \frac{i\lambda}{\hbar} x + c \Rightarrow \psi_\lambda(x) = A e^{i\frac{\lambda}{\hbar} x} \end{aligned}$$

Quando si impone (come richiesto dalle condizioni sul dominio di \tilde{P}) che $\psi_\lambda(0) = 0$ si ottiene solo una funzione nulla - che non ha senso fisico.

La ragione intuitiva per cui ciò non funziona è che, come nel caso di \mathbb{R}_+ , una situazione semiclassica in cui il momento è perfettamente definito è quella di un'onda piana, ma un'onda "confinata tra due punti" è un'onda stazionaria, che è sovrapposizione di due onde con versi opposti, e perciò anche in questo caso resta indefinita la direzione di p , e si giunge ad una contraddizione.

Tuttavia, in questo caso, possiamo "aggiustare" la situazione. Per distinguere i conti dal caso precedente, chiameremo P_0 questa "versione" del momento.

Perché sia simmetrico, come abbiamo visto, vogliamo che si annulli il primo termine dell'integrazione per parti:

$$\phi^*(2\pi) \psi(2\pi) - \phi^*(0) \psi(0) = 0$$

Perché sia così basta allora imporre:

$$D(P_0) = \{\psi \in L^2([0, 2\pi], dx) \mid \psi \text{ regolare, } \psi(0) = \psi(2\pi)\}$$

Ne segue che la condizione da imporre sulle ϕ è:

$$(\phi^*(2\pi) - \phi^*(0)) \psi(0) = 0 \Rightarrow \phi(2\pi) = \phi(0)$$

ossia la stessa che abbiamo appena imposto alle ψ . Perciò $D(P_0) = D(P_0^\dagger)$ e si ha l'autoaggiuntezza desiderata.

Matematicamente, la condizione che abbiamo appena imposto è una condizione di **periodicità**. In effetti, l'aver scelto 2π suggerisce di mappare il sistema su una circonferenza, dove allora ha senso che la particella abbia un momento ben definito (basti pensare, semiclassicamente, ad un'onda che "gira" su una circonferenza - il suo verso è definito, e non vi sono riflessioni).

Risolvendo allora l'equazione agli autovalori si giunge sempre a:

$$\psi(x) = A e^{i\frac{\lambda}{\hbar} x}$$

e imponendo le condizioni richieste nel dominio, e cioè che:

$$\psi(0) = A = \psi(2\pi) = e^{i\frac{\lambda}{\hbar}2\pi}$$

Si ottiene che $\lambda = n\hbar$, $n \in \mathbb{Z}$, e quindi lo spettro è solo puntuale: $\sigma(P_0) = n\hbar = \sigma_P(P_0)$.

$$\int_0^{2\pi} \left| e^{i\frac{\lambda}{\hbar}x} \right|^2 dx = 2\pi < \infty$$

Troviamo quindi un'importante differenza tra fisica classica e quantistica. Se consideriamo una particella che gira su un cerchio, in MC possiamo avere un qualsiasi valore del momento, ma in MQ solo certi valori “quantizzati”³³: $p = \hbar n$, $n \in \mathbb{Z}$.

In realtà la scelta che abbiamo fatto sulla restrizione del dominio di ψ per avere l'autoaggiuntezza non è l'unica che si può fare.

Se vogliamo che:

$$\phi^*(x)\psi(x)\Big|_0^{2\pi} = 0 \quad (3.34)$$

Potremmo, matematicamente, imporre anche solo che $\psi(2\pi) = e^{i\gamma}\psi(0)$ (ossia che la ψ sia una funzione “periodica” che torna nello stesso punto con uno sfasamento fisso dato da γ). Definiamo un operatore P_γ che incorpora questa condizione nel suo dominio.

Se $\gamma = 0$ ritroviamo il caso di prima, ma generalmente $\gamma \in [0, 2\pi[$. Qual è allora la condizione da imporre sulle ϕ ?

Imponendo (3.34) con la nuova condizione sulle ψ , otteniamo:

$$0 = \phi^*(2\pi)\psi(2\pi) - \phi^*(0)\psi(0) = (\phi^*(2\pi)e^{i\gamma} - \phi^*(0))\psi(0) \Rightarrow \phi(2\pi) = e^{i\gamma}\phi(0)$$

che è la stessa condizione imposta sulle ψ ! Abbiamo quindi l'autoaggiuntezza:

$$D(P_\gamma) = D(P_\gamma^\dagger)$$

Ma che osservabile corrisponde a questo operatore? Partiamo dalla solita soluzione $A \exp\left(\frac{i\lambda}{\hbar}x\right)$ dell'equazione agli autovalori e imponiamo la condizione fatta nel dominio:

$$e^{i\gamma}\psi_\lambda(0) = e^{i\gamma}Ae^{i\frac{\lambda}{\hbar}\cdot 0} = e^{i\gamma}A \stackrel{!}{=} \psi_\lambda(2\pi) = Ae^{i\frac{\lambda}{\hbar}2\pi}$$

Da cui:

$$e^{i\gamma} = e^{i\frac{\lambda}{\hbar}2\pi} \Rightarrow \exp\left(i\left(\frac{\lambda}{\hbar}2\pi - \gamma\right)\right) = 1 \quad (3.35)$$

$$\frac{\lambda}{\hbar}2\pi - \gamma = 2\pi n \Rightarrow \lambda = \hbar\left(\frac{\gamma}{2\pi} + n\right), \quad n \in \mathbb{Z} \quad (3.36)$$

E perciò il dominio di P_γ è di nuovo esclusivamente puntuale:

$$\sigma(P_\gamma) = \sigma_P(P_\gamma) = \left\{ \hbar\left(\frac{\gamma}{2\pi} + n\right), \quad n \in \mathbb{Z} \right\}$$

³³^Potremmo riconoscere in ciò la condizione trovata da de Broglie: perché ogni particella ha un comportamento ondulatorio, solo i momenti che fanno sì che “l'onda interferisca costruttivamente con se stessa” sono ammessi

Sostituendo l'espressione per λ nella soluzione:

$$\psi_\lambda(x) = A \exp\left(\frac{\lambda}{\hbar}x\right) = A \exp\left[ix\left(\frac{\gamma}{2\pi} + n\right)\right] \equiv \psi_n(x); \quad n \in \mathbb{Z}$$

Se condensiamo la parte che dipende da n in un'opportuna funzione $\tilde{\psi}_n(x)$ possiamo riscrivere:

$$\psi_n(x) = \tilde{\psi}_n(x) \exp\left(i\frac{\gamma}{2\pi}x\right); \quad \tilde{\psi}_n(0) = \tilde{\psi}_n(2\pi); \quad \lambda = \hbar\left(\frac{\gamma}{2\pi} + n\right); \quad n \in \mathbb{Z}$$

E perciò, con questa notazione, l'equazione agli autovalori diviene:

$$P_\gamma \psi_n = P_\gamma (\tilde{\psi}_n(x) e^{i\frac{\gamma}{2\pi}x}) = \hbar\left(\frac{\gamma}{2\pi} + n\right) (\tilde{\psi}_n(x) e^{i\frac{\gamma}{2\pi}x})$$

Ma allora portando l'esponenziale evidenziata in azzurro a sinistra possiamo evidenziare una nuova equazione agli autovalori:

$$(e^{-i\frac{\gamma}{2\pi}x} P_\gamma e^{i\frac{\gamma}{2\pi}x}) \tilde{\psi}_n(x) = \hbar\left(\frac{\gamma}{2\pi} + n\right) \tilde{\psi}_n(x)$$

E sostituendo la definizione di P_γ :

$$(e^{-i\frac{\gamma}{2\pi}x} P_\gamma e^{i\frac{\gamma}{2\pi}x}) \tilde{\psi}_n(x) = e^{-i\frac{\gamma}{2\pi}x} \left(-i\hbar \frac{d}{dx} (e^{i\frac{\gamma}{2\pi}x} \tilde{\psi}_n(x)) \right) = \frac{\gamma\hbar}{2\pi} \tilde{\psi}_n(x) - i\hbar \frac{d}{dx} \tilde{\psi}_n(x)$$

(Nota: la derivata in un prodotto di operatori agisce su tutto quello che sta alla sua destra).

Quindi, se consideriamo solo gli operatori (rimuovendo le $\tilde{\psi}_n(x)$) giungiamo all'uguaglianza:

$$e^{-i\frac{\gamma}{2\pi}x} \left(-i\hbar \frac{d}{dx} \right) e^{i\frac{\gamma}{2\pi}x} = \frac{\gamma\hbar}{2\pi} \underbrace{-i\hbar \frac{d}{dx}}_{P_0} = \frac{\gamma\hbar}{2\pi} + P_0$$

Perciò le $\psi_n(x)$ sono gli autovettori corrispondenti ad un operatore “momento + costante”. A che osservabile si riferisce?

Per capire il significato fisico di P_γ ci riconduciamo ad un caso classico. Consideriamo un solenoide strettamente interno alla circonferenza in cui si muove la particella perpendicolare al piano della circonferenza con flusso magnetico ϕ .

Poiché il campo magnetico è nullo su S^1 classicamente la particella non risente del flusso ϕ del campo magnetico, ma quantisticamente invece “sente” ϕ topologicamente. Infatti possiamo scrivere ϕ in termini di potenziale vettore \vec{A} , che non è nullo su S^1 :

$$\phi = \int_{\Sigma} \vec{B} \cdot d\Sigma = \oint_{S^1} \vec{A} \cdot d\vec{l} = 2\pi A$$

(dove si è applicato il teorema di Stokes). Il modulo del potenziale vettore è allora costante e dipende dal flusso:

$$A = \frac{\phi}{2\pi}$$

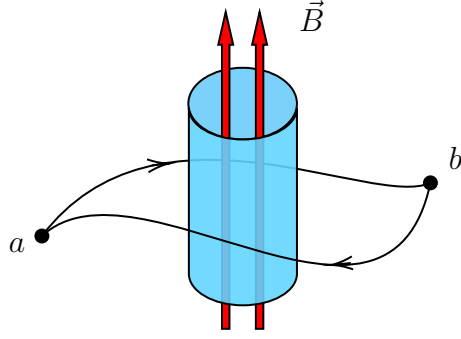


Figura 3.2: Illustrazione del setup dell'effetto Aharonov-Bohm

Classicamente la presenza di un campo magnetico nel formalismo hamiltoniano classico conduce alla sostituzione del momento \vec{p} con un termine³⁴:

$$\vec{p} \rightarrow \vec{p} + \frac{e}{c} \vec{A}$$

Nel nostro caso 1D:

$$p \rightarrow p + \frac{e}{c} A = p + \frac{e}{c} \frac{\phi}{2\pi}$$

Confrontando questo caso classico con quello quantistico, se avessimo

$$\frac{e}{c} \frac{\phi}{2\pi} = \frac{\gamma \hbar}{2\pi} \Rightarrow \gamma = \frac{e\phi}{\hbar c}$$

allora avremmo trovato il caso quantistico corrispondente a questo caso classico. In effetti ciò fu verificato sperimentalmente, anni dopo la prima descrizione matematica di questo operatore.

Pertanto P_γ descrive il momento di una particella quantistica in S^1 in presenza di un solenoide interno a S^1 con flusso $\phi(\gamma = \frac{e\phi}{\hbar c})$ che altera lo spettro $\sigma(P_\gamma) = \sigma(P_0) + \hbar\gamma$.

Tale fenomeno prende il nome di effetto Aharonov-Bohm.

Come si intuisce dalla trattazione matematica, si dimostra sperimentalmente che si tratta di un fenomeno puramente topologico: non conta la distanza tra circonferenza e solenoide, ma basta che la circonferenza *sia attorno* alla regione in cui è presente \vec{B} , ossia che ogni superficie delimitata dal circuito su cui si trova la particella intersechi una regione di campo magnetico non nullo.

Fase di Berry

Lo sfasamento che acquisisce una ψ dopo un *ciclo* è detto **fase di Berry**, e per ψ normalizzate ($\langle \psi | \psi \rangle = 1$) ha la seguente definizione:

$$e^{i\oint_C \langle \psi | d\psi \rangle}$$

³⁴ L'Hamiltoniana di una particella di carica q in un campo elettromagnetico con potenziale vettore \vec{A} e potenziale scalare ϕ è infatti data da: $H = \frac{1}{2m} \left(\vec{p} - \frac{q}{c} \vec{A} \right)^2 + q\phi$

dove \mathcal{C} è una curva chiusa nello spazio degli stati (o delle funzioni).

Con questa definizione la fase di Berry è un'osservabile. Mostriamo che il suo valore è non nullo se calcolata lungo un percorso ciclico (per esempio una circonferenza). Partiamo considerando le autofunzioni di P_γ :

$$\psi_n(x) = \frac{1}{\sqrt{2\pi}} e^{i(n+\frac{\gamma}{2\pi})x}$$

Senza modificare l'autovalore è possibile moltiplicarle per una fase “attorno al cerchio”. Giungiamo quindi alla famiglia:

$$\psi_{n,\alpha} = \frac{1}{\sqrt{2\pi}} e^{i(n+\frac{\gamma}{2\pi})(x+\alpha)}; \quad \psi_{n,2\pi}(x) = e^{i\alpha} \psi_{n,0}(x); \quad \alpha \in [0, 2\pi[$$

Possiamo ora definire \mathcal{C} :

$$C = \{\psi_{n,\alpha}(x) \mid \alpha \in [0, 2\pi]\}$$

e quindi calcolare l'integrale per la fase di Berry:

$$\begin{aligned} \oint_{\mathcal{C}} \langle \psi_{n,\alpha} | d\psi_{n,\alpha} \rangle &= \int_0^{2\pi} d\alpha \langle \psi_{n,\alpha} | \frac{d}{d\alpha} \psi_{n,\alpha} \rangle \\ \langle \psi_{n,\alpha} | \frac{d}{d\alpha} \psi_{n,\alpha} \rangle &= \int_0^{2\pi} dx \frac{1}{2\pi} e^{-i(\frac{\gamma}{2\pi}+n)(x+\alpha)} \left(i \left(\frac{\gamma}{2\pi} + n \right) \right) e^{i(\frac{\gamma}{2\pi}+n)(x+\alpha)} = i \left(\frac{\gamma}{2\pi} + n \right) \\ e^{\oint_{\mathcal{C}} \langle \psi | d\psi \rangle} &= \exp \int_0^{2\pi} d\alpha i \left(\frac{\gamma}{2\pi} + n \right) = \exp i 2\pi \left(\frac{\gamma}{2\pi} + n \right) = e^{i\gamma} \end{aligned}$$

3.6 Stati misti in MQ

Supponiamo di non conoscere in quale stato puro $|\phi\rangle$ il sistema si trovi, ma solo di sapere che può essere negli stati puri $|\phi_1\rangle, \dots, |\phi_n\rangle$ (che consideriamo normalizzati, $\langle \phi_i | \phi_i \rangle = 1$) con probabilità $p_1 \dots p_n$, $\sum_{i=1}^n p_i = 1$.

Nota: n è per forza numerabile, ma le ϕ non hanno vincoli di ortogonalità, e le p_i sono probabilità classiche.

Dalle regole della probabilità, se Σ denota lo stato di informazione non massimale (misto) sopra descritto e A è un'osservabile, allora il valor medio di A nello stato Σ è dato da una “media pesata” dei valor medi dei singoli stati puri:

$$\langle A \rangle_\Sigma = \sum_{i=1}^n p_i \langle \phi_i | A | \phi_i \rangle$$

Osserviamo che se ϕ è un qualsiasi stato di \mathcal{H} , possiamo scrivere il valor medio come:

$$\langle \phi | A | \phi \rangle = \text{Tr}(|\phi\rangle \langle \phi| A)$$

Infatti, sia $\{|\chi_j\rangle, j = 1, \dots, \dim \mathcal{H}\}$ una base ON per \mathcal{H} , con la condizione che $|\chi_1\rangle = |\phi\rangle$ (allineo il primo vettore della base con il vettore in esame). Allora:

$$\text{Tr}(|\phi\rangle \langle \phi| A) = \sum_{j=1}^{\dim \mathcal{H}} \langle \chi_j | (|\phi\rangle \langle \phi| A) | \chi_j \rangle \stackrel{(a)}{=} \sum_{j=1}^{\dim \mathcal{H}} \delta_{j1} \langle \phi | A | \chi_j \rangle = \langle \phi | A | \phi \rangle$$

dove in (a) si è usato il fatto che $\langle \chi_j | \phi \rangle = \langle \chi_j | \chi_1 \rangle = \delta_{j1}$ per l'ortonormalità (essendo $|\phi\rangle$ perpendicolare a tutti i $|\chi_j\rangle$ e parallelo al primo per costruzione).

Valor medio di uno stato misto

Valor medio in termini di traccia

Nota: in uno spazio di Hilbert \mathcal{H} , con $|\chi_j\rangle$ base ON, la **traccia**, per gli operatori per cui è finita, è definita come:

$$\text{Tr}(\cdot) = \sum_j \langle \chi_j | (\cdot) | \chi_j \rangle$$

Ed è indipendente dalla scelta della base.

In effetti, nel caso finito dimensionale di \mathbb{R}^n , se si parte da una base canonica si ritrova la somma degli elementi sulla diagonale principale.

Dalla definizione segue anche che la traccia è **lineare**.

Sostituendo questo risultato nella formula della media:

$$\langle A \rangle_\Sigma = \sum_{i=1}^n p_i \langle \phi_i | A | \phi_i \rangle = \sum_{i=1}^n p_i \text{Tr}(|\phi_i\rangle \langle \phi_i| A) \stackrel{(a)}{=} \text{Tr} \left(\sum_{i=1}^n p_i |\phi_i\rangle \langle \phi_i| A \right) \quad (3.37)$$

dove in (a) si è usata la linearità della traccia.

Il termine evidenziato è detto **matrice densità**:

Matrice di densità

$$\rho \equiv \sum_{i=1}^n p_i |\phi_i\rangle \langle \phi_i|$$

Dalla sua definizione segue che:

Proprietà della matrice densità ρ è simmetrico

1. ρ è simmetrico, ossia $\rho = \rho^\dagger$, in quanto A è autoaggiunto, p_i è reale e $(|\phi\rangle \langle \phi|)^\dagger = |\phi\rangle \langle \phi|$
2. ρ è **positivo** ($\rho \geq 0$), cioè produce solo “valor medi positivi”: $\forall \psi \langle \psi | \rho | \psi \rangle \geq 0$. Infatti:

ρ è positivo

$$\begin{aligned} \langle \psi | \rho | \psi \rangle &= \sum_i p_i \langle \psi | \phi_i \rangle \langle \phi_i | \psi \rangle = \sum_i p_i \langle \psi | \phi_i \rangle (\langle \psi | \phi_i \rangle)^* = \\ &= \sum_i p_i |\langle \psi | \phi_i \rangle|^2 \geq 0; \quad \forall |\psi\rangle \in \mathcal{H} \end{aligned}$$

(Dalla positività si ha poi subito che è autoaggiunto, come nel caso di un numero complesso, che è positivo solo se reale)

3. $\text{Tr} \rho = 1$, infatti $\text{Tr}(|\phi\rangle \langle \phi|) = \sum_j \langle \chi_j | \phi \rangle \langle \phi | \chi_j \rangle = |\langle \phi | \phi \rangle|^2 = 1$ (essendo $|\chi_1\rangle = |\phi\rangle$). Sfruttando allora la linearità della traccia, e il fatto che le p_i , esaurendo tutte le possibilità, si sommano a 1 (cosa che avevamo chiesto fin dal principio):

$\text{Tr} \rho = 1$

$$\text{Tr} \left(\sum_i p_i |\phi_i\rangle \langle \phi_i| \right) = \sum_i p_i \text{Tr}(|\phi_i\rangle \langle \phi_i|) = \sum_i p_i = 1$$

Definizione 16. Uno **stato** (misto o puro) in MQ è descritto da un operatore ρ che soddisfa 1-3.

Descrizione matematica di uno stato

Gli **stati puri** sono descritti dalle ρ che soddisfano $\rho = \rho^2$. In tal caso ρ è un proiettore, ma poiché la traccia di ρ deve fare 1, deve essere unidimensionale (cioè c'è un solo $p_i = 1$ e tutti gli altri sono 0, poiché solo i numeri 0 e 1 sono uguali al loro quadrato³⁵).

Ma allora in tal caso la somma è ridotta a un solo elemento (tutti gli altri sono annullati dalle $p_i = 0$), e quindi $\rho = |\psi\rangle\langle\psi|$ per un qualche $|\psi\rangle \in \mathcal{H}$ normalizzato ($\langle\psi|\psi\rangle = 1$).

Ci si potrebbe allora chiedere se ρ , visto che è definito in termini di $|\psi\rangle \in \mathcal{H}$ - che sono definiti a meno di una fase - presenti la stessa ambiguità. Ma se ρ è la descrizione di uno stato, che è dato da un raggio vettore in \mathcal{S} , tale ambiguità deve sparire. E infatti, si verifica che:

$\rho \leftrightarrow$ raggi vettori

$$\begin{aligned} |\psi\rangle &\rightarrow e^{i\alpha} |\psi\rangle \\ \rho &\rightarrow e^{i\alpha} |\psi\rangle\langle\psi| e^{-i\alpha} = |\psi\rangle\langle\psi| \end{aligned}$$

Ma allora $\rho^2 = \rho$ sono in corrispondenza biunivoca con i raggi vettori dello spazio di Hilbert.

Ciò è comodo sperimentalmente, perché ρ costituisce un'osservabile (e quindi potenzialmente può essere misurata).

Per esempio, in un sistema di due fotoni, lo stato delle polarizzazioni è una matrice 2×2 di numeri che si possono ricavare sperimentalmente. Perciò, per verificare se un sistema creato in laboratorio sia in uno stato puro o meno, basta determinare tale matrice A , calcolarne il quadrato e vedere se viene lo stesso risultato di partenza, ossia se si ha $A^2 = A$.

Come verificare sperimentalmente uno stato puro

Vediamo che, come nel caso classico, uno stato misto quantistico può essere scritto come una combinazione *convessa*³⁶ di stati puri, scritti come proiettori.

Attenzione: partendo dalla decomposizione di $|\psi\rangle = |\psi_1\rangle + |\psi_2\rangle$ si potrebbe essere tentati di scrivere la seguente relazione per la sovrapposizione di stati:

$$|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| \neq |\psi\rangle\langle\psi|$$

Ma tale uguaglianza non è corretta. In effetti il primo membro è (a meno di una normalizzazione) la matrice densità di uno *stato misto*, mentre il secondo è ovviamente uno *stato puro* (essendo costituito da *un solo* proiettore).

Matematicamente la faccenda risulta evidente prendendo $|\psi_1\rangle$ e $|\psi_2\rangle$ ON. Allora le espressioni del membro a sinistra sono per forza matrici diagonali, ma quella a destra generalmente non lo è.

³⁵^In particolare, se ci fosse un $p_i \neq 0, 1$ allora non si avrebbe $\rho = \rho^2$, e se vi fossero più di un p_i pari a 1 allora $\text{Tr } \rho > 1$, e nel caso non ve ne fosse nessuno $\text{Tr } \rho = 0$, che va contro alla definizione di matrice di densità

³⁶^Una combinazione convessa è una combinazione lineare di elementi fatta con coefficienti non negativi a somma 1. Per esempio: $\lambda_1 x_1 + \dots + \lambda_m x_m$ è convessa se $\lambda_i \geq 0$ per $i = 1, \dots, m$ e $\sum_{i=1}^m \lambda_i = 1$

Concretamente, siano per esempio (trascurando, di nuovo, le normalizzazioni):

$$\begin{aligned} |\psi_1\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}; |\psi_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}; |\psi\rangle = |\psi_1\rangle + |\psi_2\rangle = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ |\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ |\psi\rangle\langle\psi| &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \end{aligned}$$

Si nota quindi che nell'espressione del membro a sinistra mancano tutti i termini fuori dalla diagonale.

Nel caso **classico** possiamo scrivere uno stato come combinazione convessa (in generale non numerabile) di stati puri:

$$\rho(q, p) = \int dq_0 dp_0 \rho(q_0, p_0) \underbrace{\delta(q - q_0) \delta(p - p_0)}_{\text{Stati puri}}$$

È quindi ben definita (dalla $\rho(q, p)$) la *famiglia* degli stati puri che definiscono lo stato misto.

In MQ la situazione è diversa. Dall'autoaggiuntezza di ρ ($\rho = \rho^\dagger$) si ha che esistono autovettori $|\lambda_n\rangle$ di ρ che formano una base ON. Possiamo allora decomporre ρ spettralmente:

$$\rho = \sum_n \lambda_n |\lambda_n\rangle\langle\lambda_n|; \quad \text{Tr}(\rho) = 1 = \sum_n \lambda_n$$

Abbiamo allora ottenuto una descrizione completamente equivalente a quella che abbiamo finora usato:

$$\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i|$$

in cui, tra l'altro, le $|\phi_i\rangle$ non devono neppure essere ON.

Il problema è: quali sono gli stati che “compongono” ρ ? Sono i $|\phi_i\rangle$ o $|\lambda_i\rangle$? Le due descrizioni sono equivalenti, e quindi la risposta non è univoca.

Perciò si hanno famiglie *diverse* di stati puri che danno la stessa informazione dello stato misto!

In MC posso comunque sapere gli stati puri che contribuiscono allo stato misto – anche se su di loro ho solo informazioni probabilistiche –, ma in MQ in uno stato misto non è chiara neanche la famiglia di stati puri sottostante (è un livello decisamente più profondo di ignoranza).

(Lezione 13 del
22/10/2018)

3.7 Concetti di teoria dei gruppi

Come vedremo nei prossimi capitoli, il concetto di **simmetria** è di fondamentale importanza in MQ: la presenza di trasformazioni che “lasciano invariato” un sistema consente infatti di semplificare conti che altrimenti sarebbero infaticabili, e

anche di mettere in evidenza conseguenze teoriche di grande interesse.

In particolare, nel capitolo 3.8 useremo la simmetria data dall'omogeneità del tempo per ricavare l'evoluzione di un sistema isolato, e nei capitoli 4.6 e 4.7 esamineremo le simmetrie traslazionale e rotazionale, trovando, nel caso di quest'ultima, un fenomeno senza analogo classico.

Il *framework* più adatto per la trattazione matematica delle simmetrie è quello della **teoria dei gruppi** (in particolare dei gruppi di Lie, come accenneremo in seguito), e della **teoria delle rappresentazioni** (che permetterà di associare a simmetrie “astratte” *oggetti concreti da usare nei conti*, e di cui studiare le proprietà).

Nel seguente capitolo, perciò, introduciamo alcuni aspetti elementari di questa branca (estremamente ampia) della matematica pura.

Definizione 17. Si definisce **gruppo** $(G, *)$ un insieme di *oggetti* g per cui è definita un'operazione di gruppo, indicata con $*$, che ha le seguenti proprietà:

*Gruppo:
definizione*

- **Chiusura:** $\forall a, b \in G, a * b \in G$
- **Associatività:** $(a * b) * c = a * (b * c) \quad \forall a, b, c \in G$
- **Esistenza dell'elemento neutro:** $\exists e \in G$ t.c. $e * a = a * e = a \quad \forall a \in G$
- **Esistenza dell'inverso** $\forall a \in G \exists b \equiv a^{-1}$ t.c. $b * a = a * b = e$

Notiamo che queste sono proprio le proprietà di base che desidereremmo in una *simmetria*. Vediamo qualche esempio:

- **Traslazioni** (temporali o spaziali): consideriamo l'insieme G di tutte le traslazioni che si possono fare su un sistema, in ogni direzione. Notiamo che l'applicazione “di fila” di due traslazioni si può sempre *sintetizzare* come l'effetto di un'unica traslazione “più grande”.
Consideriamo, per esempio, un punto P sulla retta reale. Sia A la traslazione che trasla tutti i punti della retta di una certa distanza a verso destra, e analogamente B la traslazione che sposta di b nella stessa direzione. Appliciamo A e B in sequenza su P :

$$P \xrightarrow{A} P' \xrightarrow{B} P''$$

Equivalentemente, possiamo arrivare a P'' considerando una terza traslazione C , che sposti i punti della retta di $a + b$ verso destra:

$$P \xrightarrow{C} P''$$

Creiamo allora un'operazione, detta *composizione*, che a partire da A e B consente di trovare la $C = B \circ A$ che *sintetizza* la loro azione. Notiamo che, comunque scelte A e B , il risultato della loro composizione è ancora una traslazione: allora G è chiuso rispetto a \circ .

Le altre proprietà sono immediate: le traslazioni sono associative (di più, non conta l'ordine di applicazione delle traslazioni, quindi abbiamo anche la

commutatività), esiste una traslazione “che non fa nulla” - ossia quella che non sposta punti, e che funge da elemento neutro, e per ogni traslazione è possibile trovare quella “contraria” che “torna indietro” e funge da inversa. Diciamo perciò che (G, \circ) è un **gruppo** per la definizione data sopra.

- **Rotazioni:** analogamente è possibile dimostrare che anche l'insieme di tutte le possibili rotazioni, con l'operazione di gruppo data dalla composizione, è un gruppo.

Notiamo ora che le traslazioni e le rotazioni hanno però ben più proprietà rispetto a quelle conferite unicamente dalla struttura di gruppo. Sono, infatti, in un certo senso, “continue”, ossia data una trasformazione g è sempre possibile trovare una trasformazione “arbitrariamente vicina” ad essa, esattamente come succede con i punti di \mathbb{R}^n . In altre parole, un G “continuo” *assomiglia* (almeno localmente) a un aperto di \mathbb{R}^n - ossia si comporta come una *varietà differenziale*. Definiamo quindi:

Definizione 18. Un **gruppo di Lie** $(G, *)$ è un gruppo che è anche una varietà differenziale. In particolare ciò significa che l'operazione di gruppo $*$ è una applicazione *differenziabile*³⁷. Indichiamola allora come una funzione $x * y = \mu(x, y)$, per cui sono definite (e sono continue) le derivate parziali rispetto a x e y di qualsiasi ordine.

Gruppo di Lie

Un esempio di gruppo di Lie è il cosiddetto *gruppo cerchio* \mathbb{T} , che è costituito da tutti i numeri complessi di modulo unitario³⁸:

$$\mathbb{T} = \{z \in \mathbb{C} \text{ t.c. } |z| = 1\}$$

con l'operazione di gruppo data dalla *moltiplicazione complessa*.

Gli elementi di \mathbb{T} sono della forma $e^{i\theta}$, e la moltiplicazione di due complessi di norma unitaria genera un risultato che ha ancora norma unitaria, ed è associativa (di più, è anche commutativa). L'elemento neutro è $e^{i0} = 1$, e l'inverso di $e^{i\theta}$ è $e^{-i\theta}$. Inoltre, la funzione μ definita dall'operazione di gruppo:

$$\mu(x, y) = xy, \quad x, y \in \mathbb{T}$$

è chiaramente \mathcal{C}^∞ . Abbiamo allora mostrato come \mathbb{T} sia un gruppo di Lie.

Un altro gruppo di Lie molto importante è $GL(n, \mathbb{C})$, che è costituito da tutte le matrici $n \times n$ invertibili (ossia con $\det \neq 0$) di elementi complessi, considerate con l'operazione di gruppo data dalla *moltiplicazione matriciale*. Se richiediamo in più che $\det = 1$ otteniamo le matrici “speciali” indicate con $SL(n, \mathbb{C})$. Analogamente, $O(n)$ è il gruppo delle matrici ortogonali, $SO(n)$ di quelle ortogonali con $\det = 1$, e $U(n)$ di quelle unitarie (e $SU(n)$ di quelle unitarie con $\det = 1$).

³⁷^Qui useremo il termine *differenziabile* in modo vago: per la precisione, nel caso di gruppi di Lie, il grado di differenziabilità è quello delle funzioni analitiche

³⁸^Poiché le matrici 1×1 sono effettivamente numeri complessi, spesso si indica \mathbb{T} come $U(1)$, ossia il gruppo di matrici unitarie 1×1

Possiamo allora pensare di usare un gruppo di Lie per *descrivere* una simmetria: dobbiamo però specificare esattamente *in che modo*. In che senso possiamo identificare un generico $g \in (G, *)$ con una *trasformazione* di uno spazio V , che è data da una funzione $f : V \rightarrow V$ che mappa $V \ni v \rightarrow f(v) \in V$? Diamo allora la seguente definizione:

Definizione 19. Si definisce **azione** di un gruppo $(G, *)$ su un insieme X qualsiasi una funzione $\varphi : G \times X \rightarrow X$ che “applica” un elemento $g \in G$ ad un elemento $x \in X$ per ottenere un altro elemento $\varphi(g, x) \in X$.

Azione di un gruppo

Tale φ deve “essere compatibile con la struttura di gruppo”, cioè rispettare le seguenti due proprietà:

- $\varphi(e, x) = x \quad \forall x \in X$ (cioè applicare l’elemento neutro di G a x lo lascia invariato)
- $\varphi(g * h, x) = \varphi(g, \varphi(h, x))$ (ossia applicare la composizione - mediante operazione di gruppo - di g e h equivale a comporre le loro azioni)

Nella notazione è più comodo “spezzare” la $\varphi(g, x)$: definiamo cioè la trasformazione di X come $U(g) : X \rightarrow X$, con $x \mapsto U(g)(x) = \varphi(g, x)$. In altre parole, per ogni $g \in G$ è definita una trasformazione $U(g)$ che prende elementi di X e li associa ad altri elementi di X . Le proprietà di sopra divengono allora:

- $U(e) = \text{id}_X$
- $U(g * h) = U(g) \circ U(h)$

Nella pratica, questa definizione è fin troppo generica per quello che intendiamo fare. Noi ci occuperemo infatti di applicare simmetrie a spazi vettoriali (in particolare a spazi di Hilbert, \mathcal{H}) e quindi ci concentriamo su quelle trasformazioni che *preservano* la struttura di spazio vettoriale, ossia le trasformazioni lineari.

Un’azione che è anche *lineare* si dice **rappresentazione lineare** di un gruppo G . Più precisamente:

Definizione 20. Una **rappresentazione lineare** di un gruppo $(G, *)$ su uno spazio vettoriale V è una mappa $O : G \rightarrow \mathcal{L}(V)$, ove $\mathcal{L}(V)$ indica lo spazio degli operatori lineari su V , che associa ad ogni elemento di G un *operatore lineare invertibile* $O(g) \in \mathcal{L}(V)$ che soddisfa:

Rappresentazione lineare di un gruppo

- $O(e) = \text{id}_V$
- $O(g_1 * g_2) = O(g_1) \circ O(g_2)$

Si tratta cioè di quella particolare azione di G che avviene tramite operatori lineari invertibili.

Nella pratica, se V è di $\dim n$, possiamo fissare una base su di esso scegliendo n vettori linearmente indipendenti, e di conseguenza scrivere $O \in \mathcal{L}(V)$ come una matrice $n \times n$ con $\det \neq 0$ (dovendo essere invertibile).

Facciamo un esempio partendo dal gruppo \mathbb{T} introdotto in precedenza, che ricordiamo essere:

$$\mathbb{T} = \{z \in \mathbb{C} \text{ t.c. } |z| = 1\}$$

Un elemento generico di \mathbb{T} è scritto come $e^{i\theta}$. Se vogliamo *applicarlo* ad uno spazio vettoriale, per esempio \mathbb{R}^2 , dobbiamo definire una *rappresentazione lineare*. Visto che $e^{i\theta}$ fissa un angolo sulla circonferenza unitaria rispetto all'asse x , potremmo pensare alla sua azione come una *rotazione* 2D, per cui $e^{i\pi}$, per esempio, ruota i punti di \mathbb{R}^2 di 180° in senso antiorario.

Precisamente, scriviamo la mappa della rappresentazione:

$$e^{i\theta} \mapsto \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

Notiamo che e^{i0} viene mandato in \mathbb{I}_2 , e la mappa rispetta l'operazione di gruppo, cioè:

$$e^{i\theta_1} \cdot e^{i\theta_2} = e^{i(\theta_1 + \theta_2)} = \begin{pmatrix} \cos \theta_1 & -\sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{pmatrix} \begin{pmatrix} \cos \theta_2 & -\sin \theta_2 \\ \sin \theta_2 & \cos \theta_2 \end{pmatrix} = \begin{pmatrix} \cos(\theta_1 + \theta_2) & -\sin(\theta_1 + \theta_2) \\ \sin(\theta_1 + \theta_2) & \cos(\theta_1 + \theta_2) \end{pmatrix}$$

come si può verificare svolgendo i conti e utilizzando le formule di addizione/sottrazione di seno/coseno.

Riepilogando: il nostro obiettivo è descrivere matematicamente le simmetrie di un sistema, per semplificarci i conti e capirci qualcosa di più. Abbiamo così *caratterizzato matematicamente* tali simmetrie partendo dalla nozione di *gruppo*, che ci dà le proprietà fondamentali (composizione e invertibilità). Considerando che traslazioni/rotazioni possono essere “piccole a piacimento” abbiamo aggiunto al gruppo una struttura “continua”, trasformandolo in un *gruppo di Lie*. Infine, abbiamo specificato in che senso gli elementi di un gruppo di Lie “fungono da trasformazioni” tramite l'idea di *rappresentazione lineare*.

È giunto il momento di *concretizzare* tutto ciò nel campo della MQ.

Definizione 21. Definiamo una **simmetria fisica** T come una mappa che trasforma *osservabili* in altre *osservabili*, e *stati puri* in altri *stati puri*, rispettando la struttura dell'algebra (ossia una combinazione lineare di osservabili è mandata nella combinazione lineare delle osservabili *trasformate*, e stessa cosa per i prodotti (b)) e preservando i *valor medi* (a). Tali richieste corrispondono alla definizione di simmetria come una trasformazione che “non cambia i risultati degli esperimenti”. Matematicamente, chiediamo che T , che mappa osservabili $A \mapsto A'$ e stati $|\psi\rangle \mapsto |\psi'\rangle$, verifichi le seguenti proprietà:

$$(a) : \langle \psi | A | \psi \rangle = \langle \psi' | A' | \psi' \rangle; \quad (b) : \alpha A + \beta B \mapsto \alpha A' + \beta B'; \quad AB \mapsto A'B'$$

Una simmetria fisica così definita **preserva** automaticamente le **probabilità di transizione**. Infatti, se scegliamo come operatore A un proiettore $|\phi\rangle\langle\phi|$, otteniamo:

$$\langle \psi | (|\phi\rangle\langle\phi|) | \psi \rangle = \langle \psi | \phi \rangle \langle \phi | \psi \rangle \mapsto \langle \psi' | \phi' \rangle \langle \phi' | \psi' \rangle = \langle \psi' | (|\phi'\rangle\langle\phi'|) | \psi' \rangle$$

e l'ultimo termine è uguale al primo per la proprietà (a).

Indicando con U la *trasformazione di simmetria*, dalla definizione di simmetria fisica troviamo immediatamente la formula di trasformazione per un generico operatore A :

$$\begin{aligned}\langle \psi' | A' | \psi' \rangle &\stackrel{!}{=} \langle \psi | A | \psi \rangle \\ \Rightarrow \langle U\psi | A' | U\psi \rangle &= \langle \psi | U^\dagger A' U | \psi \rangle \stackrel{!}{=} \langle \psi | A | \psi \rangle \Rightarrow U^\dagger A' U = A \Rightarrow A' = U A U^\dagger\end{aligned}\quad (3.38)$$

Utilizzando i termini definiti in precedenza, T è quindi una rappresentazione del rispettivo gruppo di simmetria, e agisce nello spazio degli *stati*, che non è propriamente \mathcal{H} , ma lo spazio dei **raggi vettori** \mathcal{PH} . Infatti, dai postulati della MQ sappiamo che una $|\psi\rangle \in \mathcal{H}$ e $e^{i\gamma}|\psi\rangle$ rappresentano lo stesso stato (puro)³⁹: ciò deriva dal fatto che le *fasi* “*assolute*” γ non sono misurabili, ma lo sono solo quelle relative, e perciò “sfasare” il vettore di stato del sistema non cambia i risultati sperimentali.

Ciò è un problema: \mathcal{PH} non è neanche uno spazio vettoriale, ma uno spazio *proiettivo* su cui è difficile lavorare.

Un'idea per procedere è allora la seguente. Partiamo da una simmetria associata ad un gruppo $(G, *)$, la cui rappresentazione $T(g)$, $g \in G$ agisce su \mathcal{PH} :

$$T(g) : \mathcal{PH} \rightarrow \mathcal{PH}$$

E per $g, h \in G$ avremo la proprietà “di gruppo”:

$$T(g * h) = T(g) \circ T(h)$$

Data ora una qualsiasi $|\psi\rangle \in \mathcal{H}$, è univocamente definito il *raggio vettore* a cui corrisponde. Esiste cioè una proiezione canonica $\pi : \mathcal{H} \rightarrow \mathcal{PH}$, che associa ogni vettore di \mathcal{H} al raggio vettore a cui appartiene.

Un'idea, allora, è di prendere al posto di $T(g)$ una trasformazione $V(g) : \mathcal{H} \rightarrow \mathcal{H}$, che se *proiettata* su \mathcal{PH} tramite π viene identificata proprio con la $T(g)$. Matematicamente:

$$\pi : V(g)(v) \mapsto T(g)(\pi(v))$$

Tuttavia, poiché π è una proiezione, e quindi è *non iniettiva*, si ha che anche tutte le $e^{i\gamma}V(g)(v)$ soddisfano questa proprietà, e quindi la scelta di $V(g)$ non è unica, ma è definita a meno di un fattore moltiplicativo.

Di conseguenza, la proprietà di gruppo diverrà:

$$V(g * h) = e^{i\alpha(g,h)} V(g) \circ V(h)$$

dove α è una qualche funzione di g e h a valori reali.

Ciò fa sì che la mappa $g \mapsto V(g)$ non sia più esattamente una rappresentazione di

³⁹ In generale moltiplicare $|\psi\rangle$ per un qualsiasi complesso a non cambia lo stato. Poiché lavoriamo con vettori normalizzati, $a|\psi\rangle$ equivale, dopo alla normalizzazione, ad un certo sfasamento $e^{i\gamma}$ di $|\psi\rangle$

G , ma una rappresentazione “a meno di una fase”. Chiamiamo questo genere di mappa una **rappresentazione proiettiva**, e la $V(g)$ *definita a meno di una fase* un **raggio operatore** (dato che corrisponde a una $T(g)$ che agisce su \mathcal{PH}).

Purtroppo anche le rappresentazioni proiettive sono *scomode*, per cui questo passaggio apparentemente non ha portato a molto guadagno. Tuttavia, possiamo ora usare le proprietà di spazio vettoriale di \mathcal{H} , e, affrontando un problema alla volta, occupiamoci innanzitutto di capire di che natura siano le $V(g)$ che corrispondono a simmetrie, nella speranza che ciò ci offra qualche semplificazione.

Le $V(g)$ sono normali operatori $\mathcal{H} \rightarrow \mathcal{H}$. Ci chiediamo allora *quali* operatori su \mathcal{H} preservino le probabilità di transizione.

Definizione 22. Ricordiamo che un operatore U si dice **unitario** se verifica $U^\dagger U = \mathbb{I}$. Un tale operatore *preserva* le probabilità di transizione, infatti:

$$\begin{cases} |\phi'\rangle = U|\phi\rangle \\ |\psi'\rangle = U|\psi\rangle \end{cases} \Rightarrow |\langle\phi'|\psi'\rangle|^2 = |\langle U\phi|U\psi\rangle|^2 = |\langle\phi|U^\dagger U\psi\rangle|^2 = |\langle\phi|\psi\rangle|^2$$

Un operatore \bar{U} si dice **antiunitario** se invece soddisfa $(\bar{U}\phi, \bar{U}\psi) = (\phi, \psi)^*$. Poiché la coniugazione preserva il modulo, si ha che anche \bar{U} preserva le probabilità di transizione.

Perciò operatori unitari o antiunitari su \mathcal{H} , se considerati *a meno di una costante*, sono rappresentazioni proiettive di simmetrie fisiche.

Ci chiediamo: queste due possibilità esauriscono tutti i casi possibili? In altre parole, una qualsiasi simmetria data da un gruppo G è rappresentata proiettivamente in \mathcal{H} solo da operatori (anti)unitari?

La risposta (affermativa) è data dal seguente teorema:

Teorema 3.7.1. ⁴⁰ *Le mappe tra raggi vettori in \mathcal{PH} che preservano le probabilità di transizioni (e quindi sono **rappresentazioni** di gruppi che descrivono **simmetrie fisiche**) sono **tutte e sole** indotte dalla proiezione canonica $\pi : \mathcal{H} \rightarrow \mathcal{PH}$ o da operatori **unitari** o da operatori **antiunitari** definiti a meno di una fase.*

Teorema di
Wigner

*Perciò, una simmetria fisica descritta da un gruppo $(G, *)$ ha rappresentazione $G \ni g \mapsto T(g) : \mathcal{PH} \rightarrow \mathcal{PH}$, con $T(g)$ identificata da un **raggio operatore** **(anti)unitario** $(\hat{U}) \hat{U}$:*

$$\hat{U} = \{e^{i\alpha}, U \text{ unitario}\}; \quad \hat{\bar{U}} = \{e^{i\alpha}U, \bar{U} \text{ antiunitario}\}$$

In realtà, se vogliamo che $g \mapsto V(g)$ sia una rappresentazione proiettiva di un gruppo di Lie (ossia “continuo”), allora $V(g)$ deve essere unitaria, e non antiunitaria.

Infatti, la proprietà di compatibilità con l’operazione di gruppo è data da:

$$V(g * h) = e^{i\alpha(g,h)} V(g) \circ V(h)$$

⁴⁰ In realtà questo enunciato è un po’ semplificato. Chi volesse esaminare il teorema nella sua interezza, con tanto di dimostrazione, dia un’occhiata a <https://www.staff.science.uu.nl/~ban00101/lecnotes/repq.pdf>

Per V unitario non c'è problema, dato che la composizione di operatori unitari è ancora unitaria. Ma per $V = \bar{U}$ antiunitario:

$$\bar{U}(g * h) = e^{i\alpha(g,h)} \bar{U}(g) \circ \bar{U}(h) \quad (3.39)$$

Il membro a sinistra è antiunitario, mentre quello a destra è la composizione di operatori antiunitari, che è unitaria. Infatti, applicando due volte la definizione di antiunitarietà:

$$\langle \bar{U}(g)\bar{U}(h)x \mid \bar{U}(g)\bar{U}(h)y \rangle = \langle \bar{U}(g)x \mid \bar{U}(g)y \rangle^* = \langle \bar{U}(g)y \mid \bar{U}(g)x \rangle = \langle y \mid x \rangle^* = \langle x \mid y \rangle$$

Ma l'unico operatore che è sia unitario che antiunitario (e può quindi soddisfare l'uguaglianza (3.39)) è l'operatore nullo, che non può essere una rappresentazione (proiettiva) di un gruppo (per esempio non mappa l'elemento neutro nell'identità).

Nota: gli operatori antiunitari possono però rappresentare (proiettivamente) simmetrie “discrete”, come la *parità* e l'*inversione temporale*⁴¹.

Riepilogando: abbiamo scoperto che una simmetria fisica con una struttura di gruppo $(G, *)$ è *rappresentata proiettivamente* da una mappa $G \ni g \mapsto U(g) : \mathcal{H} \rightarrow \mathcal{H}$, con $U(g)$ operatore unitario.

Vorremmo ora *semplificare* il problema, e ottenere una rappresentazione “normale” (ossia non proiettiva). Matematicamente, stiamo chiedendo se, nell'espressione:

$$U(g * h) = e^{i\alpha(g,h)} U(g)U(h) \quad g, h \in G$$

sia possibile sempre scegliere le fasi degli $U(g)$ e $U(h)$ in modo da cancellare $e^{i\alpha(g,h)}$, e condursi quindi alla relazione di *rappresentazione*:

$$U(g * h) = U(g)U(h)$$

Fortunatamente, ciò è possibile, ma sotto opportune condizioni.

La prima è che G sia connesso⁴² (intuitivamente, se così non fosse si potrebbero avere rappresentazioni “diverse” sulle varie componenti connesse) e semplicemente connesso⁴³, e la seconda è che, nel caso di \mathcal{H} infinito-dimensionale, G abbia determinate (deboli) proprietà, che nella pratica considereremo sempre soddisfatte.

Purtroppo molti gruppi interessanti **non** sono semplicemente connessi. Ciò significa, matematicamente, che data una curva **chiusa** $f : \mathbb{R} \rightarrow G$, $t \mapsto f(t)$, non è possibile “deformarla con continuità” a un punto senza uscire da G .

⁴¹ Per approfondimenti, fare riferimento alla sezione 5 di <https://www.ias.ac.in/article/fulltext/reso/019/10/0900-0916>

⁴² In maniera *pittoresca* si immagini G come un “tutt'uno”, ossia un insieme fatto di un “solo pezzo”

⁴³ Ossia connesso per archi e tale che ogni curva in G possa essere contratta con continuità ad un punto rimanendo in G . Sempre in maniera *pittoresca*, si immagini G semplicemente connesso come “senza buchi”

Per esempio, si consideri $G = \mathbb{T}$, che è il cerchio unitario. Una qualsiasi curva chiusa che “gira attorno al cerchio” non è contraibile. Pittorescamente, si immagina di cercare di “comprimere ad un punto” un elastico che avvolge un bicchiere (il cui contorno circolare costituisce *i punti del gruppo* \mathbb{T}): le uniche possibilità sono *rompere l’elastico* (e quindi fare una trasformazione *non continua*) oppure toglierlo dal bicchiere (ossia *uscire* da G).

Un’idea per ovviare a questo problema è quella di considerare un gruppo “il più simile possibile” a G che sia però semplicemente connesso. Nel nostro esempio pittoresco, potremmo usare, al posto del contorno del bicchiere, una striscia di carta (molto lunga - infinita). Se poniamo un elastico sulla striscia, per quanto lo allunghiamo non riusciremo mai ad *avvolgerla* (ciò proprio non ha senso), e quindi potremo sempre - rimanendo sulla striscia - comprimere l’elastico a un punto. Ci serve ora un modo di *emulare* il caso del bicchiere con questa striscia. Semplicemente, arrotoliamola al bicchiere. Ora, per ogni punto del contorno di prima ve ne sono appena sopra un gran numero (infinito) di corrispondenti sulla striscia - possiamo allora dire che uno qualsiasi di questi punti viene *proiettato* su G originale. Perciò abbiamo trovato un altro gruppo \tilde{G} che è semplicemente connesso (la striscia), e che possiamo mettere in relazione con G (il contorno del bicchiere): in particolar modo, ruotare un punto attorno al bicchiere equivale a traslare i suoi rappresentativi lungo la striscia, e quindi anche le “operazioni di gruppo” sono preservate.

Chiameremo questo \tilde{G} **ricoprimento universale** di G . Formalizziamo in termini matematici tale concetto nella seguente definizione:

Definizione 23. Dato un gruppo $(G, *)$ connesso, definiamo (\tilde{G}, \cdot) , il gruppo di **ricoprimento** (o rivestimento) **universale** di G , come il più piccolo gruppo che ha le seguenti proprietà:

Ricoprimento universale di un gruppo

1. Omeomorfo a G , cioè esiste una mappa π dal ricoprimento al gruppo originario $\pi : \tilde{G} \rightarrow G$ che preserva la struttura di gruppo:

$$\pi(\tilde{g}_1 \cdot \tilde{g}_2) = \pi(\tilde{g}_1) * \pi(\tilde{g}_2) \quad \forall \tilde{g}_1, \tilde{g}_2 \in \tilde{G}; \quad \pi(\tilde{e}) = e$$

ossia proietta la composizione di gruppo di \tilde{G} nella composizione di gruppo di G dei corrispettivi elementi, e associa tra loro gli elementi neutri.

2. \tilde{G} è **semplicemente connesso**, cioè ogni cammino chiuso in \tilde{G} del tipo:

$$\{\tilde{g}(t), t \in [0, 1], \tilde{g}(0) = \tilde{g}(1)\}$$

si può deformare con continuità a un punto.

Facciamo alcuni esempi:

- Se $G = (\mathbb{R}, +)$, un cammino chiuso è uno che “va avanti e indietro” sulla retta reale. Ma allora basta “spostare” il punto estremo (quello raggiunto quando “si gira indietro”) e farlo coincidere con quello iniziale - cosa che si può fare con continuità. Allora il rivestimento universale di $(\mathbb{R}, +)$ è $(\mathbb{R}, +)$ stesso:

$$\tilde{\mathbb{R}} = (\mathbb{R}, +)$$

- Sia invece $G = U(1) = \{e^{i\alpha}, \text{ con l'operazione di moltiplicazione} \}$ (infatti $e^{i\alpha}e^{i\beta} = e^{i(\alpha+\beta)}$).

G non è semplicemente connesso: in effetti posso rappresentarlo come una circonferenza, e se considero un cammino chiuso che “si avvolge completamente su di essa”, non c'è modo di contrarlo ad un punto. Perciò questo caso non si può risolvere come prima.

Potremmo però considerare una “spirale” infinita in tre dimensioni, la cui proiezione sul piano è la circonferenza originaria. Stando sulla spirale, un percorso che è chiuso nella sua proiezione è in realtà aperto e contraibile. Tale spirale è isomorfa a \mathbb{R} , e quindi si ha:

$$\tilde{U}(1) \approx \mathbb{R}$$

Possiamo finalmente enunciare il risultato che stavamo cercando:

Teorema 3.7.2. *Sotto assunzioni deboli su G connesso, esiste una corrispondenza biunivoca tra le rappresentazioni proiettive continue di un gruppo G e le rappresentazioni unitarie del suo gruppo di ricoprimento universale \tilde{G} .*

Teorema di
Bargmann

Concretizziamo tutto ciò nel caso delle **traslazioni**.

Per esempio, una traslazione in una dimensione (spaziale o temporale) è descritta dal gruppo $G = (\mathbb{R}, +)$, dove un numero reale $a \in \mathbb{R}$ indica “di quanto traslare” i punti del sistema che studiamo.

Per teorema di Wigner sappiamo che la rappresentazione proiettiva di \mathbb{R} che agisce su \mathcal{H} è unitaria. In termini concreti, preso un $t \in \mathbb{R}$, ad esso è associato un operatore unitario $U(t) : \mathcal{H} \rightarrow \mathcal{H}$ a meno di una costante, e che è compatibile con l'operazione di gruppo:

$$U(t_1)U(t_2) = e^{i\alpha(t_1, t_2)}U(t_1 + t_2)$$

Per teorema di Bargmann, essendo \mathbb{R} semplicemente connesso (e quindi è il ricoprimento universale di se stesso), sappiamo di poter sempre scegliere le fasi di $U(t_1)$ e $U(t_2)$ in modo da cancellare il termine esponenziale, e quindi ridurci ad una rappresentazione non più proiettiva, per cui vale:

$$U(t_1)U(t_2) = U(t_1 + t_2)$$

La mappa $t \mapsto U(t)$ è continua, e gli $U(t)$ formano un gruppo, che chiamiamo **gruppo continuo a un parametro**⁴⁴ di operatori unitari.

Si dimostra che gruppi di questo tipo sono “generati” da un unico operatore, come vediamo nel seguente teorema:

Teorema 3.7.3. *Dato un gruppo continuo ad un parametro di operatori unitari $U(t)$ in \mathcal{H} , esiste un dominio denso $D(A)$ in \mathcal{H} in cui, $\forall \psi \in D(A)$, (in topologia forte):*

Teorema di Stone

$$\exists \lim_{t \rightarrow 0} \frac{U(t) - \mathbb{I}}{it} \psi \equiv \frac{1}{i} \frac{dU(t)}{dt} \psi \equiv A\psi$$

⁴⁴ ∧ Detti $g \in \{U(t), t \in \mathbb{R}\}$ si ha $g(t)g(s) = g(t+s)$, $g(0) = e$, $g(t)$ continuo in t

con A operatore autoaggiunto in $D(A)$ e $U(t)$ operatore unitario definito da:

$$U(t) = e^{itA}$$

(l'unitarietà di questa forma esponenziale era già stata dimostrata in (3.20))
 L'operatore A è detto **generatore infinitesimo** del gruppo a un parametro $U(t)_{t \in \mathbb{R}}$, e corrisponde “intuitivamente”⁴⁵ ad una “trasformazione infinitesima”. Tale A , essendo un operatore autoaggiunto con dominio denso, è un'osservabile. Stiamo perciò associando ad una grandezza fisica A un “flusso” (l'esponenziale e^{itA}) che è un'operatore che “applica una trasformazione” agli stati del sistema quantistico, lasciandolo “globalmente invariato”, ossia è una simmetria.

3.8 Evoluzione temporale

(Lezione 13 del
22/10/2018)

Occupiamoci ora di analizzare le regole che in MQ determinano l'evoluzione temporale di un sistema da un certo stato $|\psi_1\rangle$ a un altro $|\psi_2\rangle$. Per farlo useremo buona parte dei concetti di teoria dei gruppi enunciati nella sezione precedente. Vi sono due modi *equivalenti* di vedere l'evoluzione temporale di un sistema isolato: la formulazione di Heisenberg e quella di Schrödinger. Esaminiamoli.

3.8.1 Formulazione di Heisenberg

Nella formulazione di Heisenberg si parte dalla descrizione matematica della regola combinatoria di Ritz-Rydberg, per cui le osservabili di un atomo possono essere descritte come:

$$q_{mn}(t) = q_{mn} e^{i\omega_{mn}t}$$

Calcoliamone la derivata temporale:

$$\dot{q}_{mn}(t) = q_{mn} i\omega_{mn} e^{i\omega_{mn}t} \underset{(a)}{=} \frac{\mathcal{E}_n - \mathcal{E}_m}{i\hbar} q_{mn}(t) \underset{(b)}{=} \frac{[q, H]_{mn}}{i\hbar}$$

dove in (a) si è usata la relazione di Rydberg, per cui $\omega_{mn} = \omega_n - \omega_m$ (ogni frequenza è differenza di due frequenze fissate), assieme all'espressione per l'energia $\mathcal{E} = \hbar\omega$, mentre in (b) introduciamo come H la *matrice diagonale* delle energie \mathcal{E}_i . In tal modo, ricordando che q è la matrice (infinito-dim.) delle q_{mn} , possiamo usare la notazione più compatta del commutatore di matrici.

Estendendo ad una generica osservabile A troviamo che l'evoluzione temporale di A , indicata con $A^H(t)$, deve soddisfare la seguente relazione:

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

⁴⁵Un'ottima spiegazione intuitiva di tutto - e anche di alcuni argomenti dei prossimi paragrafi - è disponibile a <https://physics.stackexchange.com/questions/133758/intuitive-meaning-of-the-exponential-form-of-an-unitary-operator-in-quantum-mech>

3.8.2 Formulazione di Schrödinger

Schrödinger, invece, partì da un'analogia con l'ottica geometrica, e riutilizzò il formalismo ondulatorio per l'elettromagnetismo, giungendo a scrivere l'equazione (di Schrödinger *dipendente dal tempo*):

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

Equazione di Schrödinger dipendente dal tempo

3.8.3 Formulazione di Dirac

Abbiamo appena visto due “descrizioni dell'evoluzione temporale”. Ma qual è esattamente la relazione tra di esse? (Ci aspettiamo siano equivalenti, ma come?)

Di nuovo, l'identificazione tra i due modi di procedere fu effettuata da Dirac.

Considereremo in questo corso i soli **sistemi isolati**, per cui vale l'**omogeneità del tempo**: essendo non “perturbati”, fare le misure “prima” o “dopo” non cambia il risultato, contano solo gli intervalli di tempo *relativi* tra due misurazioni successive.

Omogeneità del tempo

Matematicamente, l'evoluzione temporale del sistema è quindi data da una traslazione temporale che forma un gruppo additivo $(\mathbb{R}, +)$ isomorfo a \mathbb{R} . Ciò corrisponde al fatto che, se un sistema nello stato $\Sigma(0)$ a $t = 0$ evolve nello stato $\Sigma(t_1)$ ad un tempo t_1 e a $\Sigma(t_2)$ dopo un ulteriore t_2 , possiamo arrivare a tale stato “finale” $\Sigma(t_2)$ in maniera completamente equivalente partendo da $\Sigma(0)$ e lasciando passare un tempo $t_1 + t_2$ (dato che non ci saranno perturbazioni nel frattempo).

Consideriamo la probabilità di transizione da $|\phi\rangle$ a $|\psi\rangle$ data da: $\langle\phi|\psi\rangle^2$. Se l'evoluzione temporale è data da $|\phi\rangle \rightarrow |\phi(t)\rangle$ dopo un certo t , $|\psi\rangle \rightarrow |\psi(t)\rangle$ dopo un certo t , allora, per omogeneità del tempo:

$$|\langle\phi(t)|\psi(t)\rangle|^2 = \langle\phi|\psi\rangle^2$$

Evoluzione temporale di sistemi isolati preserva le probabilità di transizione

Se così non fosse vi sarebbero istanti “ben riconoscibili”, e le misure dipenderebbero dal “valore assoluto del tempo”.

Allora, un operatore che descrive l'evoluzione temporale per $|\phi\rangle \rightarrow |\phi'\rangle$, $|\psi\rangle \rightarrow |\psi'\rangle$, deve preservare le probabilità di transizione:

$$\langle\phi|\psi\rangle^2 = |\langle\phi'|\psi'\rangle|^2$$

E quindi corrisponde ad una **simmetria fisica** come definita nella (Definizione 21).

Poiché il gruppo $(\mathbb{R}, +)$ da cui essa deriva è un gruppo “continuo” (di Lie), si ha che una sua *rappresentazione proiettiva* in \mathcal{H} è data da operatori **unitari** (per il teorema di Wigner, da cui escludiamo gli operatori antiunitari).

Avremo cioè una mappa (la rappresentazione) che manda elementi del gruppo $t \in \mathbb{R}$ in operatori unitari $U(t)$ che è *compatibile* con l'operazione di gruppo (che in questo caso è la normale addizione). Vale cioè:

$$U(t_1)U(t_2) = e^{i\alpha(t_1, t_2)}U(t_1 + t_2) \quad \forall t_1, t_2 \in \mathbb{R}$$

La fase $\alpha(t_1, t_2)$ è data dal fatto che si tratta di una rappresentazione *proiettiva*, e quindi gli operatori $U(t_1)$ e $U(t_2)$ sono *definiti a meno di una fase*.

Se passiamo però ad una rappresentazione del gruppo di ricoprimento universale di $(\mathbb{R}, +)$, che è $(\mathbb{R}, +)$ stesso dato che è già semplicemente connesso, possiamo eliminare la fase per il *teorema di Bargmann*, e ottenere una *rappresentazione unitaria* (non più proiettiva), data da $t \mapsto U(t)$ con:

$$U(t_1)U(t_2) = U(t_1 + t_2) \quad \forall t_1, t_2 \in \mathbb{R}$$

Notiamo allora che gli operatori $U(t)$ definiti a partire dai t del gruppo formano a loro volta un gruppo, o meglio un *gruppo unitario a un parametro*, per cui vale il **teorema di Stone**. Esiste cioè la derivata:

$$\exists \frac{1}{i} \frac{dU(t)}{dt} \psi \equiv A\psi \Rightarrow U(t) = e^{itA}$$

e tale A è un operatore autoaggiunto con dominio $D(A)$ denso in \mathcal{H} , ossia è un'osservabile, detta **generatore infinitesimo** della simmetria.

Come trovare tale A ? Un'idea è partire da un'analogia classica. In MC, infatti, sappiamo che il generatore dell'evoluzione temporale è l'Hamiltoniana:

$$\frac{df}{dt} = \{f, H\}$$

Poniamo allora $A \sim H$. Con un argomento di analisi dimensionale, poiché t ha le dimensioni di un tempo, per rendere l'esponente di e^{itA} adimensionale A deve avere le dimensioni di $[t]^{-1}$. Notiamo allora che:

$$[H] = [\text{energia}]; \quad [\hbar] = [\text{energia}] [t]$$

E quindi un'idea (confermata dagli esperimenti) è data da:

$$A = -\frac{H}{\hbar}$$

dove il segno meno si aggiunge convenzionalmente per riprodurre la notazione di Schrödinger, come vediamo nel seguente assioma.

Assioma 4. *L'evoluzione di uno stato descritto da $\psi \in \mathcal{H}$ di un **sistema isolato** in MQ è data da:*

$$\psi(t) = U(t) \psi$$

Assioma
dell'evoluzione
degli stati puri

Con:

$$U(t) = e^{-\frac{i}{\hbar} H t}$$

Ove H è l'hamiltoniana quantistica, che descrive l'energia del sistema.
Questa è l'equazione più generale possibile, che vale per qualsiasi ψ !

Se in più $\psi \in D(H)$, allora, dal teorema di Stone otteniamo l'equazione di Schrödinger, sviluppando la seguente derivata:

$$i\hbar \frac{\partial}{\partial t} \psi(t) = i\hbar \frac{\partial}{\partial t} U(t) \psi = i\hbar \frac{\partial U(t)}{\partial t} \psi \underset{(a)}{=} H U(t) \psi = H \psi(t)$$

dove in (a) si è usato il teorema di Stone per calcolare la derivata, per cui $U(t) = e^{itA}$, con $A = -H/\hbar$.

Nota: Sebbene l'evoluzione $U(t)$ sia definita per ogni $\psi \in \mathcal{H}$, solo per le $\psi \in D(H)$ ha senso l'equazione di Schrödinger. Questo, tra l'altro, risolve il problema che può dar luogo a fraintendimenti nella soluzione all'equazione di Schrödinger *indipendente dal tempo*:

$$H\psi = \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right) \psi(x)$$

Perché tale equazione sia ben definita, $\psi''(x)$ deve almeno esistere q.o., ma questo non è garantito dalla sola condizione $\psi \in L^2$! Questa non è una restrizione su $\psi \in L^2$, ma su $D(H)$. Perciò, un generico stato *fisico* $|\psi\rangle$ appartiene a L^2 , mentre l'equazione di Schrödinger può essere applicata solamente alle ψ di cui esiste la derivata seconda. In effetti, tale equazione produrrà soluzioni che saranno *generalmente* più regolari delle generiche $\psi \in L^2$.

Espandiamo la relazione $U(t) = e^{-iHt/\hbar}$ data dal teorema di Stone per l'evoluzione di uno stato nel caso generale. Dato che H è autoaggiunto in $D(H)$, una funzione di H , come la $U(t)$, è ben definita mediante la famiglia spettrale $P^H(\lambda)$:

$$U(t)\psi = \int e^{-\frac{it\lambda}{\hbar}} dP^H(\lambda) \psi$$

Allora (nel caso senza degenerazione, per semplicità):

$$\begin{aligned} U(t)|\psi\rangle &= \sum_{\mathcal{E}_n \in \sigma_P(H)} \exp\left(-\frac{i}{\hbar}\mathcal{E}_n t\right) |\mathcal{E}_n\rangle \langle \mathcal{E}_n | \psi \rangle + \\ &+ \int_{\sigma_C(H)} \exp\left(-\frac{i}{\hbar}\mathcal{E} t\right) |\mathcal{E}\rangle \langle \mathcal{E} | \psi \rangle d\mathcal{E} \quad \forall |\psi\rangle \in \mathcal{H} \end{aligned} \quad (3.40)$$

Vedremo nelle prossime sezioni cosa succede a risolvere l'equazione di Schrödinger per una $\psi \notin D(H)$, come quella definita da $\psi(x) = 1$ tra 0 e 1 (e pari a 0 altrimenti). Noteremo allora come il metodo generale, per cui $\psi(t) = U(t)\psi$, produrrà la soluzione corretta, mentre l'equazione di Schrödinger dipendente dal tempo produrrà un risultato assurdo.

Riepilogando: siamo partiti dall'**omogeneità del tempo**, da cui deriva che le *traslazioni temporali* (descritte dal gruppo $(\mathbb{R}, +)$) devono lasciare invariate le probabilità di transizione.

Partendo da tale proprietà, Wigner ha dimostrato che tali trasformazioni sono *rappresentate proiettivamente* su \mathcal{H} , e Bargmann, notando che il ricoprimento universale del gruppo di omogeneità è \mathbb{R} , ha dimostrato che sono rappresentate da un gruppo a un parametro di operatori unitari $U(t)$, e infine per teorema di Stone si giunge all'espressione:

$$U(t) = \exp\left(-\frac{i}{\hbar}tH\right)$$

Nel caso più generale, tenendo conto anche della degenerazione degli autovalori, la formula (3.40) diviene:

$$\begin{aligned}\psi(t) &= U(t)\psi = e^{-i\frac{t}{\hbar}H}\psi = \int e^{-i\frac{t}{\hbar}\lambda} dP^H(\lambda)\psi \quad \forall \psi \in \mathcal{H} \\ |\psi(t)\rangle &= \sum_{\mathcal{E}_n \in \sigma_P(H)} e^{-i\frac{t}{\hbar}\mathcal{E}_n} \sum_{r=1}^{d(\mathcal{E}_n)} |\mathcal{E}_n, r\rangle \langle \mathcal{E}_n, r | \psi \rangle + \int_{\sigma_C(H)} d\mathcal{E} e^{-i\frac{\mathcal{E}}{\hbar}t} \sum_{r=1}^{d(\lambda)} |\mathcal{E}, r\rangle \langle \mathcal{E}, r | \psi \rangle\end{aligned}\quad (3.41)$$

Se inoltre ψ appartiene al dominio dell'operatore energia ($\psi \in D(H)$), allora $\psi(t)$ si può ottenere anche risolvendo l'equazione di Schrödinger:

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

Analogamente al caso degli stati puri, uno **stato misto** $\rho = \sum_i c_i |\phi_i\rangle \langle \phi_i|$ evolve come:

*Evoluzione degli
stati misti*

$$\rho(t) = \sum_i c_i |\phi_i(t)\rangle \langle \phi_i(t)| = \sum_i c_i U(t) |\phi_i\rangle \langle \phi_i| U^\dagger(t) = U(t) \rho U^\dagger(t) \quad (3.42)$$

Se tutte le $|\phi_i\rangle \in D(H)$ allora vale anche un'equazione analoga a quella di Schrödinger. Partiamo scrivendo la relazione per le $|\phi_i\rangle$ e quella duale per le $\langle \phi_i|$ (per cui basta calcolare un complesso coniugato):

$$\begin{aligned}i\hbar \frac{\partial}{\partial t} |\phi_i(t)\rangle &= H |\phi_i(t)\rangle \\ -i\hbar \frac{\partial}{\partial t} \langle \phi_i(t)| &= \langle \phi_i(t)| H\end{aligned}$$

Perciò per $\sum_i c_i |\phi_i\rangle \langle \phi_i|$ si ottiene l'**equazione di von Neumann**:

$$i\hbar \frac{\partial}{\partial t} \rho(t) = \sum_i c_i (H |\phi_i(t)\rangle \langle \phi_i(t)| - |\phi_i(t)\rangle \langle \phi_i(t)| H) = [H, \rho(t)] \quad (3.43)$$

3.8.4 Visuali di Schrödinger e Heisenberg

Abbiamo finora discusso l'evoluzione temporale degli stati. Agli effetti potremmo "condensare" la descrizione completa dell'evoluzione di un sistema fisico nella variazione dei suoi stati, e quindi assumere che le osservabili restino sempre le stesse, ossia che gli operatori che le descrivono non dipendano dal tempo.

Tale convenzione fa parte della cosiddetta **visuale di Schrödinger** dell'evoluzione⁴⁶:

*Visuale di
Schrödinger*

$$\begin{aligned}\psi &\rightarrow \psi(t) = U(t)\psi \\ A &\rightarrow A\end{aligned}\quad (3.44)$$

Del resto, sperimentalmente abbiamo accesso esclusivamente ai valor medi, che evolvono come:

$$\langle A \rangle_{\psi(t)} = \langle \psi(t) | A | \psi(t) \rangle = \langle \psi | U^\dagger(t) A U(t) | \psi \rangle = \langle \psi | A^H(t) | \psi \rangle = \langle A^H(t) \rangle_\psi \quad (3.45)$$

ove $A^H(t) = U^\dagger(t) A U(t)$.

L'equazione (3.45) suggerisce un approccio diverso: focalizzarci sull'evoluzione delle osservabili ($A^H(t)$) e tenere fissi gli stati.

Così facendo si ottiene la **visuale di Heisenberg** dell'evoluzione⁴⁷:

$$\begin{aligned} \psi &\rightarrow \psi \\ A &\rightarrow A^H(t) = U^\dagger(t) A U(t) \end{aligned} \quad (3.46)$$

Confrontando (3.42) con (3.46) si può notare una differenza nell'ordine degli aggiunti. Ciò non sorprende: avevamo notato un qualcosa di simile anche in MC, quando scrivendo l'evoluzione degli stati in (2.17) si otteneva un segno $-$, che non c'era nel caso dell'evoluzione delle sole osservabili.

Notiamo inoltre che la formula $A^H(t) = U^\dagger(t) A U(t)$ risulta *diversa* rispetto a quella di una normale simmetria fisica vista in (3.38). Il punto chiave è che, nelle visuali di Schrödinger o Heisenberg, solo le osservabili o solo gli stati si evolvono, e mai entrambi contemporaneamente, mentre nel caso di una simmetria fisica (es. una rotazione) sia operatori che stati vengono modificati.

Se il commutatore $[A^H(t), H]$ esiste, allora otteniamo l'**equazione di Heisenberg**:

$$i\hbar \frac{\partial}{\partial t} A^H(t) = i\hbar \frac{d}{dt} (U^\dagger(t) A U(t)) = i\hbar \frac{dU^\dagger(t)}{dt} A U(t) + i\hbar U^\dagger(t) A \frac{dU(t)}{dt} \quad (3.47)$$

Sostituendo le espressioni per $U(t)$ e $U^\dagger(t)$:

$$U^\dagger(t) = \exp\left(i \frac{t}{\hbar} H\right); \quad U(t) = \exp\left(-\frac{i}{\hbar} t H\right)$$

otteniamo:

$$(3.47) = -H U^\dagger(t) A U(t) + U^\dagger(t) A U(t) H = [A^H(t), H]$$

E allora:

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

Confrontando l'espressione appena trovata con quella classica:

$$\frac{df}{dt} = \{f, H\}$$

⁴⁶^In inglese "Schrödinger picture"

⁴⁷^Chiaramente entrambi gli approcci danno luogo agli stessi risultati sperimentali - che in una teoria coerente devono rimanere fissi indipendentemente dalla descrizione.

*Evoluzione di
un'osservabile:
 $A^H(t)$*

*Visuale di
Heisenberg*

si potrebbe dire (come fece Dirac) che il commutatore $[\cdot, \cdot]/(i\hbar)$ non è altro che la versione “quantizzata” delle parentesi di Poisson classiche:

$$\{\cdot, \cdot\} \rightarrow \frac{[\cdot, \cdot]}{i\hbar}$$

Perciò, per un sistema isolato, l'evoluzione temporale in MQ procede in un modo molto "classico", e soprattutto è completamente **deterministica**.

Tuttavia, per ricavare informazioni sarà necessario fare una misura, e in quel momento il sistema non potrà più essere considerato isolato, e in particolare esibirà un comportamento quantistico (probabilistico).

Facciamo ora qualche esempio per familiarizzarci con il formalismo, e discuteremo poi la teoria della misura (cosa succede ad un sistema quantistico a seguito di una misurazione).

3.8.5 Buca infinitamente profonda in 1D

Consideriamo una buca di potenziale infinitamente profonda tra $[-\frac{a}{2}, \frac{a}{2}]$, in una dimensione⁴⁸.

L'energia della particella è data da:

$$H = \frac{p^2}{2m} + V(x); \quad V(x) = \begin{cases} 0 & |x| < \frac{a}{2} \\ V_0 \rightarrow +\infty & |x| \geq \frac{a}{2} \end{cases}$$

Consideriamo gli stati con energia fissata \mathcal{E} , e confrontiamo le previsioni di MC e MQ per il comportamento della particella.

In Meccanica Classica

Classicamente non vi è alcuna limitazione sui valori di energia che può assumere la particella, e perciò $\sigma(H) = \mathbb{R}_+$.

Supponiamo di sapere solo che l'energia è \mathcal{E} (ossia, siamo sicuri che facendo una misura di energia si troverà con certezza \mathcal{E}). Nello spazio delle fasi, ciò corrisponde a considerare tutti i punti che hanno la x tra⁴⁹ $-a/2$ e $+a/2$, e p tale che l'energia cinetica $p^2/2m$ sia esattamente pari a \mathcal{E} . Nessuno di tali punti è preferibile ad un altro, e quindi lo stato è dato da una distribuzione uniforme:

$$\rho_{\mathcal{E}}(x, p) = A \chi_{[-\frac{a}{2}, \frac{a}{2}]}(x) \delta\left(\frac{p^2}{2m} - \mathcal{E}\right) \quad (3.48)$$

dove χ è la funzione caratteristica, e A è la costante di normalizzazione, che troviamo imponendo che l'integrale di $\rho_{\mathcal{E}}(x, p)$ sullo spazio delle fasi sia esattamente

⁴⁸ΛFisicamente tale sistema descrive “pareti impenetrabili”, sia classicamente che quantisticamente

⁴⁹ΛSe così non fosse, $H \rightarrow \infty$ per come abbiamo definito il potenziale, e perciò certamente non varrebbe \mathcal{E}

1:

$$\int_{\mathbb{R}} \chi_{[-\frac{a}{2}, \frac{a}{2}]}(x) dx \int_{\mathbb{R}} dp A \delta\left(\frac{p^2}{2m} - \mathcal{E}\right) = \int_{-\frac{a}{2}}^{\frac{a}{2}} dx \int_{\mathbb{R}} dp A \delta\left(\frac{p^2}{2m} - \mathcal{E}\right) \stackrel{!}{=} 1 \quad (3.49)$$

Ricordiamo che vale la seguente identità⁵⁰ per la composizione della delta di Dirac con una funzione differenziabile $g(x)$:

$$\delta(g(x)) = \sum_i \frac{\delta(x - a_i)}{|g'(a_i)|}$$

dove a_i sono gli zeri (semplici) di $g(x)$.

Applicandola al nostro caso:

$$\begin{aligned} \delta\left(\frac{p^2}{2m} - \mathcal{E}\right) &= \frac{1}{|p|/m} \Big|_{p=\pm\sqrt{2m\mathcal{E}}} (\delta(p - \sqrt{2m\mathcal{E}}) + \delta(p + \sqrt{2m\mathcal{E}})) = \\ &= \frac{m}{\sqrt{2m\mathcal{E}}} (\delta(p - \sqrt{2m\mathcal{E}}) + \delta(p + \sqrt{2m\mathcal{E}})) \end{aligned} \quad (3.50)$$

Sostituendo nell'integrale in (3.49) otteniamo:

$$\begin{aligned} (3.49) &= a \int_{\mathbb{R}} A \underbrace{\sqrt{\frac{m}{2\mathcal{E}}} (\delta(p - \sqrt{2m\mathcal{E}}) + \delta(p + \sqrt{2m\mathcal{E}}))}_{=2} dp = 2a \sqrt{\frac{m}{2\mathcal{E}}} A = \\ &= a \sqrt{\frac{\mathcal{E}}{2m}} A \stackrel{!}{=} 1 \Rightarrow A = \frac{1}{a} \sqrt{\frac{\mathcal{E}}{2m}} \end{aligned}$$

Perciò lo stato (3.48) è:

$$\rho_{\mathcal{E}}(x, p) = \frac{1}{a} \sqrt{\frac{\mathcal{E}}{2m}} \chi_{[-\frac{a}{2}, \frac{a}{2}]}(x) \delta\left(\frac{p^2}{2m} - \mathcal{E}\right) \quad (3.51)$$

Notiamo che **non si tratta di uno stato puro** (ossia non è possibile scriverlo come una combinazione *finita* di δ di Dirac). In effetti vi sono 2 valori possibili per il momento, e ∞ per la posizione.

Possiamo naturalmente calcolare la probabilità che la particella si trovi in un certo range.

Per esempio, la densità di probabilità che il valore di posizione λ sia all'interno di $[-\frac{a}{2}, \frac{a}{2}]$ è data da:

$$\begin{aligned} w_{\rho_{\mathcal{E}}}^X(\lambda) &= \frac{d}{d\lambda} \int_{\mathbb{R}} dp \int_{-\frac{a}{2}}^{\frac{a}{2}} dx \rho_{\mathcal{E}}(x, p) H(\lambda - x) = \int_{\mathbb{R}} dp \int_{-\frac{a}{2}}^{\frac{a}{2}} dx \rho_{\mathcal{E}}(x, p) \delta(\lambda - x) = \\ &\stackrel{(a)}{=} \int_{-\frac{a}{2}}^{\frac{a}{2}} dx \chi_{[-\frac{a}{2}, \frac{a}{2}]}(x) \frac{\delta(\lambda - x)}{a} = \frac{1}{a} \end{aligned}$$

⁵⁰ΛPag. 31 di [1]

Dove in (a) si è svolto l'integrale in dp :

$$\begin{aligned}\int_{\mathbb{R}} \rho_{\mathcal{E}}(x, p) dp &= \int_{\mathbb{R}} \frac{1}{a} dp \sqrt{\frac{\mathcal{E}}{2m}} \chi_{[-\frac{a}{2}, \frac{a}{2}]}(x) \sqrt{\frac{m}{2\mathcal{E}}} \underbrace{(\delta(p - \sqrt{2m\mathcal{E}}) + \delta(p + \sqrt{2m\mathcal{E}}))}_2 = \\ &= \frac{1}{a} \chi_{[-\frac{a}{2}, \frac{a}{2}]}(x)\end{aligned}$$

Il risultato è proprio quello che ci si aspetta da una distribuzione uniforme.

Per quanto riguarda i momenti, abbiamo solo due valori possibili (uno positivo e uno negativo). Per esempio, potremmo calcolare la probabilità che una misura del sistema nello stato $\rho_{\mathcal{E}}$ trovi esattamente $\sqrt{2m\mathcal{E}}$:

$$\begin{aligned}\mathcal{P}_{\rho_{\mathcal{E}}}^p(\{\sqrt{2m\mathcal{E}}\}) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \chi_{\{\sqrt{2m\mathcal{E}}\}}(p) \rho_{\mathcal{E}}(x, p) dx dp = \\ &\stackrel{(3.51)}{=} \frac{1}{a} \sqrt{\frac{\mathcal{E}}{2m}} \underbrace{\int_{\mathbb{R}} \chi_{[-\frac{a}{2}, \frac{a}{2}]}(x) dx}_a \int_{\mathbb{R}} \chi_{\{\sqrt{2m\mathcal{E}}\}}(p) \delta\left(\frac{p^2}{2m} - \mathcal{E}\right) dp = \\ &\stackrel{(3.50)}{=} \frac{1}{a} \int_{\mathbb{R}} dp \sqrt{\frac{\mathcal{E}}{2m}} \sqrt{\frac{m}{2\mathcal{E}}} [\delta(p - \sqrt{2m\mathcal{E}}) + \delta(p + \sqrt{2m\mathcal{E}})] \chi_{\{2m\mathcal{E}\}}(p) = \\ &= \sqrt{\frac{\mathcal{E}}{2m}} \sqrt{\frac{m}{2\mathcal{E}}} = \frac{1}{2}\end{aligned}$$

In Meccanica Quantistica

Si dimostra che, analogamente al caso classico, per $V_0 \rightarrow +\infty$ la particella può avere solo posizioni in $[-\frac{a}{2}, \frac{a}{2}]$ ⁵¹. Perciò $\mathcal{H} = L^2\left([-\frac{a}{2}, \frac{a}{2}] dx\right)$.

Per prima cosa definiamo bene la H quantistica, che deve essere un operatore autoaggiunto:

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2}$$

Occorre perciò che il dominio $D(H)$ sia tale che $H = H^\dagger$.

Nota: Specificare il dominio dell'hamiltoniana è **fondamentale** per definire tale operatore. Un'errata definizione porta a risultati assurdi!

Prendiamo $\psi \in D(H^\dagger)$, $\varphi \in D(H)$ e verifichiamo:

$$(\psi, H\varphi) = (H^\dagger\psi, \varphi) \quad (3.52)$$

Espandendo i prodotti scalari e “spostando” la derivata seconda sulla ψ con un'integrazione per parti (in (a) e in (b)):

$$\begin{aligned}-\frac{\hbar^2}{2m} \int_{-\frac{a}{2}}^{\frac{a}{2}} \psi^*(x) \varphi''(x) dx &\stackrel{(a)}{=} -\frac{\hbar^2}{2m} \psi^* \varphi' \Big|_{-\frac{a}{2}}^{\frac{a}{2}} + \frac{\hbar^2}{2m} \int_{-\frac{a}{2}}^{\frac{a}{2}} \psi^{*'}(x) \varphi'(x) dx \\ &\stackrel{(b)}{=} -\frac{\hbar^2}{2m} \psi^* \varphi' \Big|_{-\frac{a}{2}}^{\frac{a}{2}} + \frac{\hbar^2}{2m} \psi^{*'} \varphi \Big|_{-\frac{a}{2}}^{\frac{a}{2}} - \underbrace{\frac{\hbar^2}{2m} \int_{-\frac{a}{2}}^{\frac{a}{2}} \psi^{*''}(x) \varphi(x) dx}_{=(H^\dagger\psi, \varphi)}\end{aligned}$$

La funzione d'onda si annulla ai margini di una buca infinita

⁵¹ΛA priori in MQ non è ovvio, e ci ritorneremo più avanti

Perciò l'uguaglianza (3.52) è verificata se i due termini evidenziati si annullano. Vogliamo allora imporre condizioni su φ in modo che si verifichi proprio questo, e che soprattutto da ciò automaticamente segua che le stesse condizioni debbano valere anche su ψ :

$$\psi^* \varphi' - \psi'^* \varphi \Big|_{-\frac{a}{2}}^{\frac{a}{2}} = 0 \Rightarrow \text{Stesse condizioni su } \psi \quad (3.53)$$

Proviamo imponendo:

$$\varphi\left(\frac{a}{2}\right) = 0 = \varphi\left(-\frac{a}{2}\right)$$

Si ottiene da (3.53):

$$\psi^* \varphi' \Big|_{-\frac{a}{2}}^{\frac{a}{2}} \Rightarrow \psi\left(\frac{a}{2}\right) = 0 = \psi\left(-\frac{a}{2}\right)$$

essendo $\varphi'\left(\frac{a}{2}\right), \varphi'\left(-\frac{a}{2}\right)$ arbitrarie.

In questo modo l'operatore H è autoaggiunto. Possiamo allora scrivere il suo dominio:

$$D(H) = \left\{ \varphi \in \mathcal{H} \mid \varphi \text{ regolare (vale l'integrazione per parti per } \varphi \text{ e } \varphi') , \varphi\left(\pm\frac{a}{2}\right) = 0 \right\}$$

Cerchiamo lo spettro discreto $\sigma_P(H)$ risolvendo l'equazione agli autovalori in $D(H)$:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \varphi_{\mathcal{E}}(x) = \mathcal{E} \varphi_{\mathcal{E}}(x)$$

che ha la forma dell'equazione differenziale di un oscillatore armonico di pulsazione k :

$$\frac{d^2}{dx^2} \varphi_{\mathcal{E}}(x) + \underbrace{\frac{2m\mathcal{E}}{\hbar^2}}_{k^2} \varphi_{\mathcal{E}}(x) = 0 \Rightarrow k = \frac{\sqrt{2m\mathcal{E}}}{\hbar} \quad (3.54)$$

E perciò l'integrale generale è dato da:

$$\varphi_{\mathcal{E}} = c_+(x) e^{ikx} + c_- e^{-ikx}$$

Imponendo che la soluzione si annulli agli estremi della buca (come richiesto nel dominio di H):

$$\varphi_{\mathcal{E}}\left(\pm\frac{a}{2}\right) = 0 \Rightarrow \begin{cases} c_+ e^{ik\frac{a}{2}} + c_- e^{-ik\frac{a}{2}} = 0 \\ c_+ e^{-ik\frac{a}{2}} + c_- e^{ik\frac{a}{2}} = 0 \end{cases} \Rightarrow \begin{pmatrix} e^{ik\frac{a}{2}} & e^{-ik\frac{a}{2}} \\ e^{-ik\frac{a}{2}} & e^{ik\frac{a}{2}} \end{pmatrix} \begin{pmatrix} c_+ \\ c_- \end{pmatrix} = 0 \quad (3.55)$$

Se la matrice delle esponenziali è invertibile, allora si ottiene immediatamente la soluzione nulla ($c_+ = 0 = c_-$), che non definisce alcuno stato fisico. L'unica possibilità è che allora tale matrice sia singolare, e ciò si verifica se:

$$\det \begin{pmatrix} e^{ik\frac{a}{2}} & e^{-ik\frac{a}{2}} \\ e^{-ik\frac{a}{2}} & e^{ik\frac{a}{2}} \end{pmatrix} = 0 \Rightarrow e^{ika} = e^{-ika} \Leftrightarrow \sin(ka) = 0 \quad (3.56)$$

Otteniamo perciò un insieme discreto di soluzioni, le cui pulsazioni soddisfano: $k_n a = n\pi$ $n \in \mathbb{Z}$. Sostituendo in (3.54) otteniamo che gli autovalori dell'energia, ossia i possibili risultati che si possono ottenere da una misura dell'energia della particella nella buca, sono:

$$\mathcal{E}_n = \frac{\hbar^2}{2m} \left(\frac{n\pi}{a} \right)^2 \quad (3.57)$$

*Autovalori di H
per la buca
infinita*

Cerchiamo ora gli autovettori $\varphi_{\mathcal{E}_n}$ (o autofunzioni) ad essi associati⁵². Una “furbizia” per ottenerli più velocemente è riscrivere la condizione per la singolarità (3.56) come:

$$\begin{aligned} \sin(k_n a) &= 2 \sin\left(k_n \frac{a}{2}\right) \cos\left(k_n \frac{a}{2}\right) = 0 \\ \Rightarrow \underbrace{\sin\left(\frac{k_n a}{2}\right)}_{(a)} &= 0 \vee \underbrace{\cos\left(\frac{k_n a}{2}\right)}_{(b)} = 0 \end{aligned}$$

Esaminiamo separatamente i casi (a) e (b).

a) $\sin k_n \frac{a}{2} = 0 \Rightarrow k_n \frac{a}{2} = n\pi \Rightarrow k_n = \frac{2n\pi}{a}$.
Riscrivendo poi il sin come:

$$\begin{aligned} \sin\left(k_n \frac{a}{2}\right) = 0 &\Rightarrow \frac{1}{2i} \left[\exp\left(ik_n \frac{a}{2}\right) - \exp\left(-ik_n \frac{a}{2}\right) \right] = 0 \\ &\Rightarrow \exp\left(ik_n \frac{a}{2}\right) = \exp\left(-ik_n \frac{a}{2}\right) \end{aligned}$$

E quindi possiamo semplificare l'espressione (3.55), da cui ricaviamo:

$$\begin{aligned} c_+ + c_- = 0 &\Rightarrow c_- = -c_+ \\ \varphi_{\mathcal{E}}^{2n}(x) &= A \sin\left(\frac{2n\pi}{a}x\right) \end{aligned}$$

Notiamo che $n \neq 0$ (se così non fosse si otterrebbe la soluzione nulla, che non è fisica), e inoltre n e $-n$ corrispondono allo stesso stato:

$$\varphi_{\mathcal{E}}^{2n}(x) = -\varphi_{\mathcal{E}}^{-2n}(x)$$

Normalizzando:

$$\begin{aligned} 1 &\stackrel{!}{=} \int_{-\frac{a}{2}}^{\frac{a}{2}} |\varphi_{\mathcal{E}}^{2n}(x)|^2 dx = \int_{-\frac{a}{2}}^{\frac{a}{2}} |A|^2 \sin^2 \frac{2n\pi x}{a} dx = |A|^2 \frac{a}{2} \\ &\Rightarrow \varphi_{\mathcal{E}}^{2n}(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{2n\pi}{a}x\right) \end{aligned}$$

essendo la media integrale di \sin^2 pari a $1/2$.

⁵²^Ometteremo nei prossimi passaggi la n a pedice per economia di notazione

b) Analogamente, per la condizione (b) avremo:

$$\cos \frac{k_n a}{2} = 0 \Rightarrow k_n = (2n + 1) \frac{\pi}{a}$$

e riscrivendo il cos tramite esponenziali si ottiene:

$$\begin{aligned} \cos \left(\frac{k_n a}{2} \right) &= \frac{1}{2} \left[\exp \left(i k_n \frac{a}{2} \right) + \exp \left(-i k_n \frac{a}{2} \right) \right] = 0 \\ \Rightarrow \exp \left(i k_n \frac{a}{2} \right) &= - \exp \left(-i k_n \frac{a}{2} \right) \end{aligned}$$

Possiamo allora semplificare la (3.55):

$$c_+ \exp \left(-i k_n \frac{a}{2} \right) + c_- \exp \left(i k_n \frac{a}{2} \right) = 0 \Rightarrow c_+ = -c_-$$

La normalizzazione è identica al caso precedente, e alla fine si ottiene:

$$\varphi_{\mathcal{E}}^{2n+1}(x) = \sqrt{\frac{2}{a}} \cos \left(\frac{(2n+1)\pi}{a} x \right)$$

Alla fine, le soluzioni dell'equazione agli autovalori in $D(H)$ sono quindi:

*Autofunzioni di
H per la buca
infinita*

$$\varphi_n(x) = \begin{cases} \sqrt{\frac{2}{a}} \cos \left(\frac{n\pi x}{a} \right) & n \text{ dispari} > 0 \\ \sqrt{\frac{2}{a}} \sin \left(\frac{n\pi x}{a} \right) & n \text{ pari} > 0 \end{cases}$$

Notiamo che tali autofunzioni sono pari per n dispari, e dispari per n pari (con $n > 0$).

Abbiamo quindi dimostrato che esiste una (sola) soluzione $\forall n \in \mathbb{N}$, per cui la degenerazione di $\sigma_P(H) = 1$.

Inoltre, poiché $\{\varphi_n\}_{n \in \mathbb{N}}$ costituisce una base ortonormale per $L^2([-\frac{a}{2}, \frac{a}{2}], dx)$, lo spettro è esclusivamente puntuale, e quindi non c'è $\sigma_C(H)$.

Differenze tra MC e MQ

- In MC l'energia assume valori **continui** $\sigma(H_{clas.}) = \mathbb{R}_+$, mentre in MQ lo spettro è **discreto**:

$$\sigma(H_q) = \{\mathcal{E}_n = \frac{\hbar^2 k_n^2}{2m} = \hbar^2 \frac{n^2 \pi^2}{2ma^2}, n \in \mathbb{N}\}$$

In questo senso i sistemi quantistici con stati legati sono “più indeformabili”: per modificarli è necessario dare un certo contributo minimo di energia, mentre nel caso classico basta qualsiasi contributo energetico, per quanto piccolo.

- Mentre in MC la conoscenza di \mathcal{E} **non** è massimale (**stato misto**), in MQ lo è (**stato puro**).

- La particella quantistica “ha dei punti in cui preferisce trovarsi”, e dei punti in cui proprio non si può trovare. Ciò è molto controintuitivo se la pensiamo come particella (“biglia”) classica, ma facile da comprendere in un’ottica ondulatoria: un’onda stazionaria presenta infatti dei nodi fissi in cui l’ampiezza dell’oscillazione è nulla.

Il momento in questo dominio compatto non è definito (come abbiamo visto nel primo esempio di sezione 3.5.3). Possiamo però calcolare la densità di probabilità della posizione:

$$w_{\varphi_n}^X(\lambda) = \int dx \delta(\lambda - x) |\varphi_n(x)|^2 = |\varphi_n(\lambda)|^2 = \frac{2}{a} \begin{cases} \cos^2\left(\frac{n\pi\lambda}{a}\right) & n \text{ dispari} \\ \sin^2\left(\frac{n\pi\lambda}{a}\right) & n \text{ pari} \end{cases}$$

Figura 3.3: Grafico delle autofunzioni $\varphi_n(x)$ per la buca infinita e delle relative densità di probabilità $w_{\varphi_n}^X(\lambda)$

Nota: se consideriamo un potenziale “smussato”, che non tende di scatto a ∞ ma “in maniera continua” con una salita ripida, classicamente ci aspettiamo che la particella “rallenti” su tali salite, e quindi ci trascorra più tempo - ci aspettiamo di trovarla più facilmente in quelle parti. In realtà, quantisticamente, quelle sono le posizioni in cui la probabilità per lo stato di energia minima è minore).

La probabilità $P_{\varphi_n}^X(\lambda)$ non è poi uniforme come nel caso classico, ma tende ad esso per $n \rightarrow \infty$, infatti, dalle formule di bisezione per sin e cos:

$$\left\{ \cos^2\left(\frac{n\pi x}{a}\right), \sin^2\left(\frac{n\pi x}{a}\right) \right\} = \frac{1}{2} \left(1 \pm \cos \frac{2n\pi x}{a} \right) \xrightarrow[n \rightarrow \infty]{\mathcal{S}'} + \frac{1}{2}$$

$$P_{\varphi_n}^X(\lambda) \xrightarrow[n \rightarrow \infty]{} \frac{2}{a} \frac{1}{2} = \frac{1}{a}$$

Principio di corrispondenza per la buca infinita

Dove si fa uso della convergenza debole⁵³ in \mathcal{S}' .

Proviamo ora ad esaminare l’evoluzione temporale del sistema partendo dallo stato iniziale $\psi(x, t_0)$ così definito:

Esempio (buca infinita)

$$\psi(x, t_0) = \cos \frac{\pi}{a} x + 2 \sin \frac{2\pi}{a} x \quad (3.58)$$

Per prima cosa dobbiamo riscrivere la $|\psi(t_0)\rangle$ nella base degli autostati di H , e quindi applicare l’evoluzione temporale:

$$|\psi(t_0)\rangle = \sum_n |\varphi_n\rangle \langle \varphi_n | \psi(t_0) \rangle$$

$$|\psi(t)\rangle = \sum_n \underbrace{\exp\left(-\frac{i}{\hbar} \mathcal{E}_n(t - t_0)\right)}_{U(t)} |\varphi_n\rangle \langle \varphi_n | \psi(t_0) \rangle \quad (3.59)$$

Riepilogando, il procedimento è questo:

⁵³ΛCFR pag. 24 di [1]

1. Normalizzare $\|\psi(t_0)\|^2 = 1$
2. Risolvere l'equazione agli autovalori per H , imponendo le condizioni al contorno date dal dominio di H , in modo da trovare le autofunzioni $\varphi_{\mathcal{E}_n}$ associate agli autovalori \mathcal{E}_n .
3. Scrivere $|\psi(t_0)\rangle$ e l'operatore di evoluzione $U(t)$ in termini di autovalori e autovettori di H , e poi applicare la formula (3.59).

In questo caso abbiamo già risolto l'equazione agli autovalori, trovando le φ_n associate agli autovalori \mathcal{E}_n . Scrivendo $|\psi(t_0)\rangle$ in tale base e normalizzando:

$$|\psi(t_0)\rangle = A\sqrt{\frac{a}{2}}(|\varphi_1\rangle + 2|\varphi_2\rangle)$$

$$\langle\psi(t_0)|\psi(t_0)\rangle = |A|^2\frac{a}{2}(1+4) \Rightarrow A = \sqrt{\frac{2}{5a}}$$

dove si è usato il fatto che le $|\varphi_i\rangle$ sono ortonormali.
Applicando allora l'evoluzione temporale:

$$|\psi(t_0)\rangle = \frac{1}{\sqrt{5}}(|\varphi_1\rangle + 2|\varphi_2\rangle)$$

$$\langle x|\psi(t)\rangle = \frac{1}{\sqrt{5}}\left(\exp\left(-\frac{i}{\hbar}\mathcal{E}_1(t-t_0)\right)\underbrace{\langle x|\varphi_1(t)\rangle}_{\varphi_1(x)} + 2\exp\left(-\frac{i}{\hbar}\mathcal{E}_2(t-t_0)\right)\underbrace{\langle x|\varphi_2\rangle}_{\varphi_2(x)}\right) \quad (3.60)$$

$$\mathcal{E}_i = \frac{\hbar^2}{2m}\left(\frac{n\pi}{a}\right)^2 \quad (3.61)$$

(Lezione 15 del
26/10/2018)

3.8.6 Calcolo di probabilità

Dato un'osservabile A e uno stato ψ , sia $P_\psi^A(\lambda)$ la famiglia spettrale di A , allora la probabilità di ottenere un risultato in un insieme $\Delta \subseteq \mathbb{R}$ misurando A nello stato ψ è data dalla formula generale:

$$P_\psi^A(\Delta) = \int_{\mathbb{R}} \chi_\Delta(\lambda) dP^A(\lambda) \quad (3.62)$$

Spettro discreto

Se lo spettro dell'operatore è unicamente discreto, espandendo la (3.15):

$$dP_\psi^A(\lambda)\Big|_{\sigma_P(A)} = \sum_{\lambda_n \in \sigma_P(A)} \delta(\lambda - \lambda_n) \sum_{r=1}^{d(\lambda_n)} \langle \psi | \lambda_n, r \rangle \langle \lambda_n, r | \psi \rangle d\lambda$$

dove gli autovettori generalizzati sono ortonormali:

$$\langle \lambda_n, r \mid \lambda_m, r' \rangle = \delta_{nm} \delta_{rr'}$$

Sostituendo tale risultato in (3.62) si ottiene, per la probabilità che una misura di A abbia un risultato pari a un elemento λ_n dello spettro $\sigma(A)$:

$$\begin{aligned} P_\psi^A(\{\lambda_n\}) &= \int_{\mathbb{R}} \chi_{\{\lambda_n\}}(\lambda) dP_\psi^A(\lambda) = \\ &= \int_{\mathbb{R}} d\lambda \chi_{\{\lambda_n\}} \sum_{\lambda_n \in \sigma_P(A)} \delta(\lambda - \lambda_m) \sum_{r=1}^{d(\lambda_m)} \langle \psi \mid \lambda_m, r \rangle \langle \lambda_m, r \mid \psi \rangle = \\ &\stackrel{(a)}{=} \sum_{r=1}^{d(\lambda_n)} |\langle \psi \mid \lambda_n, r \rangle|^2 = W_\psi^A(\lambda_n) \end{aligned} \quad (3.63)$$

Dove in (a) si è usata la definizione della funzione caratteristica:

$$\chi_{\{\lambda_n\}}(\lambda) = \begin{cases} 1 & \lambda = \lambda_n \\ 0 & \text{altrove} \end{cases}$$

per far “collassare” l’integrale.

$P_\psi^A(\{\lambda_n\})$ (denotata anche con $W_\psi^A(\lambda_n)$) è dunque la probabilità che una misura di A nello stato ψ restituisca λ_n .

Spettro continuo

Se lo spettro è continuo, la misura “non pesa i punti” (ossia i punti sono considerati insiemi Lebesgue-trascurabili) e quindi la probabilità di ottenere un certo (singolo) valore di A è nulla. L’unica cosa che possiamo fare è definire una densità di probabilità:

$$w_\psi^A(\lambda) = \frac{dP_\psi^A(\lambda)}{d\lambda} = \sum_{r=1}^{d(\lambda)} |\langle \psi \mid \lambda, r \rangle|^2 \quad (3.64)$$

dato che:

$$dP_\psi^A(\lambda) \Big|_{\sigma_C(A)} = \sum_{r=1}^{d(\lambda)} \langle \psi \mid \lambda, r \rangle \langle \lambda, r \mid \psi \rangle d\lambda$$

Stiamo qui considerando i λ normalizzati in modo che $\langle \lambda, r \mid \lambda', r' \rangle = \delta(\lambda - \lambda') \delta_{rr'}$.

Applicazione alla buca infinita

Applichiamo quanto appena trovato al caso della buca infinitamente profonda in una dimensione. In (3.60) avevamo ricavato l’evoluzione temporale del sistema a partire da uno stato iniziale, ottenendo:

$$|\psi(t)\rangle = \frac{1}{\sqrt{5}} \left(\exp\left(-i\frac{\mathcal{E}_1}{\hbar}(t-t_0)\right) |\varphi_1\rangle + 2 \exp\left(-i\frac{\mathcal{E}_2}{\hbar}(t-t_0)\right) |\varphi_2\rangle \right) \quad (3.65)$$

dove $|\varphi_1\rangle$ e $|\varphi_2\rangle$ sono autofunzioni dell'operatore hamiltoniana.

Proviamo allora a calcolare la probabilità di trovare \mathcal{E}_1 misurando H in $\psi(t)$.

Abbiamo mostrato che lo spettro dell'hamiltoniana nel caso della buca infinitamente profonda tra $[-\frac{a}{2}, \frac{a}{2}]$ è puramente discreto:

$$\sigma(H) = \sigma_P(H) = \{\mathcal{E}_n = \frac{\hbar^2}{2m} \left(\frac{n\pi}{a}\right)^2, n \in \mathbb{N}\}$$

Applicando allora la formula (3.63) otteniamo:

$$W_{\psi(t)}^H(\mathcal{E}_1) = |\langle \psi(t) | \varphi_1 \rangle|^2 \quad (3.66)$$

dove $|\varphi_1\rangle$ è l'autofunzione di H di autovalore \mathcal{E}_1 , ossia verifica, in notazione di Dirac:

$$H |\varphi_1\rangle = \mathcal{E}_1 |\varphi_1\rangle \Rightarrow |\varphi_1\rangle = |\mathcal{E}_1\rangle$$

Sostituendo allora (3.65) in (3.66) e svolgendo i conti:

$$\begin{aligned} W_{\psi(t)}^H(\mathcal{E}_1) &= \left| \langle \varphi_1 | \left(\frac{1}{\sqrt{5}} \exp\left(-\frac{i}{\hbar} \mathcal{E}_1(t-t_0)\right) |\varphi_1\rangle + 2 \langle \varphi_2 | \exp\left(-\frac{i}{\hbar} \mathcal{E}_2(t-t_0)\right) |\varphi_1\rangle \right) \right|^2 = \\ &= \left| \frac{1}{\sqrt{5}} \exp\left(-\frac{i}{\hbar} \mathcal{E}_1(t-t_0)\right) \right|^2 = \frac{1}{5} \end{aligned}$$

essendo le autofunzioni ortonormali: $\langle \varphi_1 | \varphi_1 \rangle = 1$ e $\langle \varphi_1 | \varphi_2 \rangle = 0$.

3.8.7 Note sull'equazione di Schrödinger (Esercizio 1)

Occupiamoci ora di esaminare cosa succederebbe se risolvessimo, inappropriatamente, l'equazione di Schrödinger per uno stato che non è nel dominio di H . Confronteremo quindi il risultato con la soluzione effettiva (come si era lasciato per esercizio alcune lezioni fa).

Utilizziamo come sistema su cui eseguire i conti la buca infinita tra $[-\frac{a}{2}, \frac{a}{2}]$, e consideriamo $\psi(x, t=0) = 1$, che non è nel dominio di H :

$$\psi(x, t=0) \notin D(H) = \left\{ \varphi \in L^2\left(\left[-\frac{a}{2}; \frac{a}{2}\right]\right), \varphi \text{ regolare}, \varphi\left(-\frac{a}{2}\right) = 0 = \varphi\left(\frac{a}{2}\right) \right\}$$

Normalizzando la ψ :

$$\psi(x, t=0) = \frac{1}{\sqrt{a}}$$

Risolvendo l'equazione di Schrödinger:

$$i\hbar \frac{\partial \psi}{\partial t}(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x, t)$$

otteniamo:

$$\psi(x, t) = \frac{1}{\sqrt{a}}$$

che è una soluzione sbagliata! (E in effetti se calcolassimo il valore dell'energia otterremmo un infinito⁵⁴).

La giusta risoluzione si ottiene decomponendo la $|\psi\rangle$ sulla base delle autofunzioni di H :

$$|\psi\rangle = \sum_n |\varphi_n\rangle \langle\varphi_n|\psi\rangle$$

dove lo stato iniziale, in rappresentazione in posizioni, è dato da:

$$\langle x|\psi\rangle = \psi(x, t=0) = \frac{1}{\sqrt{a}} \equiv \psi$$

Ricordando che le autofunzioni $\psi_n(x)$ sono:

$$\psi_n(x) = \begin{cases} \sqrt{\frac{2}{a}} \cos\left(\frac{n\pi x}{a}\right) & n \text{ dispari} \\ \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) & n \text{ pari} \end{cases}$$

possiamo calcolare le “proiezioni” $\langle\varphi_n|\psi\rangle$ sulla base ortonormale delle autofunzioni di H per n pari e dispari:

$$\langle\varphi_n|\psi\rangle = \sqrt{\frac{1}{a}} \sqrt{\frac{2}{a}} \int_{-\frac{a}{2}}^{\frac{a}{2}} \cos \frac{n\pi x}{a} dx = \frac{\sqrt{2}}{\sqrt{a}} \frac{1}{n\pi} \underbrace{\sin \frac{n\pi x}{a} \Big|_{-\frac{a}{2}}^{\frac{a}{2}}}_{=2} = \frac{2\sqrt{2}}{n\pi} \quad n \text{ dispari}$$

$$\langle\varphi_n|\psi\rangle = \frac{1}{\sqrt{a}} \sqrt{\frac{2}{a}} \int_{-\frac{a}{2}}^{\frac{a}{2}} \sin \frac{n\pi x}{a} dx = \frac{\sqrt{2}}{\sqrt{a}} \frac{1}{n\pi} \left(-\cos \frac{n\pi x}{a} \right) \Big|_{-\frac{a}{2}}^{\frac{a}{2}} = 0 \quad n \text{ pari}$$

Notiamo perciò che rimangono solamente le proiezioni per le n dispari.

L'evoluzione temporale sarà allora data da:

$$|\psi(t)\rangle = \sum_n \exp\left(-\frac{i}{\hbar} \mathcal{E}_n t\right) |\varphi_n\rangle \langle\varphi_n|\psi\rangle$$

che, in rappresentazione in posizioni, porta ad un risultato del tutto diverso da quello ottenuto applicando erroneamente Schrödinger:

$$\langle x|\psi(t)\rangle = \sum_{n \text{ dispari}} \exp\left(-\frac{i}{\hbar} \mathcal{E}_n t\right) \frac{2\sqrt{2}}{n\pi} \langle x|\varphi_n\rangle \neq \langle x|\psi(0)\rangle = \frac{1}{\sqrt{a}}$$

Perciò bisogna sempre “spacchettare” la funzione d'onda nelle autofunzioni di H e procedere in accordo.

⁵⁴Ma potremmo usare Schrödinger per calcolare la posizione, o il valor medio di un operatore limitato, basta che il suo dominio contenga ψ

3.8.8 Teoria della misura

Occupiamoci ora di definire cosa succede a seguito di una misurazione su un sistema.

Abbiamo visto che l'evoluzione temporale in un sistema isolato in MQ è deterministica (come in MC). Tuttavia, non appena vogliamo ricavare informazioni dobbiamo eseguire misure, e perciò non possiamo più considerare isolato il sistema (dobbiamo interagire con esso!) e, come discusso nel primo capitolo, un disturbo causato dalla misura **non** può essere ridotto arbitrariamente come in MC.

Ma allora come cambia uno stato a seguito di una misura?

Per rendere le cose semplici, consideriamo il caso idealizzato di una **misura istantanea**⁵⁵.

Per definizione, una **misura istantanea**⁵⁶ dell'osservabile A è detta di **prima specie** se è tale che, ripetendo immediatamente dopo la prima misura un'altra misura di A , otteniamo lo stesso risultato con certezza (probabilità 1). (Ciò è quello che succede sempre in MC)

Misure di prima specie

Ad esempio la misura di posizione di un elettrone con una lastra fotografica nell'esperimento delle due fenditure non è di prima specie, poiché dopo la misura l'elettrone è “distrutto” (dopo la misura “è sparito” - viene assorbito da un atomo - e perciò non ha senso chiedere di “rimisurarlo”, dato che non è più “l'elettrone di prima”).

Ogni misura che non è di prima specie è detta di **seconda specie** (e non ci occuperemo di esse).

Notiamo che in una misura di prima specie lo stato dopo la misura dipende sia dalla misura stessa che dal risultato, in completa opposizione con quanto succede in MC, in cui fare una misura è *irrilevante* per l'evoluzione del sistema, e il risultato è univoco e predeterminato (per uno stato puro). Qual è tale nuovo stato?

Nel caso in cui lo spettro di A sia discreto $\sigma(A) = \sigma_P(A) = \{\lambda_n\}$ e non degenerare, se una misura di prima specie di A dà il risultato λ_0 , dovendo una seconda misura dare lo stesso risultato con probabilità 1, si ha immediatamente che lo stato dopo la prima misura deve essere un autostato di A appartenente all'autovalore λ_0 , ossia, in notazione di Dirac, uno stato $|\lambda_0\rangle$.

In altri termini, la misura ha “distrutto” tutte le informazioni dello stato precedente, dato che la misura di λ_0 potrebbe essere stata originata da un insieme enormemente grande di stati.

Per questo il risultato di una misura viene chiamato “riduzione del pacchetto d'onda”:

$$|\psi\rangle = \sum_{\lambda_n \in \sigma(A)} c_n |\lambda_n\rangle \rightarrow c_0 |\lambda_0\rangle = |\lambda_0\rangle \langle \lambda_0 | \psi \rangle$$

Questo processo distrugge le informazioni sullo stato precedente. Matematicamente, se $|\varphi\rangle \neq |\psi\rangle$ sono stati iniziali diversi che portano alla misura λ_0 , gli stati finali saranno rispettivamente $c_0 |\lambda_0\rangle$ e $c'_0 |\lambda_0\rangle$, e una volta normalizzati si otterrà

⁵⁵^Faremo qui alcune assunzioni che sembrano “limite”, ma che in realtà funzionano sperimentalmente, e ci permettono di dare risultati significativi

⁵⁶^D'ora in poi, quando parleremo di misure intenderemo che sono istantanee

$c_0 = c'_0$, perdendo la relazione biunivoca che collega uno stato al suo evoluto. Infatti, una misura è:

Proprietà di una misura in MQ

- **Non unitaria:** $|\lambda_0\rangle \langle \lambda_0|$ è un operatore non unitario, e quindi *non invertibile*. In altre parole non è possibile ricalcolare uno stato passato a partire da uno stato presente, ottenuto a seguito di una misura.
- **Indeterministica:** L'esito di una misura è scelto *casualmente* tra i possibili autovalori λ_n ammessi da tale osservabile
- **Completamente diversa dall'evoluzione di un sistema isolato.** In effetti un sistema isolato si evolve in maniera deterministica⁵⁷, e l'operatore evoluzione è unitario (e quindi invertibile - si può ricalcolare il passato dal presente):

$$|\psi(t)\rangle = \exp\left(-\frac{i}{\hbar}Ht\right)|\psi\rangle$$

In altre parole, la misura modifica in modo irreversibile e drastico l'evoluzione del sistema fisico, che se fosse rimasto isolato sarebbe completamente diversa.

Quindi se “osserviamo” ossia eseguiamo delle misure (di prima specie) nel sistema per estrarre informazione ne “distruggiamo” l'evoluzione in modo drastico, e poiché:

$$A|\lambda_0\rangle = \lambda_0|\lambda_0\rangle$$

A su $|\lambda_0\rangle$ “agisce” come un numero, e quindi A si “comporta” come un valore. In questo modo recuperiamo il concetto di “valore di una osservabile” nel processo di misura. *Possiamo pensare alla descrizione di un'osservabile come una “matrice infinito-dimensionale”. I valori dell'osservabile non sono i “valori della matrice” (non ha senso), ma i valori dello spettro della matrice - ossia quei particolari numeri che rimangono invariati a seguito di particolari “rotazioni” geometriche. Immediatamente dopo una misura, quella particolare osservabile si comporterà classicamente “come il valore trovato”.*

Queste considerazioni si generalizzano allo spettro discreto degenerare come segue: supponiamo che λ_0 abbia degenerazione $d(\lambda_0) > 1$, allora vi sono $d(\lambda_0)$ autovettori ortonormali: $\{|\lambda_0, r\rangle, r = 1, \dots, d(\lambda_0)\}$.

Eseguito una misura di prima specie di A , che dia come risultato λ_0 , possiamo concludere che immediatamente dopo la misura lo stato è nel sottospazio generato da $\{|\lambda_0, r\rangle, r = 1, \dots, d(\lambda_0)\}$ e non possiamo sapere di più.

Dato che vorremmo uno stato univocamente definito, introduciamo un'ulteriore idealizzazione.

Definiamo allora una **misura ideale di prima specie** una misura che “disturba”

Misura ideale

⁵⁷ Si potrebbe pensare di *estendere* il sistema a comprendere anche l'osservatore, ottenendo per ciò un nuovo sistema che stavolta è isolato, e che quindi si evolve in maniera deterministica. Ma come è possibile ciò, se all'interno di esso atti di misura sconvolgono irrimediabilmente l'evoluzione delle sue componenti? Le due descrizioni non sembrano essere per nulla compatibili! In effetti, questa situazione rappresenta parte del **problema della misura**, forse l'ultimo grande problema aperto in MQ. Si tratta comunque di una questione che ha poca importanza sperimentale: nonostante questa ambiguità di fondo, la MQ produce sempre, in ogni circostanza, risultati confrontabili con gli esperimenti, e che finora non sono mai stati confutati.

il meno possibile lo stato iniziale compatibilmente con il risultato trovato.

Per definirla matematicamente facciamo ricorso alla nozione di distanza nello spazio \mathcal{S} degli stati. Dati due stati normalizzati $|\psi\rangle$ e $|\phi\rangle$, con $\langle\psi|\psi\rangle = 1 = \langle\phi|\phi\rangle$, la distanza tra di essi è data da:

$$d(|\psi\rangle, |\phi\rangle) = (1 - |\langle\psi|\phi\rangle|^2)^{1/2}$$

Per definizione una misura di prima specie è **ideale** se la distanza tra lo stato $|\psi_i\rangle$ immediatamente prima della misura e lo stato $|\psi_f\rangle$ immediatamente dopo la misura è la **minima** compatibilmente con il risultato trovato.

Partendo allora da $|\psi\rangle_i$ ed espandendo nella base degli autovettori (generalizzati) di A :

$$|\psi_i\rangle = \sum_{\lambda_n \in \sigma(A)} \sum_{r=1}^{d(\lambda_n)} \underbrace{c(\lambda_n, r)}_{\langle\lambda_n, r|\psi\rangle} |\lambda_n, r\rangle \quad t = 0^-$$

ed eseguendo a $t = 0$ una misura ideale di prima specie, da cui si ottiene come risultato λ_0 , lo stato immediatamente successivo $|\psi_f\rangle$ a $t = 0^+$ è dato da:

$$|\psi_f\rangle = \sum_{r=1}^{d(\lambda_0)} c(\lambda_0, r) |\lambda_0, r\rangle = \sum_{r=1}^{d(\lambda_0)} |\lambda_0, r\rangle \langle\lambda_0, r|\psi_1\rangle$$

Questa definizione corrisponde al concetto di “disturbo minimo”. Vediamolo matematicamente.

Sia \mathcal{M}_{λ_0} il sottospazio generato dagli autovettori di autovalore λ_0 : $\{|\lambda_0, r\rangle, r = 1, \dots, d(\lambda_0)\}$.

Decomponiamo lo stato iniziale nelle sue componenti parallela e perpendicolare a \mathcal{M}_{λ_0} :

$$|\psi_i\rangle = \underbrace{|\psi_{\parallel}\rangle}_{\in \mathcal{M}_{\lambda_0}} + \underbrace{|\psi_{\perp}\rangle}_{\perp \mathcal{M}_{\lambda_0}} \quad (3.67)$$

e tale decomposizione è univoca in \mathcal{H} .

Lo stato finale sarà dato dal vettore $\tilde{\psi}_{\parallel} \in \mathcal{M}_{\lambda_0}$ (compatibile con λ_0) che si trova a distanza minima da $|\psi_i\rangle$. Per semplicità consideriamo direttamente il quadrato della distanza:

$$(1 - |\langle\psi_i|\tilde{\psi}_{\parallel}\rangle|^2) \stackrel{(3.67)}{=} (1 - |\langle\psi_{\parallel}|\tilde{\psi}_{\parallel}\rangle + \langle\psi_{\perp}|\tilde{\psi}_{\parallel}\rangle|^2) \stackrel{(a)}{=} (1 - |\langle\psi_{\parallel}|\tilde{\psi}_{\parallel}\rangle|^2)$$

dove in (a) si è usata l'ortogonalità tra $|\psi_{\perp}\rangle$ e $|\tilde{\psi}_{\parallel}\rangle$.

Perciò la distanza (o il suo quadrato) è minimizzata dalla $|\tilde{\psi}_{\parallel}\rangle$ che massimizza il termine evidenziato. Applicandovi Schwartz:

$$|\langle\psi_{\parallel}|\tilde{\psi}_{\parallel}\rangle|^2 \leq \|\psi_{\parallel}\|^2 \|\tilde{\psi}_{\parallel}\|^2$$

La distanza è perciò minimizzata quando si ha l'uguaglianza, ossia solo se $|\tilde{\psi}_{\parallel}\rangle$ e $|\psi_{\parallel}\rangle$ sono paralleli:

$$\tilde{\psi}_{\parallel} = \alpha \psi_{\parallel} = \frac{\psi_{\parallel}}{\|\psi_{\parallel}\|}$$

Nota: Se $|\psi_i\rangle$ è normalizzato, in generale $|\psi_{\parallel}\rangle$ e $|\psi_{\perp}\rangle$ non lo sono, ed ecco perché serve dividere per la norma nell'equazione di sopra, dovendo essere lo stato finale $|\tilde{\psi}_{\parallel}\rangle$ normalizzato: $\langle\tilde{\psi}_{\parallel}|\tilde{\psi}_{\parallel}\rangle = 1$.

In altre parole, se lo stato finale $|\psi_f\rangle$ corrisponde a quello *a distanza minima* $|\tilde{\psi}_{\parallel}\rangle$ compatibilmente con l'esito della misura λ_0 , allora si *preserva* la componente dello stato iniziale $|\psi_i\rangle$ che è parallela all'autospazio associato a λ_0 , che è il massimo che si possa fare (tutta la componente perpendicolare si perde per forza, altrimenti misure successive non darebbero con certezza λ_0). Questo è in linea con la definizione da cui eravamo partiti, per cui una misura ideale è quella che “disturba” meno il sistema.

Notiamo infine che $\sum_{r=1}^{d(\lambda_0)} |\lambda_0, r\rangle \langle\lambda_0, r|$ è un proiettore. In effetti, potremmo vedere l'azione di una misura ideale di prima specie come la proiezione dello stato iniziale $|\psi_i\rangle$ su \mathcal{M}_{λ_0} . In termini della famiglia spettrale di A vale:

$$\sum_{r=1}^{d(\lambda_0)} |\lambda_0, r\rangle \langle\lambda_0, r| = \int \chi_{\{\lambda_0\}}(\lambda) dP^A(\lambda) \quad (3.68)$$

e questo suggerisce il **postulato di proiezione di von Neumann**:

*Postulato di
proiezione di von
Neumann*

Assioma 5. *Se una misura ideale di prima specie della osservabile A nello stato $|\psi\rangle$ dà un risultato nell'insieme (misurabile) Δ , immediatamente dopo la misura lo stato del sistema è descritto dallo stato calcolato generalizzando (3.68) ad uno spettro qualsiasi:*

$$\begin{aligned} P^A(\Delta) |\psi\rangle &\equiv \int \chi_{\Delta}(\lambda) dP^A(\lambda) |\psi\rangle = \int_{\Delta \cap \sigma_C(A)} d\lambda \sum_{r=1}^{d(\lambda)} |\lambda, r\rangle \langle\lambda, r| \psi\rangle + \\ &+ \sum_{\{\sigma_P(A) \ni \lambda_n\} \cap \Delta} \sum_{r=1}^{d(\lambda_n)} |\lambda_n, r\rangle \langle\lambda_n, r| \psi\rangle \end{aligned}$$

In particolare lo stato finale normalizzato è:

$$\frac{P^A(\Delta) |\psi\rangle}{\sqrt{\langle\psi|P^A(\Delta)|\psi\rangle}}$$

Poiché $P^A(\Delta)$ è un proiettore, e quindi possiamo scriverne la norma al quadrato come:

$$P^A(\Delta)^\dagger P^A(\Delta) = (P^A(\Delta))^2 = P^A(\Delta)$$

Vediamo ora, graficamente, la differenza tra una misura ideale e una non ideale. Supponiamo che, prima della misura, lo stato del sistema sia descritto da una funzione $\langle x | \psi_i \rangle = \psi(x, 0^-)$, rappresentata in figura 3.4a.

A $t = 0$ viene eseguita una misura (di prima specie) di posizione, e la particella viene trovata nell'intervallo $\Delta \subset \mathbb{R}$.

Poiché tutte le misure di posizione seguenti dovranno dare lo stesso risultato,

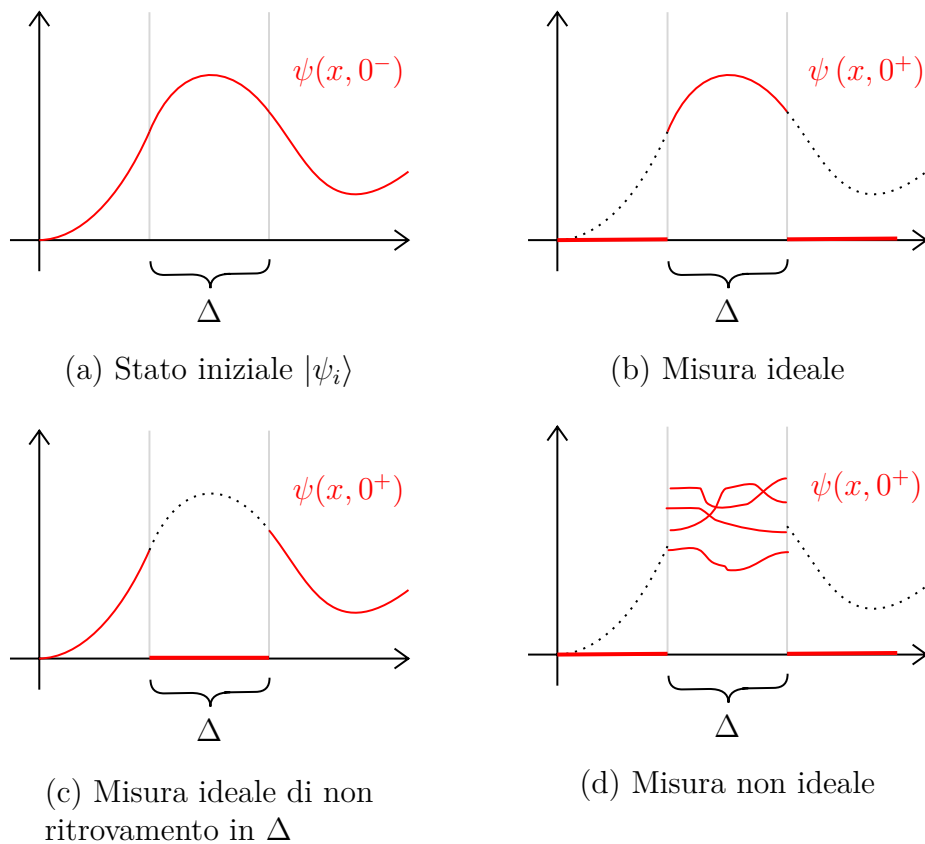


Figura 3.4: Differenti risultati di una misura.

$\langle x | \psi_i(0^+) \rangle$ è nulla su tutto \mathbb{R} , tranne che in Δ : sappiamo con certezza che la particella deve trovarsi in Δ .

Se la misura non è stata ideale non sappiamo nient'altro: i valori assunti in Δ potrebbero essere *qualsiasi*. Se invece la misura è stata ideale sappiamo di non aver “deturpato eccessivamente” la situazione originaria, e perciò la nuova funzione d'onda in Δ ha lo stesso andamento che aveva prima della misura, come rappresentato in figura 3.4b e 3.4d.

Cosa sarebbe successo se invece la misura **non** avesse trovato la particella in Δ ? Anche in questo caso, lo stato del sistema sarebbe cambiato drasticamente. Anche facendo una misura ideale, infatti, la nuova situazione è quella illustrata in figura 3.4c.

Classicamente ciò è assurdo: non avendo trovato la particella, effettivamente non vi è stata alcuna *interazione* tra rivelatore e particella, e quindi ci aspettiamo di non aver modificato nulla.

In MQ, invece, basta semplicemente *avere un'informazione in più* - che ci sia o non ci sia una particella in un certo range - per modificare l'intera funzione d'onda.

Questa è parte della caratteristica non località della MQ: a distanze arbitrariamente grandi la funzione d'onda risente di una misura fatta in un dato range!

(Lezione 16 del
5/11/2018)

3.8.9 Differenze tra MC e MQ

(Lezione 16 del
5/11/2018)

Differenze
qualitative

Nonostante l'apparenza, la proiezione di von Neumann è ben diversa dalla “proiezione” di uno stato misto classico generata da una misura che riduce l'indeterminazione di un'osservabile, aumentando l'informazione disponibile allo sperimentatore. Per esempio, consideriamo un sistema “classico”, in cui una *formica* è lasciata libera di spostarsi all'interno di una stanza. Prima di un'osservazione non sappiamo esattamente in che punto si trovi - possiamo al più costruire una *distribuzione di probabilità* per la sua posizione. Per esempio, se la formica è partita al centro della stanza, dato che si muove lentamente, dopo un minuto sarà probabilmente ancora nei pressi del centro.

Se ora entriamo nella stanza e osserviamo una certa regione A (compiendo così una misura) avremo una certa probabilità di trovare o meno la formica in A .

Se la troviamo lì, immediatamente la distribuzione di probabilità “collapsa” su A - dato che siamo certi che la posizione della formica a quell'istante è in A .

Allo stesso modo, se non avessimo trovato la formica in A , immediatamente la densità di probabilità in A si annullerebbe, poiché sapremmo con certezza che la formica non è in A .

Tutto ciò sembra ovvio e banale, e non crea problemi poiché abbiamo fatto uso di una distribuzione di probabilità in primo luogo a seguito di una nostra ignoranza - si tratta di un *artificio* matematico che non ha alcun significato fisico: il sistema si evolve *classicamente* indipendentemente dalle misure che facciamo, e in particolare indipendentemente da come *evolve* a seguito delle misure la distribuzione di probabilità. In effetti, il sistema prima della misura è in uno **stato misto**: abbiamo tante possibilità per la posizione della formica, ciascuna *pesata* dalla sua probabilità. Dopo la misura, se sappiamo la posizione della formica, l'informazione sulla formica si è accresciuta. Nel caso limite in cui conosciamo sia posizione che momento della formica, lo stato del sistema sarà addirittura puro: l'unica differenza sta perciò nello sperimentatore, che dopo la misura ha più informazioni sul sistema.

In MQ, invece, l'evoluzione del sistema comporta l'evoluzione della sua distribuzione di probabilità - che emerge anche per **stati puri**. In effetti, se consideriamo l'analogo problema di una particella a energia fissata in una buca unidimensionale infinita, abbiamo a che fare con un sistema descritto da uno stato puro, e dopo la misura (ideale di prima specie) si ottiene ancora uno stato puro, diverso dal primo. Non abbiamo quindi *acquisito informazione e basta*, ma *diminuito* l'incertezza su alcune osservabili a discapito di altre, mantenendo comunque la stessa “quantità” (massimale) di informazione. Il sistema dopo la misura, tuttavia, si comporterà in modo diverso da quello prima: l'atto di misura *ha perturbato* l'evoluzione del sistema.

Rendiamo più concreto (matematicamente) tale esempio.

Consideriamo *classicamente* una particella libera di muoversi lungo una dimensione, di cui supponiamo conoscere l'impulso p_0 classico, e di sapere che la sua posizione ha una distribuzione di probabilità $\rho(x)$.

A $t = 0$ eseguiamo una “misura ideale di prima specie” nel senso sopra descritto,

Esempio
numerico: misura
in MC

ma a livello classico, volta a verificare se la particella è in Δ .

Se la misura trova la particella in Δ , lo stato del sistema appena dopo (a $t = 0^+$) sarà dato dalla *proiezione*:

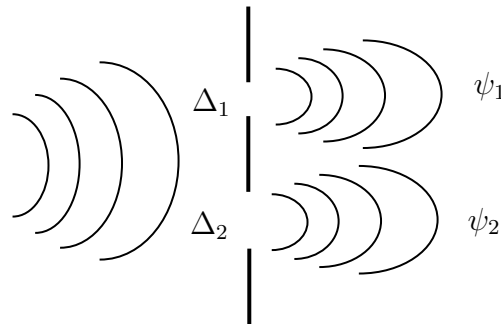
$$f(x, p)_{t=0} = \rho(x) \delta(p - p_0) \rightarrow f(x, p)_{t=0^+} = \frac{\rho(x) \chi_{\Delta}(x) \delta(p - p_0)}{\int_{\Delta} \rho(x) dx}$$

per cui a $t = 0^+$ la distribuzione di probabilità è annullata al di fuori di Δ dalla funzione caratteristica χ_{Δ} , mentre per $x \in \Delta$ ha lo stesso andamento di prima (a meno del fattore di normalizzazione dato dal denominatore).

Notiamo allora che l'indeterminazione in x è diminuita **senza modificare** quella in p , cioè e la “fisica del sistema non è cambiata” (ossia il suo comportamento) con l'acquisizione di informazione.

Nel caso quantistico, invece, sia $|\psi\rangle$ che la sua proiezione di von Neumann $P^X(\Delta) |\psi\rangle$ sono entrambi **stati puri**, e quindi di conoscenza **massimale**, pertanto il miglioramento della conoscenza di X deve comportare il peggioramento della conoscenza di altre osservabili (sicuramente $|\psi\rangle \langle \psi|$, dato che non sappiamo più se la particella sia nello stato $|\psi\rangle$ - e anzi, nel caso di una misura ideale proprio non ci sta - e probabilmente anche di p per il principio di indeterminazione). Diversamente dal caso classico abbiamo perciò modificato la “fisica” (il comportamento) del sistema.

*Esempio
numerico: misura
in MQ*



Ad esempio, nel caso dell'**esperimento delle due fenditure**, se spariamo un elettrone alla volta contro lo schermo, e non lo osserviamo tra le fenditure e lo schermo, lo stato del sistema sarà dato da:

$$|\psi\rangle_{\text{oss}} = \frac{1}{\sqrt{2}}(|\psi_1\rangle + |\psi_2\rangle)$$

dove $|\psi\rangle_1$ è lo stato costituito dall'onda che passa dalla prima fenditura, e $|\psi\rangle_2$ quello dell'onda che passa dalla seconda. Lo stato finale del sistema sarà quindi una *combinazione* dei due.

Piazziamo ora un rivelatore immediatamente dopo la prima fenditura. Se osserviamo (con una misura ideale di prima specie) il passaggio da Δ_1 , il nuovo stato dovrà essere compatibile con tale osservazione, e sarà quindi dato dalla proiezione di von Neumann:

$$P^X(\Delta_1) |\psi\rangle \approx |\psi_1\rangle$$

*Esperimento delle
due fenditure e
misure
quantistiche*

In quanto, se il rilevatore è molto prossimo alla fenditura 1, è praticamente impossibile che il passaggio dalla 2 sia *compatibile* con tale osservazione.

Del resto, se invece non troviamo la particella in 1, il nuovo stato sarà:

$$P^X(\Delta_2) |\psi\rangle \approx |\psi_2\rangle$$

In altre parole, l'osservazione *annulla totalmente* l'azione di una delle due funzioni d'onda.

Ripetendo la misura molte volte ciascuna delle due possibilità si presenterà in circa metà dei casi, ossia a $|\psi\rangle_1$ e $|\psi\rangle_2$ è associata una probabilità di $1/2$. Ciò significa che lo stato ρ_{oss} è **misto**⁵⁸ e nettamente diverso da quello ρ_{oss} che avrebbe assunto il sistema se non fosse stato osservato:

$$\rho_{\text{oss}} = \frac{1}{2} |\psi_1\rangle \langle \psi_1| + \frac{1}{2} |\psi_2\rangle \langle \psi_2| \neq \frac{1}{2} (|\psi_1\rangle + |\psi_2\rangle)(\langle \psi_1| + \langle \psi_2|) = |\psi\rangle \langle \psi| = \rho_{\text{oss}}$$

In particolare, dopo la misura cambia la probabilità $\rho(x)$ di trovare l'elettrone in una certa posizione x :

$$\rho(x)_{\text{oss}} = \frac{1}{2} \langle x | \psi_1 \rangle \langle \psi_1 | x \rangle + \frac{1}{2} \langle x | \psi_2 \rangle \langle \psi_2 | x \rangle = \frac{1}{2} |\psi_1(x)|^2 + \frac{1}{2} |\psi_2(x)|^2$$

dove non vi è il termine di interferenza, che invece compariva facendo il quadrato della funzione d'onda nel caso senza osservazione:

$$\begin{aligned} \psi(x) &= \langle x | \psi \rangle = \frac{1}{\sqrt{2}} \langle x | \psi_1 \rangle + \langle x | \psi_2 \rangle = \frac{1}{\sqrt{2}} (\psi_1(x) + \psi_2(x)) \\ \rho(x)_{\text{oss}} &= \frac{1}{2} \psi_1(x)^2 + \frac{1}{2} \psi_2(x)^2 + \psi_1(x) \cdot \psi_2(x) \neq \rho(x)_{\text{oss}} \end{aligned}$$

D'altro canto, se esaminiamo il sistema facendo evolvere le osservabili invece che gli stati, si ha che un'evoluzione unitaria non modifica lo spettro di un operatore autoaggiunto, ma solo le probabilità che ciascun autovalore - per esempio una qualsiasi posizione - venga ottenuto da una misura. Così facendo, la particella (se non osservata) “esplora” tutto lo spettro, e passa da “entrambe” le fenditure. Quando si fa una misura l'evoluzione non è più unitaria, e lo spettro viene modificato: in particolare, tutti i valori dello spettro che non sono compatibili con la misura ottenuta non sono più “fisici”, e vengono “terminati”. Ci occuperemo di esaminare questa visuale di Heisenberg nel prossimo esempio.

3.9 Esempi di evoluzione di sistemi quantistici isolati

3.9.1 Particella libera in 1D

Occupiamoci ora di esaminare l'evoluzione nel tempo delle medie e delle fluttuazioni degli operatori posizione X e momento P applicati ad una particella libera

⁵⁸In particolare è *non deterministico* e di informazione *non massimale*. L'atto di misura fa perdere informazione sul sistema

di muoversi in una dimensione.

L'hamiltoniana del sistema è data da:

$$H = \frac{P^2}{2m}; \quad X^H(t) = U^\dagger(t) X U(t)$$

Adottiamo, per questo esempio, la visuale di Heisenberg (evoluzione delle osservabili, stati costanti), per cui la media di un operatore (per esempio X) evolve come:

$$\langle X \rangle_{\psi(t)} = \langle X^H(t) \rangle_\psi$$

dove $X^H(t)$ obbedisce all'equazione di Born-Heisenberg-Jordan⁵⁹:

$$\frac{dX^H(t)}{dt} = \frac{[X^H(t), H]}{i\hbar} = \frac{U^\dagger(t) X \textcolor{yellow}{U(t) H} - \textcolor{blue}{H U^\dagger(t)} X U(t)}{i\hbar} \quad (3.69)$$

Dato che $U(t)$ è una funzione di H , infatti dalla definizione:

$$U(t) = \exp\left(-\frac{i}{\hbar}tH\right)$$

si ha che $U(t)$ commuta⁶⁰ con H , ossia $H U(t) = U(t) H$, e allora possiamo riarrangiare le U e “portarle fuori” da (3.69):

$$\frac{dX^H(t)}{dt} = \frac{U^\dagger(t) X \textcolor{yellow}{H U(t)} - \textcolor{blue}{U^\dagger(t) H} X U(t)}{i\hbar} = U^\dagger(t) \frac{[X, H]}{i\hbar} U(t) \quad (3.70)$$

In particolare:

$$[X, H] = \left[x, \frac{P^2}{2m}\right] = \frac{1}{2m}[X, P \cdot P] \underset{(a)}{=} \frac{1}{2m}([X, P]P + P[X, P]) = \frac{i\hbar}{m}P \quad (3.71)$$

dove in (a) abbiamo usato la formula per il commutatore:

$$[A, BC] = [A, B]C + B[A, C] = ABC - BAC + BAC - BCA$$

Sostituendo (3.71) in (3.70):

$$\frac{dX^H(t)}{dt} = U^\dagger(t) \frac{i\hbar P}{m i\hbar} U(t) = \frac{P^H(t)}{m} \quad (3.72)$$

Come ci aspettiamo, la *velocità* è data dal momento diviso la massa.

Esaminiamo quindi l'evoluzione del momento:

$$\frac{dP^H(t)}{dt} = \frac{[P^H(t), H]}{i\hbar} = U^\dagger(t) \left[P, \frac{P^2}{2m}\right] U(t) = 0$$

*Formula
“classica” per la
velocità di una
particella libera*

⁵⁹^Si noti che la H all'apice della $X^H(t)$ sta per *Heisenberg*, mentre l'altra H che compare nella formula indica l'hamiltoniana del sistema

⁶⁰^Se non commutasse non sarebbe possibile calcolarlo, dato che farlo modificherebbe H , e di conseguenza anche $U(t)$ che dipende da H , procedendo così in un loop infinito.

dato che P commuta con P^2 .

Ma allora $P^H(t) = P$ è costante (rimane lo stesso), e quindi, sostituendo in (3.72) e integrando tra gli istanti 0 e t :

$$\frac{dX^H(t)}{dt} = \frac{P}{m} \Rightarrow X^H(t) = x_0 + t \frac{p}{m}$$

dove x_0 è l'operatore posizione al tempo iniziale $t = 0$. Passando ai valor medi si ritrova la *legge classica*:

$$\langle x^H(t) \rangle_\psi = \langle x_0 \rangle_\psi + t \frac{\langle p \rangle_\psi}{m}$$

Calcolando invece la fluttuazione di X :

$$\begin{aligned} (\Delta X)_{\psi(t)}^2 &\stackrel{(a)}{=} \langle X^2 \rangle_{\psi(t)} - \langle X \rangle_{\psi(t)}^2 = \langle x^H(t)^2 \rangle_\psi - \langle x^H(t) \rangle_\psi^2 = \\ &= \left\langle \left(x_0 + \frac{p}{m} t \right)^2 \right\rangle_\psi - \left\langle x_0 + \frac{p}{m} t \right\rangle_\psi^2 = \\ &= \left(\langle x_0^2 \rangle_\psi + \frac{t}{m} \langle x_0 p + p x_0 \rangle_\psi + \frac{\langle p^2 \rangle_\psi}{m^2} t^2 \right) - \left(\langle x_0 \rangle_\psi^2 + \frac{\langle p \rangle_\psi^2}{m^2} t^2 + \frac{2t}{m} \langle x_0 \rangle_\psi \langle p \rangle_\psi \right) = \\ &\stackrel{(b)}{=} \langle \Delta X \rangle_\psi^2 + \frac{1}{m^2} \langle \Delta P \rangle_\psi^2 t^2 + \frac{t}{m} \left[\langle X P + P X \rangle_\psi - 2 \langle X \rangle_\psi \langle P \rangle_\psi \right] \end{aligned} \quad (3.73)$$

dove si è sfruttata la linearità delle medie per “estrarre” le costanti, e in (a)-(b) si è usata la formula⁶¹:

$$\langle \Delta A \rangle_\psi^2 = \langle (A - \langle A \rangle_\psi)^2 \rangle_\psi = \langle A^2 \rangle_\psi - \langle A \rangle_\psi^2$$

Se richiediamo che all'inizio la fluttuazione sia minima, ossia:

$$\begin{aligned} \frac{d}{dt} (\Delta X)_{\psi(t)}^2 \Big|_{t=0} &\stackrel{!}{=} 0 \Rightarrow \frac{2t}{m^2} \langle \Delta P \rangle_\psi^2 + \frac{1}{m} [\langle X P + P X \rangle_\psi - 2 \langle X \rangle_\psi \langle P \rangle_\psi] \Big|_{t=0} = \\ &= \frac{1}{m} [\langle X P + P X \rangle_\psi - 2 \langle X \rangle_\psi \langle P \rangle_\psi] = 0 \quad (3.74) \\ \Rightarrow (\Delta X)_{\psi(t)}^2 &\stackrel{(3.73)}{=} \langle \Delta X \rangle_\psi^2 + \frac{t^2}{m^2} \langle \Delta P \rangle_\psi^2 \stackrel{(3.74)}{\geq} \langle \Delta X \rangle_\psi^2 + \frac{t^2 \hbar^2}{4m^2} \frac{1}{\langle \Delta X \rangle_\psi^2} \end{aligned}$$

Dove in (b) si è applicato il principio di indeterminazione:

$$(\Delta P)_\psi \geq \frac{\hbar}{2} \frac{1}{(\Delta X)_\psi}$$

Notiamo allora che riducendo l'incertezza su X aumenta quella su P . Infatti, se $(\Delta X)_\psi$ decresce, allora $(\Delta P)^2(t)$ cresce rapidamente come:

$$\frac{\hbar^2 t^2}{4m^2} \frac{1}{(\Delta x)_\psi^2}$$

⁶¹ΛGià vista in Statistica (Sperimentazioni 1) per la Varianza

3.9.2 Buca infinitamente profonda in \mathbb{R}^3

Consideriamo una particella nello spazio tridimensionale, racchiusa tra pareti impenetrabili da un potenziale:

$$V_i(x_i) = \begin{cases} 0 & 0 < x_i < a_i \\ \infty & \text{altrove} \end{cases} \quad i = 1, 2, 3$$

L'hamiltoniana è quindi data da:

$$H = \frac{p^2}{2m} + V_1(x) + V_2(y) + V_3(z)$$

Con dominio:

$$D(H) = \{\psi(x, y, z) \text{ regolari}, \psi(0, y, z) = 0 = \psi(a_1, y, z), \\ \psi(x, 0, z) = 0 = \psi(x, a_2, z), \\ \psi(x, y, 0) = 0 = \psi(x, y, a_3)\}$$

Cerchiamo lo spettro $\sigma(H)$ e le autofunzioni $\psi_{\mathcal{E}}(x, y, z)$ di H .

Scriviamo l'equazione di Schrödinger stazionaria⁶²:

$$\left[-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + V_1(x) + V_2(y) + V_3(z) \right] \psi_{\mathcal{E}}(x, y, z) = \mathcal{E} \psi_{\mathcal{E}}(x, y, z) \quad (3.75)$$

Notiamo che le dipendenze da x, y, z nell'operatore H sono ben separate. Risolviamo allora separando le variabili, cioè cercando le soluzioni nella forma:

$$\psi_{\mathcal{E}}(x, y, z) = \psi_1(x) \psi_2(y) \psi_3(z)$$

che inserita in (3.75) porta a:

$$\frac{\psi_1''}{\psi_1}(x) + \frac{\psi_2''}{\psi_2}(y) + \frac{\psi_3''}{\psi_3}(z) + \frac{2m}{\hbar^2} (\mathcal{E} - V_1(x) - V_2(y) - V_3(z)) = 0$$

Separatamente ognuno dei tre termini evidenziati deve essere costante⁶³:

$$\frac{\psi_i''}{\psi_i}(x_i) + \frac{2m}{\hbar^2} (\mathcal{E}_i - V_i(x_i)) = 0 \quad i = 1, 2, 3 \quad (3.76)$$

dove i \mathcal{E}_i sono costanti, con $\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 = \mathcal{E}$.

Notiamo allora che ciascuna delle (3.76) è l'equazione di Schrödinger stazionaria

⁶²^Possiamo farlo poiché, avendo specificato il dominio di H non rischiamo di applicarlo ad una ψ al di fuori di esso

⁶³^Infatti si può esplicitare uno in funzione degli altri 2. Isolando per esempio i termini con la x , si ha che devono essere pari ad una espressione che non dipende dalla x , e che assume un valore costante una volta fissate y e z . Tale ragionamento si ripete analogamente per le altre due variabili.

per la buca infinitamente profonda in 1D, la cui soluzione (per esempio per la x) è:

$$\psi_1(x) = A \sin(k_1 x + \delta); \quad k_i = \frac{1}{\hbar} \sqrt{2m\mathcal{E}_i} \quad (3.77)$$

A cui dobbiamo imporre la condizione iniziale: $\psi_i(0) = 0 = \psi_i(a_i)$.

Si ha allora:

$$\begin{aligned} \psi_1(0) &\stackrel{!}{=} 0 \Rightarrow A \sin \delta \Rightarrow \delta = 0 \\ \psi_1(a_1) &\stackrel{!}{=} 0 \Rightarrow A \sin(k_1 a_1) = 0 \Rightarrow k_1 a_1 = n_1 \pi; \quad n_1 \in \mathbb{N} \end{aligned} \quad (3.78)$$

Normalizzando:

$$\begin{aligned} \int_0^{a_1} dx |A|^2 \sin^2 k_1 x &\stackrel{!}{=} 1 \\ A^2 \int_0^{a_1} \left(\frac{1}{2} - \frac{\cos(2k_1 x)}{2} \right) dx &= A^2 \left[\frac{1}{2}x - \frac{1}{4k_1} \sin(2k_1 x) \right]_0^{a_1} = \\ &= A^2 \left(\frac{1}{2}a_1 + \underbrace{\sin(2n_1 \pi)}_{=0} - \sin 0 \right) \\ &= \frac{a_1 A^2}{2} = 1 \Rightarrow A = \sqrt{\frac{2}{a_1}} \end{aligned}$$

Ricavando k_i da (3.78) e sostituendo nella definizione di k_i (3.77) troviamo:

$$k_i = \frac{n_i \pi}{a_i} = \frac{1}{\hbar} \sqrt{2m\mathcal{E}_i}$$

Ricavando le energie:

$$\mathcal{E}_i = \frac{n_i^2}{a_i^2} \frac{\hbar^2}{2m} \pi^2$$

E allora, dalla condizione di energia del sistema, per cui $\mathcal{E}_{n_1, n_2, n_3} = \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3$, otteniamo lo spettro di H , che è unicamente discreto:

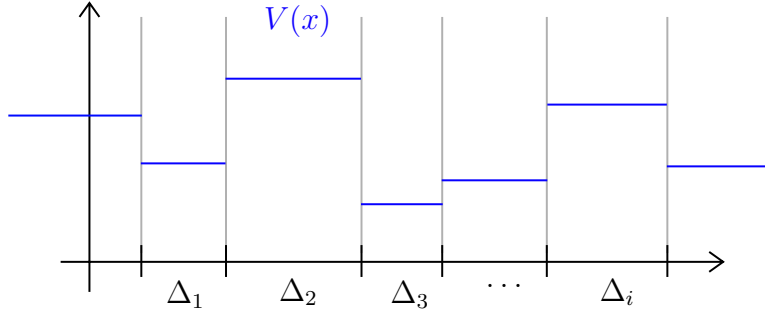
$$\sigma(H) = \left\{ \frac{\hbar^2}{2m} \pi^2 \left(\frac{n_1^2}{a_1^2} + \frac{n_2^2}{a_2^2} + \frac{n_3^2}{a_3^2} \right) \right\}; \quad n_i \in \mathbb{N}$$

ossia non tutte le energie sono permesse, ma solo quelle associate ad una scelta di tre numeri naturali nella formula di sopra.

A ciascun autovalore è associata l'**autofunzione** ottenuta sostituendo in $\psi(x, y, z) = \psi(x)\psi(y)\psi(z)$ le soluzioni trovate in (3.77):

$$\psi_{\mathcal{E}_{n_1, n_2, n_3}}(x) = \sqrt{\frac{8}{a_1 a_2 a_3}} \sin \frac{n_1 \pi x}{a_1} \sin \frac{n_2 \pi y}{a_2} \sin \frac{n_3 \pi z}{a_3}$$

3.9.3 Evoluzione di una particella in 1D con potenziale costante a tratti: caso generale



Consideriamo lo spazio $\mathcal{H} = L^2(\mathbb{R})$. L'hamiltoniana della particella è data da:

$$H = \frac{p^2}{2m} + V(x)$$

Dove $V(x)$ è un potenziale costante “a tratti”, del tipo:

$$V(x) = \begin{cases} V_1 & x \in \Delta_1 \\ V_2 & x \in \Delta_2 \\ \vdots & \vdots \end{cases}$$

Vogliamo trovare le soluzioni all'equazione agli autovalori per H . Detta $\varphi_{\mathcal{E}}(x)$ l'autofunzione corrispondente all'autovalore \mathcal{E} , procediamo per gradi:

1. Risolviamo l'equazione in ciascuno degli intervalli Δ_i dove il potenziale è costante.
2. Incolliamo le soluzioni φ_i ottenute da ciascun Δ_i con le seguenti richieste:

- Le $\varphi_{\mathcal{E}}$ e $\varphi'_{\mathcal{E}}$ devono “incontrarsi al bordo comune” con continuità. Infatti se $\mathcal{E} \in \sigma_P(H)$ allora la soluzione totale - ottenuta “fondendo” le $\varphi_{\mathcal{E}}$ - deve essere integrabile per parti, come richiesto nel definire $D(H)$, e perciò, come si dimostra, continua. Consideriamo invece il caso di $\mathcal{E} \in \sigma_C(H)$. Dato che ora parliamo di autofunzioni *generalizzate* nello spazio di Schwartz \mathcal{S}' non possiamo più usare la richiesta di regolarità (che si applica in \mathcal{H}) come prima. Ipotizziamo allora per assurdo che “fondendo” due $\varphi_{\mathcal{E}}$ si ottenga una funzione non continua. Essendo le singole $\varphi_{\mathcal{E}}$ limitate, l'unico modo perché ciò avvenga è che ci sia un salto, ossia per esempio si abbia $\varphi_{\mathcal{E}}(x_i^-) \neq \varphi_{\mathcal{E}}(x_i^+)$. Ma un salto ha localmente l'andamento di una “funzione di Heaviside”, la cui derivata (distribuzionale) è la delta di Dirac. Perciò avremo (localmente) $\varphi'_{\mathcal{E}} \sim \delta$ e $\varphi''_{\mathcal{E}} \sim \delta'$, ma il potenziale $V(x)$ non presenta alcuna δ o δ' , e quindi l'equazione di Schrodinger non potrebbe essere

soddisfatta da una φ del genere.

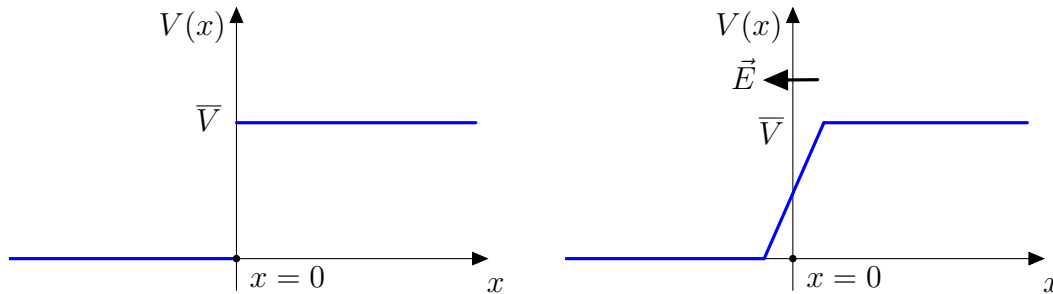
Stessa cosa se ci fosse una discontinuità nella derivata, ossia $\varphi'_\mathcal{E}(x_i^-) \neq \varphi'_\mathcal{E}(x_i^+)$, dato che $\varphi''_\mathcal{E} \sim \delta$. Perciò l'unica è che la “fusione” delle $\varphi_\mathcal{E}$ produca una funzione *continua*.

- $\varphi_\mathcal{E}(x)$ non può crescere esponenzialmente per $x \rightarrow \pm\infty$. Infatti, non può essere $\mathcal{E} \in \sigma_P(H)$, dato che $\varphi_\mathcal{E} \notin \mathcal{H}$ (essendo non normalizzabile), e neanche $\mathcal{E} \in \sigma_C(H)$, dato che $\varphi_\mathcal{E} \notin \mathcal{S}'(\mathbb{R})$ (dato che cresce esponenzialmente)

(Lezione 17 del
7/11/2018)

3.9.4 Il potenziale a gradino

Applichiamo la procedura generale vista nella sottosezione precedente ad un esempio concreto.



Consideriamo il *potenziale a gradino* $V(x) = \bar{V}H(x)$.

Fisicamente si tratta di un’*idealizzazione* del potenziale di un campo elettrico frenante che agisce su una regione molto piccola, per cui la crescita (che comunque resta lineare) può essere vista come un gradino.

Cominciamo considerando il sistema in **Meccanica Classica**. In particolare, esaminiamo il comportamento di una particella che si trova inizialmente a $x(t = 0) = x_0 < 0$ (è cioè a *sinistra* del gradino) e si muove verso di esso con un momento iniziale $p(t = 0) = p_0 > 0$. La sua hamiltoniana H (pari all’energia \mathcal{E} della particella) è quindi data da⁶⁴:

Comportamento
in MC

$$H = \frac{p^2}{2m} + \bar{V}H(x) = \mathcal{E}$$

Poiché p deve essere un numero reale, si ha che, per certi valori dell’energia \mathcal{E} (che rimane costante) la particella non può accedere a tutte le posizioni. Infatti, se risolviamo per p :

$$p = \pm \sqrt{2m(\mathcal{E} - \bar{V}H(x))} \Rightarrow \begin{cases} \mathcal{E} > \bar{V} \Rightarrow p \in \mathbb{R} \forall x \in \mathbb{R} \\ \mathcal{E} < \bar{V} \Rightarrow p \in \mathbb{R} \text{ solo per } x < 0 \end{cases}$$

⁶⁴^Attenzione a non confondere H dell’hamiltoniana con $H(x)$ funzione di Heaviside!

dato che, se $\mathcal{E} < \bar{V}$, per $x > 0$ l'espressione sotto radice è negativa. Fisicamente ciò corrisponde al fatto che una particella con troppa poca energia *non* può attraversare il gradino del potenziale.

Esaminiamo l'evoluzione temporale (classica) del sistema.

Derivazione dell'equazione del moto La particella parte da x_0 , e si muove (inizialmente verso destra) a velocità:

$$\dot{x} = \frac{\partial H}{\partial p} = \frac{p}{m} = \frac{\pm \sqrt{2m(\mathcal{E} - \bar{V}H(x))}}{m} = \pm \sqrt{\frac{2}{m}(\mathcal{E} - \bar{V}H(x))}$$

Basta allora integrare per trovare l'equazione del moto. Per $x < 0$, per cui $H(x) = 0$, si trova:

$$x(t) = x(t=0) \pm t \sqrt{\frac{2}{m}\mathcal{E}}$$

Che in termini delle condizioni iniziali (e imponendo $p_0 > 0$) diviene:

$$x_1(t) = x_0 + \frac{p_0}{m}t$$

Il dominio di questa soluzione non è tutto $t \in \mathbb{R}$, perché siamo partiti supponendo $x < 0$:

$$x_1(t) < 0 \Rightarrow t < -x_0 \frac{m}{p_0} = \frac{|x_0|m}{p_0} \equiv t_0$$

Quindi $x_1(t)$ vale per $t \in (-\infty, t_0)$.

Per $x > 0$, d'altro canto, $H(x) = 1$, e quindi l'equazione da integrare è:

$$\dot{x} = \pm \sqrt{\frac{2}{m}(\mathcal{E} - \bar{V})}$$

che, come già osservato, ha soluzione reale solo per $\mathcal{E} > \bar{V}$. Integrando otteniamo:

$$x_2(\tau) = x(\tau=0) \pm \tau \sqrt{\frac{2}{m}(\underbrace{\mathcal{E}}_{\frac{p_0^2}{2m}} - \bar{V})} = x(\tau=0) \pm \frac{1}{m} \sqrt{(p_0^2 - 2m\bar{V})} \tau$$

Dove $x(\tau=0) = 0$ per raccordo, e $\tau = t - t_0$ (dato che la particella che parte a $t = 0$ in x_0 raggiunge a $t = t_0$ il gradino, dove si ha il *passaggio* all'altra soluzione). Notiamo poi che solo la soluzione con il $+$ ha senso, dato che l'altra ha un dominio temporale nullo (rappresenta una particella che “torna indietro”, ma allora $x_2(\tau = 0^+) < 0$, e quindi esce dall'ipotesi che abbiamo fatto prima di integrare).

Perciò, se $\mathcal{E} > \bar{V}$, con $x_0 < 0$, $p_0 > 0$, la soluzione finale si ottiene unendo $x_1(t)$ e $x_2(\tau)$:

$$x(t)_{\mathcal{E} > \bar{V}} = \begin{cases} x_0 + \frac{p_0}{m}t & t < t_0 \\ \frac{1}{m} \sqrt{p_0^2 - 2m\bar{V}}(t - t_0) & t \geq t_0 \end{cases}$$

D'altro canto, se $\mathcal{E} < \bar{V}$, la regione $x > 0$ è proibita. Quando la particella vi arriva può perciò solo “tornare indietro”^a, e quindi passare alla soluzione per $x < 0$ che è compatibile con un momento iniziale negativo. Avremo quindi:

$$x(t)_{\mathcal{E} < \bar{V}} = \begin{cases} x_0 + \frac{p_0}{m}t & t < t_0 \\ -\frac{p_0}{m}(t - t_0) & t \geq t_0 \end{cases}$$

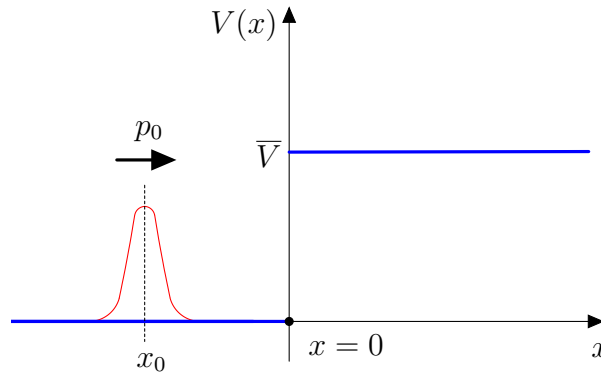
^aSe non si è convinti di ciò, si può provare a considerare un potenziale *continuo*, che cresce linearmente tra $-a$ e a da 0 a \bar{V} , quindi risolvere il moto e prendere il limite per $a \rightarrow 0$. In tal caso nella regione di *crescita lineare* la particella subisce un'accelerazione negativa. Se non ha abbastanza energia per superare tale regione, allora ritornerà in $x = 0$ con momento $-p_0$.

Osservando che $-p_0 t_0 / m = x_0$, possiamo riscrivere in maniera più sintetica le soluzioni:

$$\begin{aligned} x(t) &= -\frac{p_0}{m}|t - t_0| & \mathcal{E} < \bar{V} \\ x(t) &= -\frac{p_0}{m}(t_0 - t)H(t_0 - t) + \frac{1}{m}\sqrt{p_0^2 - 2m\bar{V}}(t - t_0)H(t - t_0) & \mathcal{E} > \bar{V} \end{aligned}$$

Perciò per $\mathcal{E} < \bar{V}$ la particella *sbatte* contro il gradino di potenziale e **immediatamente** torna indietro, mentre per $\mathcal{E} > \bar{V}$ la particella *supera* **immediatamente** il gradino.

Vedremo che in MQ non si verifica nulla di tutto ciò: particelle di alta energia avranno una certa probabilità di essere riflesse, e in ogni caso il gradino avrà un *tempo di attraversamento* non istantaneo.



Consideriamo ora l'analogo sistema **quantistico**. La particella classica è sostituita da un **pacchetto d'onda** inizialmente *piccato* attorno a x_0 e con momento *piccato* attorno a p_0 .

Comportamento in MQ

Richiami sulla nozione di pacchetto d'onda

Un **pacchetto d'onda** non è altro che “un'onda di dimensioni finite”, ossia una perturbazione dello spazio che ha valori $\neq 0$ solo in una **regione ristretta**, come in genere si verifica sperimentalmente, dato che non si ha a che fare con segnali infiniti.

Matematicamente, un pacchetto d'onda si ottiene dalla sovrapposizione di un nu-

mero infinito di onde ideali (es. onde piane), ciascuna con una certa lunghezza d'onda λ e un relativo numero d'onda (*angolare*) $k = 2\pi/\lambda$. Scegliendo bene i termini della “somma” (che, in realtà, è realizzata con un integrale), si ottiene un'interferenza costruttiva in una regione limitata, e distruttiva in tutto il resto dello spazio.

Nota: per *localizzare* il pacchetto è necessario incrementare le regioni di interferenza distruttiva, e quindi considerare nella somma un numero maggiore di termini, ossia più possibili scelte per k . Al diminuire dell'*incertezza sulla posizione* del pacchetto corrisponde il crescere dell'*incertezza sul suo momento*, e viceversa.

In generale, partendo da un pacchetto con andamento arbitrario $\psi(x, t)$, la sua rappresentazione in momenti (o k , o frequenze, dato che ω è funzione di k) si ottiene per trasformata di Fourier:

$$\tilde{\psi}(k, t) = \mathcal{F}[\psi(x, t)](k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \psi(x, t) e^{-ikx} dx$$

E il passaggio inverso si ottiene mediante l'antitrasformata:

$$\psi(x, t) = \mathcal{F}^{-1}[\tilde{\psi}(k, t)](x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \tilde{\psi}(k, t) e^{ikx} dk$$

Perciò basta fissare $\psi(x, t)$ o $\tilde{\psi}(k, t)$ per ricavare l'altro. Per esempio potremmo partire fissando l'*energia* del pacchetto, ossia il range delle k delle onde che lo compongono. Fissiamo allora la funzione $\tilde{\psi}(k, t)$ che dà l'ampiezza delle singole onde-componenti a seconda del loro k . Se separiamo componente dell'ampiezza da quella dell'evoluzione temporale:

$$\tilde{\psi}(k, t) = \tilde{f}(k) e^{-i\omega(k)t}$$

giungiamo alla scrittura di un generico pacchetto come:

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \tilde{f}(k) e^{i(kx - \omega(k)t)} dk \quad (3.79)$$

Se la funzione $\omega = \omega(k)$ è lineare, allora tutte le onde che compongono il pacchetto *si muovono alla stessa velocità*, che è detta *velocità di fase* ed è data dalla *pendenza* di $\omega(k)$:

$$v = \frac{\omega(k)}{k}$$

(per cui se $\omega(k) = \bar{v}k$ si ha immediatamente $v \equiv \bar{v}$).

In tal caso il mezzo in cui si propaga l'onda è detto **non dispersivo** (un esempio è il vuoto per un'onda classica): dato che tutte le componenti del pacchetto si muovono in maniera *sincrona*, il pacchetto “non si sfalda” e mantiene lo stesso “profilo” (envelope). In particolare la velocità del pacchetto è la stessa delle onde che lo compongono. Ecco perché la funzione $\omega = \omega(k)$ è detta **relazione di dispersione**.

Se invece $\omega = \omega(k)$ non è lineare, si ha **dispersione**. In tal caso il pacchetto,

ossia “l’inviluppo” della sovrapposizione delle onde, in genere **non** viaggia alla stessa velocità delle onde che lo compongono, ma a una certa v_g , detta **velocità di gruppo**, che ora deriviamo.

Chiaramente, ha senso parlare di “velocità di un pacchetto” è “ben localizzato”, ossia se la sovrapposizione avviene tra onde di simili k . Ciò si ha se scegliamo la $\tilde{f}(k)$ con un picco *sharp* attorno a k_0 . Possiamo allora *linearizzare* $\omega(k)$ sviluppandola in serie di Taylor attorno a k_0 (che è il *massimo* che ci interessa) e troncando al primo ordine:

$$\omega(k) \approx \underbrace{\omega(k_0)}_{\omega_0} + \underbrace{\left. \frac{\partial \omega(k)}{\partial k} \right|_{k=k_0}}_{\omega'_0} (k - k_0) \quad (3.80)$$

Sostituendo (3.80) in (3.79) otteniamo (omettendo il fattore di normalizzazione per semplicità):

$$\begin{aligned} \psi(x, t)_{lin} &= \int_{-\infty}^{+\infty} \tilde{f}(k) e^{i[kx - (\omega_0 + \omega'_0(k - k_0))t]} dk = \\ &= \int_{-\infty}^{+\infty} \tilde{f}(k) \exp[i(+k_0x - \omega_0t - k_0x + k_0\omega'_0t + kx - k\omega'_0t)] = \\ &= e^{i(k_0x - \omega_0t)} \int_{-\infty}^{+\infty} dk \tilde{f}(k) e^{i(k - k_0)(x - \omega'_0t)} \end{aligned}$$

e $\psi(x, t)_{lin}$ è un'approssimazione molto buona di $\psi(x, t)$ se $\tilde{f}(k)$ ha un picco molto stretto attorno a k_0 . In quest'approssimazione il pacchetto è dato da una componente localizzata di *segnale* (in azzurro) che costituisce l'*inviluppo* (envelope) del pacchetto, ossia la sua “sagoma”. Tale espressione è moltiplicata per il termine in giallo, che è una normalissima onda piana con numero d'onda k_0 , che si muove alla velocità di fase ω_0/k_0 . L'espressione azzurra, invece, dipende da t per una $e^{i(x - \omega'_0t)}$, e quindi i suoi massimi si propagano alla **velocità di gruppo** (in quanto è quella dell'inviluppo) che sarà proprio la ω'_0 :

$$\omega'_0 = \left. \frac{d\omega}{dk} \right|_{k=k_0}$$

Nota: se $\omega(k)$ è decrescente (cosa che può succedere localmente in un mezzo irregolare) è possibile che la velocità di gruppo sia di segno opposto a quella di fase: si ha allora che le onde componenti si muovono in una direzione, mentre l'inviluppo *risale* lungo il verso opposto!^a

^a Delle animazioni carine dei vari casi si trovano su wikipedia, a http://en.wikipedia.org/wiki/Group_velocity e http://en.wikipedia.org/wiki/Wave_packet

Utilizzando la relazione (già vista per il fotone):

$$p = \hbar k \Rightarrow k = \frac{p}{\hbar} \quad (3.81)$$

possiamo usare le k come variabili, che sono più comode per una trattazione

ondulatoria. Da $\mathcal{E} = p^2/(2m)$ si ha poi $p = \sqrt{2m\mathcal{E}}$, e giungiamo quindi a:

$$k = \frac{\sqrt{2m\mathcal{E}}}{\hbar} \Rightarrow \mathcal{E}(k) = \frac{k^2 \hbar^2}{2m} \quad (3.82)$$

Un'onda piana in 1D ha la forma generale di $u(x, t) = e^{i(kx - \omega t)}$. Un pacchetto d'onda è la combinazione di *infinite* onde, la cui ampiezza dipende da k ed è regolata da una funzione $A(k)$ che descrive la “forma” del pacchetto $\bar{\psi}$. Avremo quindi:

$$\bar{\psi}(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} A(k) u(x, t) dk$$

dove il coefficiente $1/\sqrt{2\pi}$ compare per convenzioni legate alle trasformate di Fourier⁶⁵. Nel caso quantistico, per un pacchetto d'onde che cerchi di approssimare un comportamento particellare, non tutte le $u(x, t)$ sono ammesse, ma solo quelle che sono autofunzioni di H , e che quindi indichiamo come $\varphi_{\mathcal{E}}(x)$. Per le ampiezze usiamo invece la notazione $\tilde{f}_{k_0}(k)$, dove il pedice indica la posizione del *picco*, e la \sim ne evidenzia la natura di trasformata di Fourier.

Un'ultima importante modifica sta nel variare gli estremi di integrazione: vogliamo infatti considerare la particella che inizialmente si muove verso il gradino, e che quindi ha p (e di conseguenza k) positivo. Basta allora partire da 0 invece che da $-\infty$.⁶⁶

Definiamo allora il pacchetto d'onda iniziale come:

$$\psi(x, t = 0) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} dk \tilde{f}_{k_0}(k) \varphi_{\mathcal{E}}(x) \quad (3.83)$$

La $\tilde{f}_{k_0}(k)$ è nello specifico una funzione positiva piccata su k_0 (ossia il valor medio $\langle k \rangle_{\tilde{f}_{k_0}} \approx k_0$) - e perciò è analoga alla particella classica di momento iniziale p_0 . Portando avanti l'analogia classica abbiamo ora due casi: il pacchetto d'onda ha “nel complesso” un'energia più bassa della barriera di potenziale, oppure ne ha una più alta. Esaminiamoli separatamente.

Energia più bassa del gradino Scegliamo in questo caso $\tilde{f}_{k_0}(k)$ con un *supporto*⁶⁷ Δ tale che:

$$\mathcal{E}(k) = \frac{\hbar^2 k^2}{2m} < \bar{V} \quad \forall k \in \Delta$$

In questo modo anche le onde “più energetiche” che compongono il pacchetto non hanno comunque abbastanza energia (almeno classicamente) per superare il gradino (come per la particella con $\mathcal{E} < V$).

Per esempio potremmo prendere:

$$\tilde{f}_{k_0}(k) = \chi_{[k_0 - \frac{a}{2}, k_0 + \frac{a}{2}]}(k)$$

⁶⁵ΛIn effetti, la $u(x, t)$ così scritta ha la forma di una trasformata di Fourier. Per approfondimenti vedere i *richiami* all'inizio di questa sezione

⁶⁶ΛCome vedremo, le funzioni d'onda si evolvono in maniera *dispersiva*. La presenza di $k < 0$ provocherebbe quindi un veloce “sfaldarsi” dell'onda, che - seppur fisicamente accettabile - non corrisponde in questo caso all'analogia classica che stiamo portando avanti.

⁶⁷ΛDove il supporto di una funzione è il range delle coordinate nel dominio per cui ha valore non nullo.

Per studiare l'evoluzione temporale di $\psi(x, t)$ dobbiamo prima trovare le $\varphi_{\mathcal{E}}(x)$, che saranno autovettori (eventualmente generalizzati) dell'operatore:

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \bar{V}H(x)$$

con una *fase* per i coefficienti di $\varphi_{\mathcal{E}}(x)$ scelta opportunamente perché $\psi(x, 0)$ sia piccata attorno a x_0 .

Risolviamo quindi l'equazione agli autovalori:

$$\begin{aligned} H\varphi_{\mathcal{E}}(x) &= \mathcal{E}\varphi_{\mathcal{E}}(x) \\ \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \bar{V}H(x)\right) \varphi_{\mathcal{E}}(x) &= \mathcal{E}\varphi_{\mathcal{E}}(x) \end{aligned} \quad (3.84)$$

Per risolverla conviene spezzare il dominio e raccordare poi le soluzioni. Per $x < 0$ si ha che $H(x) = 0$, e quindi (3.84) diviene:

$$\begin{aligned} \varphi_{\mathcal{E}}''(x) + \frac{2m\mathcal{E}}{\hbar^2} \varphi_{\mathcal{E}}(x) &= 0 \Rightarrow t^2 + \frac{2m\mathcal{E}}{\hbar^2} = 0 \Rightarrow t_{1,2} = \pm \frac{i}{\hbar} \sqrt{2m\mathcal{E}} = \pm ik \\ \Rightarrow \varphi_{\mathcal{E}}(x) &= c_+^1 e^{ikx} + c_-^1 e^{-ikx}, \quad x < 0 \end{aligned} \quad (3.85)$$

Analogamente, per $x \geq 0$ si ha che $H(x) = 1$, e quindi l'equazione diviene:

$$\begin{aligned} \varphi_{\mathcal{E}}''(x) + \frac{2m}{\hbar^2} \underbrace{(\mathcal{E} - \bar{V})}_{<0} \varphi_{\mathcal{E}}(x) &= 0 \Rightarrow t_{1,2} = \pm \frac{\overbrace{\sqrt{2m(\bar{V} - \mathcal{E})}}^{\chi}}{\hbar} \\ \Rightarrow \varphi_{\mathcal{E}}(x) &= c_+^2 e^{\chi x} + c_-^2 e^{-\chi x}, \quad x \geq 0 \end{aligned} \quad (3.86)$$

Riepilogando:

$$\varphi_{\mathcal{E}}(x) = \begin{cases} c_+^1 e^{ikx} + c_-^1 e^{-ikx} & x < 0 \\ c_+^2 e^{\chi x} + c_-^2 e^{-\chi x} & x \geq 0 \end{cases} \quad \chi = \frac{\sqrt{2m(\bar{V} - \mathcal{E})}}{\hbar}; \quad k = \frac{\sqrt{2m\mathcal{E}}}{\hbar} \quad (3.87)$$

*Autofunzioni del
potenziale a
gradino*

Dalla teoria sappiamo che le autofunzioni $\varphi_{\mathcal{E}}(x)$ appartengono a \mathcal{H} se corrispondono a un autovalore $\mathcal{E} \in \sigma_P(H)$ dello spettro discreto, o al più a \mathcal{S}' se invece $\mathcal{E} \in \sigma_C(H)$. In entrambi i casi **non** possono perciò divergere esponenzialmente. Ne segue allora che $c_+^2 = 0$, sennò la seconda soluzione divergerebbe per $x \rightarrow \infty$. Notiamo poi che H non ha spettro discreto in questo caso: infatti $\varphi_{\mathcal{E}}(x) \notin L^2(\mathbb{R})$, poiché l'integrale del modulo quadro diverge:

$$\int_{-\infty}^{\Lambda} |\varphi_{\mathcal{E}}(x)|^2 \xrightarrow{\Lambda \rightarrow +\infty} +\infty$$

data la presenza di termini e^{ikx} che oscillano all'infinito.

Perciò avremo che:

$$\sigma(H) \Big|_{\mathcal{E} < \bar{V}} = \sigma_C(H) \Big|_{0 < \mathcal{E} < \bar{V}}$$

Notiamo che le φ dipendono da 3 coefficienti, e imponendo le 2 equazioni di *raccordo* (per continuità e derivabilità) resta un coefficiente non determinato, e quindi la degenerazione degli autovalori è 1.

Le equazioni di raccordo per $\varphi_{\mathcal{E}}$ e $\varphi'_{\mathcal{E}}$ portano a:

$$\lim_{x \rightarrow 0^-} \varphi_{\mathcal{E}}(x) \stackrel{!}{=} \lim_{x \rightarrow 0^+} \varphi_{\mathcal{E}}(x) \Rightarrow c_+^1 + c_-^1 = c_-^2 \quad (3.88)$$

$$\lim_{x \rightarrow 0^-} \varphi'_{\mathcal{E}}(x) \stackrel{!}{=} \lim_{x \rightarrow 0^+} \varphi'_{\mathcal{E}}(x) \Rightarrow ik(c_+^1 - c_-^1) = -\chi c_-^2 \quad (3.89)$$

Dividendo per ik la (3.89) e sommando membro a membro con (3.88) si può ricavare immediatamente c_-^2 :

$$c_-^2 = \frac{2}{1 + i\frac{\chi}{k}} c_+^1 \Rightarrow c_-^1 = c_-^2 - c_+^1 = \frac{1 - i\frac{\chi}{k}}{1 + i\frac{\chi}{k}} c_+^1$$

Scriviamo le costanti in rappresentazione polare (più comoda per le trasformate di Fourier). Notando che il termine $A_+ = 1 + i\frac{\chi}{k}$ si ripete, partiamo da quello:

$$\begin{aligned} A_+ = 1 + i\frac{\chi}{k} &= \left| 1 + i\frac{\chi}{k} \right| e^{i\theta(k)} \\ \tan \theta(k) &= \frac{\chi}{k} = \frac{\sqrt{2m(\bar{V} - \mathcal{E})}}{\hbar} \frac{\hbar}{\sqrt{2m\mathcal{E}}} = \sqrt{\frac{\bar{V} - \mathcal{E}}{\mathcal{E}}} = \sqrt{\frac{\bar{V}}{\mathcal{E}}} - 1 = \\ &= \sqrt{\frac{\bar{k}^2}{k^2} - 1}; \quad \bar{k} = \frac{\sqrt{2m\bar{V}}}{\hbar} \end{aligned}$$

dove \bar{k} è il numero dell'onda che ha la stessa energia \bar{V} del gradino. Il complesso coniugato $A_- = 1 - i\frac{\chi}{k}$ ha quindi lo stesso modulo e angolo $\theta(k)$ opposto.

Le costanti c_-^1 e c_-^2 divengono quindi:

$$c_-^2 = \frac{2}{|1 + i\frac{\chi}{k}| e^{i\theta(k)}} c_+^1 = \frac{2c_+^1}{|1 + i\frac{\chi}{k}|} e^{-i\theta(k)} \quad (3.90)$$

$$c_-^1 = \frac{|1 - i\frac{\chi}{k}| e^{-i\theta(k)}}{|1 + i\frac{\chi}{k}| e^{i\theta(k)}} c_+^1 = c_+^1 e^{-2i\theta(k)} \quad (3.91)$$

Conviene ora *collegare* le costanti alle condizioni iniziali. In particolare, possiamo scegliere c_+^1 (che è un *parametro libero* dato dalla degenerazione degli stati) in modo che il pacchetto sia piccato attorno a x_0 all'istante iniziale. In particolare poniamo:

$$c_+^1 = e^{-ikx_0} \quad (3.92)$$

Verifichiamo che in tal modo il pacchetto d'onda sia davvero piccato in x_0 . Calcoliamo la (3.83) per $\varphi_{\mathcal{E}}(x)$ in (3.87) in $x < 0$. Con la scelta fatta per c_+^1 otteniamo:

$$\psi(x, t = 0) = \frac{1}{\sqrt{2\pi}} \int_0^\infty dk \tilde{f}_{k_0}(k) e^{-ikx_0} (e^{ikx} + e^{-2i\theta(k)} e^{-ikx}) \quad (3.93)$$

Degenerazione = 1

Condizioni di raccordo

Costanti in coordinate polari

Condizioni iniziali

Concentriamoci sul termine evidenziato in giallo, che corrisponde all'onda *incidente*. A meno di un fattore costante si ha:

$$\psi(x, 0)_i \propto \int_0^{+\infty} \tilde{f}_{k_0}(k) e^{ik(x-x_0)} dk = \int_0^{+\infty} \tilde{f}_0(k') e^{i(k'+k_0)(x-x_0)} dk' \quad (3.94)$$

dove in (a) si è effettuato il cambio di variabile $k \rightarrow k' + k_0$, e si è *traslata* la funzione $\tilde{f}_{k_0}(k)$:

$$\tilde{f}_{k_0}(k + k_0) = \tilde{f}_0(k')$$

che, nella nuova variabile k' , è ora piccata in 0 (nei prossimi passaggi rimuoveremo l'apice per alleggerire la notazione).

Notiamo ora la presenza di un termine $e^{ik_0(x-x_0)}$ che corrisponde all'onda piana “di fondo” del pacchetto d'onda⁶⁸ che non dipende da k e quindi possiamo estrarre dall'integrale:

$$(3.94) = e^{ik_0(x-x_0)} \int_0^{+\infty} \tilde{f}_0(k) e^{ik(x-x_0)} dk \quad (3.95)$$

Il fattore evidenziato *non* modifica la posizione dei massimi, e quindi ce ne disinteressiamo. Giungiamo allora a:

$$\psi(x, 0)_i \propto \int_0^{+\infty} \tilde{f}_0(k) e^{ik(x-x_0)} dk \quad (3.96)$$

che ha un picco in x_0 . Infatti, calcolandone derivata prima:

$$\begin{aligned} \frac{d}{dx} \int_0^{+\infty} \tilde{f}_0(k) e^{ik(x-x_0)} dk \Big|_{x=x_0} &= \int_0^{+\infty} \tilde{f}_0(k) ik e^{ik(x-x_0)} dk \Big|_{x=x_0} = \\ &= \int_0^{+\infty} \tilde{f}_0(k) ik dk \underset{(a)}{\approx} 0 \end{aligned}$$

dove l'integrale in (a) corrisponde, a meno di un fattore i , al valor atteso di $\tilde{f}_0(k)$ su $[0, +\infty)$. Ma la $\tilde{f}_0(k)$ (limitata e a supporto compatto) è piccata in $k = 0$, per cui gli unici valori per cui è $\neq 0$ sono per k piccoli, e perciò l'integrale va a 0.

Esaminando la derivata seconda otteniamo:

$$\frac{d^2}{dx^2} \int_0^{+\infty} \tilde{f}_0(k) e^{ik(x-x_0)} dk \Big|_{x=x_0} = \int_0^{+\infty} dk (-k^2) \tilde{f}_0(k) < 0$$

per cui $x = x_0$ è proprio un massimo per $\psi(x, 0)_i$.

Notiamo che tutti i ragionamenti effettuati valgono indipendentemente dal segno di k , e perciò si giunge alla stessa conclusione anche esaminando il termine evidenziato in azzurro nella (3.93).

Perciò l'intera $\psi(x, 0)$ ha massimo in $x = x_0$, come volevamo.

Sostituendo allora (3.87) in (3.83) assieme alle espressioni trovate per le costanti in (3.91) e (3.92), giungiamo all'espressione per il pacchetto a $t = 0$:

$$\psi(x, 0) = \begin{cases} \frac{1}{\sqrt{2\pi}} \int_0^\infty dk \tilde{f}_{k_0}(k) [e^{ik(x-x_0)} + e^{-ik(x+x_0)} e^{-2i\theta(k)}] & x < 0 \\ \frac{1}{\sqrt{2\pi}} \int_0^\infty dk \tilde{f}_{k_0}(k) \frac{2}{|1 + i\frac{x}{k}|} e^{-\chi x} e^{-ikx_0} e^{-i\theta(k)} & x \geq 0 \end{cases} \quad (3.97)$$

⁶⁸ Possiamo vedere le costituenti del pacchetto come una componente di *segnale* data da $e^{ik(x-x_0)}$ e un fondo sinusoidale comune a tutte dato da $e^{ik_0(x-x_0)}$

L'evoluzione di uno stato puro, come da assioma, è data da:

$$\varphi_{\mathcal{E}}(x) \xrightarrow{t} \varphi_{\mathcal{E}}(x) e^{-i \frac{\mathcal{E}}{\hbar} t}$$

In analogia con l'onda piana classica $u(x, t) = e^{i(kx - \omega t)}$, poniamo:

$$\omega \equiv \frac{\mathcal{E}}{\hbar} = \frac{k^2 \hbar}{2m} \quad (3.98)$$

Nota: ω non è una funzione lineare di k , e quindi il pacchetto d'onda subisce un effetto di *dispersione*: le onde che lo compongono viaggiano a *velocità di fase* differenti (e che dipendono da k) e quindi ci aspettiamo che il pacchetto “non mantenga la stessa forma” nel muoversi verso il gradino.

L'evoluzione temporale è allora data dalla sostituzione $\varphi_{\mathcal{E}} \rightarrow \varphi_{\mathcal{E}}(x) e^{-i\omega t}$, che, in (3.97) porta a:

$$\psi(x, t) = \begin{cases} \frac{1}{\sqrt{2\pi}} \int_0^\infty dk \tilde{f}_{k_0}(k) \left[\overbrace{e^{i[k(x-x_0) - \omega t]}}^{\psi_i \text{ incidente}} + \overbrace{e^{-i[k(x+x_0) + \omega t]} e^{-2i\theta(k)}}^{\psi_r \text{ riflessa}} \right] & x < 0 \\ \frac{1}{\sqrt{2\pi}} \int_0^\infty dk \tilde{f}_{k_0}(k) \frac{2}{|1 + i \frac{\chi}{k}|} e^{-\chi x} e^{-ikx_0} e^{-i[\theta(k) + \omega t]} & x \geq 0 \end{cases}$$

Ritroviamo perciò per $x < 0$ la sovrapposizione di un'onda incidente con una riflessa, e per $x \geq 0$ si ha un'onda che decresce esponenzialmente (a causa del termine $e^{-\chi x}$).

Poiché il picco del pacchetto d'onda rappresenta la “posizione più probabile” della particella, possiamo calcolare il suo andamento per confrontarlo con il caso classico. Poiché \tilde{f}_{k_0} è piccata su k_0 , esaminiamo la funzione integranda per quel valore di k . Per $x < 0$ e ψ_i incidente, il massimo di ordine 0 è dato dall'annullarsi della derivata:

$$\begin{aligned} \frac{\partial}{\partial k} \Big|_{k_0} \psi_i &\stackrel{!}{=} 0 \Rightarrow \cancel{\not\equiv} \left[(x - x_0) - \frac{d\omega}{dk} t \right] \cancel{e^{i[k(x-x_0) - \omega t]}} = 0 \\ &\Rightarrow x - x_0 - \frac{d\omega}{dk} \Big|_{k_0} t = 0 \end{aligned} \quad (3.99)$$

Ricordando la definizione di ω in (3.98) si ha:

$$\omega = \frac{k^2 \hbar}{2m} \Rightarrow \frac{d\omega}{dk} = \frac{k \hbar}{m}$$

Sostituendo in (3.99) si ottiene allora l'equazione del moto *classica*:

$$x = x_0 + \frac{\hbar k_0}{m} t \stackrel{(3.81)}{=} x_0 + \frac{p_0}{m} t = \frac{p_0}{m} (t - t_0)$$

Perciò, prima dell'urto, il sistema si comporta “classicamente”, con il picco che arriva al gradino a $t = t_0$.

Se invece esaminiamo, per $x < 0$, il picco per ψ_r (sempre calcolando il massimo di ordine 0 per k_0), dobbiamo considerare anche un altro termine:

$$x = -x_0 - \frac{\hbar k_0}{m}t - 2 \frac{d\theta}{dk}(k_0) = \frac{p_0}{m}(t_0 - t) - \underbrace{2 \frac{d\theta}{dk}(k_0)}_{\text{offset della riflessione}} \quad (3.100)$$

Perciò è come se l'onda riflessa non ripartisse da $x = 0$, ma da un punto diverso. Calcolando la derivata:

$$\begin{aligned} \tan \theta(k) = \frac{\chi}{k} = \sqrt{\frac{\bar{k}^2}{k^2} - 1} &\Rightarrow \theta(k) = \arctan \left(\sqrt{\frac{\bar{k}^2}{k^2} - 1} \right) \\ &\Rightarrow \frac{d\theta}{dk}(k_0) = -\frac{1}{\sqrt{\bar{k}^2 - k_0^2}} < 0 \end{aligned}$$

Sostituendo in (3.100) si giunge a:

$$x = \frac{p_0}{m}(t_0 - t) + \underbrace{\frac{2}{\sqrt{\bar{k}^2 - k_0^2}}}_{\bar{x} > 0}$$

Perciò il picco dell'onda riflessa riparte da un punto $\bar{x} > 0$ che si trova “oltre” il gradino di potenziale. Ciò significa che l'onda “riemerge” dal gradino con un certo ritardo, dato dividendo l'offset per la *velocità*:

$$\tau = \frac{-2 \frac{d\theta}{dk}(k_0)}{\frac{\hbar k_0}{m}} = \frac{2}{\sqrt{\bar{k}^2 - k_0^2}} \frac{m}{\hbar k_0} > 0$$

Classicamente ciò non ha senso: è come se una pallina che sbatte contro il muro si fermasse per un certo τ prima di rimbalzare!

Ancora più stupefacente: vi è una soluzione anche per $x > 0$, dove classicamente l'energia cinetica è negativa, sebbene questo contributo sia esponenzialmente depresso ($e^{-\chi x}$). Tale $\psi(t)$ corrisponde a un'onda di probabilità evanescente con probabilità massima a $\tau/2$.

Perciò, graficamente, il pacchetto *entra leggermente* nel gradino di potenziale, e quando *non riesce più ad andare avanti rimbalza indietro*. Questo *tentativo di entrata* produce il ritardo che viene osservato.

Differenze con il caso classico per $\mathcal{E} < \bar{V}$

Energia più alta del gradino Cosa succede invece se l'energia del pacchetto è maggiore del potenziale \bar{V} del gradino?

Ripetendo i conti dall'equazione agli autovalori, il risultato in (3.85) vale ancora, ma stavolta in (3.86) si ha un'espressione negativa dentro la radice, e perciò la parte per $x > 0$ è differente:

$$\varphi_{\mathcal{E} > \bar{V}}(x) = \begin{cases} c_+^1 e^{ikx} + c_-^1 e^{-ikx} & x < 0 \\ c_+^2 e^{ik_2 x} + c_-^2 e^{-ik_2 x} & x \geq 0 \end{cases} \quad k_2 = \frac{\sqrt{2m(\mathcal{E} - \bar{V})}}{\hbar} \quad (3.101)$$

Le condizioni di raccordo sono:

$$\varphi_{\mathcal{E}}(0^-) = \varphi_{\mathcal{E}}(0^+) \Rightarrow c_+^1 + c_-^1 = c_+^2 + c_-^2 \quad (3.102)$$

$$\varphi'_{\mathcal{E}}(0^-) = \varphi'_{\mathcal{E}}(0^+) \Rightarrow ik(c_+^1 - c_-^1) = ik_2(c_+^2 - c_-^2) \quad (3.103)$$

Lo spettro è nuovamente solo continuo (le soluzioni oscillano all'infinito, e quindi non sono in \mathcal{H}):

$$\sigma(H)\Big|_{\mathcal{E} > \bar{V}} = \sigma_C(H)\Big|_{\mathcal{E} > \bar{V}}$$

Abbiamo 4 coefficienti e 2 equazioni di raccordo, per cui la degenerazione degli autovalori è $4 - 2 = 2$.

Degenerazione =
2

Notiamo che $c_-^2 \neq 0$ conduce alla soluzione di una particella che, giungendo da $x = +\infty$, attraversa il gradino da destra a sinistra. Poiché ci interessa invece la dinamica di attraversamento da sinistra a destra, imponiamo $c_-^2 = 0$.

Allora dividendo (3.103) per ik e sommando membro a membro con (3.102), si ricava immediatamente c_+^2 , e di conseguenza c_-^1 :

$$c_+^2 = \frac{2}{1 + \frac{k_2}{k}} c_+^1; \quad c_-^1 = c_+^2 - c_+^1 = \frac{1 - \frac{k_2}{k}}{1 + \frac{k_2}{k}} c_+^1 \quad (3.104)$$

Sostituendo in (3.101) e in (3.83):

$$\psi_{\mathcal{E}}(x, t) = \begin{cases} \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} dk \tilde{f}_{k_0}(k) \left[\underbrace{e^{i[k(x-x_0)-\omega t]}}_{\psi_i} + \frac{1 - \frac{k_2}{k}}{1 + \frac{k_2}{k}} \underbrace{e^{-i[k(x+x_0)+\omega t]}}_{\psi_r} \right] & x < 0 \\ \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} dk \tilde{f}_{k_0}(k) \frac{2}{1 + \frac{k_2}{k}} \underbrace{e^{i[k_2 x - k x_0 - \omega t]}}_{\psi_t} & x \geq 0 \end{cases}$$

Notiamo che, nonostante la particella abbia *classicamente* abbastanza energia per superare il gradino di potenziale, vi è ancora una soluzione ψ_r corrispondente ad una riflessione, a cui è associata una probabilità finita.

Di nuovo, ciò va contro l'intuizione: nonostante il potenziale *non basti a frenare* una particella troppo energetica, c'è comunque una certa probabilità che essa venga riflessa!

L'equazione di continuità

Cerchiamo di quantificare la componente riflessa e trasmessa dell'onda. Per farlo ci servirà un modo per esaminare l'evoluzione non della funzione d'onda, ma del suo modulo quadro, che esprime le probabilità che ci interessano.

Partendo dall'equazione di Schrodinger possiamo giungere alla relazione che ci interessa sul modulo quadro, derivando la cosiddetta **equazione di continuità** per la densità di probabilità. Limitiamoci al caso unidimensionale, in cui l'equazione di Schrodinger è:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle \quad (3.105)$$

$$i\hbar \frac{\partial \psi}{\partial t}(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x, t) + V(x) \psi(x, t) \quad (3.106)$$

e la sua (duale) complessa coniugata è data da:

$$-i\hbar \frac{\partial \psi^*}{\partial t}(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi^*(x, t) + V(x) \psi^*(x, t) \quad (3.107)$$

Moltiplicando (3.106) per ψ^* e (3.107) per ψ e sottraendo membro a membro i contributi dei V si semplificano, e si giunge a:

$$\begin{aligned} i\hbar \underbrace{\left(\psi^* \frac{\partial \psi}{\partial t} + \frac{\partial \psi^*}{\partial t} \psi \right)}_{\frac{\partial}{\partial t}(\psi^* \psi) = \frac{\partial}{\partial t} |\psi(x, t)|^2} &= -\frac{\hbar^2}{2m} \left[\psi^* \left(\frac{\partial^2}{\partial x^2} \right) \psi - \psi \left(\frac{\partial^2}{\partial x^2} \right) \psi^* \right] \\ \Rightarrow \frac{\partial}{\partial t} |\psi(x, t)|^2 + \frac{\hbar}{2mi} \left[\underbrace{\psi^* \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial \psi^*}{\partial x} \frac{\partial \psi}{\partial x}}_{\frac{\partial}{\partial x}(\psi^* \frac{\partial \psi}{\partial x})} - \underbrace{\frac{\partial^2 \psi^*}{\partial x^2} \psi - \frac{\partial \psi^*}{\partial x} \frac{\partial \psi}{\partial x}}_{\frac{\partial}{\partial x}(\psi \frac{\partial \psi^*}{\partial x})} \right] &= 0 \\ \Rightarrow \frac{\partial}{\partial t} |\psi(x, t)|^2 + \frac{\partial}{\partial x} \underbrace{\left[\frac{\hbar}{2mi} \left(\psi^*(x, t) \frac{\partial \psi(x, t)}{\partial x} - \frac{\partial \psi^*(x, t)}{\partial x} \psi(x, t) \right) \right]}_{j(x, t)} &= 0 \\ \Rightarrow \frac{\partial}{\partial t} |\psi(x, t)|^2 + \frac{\partial}{\partial x} j(x, t) &= 0 \end{aligned} \quad (3.108)$$

L'equazione (3.108) è detta **equazione di continuità** per la probabilità. Si tratta di una relazione analoga a quella già ricavata per i fluidi o per la conservazione locale di carica dalle leggi di Maxwell. Intuitivamente, significa che la *probabilità* che una particella si trovi in un dato punto x al tempo t (codificata dal modulo quadro della sua funzione d'onda $\psi(x, t)$) può variare solamente a seguito della variazione delle probabilità dei punti circostanti. In altre parole, se la probabilità di trovarsi a $x = 0$ aumenta, quella nei punti **limitrofi** deve essere diminuita, cioè la probabilità “non fa salti”, ma si propaga come un *fluido* che scorre e si addensa nelle regioni “più probabili”. La variazione di probabilità in una regione sarà quindi pari al “flusso” di probabilità che attraversa il suo bordo: in particolare la probabilità salirà se il flusso è entrante (negativo), e scenderà se il flusso è uscente (positivo) e rimarrà costante se il flusso è nullo. Tale idea, espressa matematicamente in forma differenziale, è in sintesi quella dell'equazione (3.108).

*Equazione di
continuità*

Nel caso particolare di una autofunzione di H :

$$\varphi_{\mathcal{E}}(x, t) = \exp\left(-\frac{i}{\hbar} \mathcal{E} t\right) \varphi_{\mathcal{E}}(x) \Rightarrow |\varphi_{\mathcal{E}}(x, t)|^2 = \underbrace{\left| \exp\left(-\frac{i}{\hbar} \mathcal{E} t\right) \right|}_{=1} |\varphi_{\mathcal{E}}(x)|^2 = |\varphi_{\mathcal{E}}(x)|^2$$

Perciò la densità di probabilità associata alle autofunzioni è *costante* nel tempo (e infatti le autofunzioni di H sono dette *stati stazionari*). Sostituendo nell'equazione di continuità:

$$\frac{\partial}{\partial t} |\varphi_{\mathcal{E}}(x, t)|^2 = 0 = \frac{d}{dx} j(x)$$

Dove la corrente di probabilità $j(x)$ ha la forma vista nei passaggi per giungere a

*(Lezione 18 del
8/11/2018)*

(3.108):

$$j(x) = \frac{\hbar}{2mi} \left(\varphi_{\mathcal{E}}(x)^* \frac{d\varphi_{\mathcal{E}}(x)}{dx} - \frac{d\varphi_{\mathcal{E}}(x)^*}{dx} \varphi_{\mathcal{E}}(x) \right)$$

In cui sostituiamo la $\varphi_{\mathcal{E}}(x)$ ricavata in (3.101):

$$\varphi_{\mathcal{E}}(x) = \begin{cases} c_+^1 e^{ikx} + c_-^1 e^{-ikx} & x < 0 \\ c_+^2 e^{ik_2x} + c_-^2 e^{-ik_2x} & x \geq 0 \end{cases} \quad k = \frac{\sqrt{2m\mathcal{E}}}{\hbar}; \quad k_2 = \frac{\sqrt{2m(\mathcal{E} - \bar{V})}}{\hbar}$$

Svolgendo allora i conti per ricavare $j(x)$, per $x < 0$ otteniamo:

$$\begin{aligned} \varphi_{\mathcal{E}}(x) &= c_+^1 e^{ikx} + c_-^1 e^{-ikx} & \varphi_{\mathcal{E}}^*(x) &= c_+^{1*} e^{-ikx} + c_-^{1*} e^{ikx} \\ \varphi'_{\mathcal{E}}(x) &= ik(c_+^1 e^{ikx} - c_-^1 e^{-ikx}) & \varphi'^*_{\mathcal{E}}(x) &= ik(-c_+^{1*} e^{-ikx} + c_-^{1*} e^{ikx}) \end{aligned}$$

$$\begin{aligned} \Rightarrow j(x < 0) &= \frac{\hbar}{2mi} ik [(c_+^{1*} e^{-ikx} + c_-^{1*} e^{ikx})(c_+^1 e^{ikx} - c_-^1 e^{-ikx}) - \\ &\quad - (c_+^{1*} e^{-ikx} + c_-^{1*} e^{ikx})(c_+^1 e^{ikx} + c_-^1 e^{-ikx})] = \\ &= \frac{\hbar k}{2m} [c_+^{1*} c_+^1 - \cancel{c_+^{1*} c_-^1 e^{-2ikx}} + \cancel{c_-^{1*} c_+^1 e^{2ikx}} - c_-^{1*} c_-^1 \\ &\quad - (-c_+^{1*} c_+^1 - \cancel{c_+^{1*} c_-^1 e^{-2ikx}} + \cancel{c_-^{1*} c_+^1 e^{2ikx}} + c_-^{1*} c_-^1)] = \\ &= \frac{\hbar k}{2m} [2|c_+^1|^2 - 2|c_-^1|^2] = \frac{\hbar k}{m} (\underbrace{|c_+^1|^2}_{j_i} - \underbrace{|c_-^1|^2}_{j_r}) \end{aligned}$$

D'altro canto, per $x \geq 0$:

$$\begin{aligned} \varphi_{\mathcal{E}}(x) &= c_+^2 e^{ik_2x} & \varphi_{\mathcal{E}}^*(x) &= c_+^{2*} e^{-ik_2x} \\ \varphi'_{\mathcal{E}}(x) &= ik_2 c_+^2 e^{ik_2x} & \varphi'^*_{\mathcal{E}}(x) &= -ik_2 c_+^{2*} e^{-ik_2x} \end{aligned}$$

$$\begin{aligned} \Rightarrow j(x > 0) &= \frac{\hbar}{2mi} ik_2 [(c_+^{2*} e^{-ik_2x})(c_+^2 e^{ik_2x}) - (-c_+^{2*} e^{-ik_2x})(c_+^2 e^{ik_2x})] = \\ &= \frac{\hbar k_2}{2m} [2|c_+^2|^2] = \frac{\hbar k_2}{m} \underbrace{|c_+^2|^2}_{j_t} \end{aligned}$$

Ma dall'equazione di continuità sappiamo che:

$$\frac{d}{dx} j(x) = 0$$

ossia $j(x) \equiv j$ è costante e non dipende dalla x . Ma allora deve essere:

$$j(x < 0) = j(x > 0) \Rightarrow |c_+^1|^2 - |c_-^1|^2 = \frac{k_2}{k} |c_+^2|^2 \quad (3.109)$$

Definiamo allora il **coefficiente di riflessione** R come il rapporto tra l'intensità della corrente di probabilità in $x < 0$ che va verso destra ($\hbar k |c_+^1|^2 / m$) e quella che va verso sinistra ($\hbar k |c_-^1|^2 / m$):

$$R = \left| \frac{j_r}{j_i} \right| = \left| \frac{c_-^1}{c_+^1} \right|^2$$

Il **coefficiente di trasmissione** T è invece dato dal rapporto tra l'intensità di corrente in $x > 0$ che ha superato il gradino e va verso destra ($\hbar k_2 |c_+^2|^2/m$) e quella che in $x < 0$ si avvicina al gradino andando verso destra ($\hbar k |c_+^1|^2/m$):

$$T = \left| \frac{j_t}{j_i} \right| = \frac{k_2}{k} \left| \frac{c_+^2}{c_+^1} \right|^2$$

L'equazione (3.109) collega i due termini R e T :

$$|c_+^1|^2 - |c_-^1|^2 = \frac{k_2}{k} |c_+^2|^2 \Rightarrow 1 = \underbrace{\left| \frac{c_-^1}{c_+^1} \right|^2}_R + \underbrace{\frac{k_2}{k} \left| \frac{c_+^2}{c_+^1} \right|^2}_T \Rightarrow 1 = R + T$$

Per la scelta di $c_+^1 = e^{-ikx_0}$, con $|c_+^1| = 1$, dalle espressioni dei coefficienti ricavate in (3.104) si ha:

$$R = \left(\frac{k - k_2}{k + k_2} \right)^2; \quad T = \frac{4k^2}{(k + k_2)^2}$$

(Lezione 18 del
8/11/2018)

3.9.5 Barriera di potenziale

Consideriamo ora il caso di una *barriera di potenziale* data da $V(x) = \bar{V} \chi_{[-\frac{a}{2}, \frac{a}{2}]}$

L'effetto quantistico più rilevante si ha per $\mathcal{E} < \bar{V}$. In tal caso l'onda evanescente di probabilità vista nel gradino dà luogo all'**effetto tunnel**, cioè a una probabilità finita di trovare la particella che proviene dalla regione 1 nella regione 3, passando in una regione 2 proibita classicamente (energia cinetica < 0).

Nella risoluzione partiamo tuttavia dal caso $\mathcal{E} > \bar{V}$, che potremo poi adattare a quello $\mathcal{E} < \bar{V}$ che ci interessa.

Caso: $\mathcal{E} > \bar{V}$

Energia più alta del gradino

Se $\mathcal{E} > \bar{V}$ la soluzione dell'equazione agli autovalori è, nelle regioni 1 e 2 esattamente la stessa già trovata per il gradino, mentre la regione 3 è analoga alla 1, visto che in entrambe $V(x) = 0$. Si ha quindi:

1. Autofunzioni
 $\varphi_{\mathcal{E} > \bar{V}}(x)$

$$\varphi_{\mathcal{E}}(x) = \begin{cases} c_+^1 e^{ik_1 x} + c_-^1 e^{-ik_1 x} & x < -\frac{a}{2} \quad (1) \\ c_+^2 e^{ik_2 x} + c_-^2 e^{-ik_2 x} & -\frac{a}{2} < x < \frac{a}{2} \quad (2) \\ c_+^3 e^{ik_3 x} + c_-^3 e^{-ik_3 x} & x > \frac{a}{2} \quad (3) \end{cases} \quad (3.110)$$

con:

$$k_1 = \frac{\sqrt{2m\mathcal{E}}}{\hbar}; \quad k_2 = \frac{\sqrt{2m(\mathcal{E} - \bar{V})}}{\hbar}; \quad k_3 = k_1$$

Avremo quindi due condizioni di **raccordo**, una in $-\frac{a}{2}$ e una in $\frac{a}{2}$, che vanno applicate sia per $\varphi_{\mathcal{E}}$ che per la derivata $\varphi'_{\mathcal{E}}$, per un totale di 4 equazioni di raccordo. Dato che nella soluzione compaiono 6 costanti, la **degenerazione** degli autovalori

2. Raccordo delle
soluzioni

sarà $6 - 4 = 2$.

L'equazione di raccordo a $-\frac{a}{2}$ è:

$$\begin{cases} \varphi_{\mathcal{E}}(-\frac{a}{2}^-) = \varphi_{\mathcal{E}}(-\frac{a}{2}^+) \\ \varphi'_{\mathcal{E}}(-\frac{a}{2}^-) = \varphi'_{\mathcal{E}}(\frac{a}{2}^+) \end{cases} \Rightarrow \begin{cases} c_+^1 e^{-ik_1 \frac{a}{2}} + c_-^1 e^{+ik_1 \frac{a}{2}} = c_+^2 e^{-ik_2 \frac{a}{2}} + c_-^2 e^{ik_2 \frac{a}{2}} \\ ik_1 (c_+^1 e^{-ik_1 \frac{a}{2}} - c_-^1 e^{ik_1 \frac{a}{2}}) = ik_2 (c_+^2 e^{-ik_2 \frac{a}{2}} - c_-^2 e^{ik_2 \frac{a}{2}}) \end{cases}$$

Dividendo la seconda equazione per k_2 , per poi sommare/sottrarre membro a membro, si possono ricavare c_+^2 e c_-^2 :

$$\begin{cases} 2c_+^2 = c_+^1 \left(1 + \frac{k_1}{k_2}\right) e^{i\frac{a}{2}(k_2-k_1)} + c_-^1 \left(1 - \frac{k_1}{k_2}\right) e^{i\frac{a}{2}(k_2+k_1)} \\ 2c_-^2 = c_+^1 \left(1 - \frac{k_1}{k_2}\right) e^{-i\frac{a}{2}(k_2+k_1)} + c_-^1 \left(1 + \frac{k_1}{k_2}\right) e^{-i\frac{a}{2}(k_2-k_1)} \end{cases} \quad (3.111)$$

Possiamo allora riscrivere il sistema in forma matriciale:

$$\begin{pmatrix} c_+^2 \\ c_-^2 \end{pmatrix} = M \begin{pmatrix} c_+^1 \\ c_-^1 \end{pmatrix} \quad (3.112)$$

dove M è una matrice (2×2) data da:

$$M = \frac{1}{2} \begin{pmatrix} \left(1 + \frac{k_1}{k_2}\right) e^{i(k_2-k_1)\frac{a}{2}} & \left(1 - \frac{k_1}{k_2}\right) e^{i(k_2+k_1)\frac{a}{2}} \\ \left(1 - \frac{k_1}{k_2}\right) e^{-i(k_2+k_1)\frac{a}{2}} & \left(1 + \frac{k_1}{k_2}\right) e^{-i(k_2-k_1)\frac{a}{2}} \end{pmatrix}$$

Per quanto riguarda le equazioni di raccordo per $\frac{a}{2}$ basta riadattare quanto ottenuto, partendo dal cambiare il segno alle a , ossia sostituendo $a \leftrightarrow -a$ nelle (3.111).

Nel nuovo caso avremo le c_{\pm}^2 a sinistra, e le c_{\pm}^3 a destra.

Essendo però $k_3 = k_1$, stavolta k_2 compare a sinistra e k_1 compare a destra: ciò non è altro che il caso *simmetrico* a prima, e quindi nelle (3.111) basta scambiare $k_1 \leftrightarrow k_2$.

Possiamo perciò porre anche questo sistema in forma matriciale (stiamo sempre ricavando i termini a destra, che in questo caso sono le c_{\pm}^3):

$$\begin{pmatrix} c_+^3 \\ c_-^3 \end{pmatrix} = M' \begin{pmatrix} c_+^2 \\ c_-^2 \end{pmatrix} \quad (3.113)$$

dove la matrice M' si ottiene da M dalle sostituzioni elencate di sopra:

$$M' = M[a \leftrightarrow -a, k_2 \leftrightarrow k_1]$$

Combinando (3.112) e (3.113) otteniamo:

$$\begin{pmatrix} c_+^3 \\ c_-^3 \end{pmatrix} = M' M \begin{pmatrix} c_+^1 \\ c_-^1 \end{pmatrix}; \quad \begin{pmatrix} c_+^1 \\ c_-^1 \end{pmatrix} = (M' M)^{-1} \begin{pmatrix} c_+^3 \\ c_-^3 \end{pmatrix} \quad (3.114)$$

Che è la relazione che ci interessa, dato che il nostro obiettivo finale sarà quello di calcolare i coefficienti di *trasmissione* e *riflessione*, in modo da confrontare il

comportamento del sistema con l'analogo classico, e perciò non ci interessa il comportamento della regione di mezzo, ma solo ciò che avviene prima o dopo di essa.

Le costanti che ci servono possono essere ricavate dalle relazioni (3.114), ma farlo direttamente è una procedura lunga e tediosa. Possiamo di gran lunga semplificarla notando che se $\varphi_{\mathcal{E}}(x)$ è soluzione dell'equazione agli autovalori:

$$H\varphi_{\mathcal{E}}(x) = \mathcal{E}\varphi_{\mathcal{E}}(x)$$

Allora (essendo H autoaggiunto e \mathcal{E} reale), anche $\varphi_{\mathcal{E}}^*(x)$ è soluzione:

$$H\varphi_{\mathcal{E}}^*(x) = \mathcal{E}\varphi_{\mathcal{E}}^*(x)$$

Nel caso in questione vi è una particolare *simmetria* tra le due soluzioni complesse-coniugate. Se calcoliamo $\varphi_{\mathcal{E}}^*(x)$:

$$\begin{aligned}\varphi_{\mathcal{E}}^*(x) &= (c_+^1 e^{ik_1 x} + c_-^1 e^{-ik_1 x})^* = c_+^{1*} e^{-ik_1 x} + c_-^{1*} e^{ik_1 x} = \\ &= (c_-^{1*} e^{ik_1 x} + c_+^{1*} e^{-ik_1 x})\end{aligned}$$

Notiamo che possiamo passare da $\varphi_{\mathcal{E}}(x)$ alla coniugata $\varphi_{\mathcal{E}}^*(x)$ semplicemente sostituendo i parametri evidenziati:

$$c_+^1 \leftrightarrow c_-^{1*} \quad c_-^1 \leftrightarrow c_+^{1*} \quad (3.115)$$

Ciò ci permette di usare (3.112), (3.113), (3.114) sia per $\varphi_{\mathcal{E}}(x)$ che per $\varphi_{\mathcal{E}}^*(x)$ mantenendo le **stesse** matrici M e M' in entrambi i casi, e utilizzando le sostituzioni per “convertire” le relazioni per la $\varphi_{\mathcal{E}}^*(x)$ in relazioni per una $\varphi_{\mathcal{E}}(x)$ con opportuni coefficienti. Perciò le matrici M e M' devono avere una qualche *struttura particolare* per accomodare questa simmetria. Vedremo ora come sfruttarla per dimezzare il numero di conti necessari.

Concentriamoci sulla relazione (3.114) e diamo un nome ai termini delle matrici:

$$M'M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; \quad (M'M)^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

La relazione (3.114) esplicitata diviene:

$$\begin{pmatrix} c_+^3 \\ c_-^3 \end{pmatrix} = M'M \begin{pmatrix} c_+^1 \\ c_-^1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} c_+^1 \\ c_-^1 \end{pmatrix} \quad (3.116)$$

Effettuando le sostituzioni (3.115) otteniamo poi:

$$\begin{pmatrix} c_-^{3*} \\ c_+^{3*} \end{pmatrix} = M'M \begin{pmatrix} c_-^{1*} \\ c_+^{1*} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} c_-^{1*} \\ c_+^{1*} \end{pmatrix} \quad (3.117)$$

Esplicitiamo ora una delle due relazioni da (3.116):

$$c_-^3 = cc_+^1 + dc_-^1$$

Calcolo veloce di $M'M$ e $(M'M)^{-1}$

e la corrispettiva *coniugata* da (3.117):

$$c_-^{3*} = ac_-^{1*} + bc_+^{1*} \xrightarrow{*} c_-^3 = a^*c_-^1 + b^*c_+^1$$

Ma allora abbiamo ottenuto un'uguaglianza:

$$c_-^3 = \textcolor{yellow}{c} c_+^1 + \textcolor{blue}{d} c_-^1 = \textcolor{yellow}{b^*} c_+^1 \textcolor{blue}{a^*} c_-^1 \quad \forall c_-^1, c_+^1 \in \mathbb{C}$$

che è soddisfatta solamente se:

$$c = b^*; \quad d = a^*$$

Analogamente, ripetendo lo stesso ragionamento per $(M'M)^{-1}$ si giunge a risultati analoghi (qui ripetuti per riepilogo):

$$\begin{aligned} \begin{pmatrix} c_+^1 \\ c_-^1 \end{pmatrix} &= (M'M)^{-1} \begin{pmatrix} c_+^3 \\ c_-^3 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} c_+^3 \\ c_-^3 \end{pmatrix} \\ \begin{pmatrix} c_-^{1*} \\ c_+^{1*} \end{pmatrix} &= (M'M)^{-1} \begin{pmatrix} c_-^{3*} \\ c_+^{3*} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} c_-^{3*} \\ c_+^{3*} \end{pmatrix} \\ c_-^1 &= Cc_+^3 + Dc_-^3; \quad c_-^{1*} = Ac_-^{3*} + Bc_+^{3*} \xrightarrow{*} c_-^1 = A^*c_-^3 + B^*c_+^3 \\ C &= B^*; \quad D = A^* \end{aligned}$$

Sostituendo nelle espressioni esplicite per $M'M$, e $(M'M)^{-1}$:

$$M'M = \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix}; \quad (M'M)^{-1} = \begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} \quad (3.118)$$

Notando poi che $M'M$ ha $\det = 1$, applicando la formula per l'inversa di una matrice (2×2) :

$$M'M^{-1} = \begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} = \begin{pmatrix} a^* & -b \\ -b^* & a \end{pmatrix}$$

Basta per ciò calcolare due termini di $M'M$, per esempio a e b , per avere la visione completa. Partiamo scrivendo M e M' (che si ricava tramite le sostituzioni viste sopra):

$$\mathbf{M}' = \frac{1}{2} \begin{pmatrix} \left(1 + \frac{k_2}{k_1}\right) e^{i(k_2-k_1)\frac{a}{2}} & \left(1 - \frac{k_2}{k_1}\right) e^{-i(k_1+k_2)\frac{a}{2}} \\ \left(1 - \frac{k_2}{k_1}\right) e^{i(k_1+k_2)\frac{a}{2}} & \left(1 + \frac{k_2}{k_1}\right) e^{-i(k_2-k_1)\frac{a}{2}} \end{pmatrix} \quad (3.119)$$

$$\mathbf{M} = \frac{1}{2} \begin{pmatrix} \left(1 + \frac{k_1}{k_2}\right) e^{i(k_2-k_1)\frac{a}{2}} & \left(1 - \frac{k_1}{k_2}\right) e^{i(k_2+k_1)\frac{a}{2}} \\ \left(1 - \frac{k_1}{k_2}\right) e^{-i(k_2+k_1)\frac{a}{2}} & \left(1 + \frac{k_1}{k_2}\right) e^{-i(k_2-k_1)\frac{a}{2}} \end{pmatrix} \quad (3.120)$$

Dal prodotto scalare tra la prima riga di M' e la prima colonna di M otteniamo a :

$$\begin{aligned}
\mathbf{a} &= \frac{1}{4} \left(1 + \frac{k_2}{k_1}\right) \left(1 + \frac{k_1}{k_2}\right) e^{ia(k_2-k_1)} + \frac{1}{4} \left(1 - \frac{k_2}{k_1}\right) \left(1 - \frac{k_1}{k_2}\right) e^{-ia(k_1+k_2)} = \\
&= \frac{(k_1+k_2)^2 e^{ia(k_2-k_1)} - (k_1-k_2)^2 e^{-ia(k_1+k_2)}}{4k_1k_2} = \\
&= \frac{(k_1+k_2)^2 - (k_1-k_2)^2 e^{-2iak_2}}{4k_1k_2 e^{-ia(k_2-k_1)}} \quad (3.121)
\end{aligned}$$

Immediatamente abbiamo $A = a^*$:

$$\mathbf{A} = a^* = \frac{(k_1+k_2)^2 - (k_1-k_2)^2 e^{2iak_2}}{4k_1k_2 e^{-ia(k_1-k_2)}}$$

Fattorizzando l'esponenziale al denominatore e spezzando la frazione possiamo riscriverlo:

$$\begin{aligned}
A &= e^{iak_1} \left(\frac{(k_1+k_2)^2}{4k_1k_2} e^{-iak_2} - \frac{(k_1-k_2)^2}{4k_1k_2} e^{iak_2} \right) = \\
&= e^{iak_1} \left(\frac{(k_1+k_2)^2}{4k_1k_2} e^{-iak_2} - \frac{(k_1+k_2)^2 - 4k_1k_2}{4k_1k_2} e^{iak_2} \right) = \\
&= e^{iak_1} \left(\frac{-2i(k_1+k_2)^2}{4k_1k_2} \left[\frac{e^{-iak_2} - e^{iak_2}}{-2i} \right] + e^{iak_2} \right) = \\
&= e^{iak_1} \left(i \sin(k_2a) \left[-\frac{(k_1+k_2)^2}{2k_1k_2} + 1 \right] + \cos(k_2a) \right) = \\
&= e^{ik_1a} \left[\cos(k_2a) - i \frac{k_1^2 + k_2^2}{2k_1k_2} \sin(k_2a) \right] \quad (3.122)
\end{aligned}$$

Analogamente, il prodotto scalare tra la prima riga di M' e la seconda colonna di M restituisce b :

$$\begin{aligned}
\mathbf{b} &= \frac{1}{4} \left(1 + \frac{k_2}{k_1}\right) \left(1 - \frac{k_1}{k_2}\right) e^{ik_2a} + \frac{1}{4} \left(1 - \frac{k_2}{k_1}\right) \left(1 + \frac{k_1}{k_2}\right) e^{-ik_2a} = \\
&= -\frac{1}{4} \frac{k_1^2 - k_2^2}{k_1k_2} (e^{ik_2a} - e^{-ik_2a})
\end{aligned}$$

Da cui $B = -b$:

$$\begin{aligned}
\mathbf{B} &= \frac{1}{4} \frac{k_1^2 - k_2^2}{k_1k_2} (e^{ik_2a} - e^{-ik_2a}) = \\
&= \frac{2i}{4k_1k_2} \frac{k_1^2 - k_2^2}{2i} (e^{ik_2a} - e^{-ik_2a}) = \\
&= \frac{k_1^2 - k_2^2}{2k_1k_2} i \sin(k_2a)
\end{aligned}$$

A e B così derivati ci sono utili per determinare i parametri di *trasmissione* e *riflessione*.

3. Coefficienti di riflessione e trasmissione

Partiamo fissando $c_-^3 = 0$ poiché nel caso in esame non vi sono particelle che sorpassano il gradino “al contrario”, cioè da $3 \rightarrow 1$.

Avremo allora che, analogamente all’esempio precedente, i parametri sono definiti da:

$$\mathbb{T} = \left| \frac{c_+^3}{c_+^1} \right|^2; \quad \mathbb{R} = \left| \frac{c_-^1}{c_+^1} \right|^2$$

E fissiamo, come nel caso precedente, c_+^1 in modo che $|c_+^1| = 1$ e le formule si semplifichino. Introduciamo per semplicità la notazione $c_+^3 = t$ e $c_-^1 = r$, in modo che:

$$\mathbb{T} = |c_+^3|^2 \equiv |t|^2; \quad \mathbb{R} = |c_-^1|^2 \equiv |r|^2$$

Per trovare t e r così definiti possiamo usare una delle due relazioni (3.114), ma ci accorgiamo subito che quella con $(M'M)^{-1}$ è più comoda. L'altra, infatti, restituisce:

$$\begin{pmatrix} t \\ 0 \end{pmatrix} = M'M \begin{pmatrix} 1 \\ r \end{pmatrix} \Rightarrow \begin{cases} t = a + br \\ 0 = b^* + a^*r \end{cases}$$

con a e b che hanno una forma piuttosto complessa. Conviene piuttosto:

$$\begin{pmatrix} 1 \\ r \end{pmatrix} = (M'M)^{-1} \begin{pmatrix} t \\ 0 \end{pmatrix} \Rightarrow \begin{cases} 1 = At \\ r = B^* t \end{cases} \Rightarrow \begin{cases} t = \frac{1}{A} \\ r = \frac{B^*}{A} \end{cases}$$

Da cui ricaviamo direttamente i coefficienti desiderati:

Calcolo di \mathbb{T} , \mathbb{R}

$$\begin{aligned} \mathbb{T} &= \frac{1}{|A|^2} = \left(|e^{ik_1 a}|^2 \left| \cos(k_2 a) - i \frac{k_1^2 + k_2^2}{2k_1 k_2} \sin(k_2 a) \right|^2 \right)^{-1} = \\ &= \left(\cos^2(k_2 a) + \frac{(k_1^2 + k_2^2)}{4k_1^2 k_2^2} \sin^2(k_2 a) \right)^{-1} = \\ &= \frac{4k_1^2 k_2^2}{4k_1^2 k_2^2 \cos^2(k_2 a) + (k_1^2 + k_2^2)^2 \sin^2(k_2 a)} = \\ &= \frac{4k_1^2 k_2^2}{4k_1^2 k_2^2 \cos^2(k_2 a) + [(k_1^2 - k_2^2) + 4k_1^2 k_2^2] \sin^2(k_2 a)} \\ &= \frac{4k_1^2 k_2^2}{4k_1^2 k_2^2 + (k_1^2 - k_2^2)^2 \sin^2(k_2 a)} \\ \mathbb{R} &= \left| \frac{B}{A} \right|^2 = \frac{(k_1^2 - k_2^2)^2 \sin^2(k_2 a)}{4k_1^2 k_2^2} \frac{4k_1^2 k_2^2}{4k_1^2 k_2^2 + (k_1^2 - k_2^2)^2 \sin^2(k_2 a)} = \\ &= \frac{(k_1^2 - k_2^2) \sin^2 k_2 a}{4k_1^2 k_2^2 + (k_1^2 - k_2^2)^2 \sin^2(k_2 a)} = \frac{1}{1 + \frac{4k_1^2 k_2^2}{(k_1^2 - k_2^2)^2 \sin^2(k_2 a)}} \end{aligned}$$

E notiamo che vale (come ci aspettiamo):

$$\mathbb{T} + \mathbb{R} = 1$$

Studiamo il coefficiente di trasmissione \mathbb{T} . Si ha che:

Massimi e minimi di \mathbb{T}

- \mathbb{T} è **massimo** se il denominatore è minimo, ossia se $\sin^2(k_2 a) = 0$. In tal caso si ha una semplificazione, da cui $\mathbb{T}_{max} = 1$, ossia si ha certezza di trasmissione. I punti di massimo sono dati da $\sin(k_2 a) = 0$, per cui:

$$k_2 a = n\pi \Rightarrow \frac{2\pi}{\lambda_2} a = n\pi \Rightarrow 2a = n\lambda_2$$

dove $\lambda_2 = (2\pi)/k_2$ è la lunghezza d'onda (di de Broglie) associata alla particella quando è nella regione intermedia a potenziale \bar{V} . Notiamo quindi che la trasmissione è certa quando $2a$ è esattamente un multiplo intero di λ_2 . Se consideriamo un'onda che entra nella regione (2) e “rimbalza” tra le pareti di potenziale, $2a$ ammonta proprio alla lunghezza del percorso di *andata e ritorno*. La condizione $2a = n\lambda_2$ è quindi quella che permette l'esistenza di *onde stazionarie* all'interno del potenziale rettangolare. Tali onde sono *in fase* con l'onda uscente, che quindi viene *favorita*.

- \mathbb{T} è **minimo** se il denominatore è massimo, ossia se $\sin^2(k_2 a) = 1$. In tal caso il minimo è:

$$\mathbb{T}_{min} = (1 - \mathbb{R}) = 1 - \frac{(k_1^2 - k_2^2)^2}{(k_1^2 + k_2^2)^2} \quad (3.123)$$

Tale minimo è raggiunto quando lo spessore a del gradino di potenziale è tale che il $\sin(k_2 a) = \pm 1$, ossia quando:

$$\begin{aligned} k_2 a = \frac{\pi}{2} + \pi n = (2n + 1) \frac{\pi}{2} &\Rightarrow a = \frac{(2n + 1) \pi}{k_2} \frac{1}{2} \stackrel{(a)}{=} \frac{\lambda_2 (2n + 1) \pi}{2\pi} \frac{1}{2} = \\ &= \lambda_2 \left(\frac{1}{4} + \frac{n}{2} \right) \end{aligned}$$

dove in (a) si è usata la relazione $\lambda_2 = (2\pi)/k_2$.

Notiamo allora che in questo caso si ha:

$$2a = \frac{\lambda_2}{2} (2n + 1); \quad n \in \mathbb{N}$$

ossia la probabilità di trasmissione è minima quando all'interno del potenziale rettangolare *non* possono formarsi onde stazionarie.

Figura 3.5: Grafico di \mathbb{T} in funzione della larghezza a della barriera

Figura 3.6: Valore assoluto dell'autofunzione $|\psi_{\mathcal{E}}(x)|$ per valori di \mathcal{E} tali che \mathbb{T} sia massimo/minimo ($a = 10$, $\hbar = 1$, $m = 1$, $\bar{V} = 1$)

A differenza del caso classico, dove per $\mathcal{E} > \bar{V}$ la trasmissione è sempre certa, qui a seconda del valore specifico di k_2 e della larghezza a del potenziale c'è una certa probabilità di **riflessione**, che è nulla solo in determinate circostanze.

4. *Confronto con il caso classico*

L'effetto **non classico** è **dominante** quando $\mathbb{T}_{min} \ll 1$. Sostituendo le espressioni per k_1 e k_2 in (3.123) otteniamo:

$$\begin{aligned} k_1^2 - k_2^2 &= \frac{2m\bar{V}}{\hbar^2}; \quad k_1^2 + k_2^2 = \frac{2m}{\hbar^2}(2\mathcal{E} - \bar{V}) \\ \Rightarrow \mathbb{T}_{min} &= 1 - \left(\frac{\bar{V}}{2\mathcal{E} - \bar{V}} \right)^2 \end{aligned} \quad (3.124)$$

Essendo $\mathbb{T}_{min} > 0$, i valori più bassi sono raggiunti quando il fattore evidenziato è vicino a 1:

$$\frac{\bar{V}}{2\mathcal{E} - \bar{V}} \sim 1 \Rightarrow \bar{V} \approx 2\mathcal{E} - \bar{V} \Rightarrow \bar{V} \approx \mathcal{E}$$

Perciò se l'energia \mathcal{E} della particella è vicina a quella della *barriera* \bar{V} , piccole variazioni di \mathcal{E} o della larghezza a fanno oscillare *di molto* la probabilità di trasmissione, che va dalla quasi impossibilità alla certezza.

Dividendo per \bar{V} sopra e sotto nella (3.124) otteniamo:

$$\mathbb{T}_{min} = 1 - \left(\frac{1}{2\mathcal{E}/\bar{V} - 1} \right)^2$$

Perciò per $\mathcal{E}/\bar{V} \rightarrow +\infty$ si ha che $\mathbb{T}_{min} \rightarrow 1$, e dato che $\mathbb{T}_{max} \equiv 1$ si avrà in generale $\mathbb{T} \rightarrow 1$. Per le alte energie, perciò, si ritrova il caso classico in cui la particella energetica *oltrepassa sempre* il potenziale.

Nell'esempio precedente del gradino di lunghezza infinita avevamo notato la presenza di un *ritardo* nella riflessione/trasmissione. Proviamo a stimarlo anche in questo caso.

Partiamo determinando il tempo di attraversamento della regione 2 (per $\mathcal{E} > \bar{V}$). Nel caso **classico**, la particella, di momento iniziale p_0 , percorre una distanza a ad una velocità v_2 , per cui il tempo di attraversamento τ è dato da:

$$\tau_{classico} = \frac{a}{v_2} = \frac{a m}{\sqrt{p_0^2 - 2m\bar{V}}} \quad (3.125)$$

Se il superamento avviene a $t = 0$, il moto della particella dopo l'attraversamento è allora dato da:

$$\mathbf{x}_{cl.}(\mathbf{t}) = x_0 + a + \frac{p_0}{m}(t - \tau_{cl.}) \quad (3.126)$$

Esaminiamo ora il caso **quantistico**. Come nell'esempio precedente descriveremo la particella con un pacchetto d'onda *piccato*, nello spazio dei momenti, attorno a $k_0 = \sqrt{2m\mathcal{E}}/\hbar$, con il supporto Δ di $\tilde{f}_{k_0}(k)$ tale che:

$$\mathcal{E}(k) = \frac{\hbar^2 k^2}{2m} > \bar{V} \quad \forall k \in \Delta$$

Concentriamoci sulla regione (3) successiva all'attraversamento, dove $x > a/2$. Se eliminiamo le soluzioni corrispondenti a particelle che giungono da $+\infty$ ponendo $c_-^3 = 0$, l'autofunzione $\varphi_{\mathcal{E}}(x)$ ha la forma:

$$\varphi_{\mathcal{E} > \bar{V}}(x) = c_+^3 e^{ikx}$$

5. Ritardo nella trasmissione

Ritardo in MC

Moto in MC

Calcoli in MQ

dove siamo passati da k_1 a k per alleggerire la notazione (dato che ora k_1 non è più unico, ma assume i valori nel *range* del supporto di $\tilde{f}_{k_0}(k)$).

L'espressione per il pacchetto d'onda diviene quindi:

$$\psi(\mathbf{x}, \mathbf{0}) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} dk \tilde{f}_{k_0}(k) \underbrace{c_+^3 e^{ikx}}_{\varphi_{\mathcal{E}}(x)} \quad x > +\frac{a}{2} \quad (3.127)$$

Conviene allora porre il coefficiente c_+^3 in forma esponenziale. Ricordando la relazione (3.114) e la (3.118), si ha che:

$$\begin{pmatrix} c_+^1 \\ c_-^1 \end{pmatrix} = \begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} \begin{pmatrix} c_+^3 \\ 0 \end{pmatrix} \Rightarrow c_+^1 = A c_+^3 \Rightarrow c_+^3 = \frac{1}{A} c_+^1$$

Ponendo A in forma esponenziale:

$$\mathbf{A} = e^{ika} \left[\cos(k_2 a) - i \frac{k^2 + k_2^2}{2k k_2} \sin(k_2 a) \right] = |A| e^{i\theta(k)}$$

Dove $\theta(k)$ è data da⁶⁹:

$$\theta(\mathbf{k}) = \arctan \left(\frac{\text{Im}(A)}{\text{Re}(A)} \right) = ka - \arctan \left(\frac{k^2 + k_2^2}{2k k_2} \tan(k_2 a) \right) \quad (3.128)$$

E k_2 è una funzione di k data da:

$$\mathbf{k}_2 = \frac{\sqrt{2m(\mathcal{E} - \bar{V})}}{\hbar} = \sqrt{k^2 - \bar{k}^2}; \quad \bar{k} = \frac{\sqrt{2m\bar{V}}}{\hbar} \quad (3.129)$$

Possiamo allora scrivere c_+^3 in forma esponenziale:

$$c_+^3 = \frac{1}{|A| e^{i\theta(k)}} c_+^1 = \underbrace{\frac{1}{|A|}}_{\sqrt{\mathbb{T}}} e^{-i\theta(k)} c_+^1$$

Ponendo $c_+^1 = e^{-ikx_0}$ si ha che il pacchetto d'onda a $t = 0$ è centrato in x_0 (come dimostrato nell'esempio precedente).

Inseriamo infine l'**evoluzione temporale**, per cui $\varphi_{\mathcal{E}}(x) \rightarrow \varphi_{\mathcal{E}}(x) e^{-i\omega t}$, ricordando:

$$\omega \equiv \frac{\mathcal{E}}{\hbar} = \frac{k^2 \hbar}{2m}$$

Sostituendo allora quanto appena trovato in (3.127) otteniamo:

$$\psi(x, 0) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} dk \tilde{f}_{k_0}(k) \sqrt{\mathbb{T}(k)} e^{i[k(x-x_0) - \theta(k) - \omega t]}$$

La posizione del picco si ottiene, come già visto, linearizzando l'esponenziale attorno a k_0 , da cui si ha:

Moto in MQ

⁶⁹ Ricordiamo che *ovviamente* nel prodotto di due numeri complessi in forma esponenziale gli angoli *si sommano*

$$\mathbf{x}_q(t) = x_0 + \left. \frac{d\omega}{dk} \right|_{k=k_0} t + \left. \frac{d\theta}{dk} \right|_{k=k_0} \quad (3.130)$$

Sappiamo già che, come nell'esempio precedente:

$$\left. \frac{d\omega}{dk} \right|_{k=k_0} = \frac{p_0}{m}$$

Possiamo allora confrontare le equazioni del moto nel caso *classico* (3.126) e *quantistico* (3.130):

$$\begin{aligned} x(t) &= x_0 + \frac{p_0}{m} t + a - \frac{p_0}{m} \tau_{cl.} && \text{Classico} \\ x(t) &= x_0 + \frac{p_0}{m} t + \left. \frac{d\theta}{dk} \right|_{k=k_0} && \text{Quantistico} \end{aligned}$$

Partendo dalla corrispondenza dei due termini evidenziati, definiamo τ_q quel valore di tempo che, sostituito al posto di $\tau_{cl.}$, rende i due termini uguali (ed è quindi il corrispettivo quantistico del *ritardo di trasmissione* classico):

$$a - \frac{p_0}{m} \tau_q = \left. \frac{d\theta}{dk} \right|_{k=k_0} \Rightarrow \tau_q = \frac{m}{p_0} \left(a - \left. \frac{d\theta}{dk} \right|_{k=k_0} \right) \quad (3.131)$$

Tempo di
attraversamento
in MQ

Resta solo da calcolare il termine $d\theta/dk$, cosa che faremo solo nei casi speciali in cui il coefficiente di trasmissione è *massimo* o *minimo*.

τ_q per \mathbb{T}_{max} e
 \mathbb{T}_{min}

- Per \mathbb{T}_{max} si ha $\sin(k_2 a) = 0$, da cui $\tan(k_2 a) = 0$ e $\cos(k_2 a) = \pm 1 \Rightarrow \cos^2(k_2 a) = 1$. Ciò semplifica di molto il calcolo della derivata di (3.128), che (ricordando che anche k_2 è funzione di k , come visto in (3.129)) conduce a:

$$\left. \frac{d\theta(k)}{dk} \right|_{\sin(k_2 a)=0} = a - a \frac{k^2 + k_2^2}{2k k_2} \frac{dk_2(k)}{dk}$$

Notiamo che:

$$\begin{aligned} \frac{dk_2(k)}{dk} &= \frac{k}{\sqrt{k^2 - \bar{k}^2}} \quad k_0 = \frac{\sqrt{2m\mathcal{E}}}{\hbar}; \quad \bar{k} = \frac{\sqrt{2m\bar{V}}}{\hbar}; \quad p_0 = \sqrt{2m\mathcal{E}} \\ \frac{dk_2(k_0)}{dk} &= \frac{k_0}{\sqrt{k_0^2 - \bar{k}^2}} = \frac{\frac{\sqrt{2m\mathcal{E}}}{\hbar}}{\sqrt{\frac{2m\mathcal{E}}{\hbar^2} - \frac{2m\bar{V}}{\hbar^2}}} = \frac{p_0}{\sqrt{p_0^2 - 2m\bar{V}}} \end{aligned}$$

Sostituendo quanto trovato in (3.131) otteniamo:

$$\begin{aligned} \tau_q &= \frac{m}{p_0} \left(a - \frac{k_0^2 + k_2^2}{2k_0 k_2} \frac{dk_2(k_0)}{dk} \right) = \frac{m}{p_0} \left(\frac{k_0^2 + k_2^2}{2k_0 k_2} \frac{a}{\sqrt{p_0^2 - 2m\bar{V}}} \right) = \\ &= \frac{\frac{k_0^2 + k_2^2(k_0)}{2k_0 k_2(k_0)}}{\underbrace{\frac{m a}{\sqrt{p_0^2 - 2m\bar{V}}}}_{\tau_{cl.} \text{ (3.125)}}} \quad (3.132) \end{aligned}$$

Il termine evidenziato generalmente non è pari a 1, perciò $\tau_q \neq \tau_{cl}$. In effetti, per $k_2 \ll k_0$, ossia quando $\mathcal{E} - \bar{V} \ll \mathcal{E}$, o $\mathcal{E} \sim \bar{V}$, tale fattore è $\gg 1$.

Perciò una particella quantistica che ha poco più dell'energia necessaria per superare la barriera impiega molto più tempo per farlo rispetto al caso di una particella classica (di nuovo, l'effetto quantistico è massimo quando energia della particella e della barriera sono simili).

- Nel caso di \mathbb{T}_{min} , invece, $\sin^2(k_2 a) = 1$, da cui $\cos(k_2 a) = 0$ e $|\tan(k_2 a)| = +\infty$. Svolgendo i conti:

$$\begin{aligned} \frac{d\theta(k)}{dk} &= a - \frac{1}{\tan^2(k_2 a)} \frac{1}{\frac{1}{\tan^2(k_2 a)} + \left(\frac{k^2 + k_2^2}{2kk_2}\right)^2} \left[\frac{d}{dk} \left(\frac{k^2 + k_2^2}{2kk_2} \right) \tan(k_2 a) + \right. \\ &\quad \left. + \left(\frac{k^2 + k_2^2}{2kk_2} \right) \frac{1}{\cos^2(k_2 a)} a \frac{dk_2(k)}{dk} \right] = \\ &= \overbrace{a - \frac{\cos^2(k_2 a)}{\sin^2(k_2 a)} \frac{1}{\left(\frac{1}{\tan^2(k_2 a)}\right) + \left(\frac{k^2 + k_2^2}{2kk_2}\right)^2} \frac{d}{dk} \left(\frac{k^2 + k_2^2}{2kk_2} \right) \frac{\sin(k_2 a)}{\cos(k_2 a)}}^{\rightarrow 0} + \\ &\quad + \frac{\cos^2(k_2 a)}{\sin^2(k_2 a)} \frac{1}{\underbrace{\left(\frac{1}{\tan^2(k_2 a)}\right) + \left(\frac{k^2 + k_2^2}{2kk_2}\right)^2}_{\rightarrow 0}} \cancel{\left(\frac{k^2 + k_2^2}{2kk_2}\right)} \frac{1}{\cancel{\cos^2(k_2 a)}} a \frac{dk_2}{dk} = \\ &= a - a \frac{dk_2(k_0)}{dk} \frac{2k_0 k_2(k_0)}{k_0^2 + k_2^2(k_0)} \end{aligned}$$

Sostituendo nell'espressione per τ_q si ottiene:

$$\tau_q = \underbrace{\frac{ma}{p_0} \frac{dk_2(k_0)}{dk}}_{\tau_{cl.}} \frac{2k_0 k_2(k_0)}{k_0^2 + k_2^2(k_0)} \quad (3.133)$$

Il termine evidenziato in (3.133) è esattamente il reciproco di quello evidenziato in (3.132). Si avrà perciò che $\tau_q \ll \tau_{cl.}$ per particelle con energia $\mathcal{E} \sim \bar{V}$ quando la probabilità di trasmissione è minima.

Energia più bassa del gradino

Esaminiamo ora il caso quantistico analogo alla particella classica che non ha sufficiente energia per oltrepassare il potenziale rettangolare, ossia per cui $\mathcal{E} < \bar{V}$. Partendo dalle autofunzioni calcolate in (3.110), notiamo che per nelle regioni (1) e (3) non si ha alcuna modifica, mentre per la (2) l'equazione differenziale ha come soluzione:

$$c_+^2 e^{ik_2 x} + c_-^2 e^{-ik_2 x} \rightarrow c_+^2 e^{\chi_2 x} + c_-^2 e^{-\chi_2 x}$$

dove χ_2 è dato da:

$$\chi_2 = \frac{\sqrt{2m(\bar{V} - \mathcal{E})}}{\hbar} = ik_2 \Rightarrow k_2 = -i\chi_2$$

Caso: $\mathcal{E} < \bar{V}$

Basta allora sostituire nei conti precedenti $k_2 \rightarrow -i\chi_2$. In particolare, il coefficiente di trasmissione diviene:

$$\begin{aligned} \mathbb{T}_{\mathcal{E} > \bar{V}} &= \frac{4k_1^2 k_2^2}{4k_1^2 k_2^2 + (k_1^2 - k_2^2) \sin^2 k_2 a} \Rightarrow \mathbb{T}_{\mathcal{E} < \bar{V}} \stackrel{(a)}{=} \frac{-4k_1^2 \chi_2^2}{-4k_1^2 \chi_2^2 - (k_1^2 + \chi_2^2) \sinh^2(\chi_2 a)} = \\ &= \frac{4k_1^2 \chi_2^2}{4k_1^2 \chi_2^2 + (k_1^2 + \chi_2^2) \sinh^2(\chi_2 a)} \end{aligned}$$

Dove in (a) si è usata l'identità delle funzioni iperboliche:

$$\sinh(x) = -i \sin(ix) \Rightarrow \sin(k_2 a) \rightarrow \sin(-i\chi_2 a) = \frac{\overbrace{-i \sin(\chi_2 a)}^{\sinh(\chi_2 a)}}{i} = -i \sin(\chi_2 a)$$

Notiamo che $\mathbb{T} > 0$, e perciò anche se $\mathcal{E} < \bar{V}$ c'è sempre una probabilità, per la particella quantistica, di *superare* la barriera, contrariamente al caso classico: è come se la particella quantistica “passasse attraverso un tunnel scavato nel potenziale”. La trasmissione di una particella attraverso una regione di potenziale che classicamente sarebbe *proibita* viene detta **effetto tunnel**.

Effetto tunnel

Come ci si aspetta, la probabilità di trasmissione cresce monotonamente con \mathcal{E}/\bar{V} . Se invece consideriamo energie \mathcal{E} sempre più basse relativamente a \bar{V} , tale probabilità decresce esponenzialmente.

Infatti, $\chi_2 a \gg 1$ (ossia se il divario energetico è alto $\bar{V} \gg \mathcal{E}$ o la regione proibita è larga $a \gg 1$), possiamo approssimare \sinh :

$$\sinh(x) = \frac{e^x - e^{-x}}{2} \underset{x \gg 1}{\sim} \frac{e^x}{2} \Rightarrow \mathbb{T} \approx \frac{4k_1^2 \chi_2^2}{4k_1^2 \chi_2^2 + \frac{1}{4}e^{2\chi_2 a}} \approx \frac{16k_1^2 \chi_2^2}{(k_1^2 + \chi_2^2)^2} e^{-2\chi_2 a} \quad (3.134)$$

ossia la probabilità di trasmissione decade esponenzialmente con una velocità che dipende da a (lunghezza della buca) e da $\sqrt{\bar{V} - \mathcal{E}}$ (differenza di energie).

Figura 3.7: \mathbb{T} in funzione di \mathcal{E}/\bar{V} ($a = 10$, $\hbar = 1$, $m = 1$, $\bar{V} = 1$). Si noti che la funzione è > 0 anche per $\mathcal{E} < \bar{V}$, e ciò da luogo all'effetto tunnel. I massimi “periodici” corrispondono al caso della **risonanza**, e tale effetto *non classico* è notevolmente amplificato per valori di $\mathcal{E} \sim \bar{V}$, come discusso

Applicazioni del potenziale rettangolare

Il sistema costituito da una particella quantistica che si scontra contro un potenziale rettangolare è un modello ideale che, nella sua semplicità, può spiegare qualitativamente fenomeni ben più complessi, e ha applicazioni in tecniche avanzate di rilevazione. Facciamo qualche esempio.

Esempi di utilizzo del potenziale rettangolare

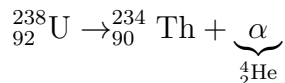
- **Radioattività** Consideriamo un nucleo di uranio ${}_{92}^{238}\text{U}$, dove i due numeretti

Range dei tempi di decadimento α

ad apice e pedice indicano rispettivamente il numero di nucleoni e di protoni:



Sperimentalmente si trova che tale atomo decade spontaneamente secondo la reazione:



Ossia, spontaneamente, una coppia di protoni e neutroni “si libera” dal nucleo.

Possiamo schematizzare il nucleo come una *buca di potenziale* molto profonda rispetto al *potenziale coulombiano* che vi è attorno ad essa⁷⁰.

La differenza di energia tra il fondo della buca e l'esterno è dell'ordine di 25 MeV.

In questa buca si trovano i nucleoni. Poiché, prima o poi, una particella α si libera dal nucleo, la sua funzione d'onda all'esterno oscilla all'infinito, e quindi tale $\varphi_{\mathcal{E}} \notin \mathcal{H}$, per cui gli autovalori \mathcal{E} di H delle soluzioni *uscenti* devono appartenere allo spettro continuo di H : $\mathcal{E} \in \sigma_C(H)$.

Perciò una particella all'interno del nucleo si comporta come un *pacchetto d'onda* (in analogia a quanto visto nell'esempio), che ha un'energia concentrata attorno ad un certo valore \mathcal{E}_0 , che si può misurare sperimentalmente una volta che tale particella è uscita dal nucleo e si trova sufficientemente lontana da ogni sorgente di campi elettrici. Da queste misure scopriamo che il salto tra \mathcal{E}_0 e il potenziale esterno è in generale dell'ordine di 4-9 MeV. Tale range relativamente piccolo deve però spiegare l'estrema varietà di *tempi di decadimento*, che vanno dai 10^{-7} s, ai 10^{10} anni. Ci aspettiamo quindi che una *piccola* variazione nel dislivello tra \mathcal{E}_0 e \bar{V} provochi una *grandissima* variazione nella *probabilità* di uscita dal nucleo: e ciò è proprio quello che si nota osservando la formula (3.134) e l'andamento *esponenziale* di tale probabilità in funzione del dislivello energetico.

- **Microscopio a effetto tunnel** (STM: Scanning Tunnelling Microscope).

*Microscopio a
effetto tunnel*

⁷⁰^Ciò si ha poiché l'interazione nucleare forte è “molto più forte” dell'interazione elettromagnetica, ma è anche a corto raggio, mentre quest'ultima ha raggio infinito.

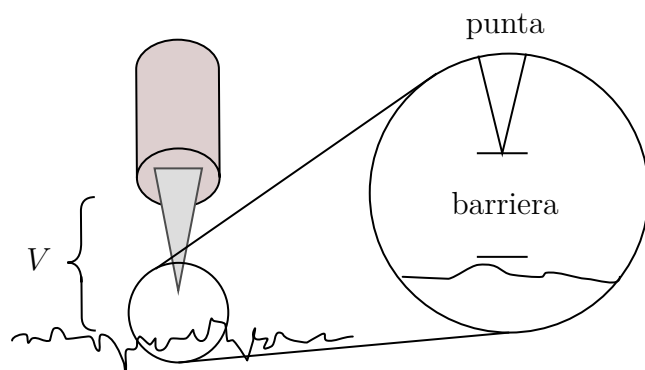


Figura 3.8: Schema del microscopio a effetto tunnel

Si può sfruttare l'effetto tunnel per ottenere microscopi sensibilissimi.

Consideriamo una punta conduttrice molto fine (di poche decine di atomi di larghezza), che viene avvicinata all'oggetto in esame.

Si applica una differenza di potenziale tra materiale (ricoperto di un sottile strato conduttore) e punta del microscopio. Si avvicina la punta al materiale, lasciando un certo spazio vuoto tra essa e la superficie. Tale distanza funge da *barriera di potenziale*: classicamente gli elettroni dell'oggetto rimarrebbero sull'oggetto, e quelli della punta sulla punta.

Tuttavia, per effetto tunnel, c'è una probabilità che gli elettroni passino dalla superficie alla punta, generando una corrente che può essere misurata.

Supponiamo ora di muovere la punta in modo da mantenere questa corrente *di tunnel* costante: per farlo dovremo periodicamente avvicinare o allontanare la punta stessa, a seconda delle caratteristiche della superficie in esame.

Otteniamo perciò una procedura per mappare le aree del materiale a densità di carica costante (*local density of states* - LDOS). Individuiamo quindi la "nuvola di carica" elettronica, e misurandola possiamo, in un certo senso, "vedere" gli atomi.

(Lezione 19 del
9/11/2018)

3.9.6 Buca finitamente profonda

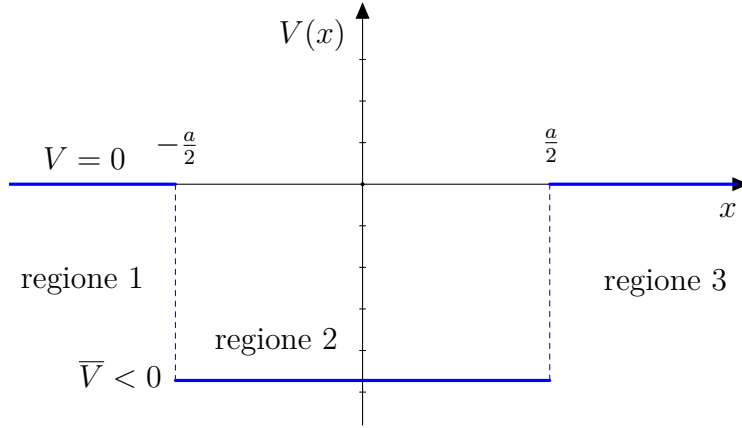


Figura 3.9: Potenziale della buca finitamente profonda

Consideriamo ora il caso della *buca di potenziale finita*, ovvero il sistema quantistico 1-dimensionale con potenziale

$$V(x) = \bar{V} \chi_{[-\frac{a}{2}, \frac{a}{2}]}(x)$$

Se $\mathcal{E} > 0$ la situazione è analoga a quella della barriera di potenziale per $\mathcal{E} > \bar{V}$, solo che in questo caso $\bar{V} < 0$ mentre per la barriera $\bar{V} > 0$. Per le autofunzioni, infatti, valgono le stesse espressioni viste in (3.110):

$$\varphi_{\mathcal{E}}(x) = \begin{cases} c_+^1 e^{ik_1 x} + c_-^1 e^{-ik_1 x} & x < -\frac{a}{2} \\ c_+^2 e^{ik_2 x} + c_-^2 e^{-ik_2 x} & -\frac{a}{2} < x < \frac{a}{2} \\ c_+^3 e^{ik_1 x} + c_-^3 e^{-ik_1 x} & x > \frac{a}{2} \end{cases}$$

$$k_1 = \frac{\sqrt{2m\mathcal{E}}}{\hbar}; \quad k_2 = \frac{\sqrt{2m(\mathcal{E} - \bar{V})}}{\hbar}$$

In entrambi i casi $\sigma(H)$ è continuo (dato che le soluzioni oscillano all'infinito, e quindi non sono in \mathcal{H}) con degenerazione 2 (vi sono 6 costanti e 4 condizioni di raccordo - 2 per la $\varphi_{\mathcal{E}}(x)$ e 2 per la derivata $\varphi'_{\mathcal{E}}(x)$, e perciò $6 - 4 = 2$ parametri liberi).

In particolare vale quanto ricavato per il coefficiente di trasmissione \mathbb{T} . Ritroviamo quindi il fenomeno di *risonanza di trasmissione*, per cui una particella viene trasmessa con *certezza* solo per certi valori di energia (tali che $\sin(k_2 a) = 0$), ma in generale $\mathbb{T} \neq 1$, e perciò può capitare che una particella *con abbastanza energia* per oltrepassare la barriera venga invece riflessa.

In particolare per determinate energie la particella *trascorre* più tempo nella regione della buca, e per altri di meno⁷¹. Questo fenomeno si può intuire se *estendiamo* gli stati discreti presenti nella buca “al di sopra” di essa. Quando una particella

⁷¹rispettivamente al tempo calcolato classicamente

ha un'energia che *corrisponderebbe* a quella di uno stato discreto “esteso”, allora tale particella “è in risonanza” e trascorre più tempo nella regione della buca.

D'altro canto, per $\bar{V} < \mathcal{E} < 0$ riadattiamo l'espressione per $\varphi_{\mathcal{E}}(x)$, seguendo la stessa idea vista per il caso della barriera rettangolare con $\mathcal{E} < \bar{V}$, solo che qui la modifica va fatta nelle regioni (1) e (3):

$$c_+^1 e^{ik_1 x} + c_-^1 e^{-ik_1 x} \rightarrow c_+^1 e^{\chi x} + c_-^1 e^{-\chi x}$$

$$\chi = \frac{\sqrt{-2m\mathcal{E}}}{\hbar} = ik_1 \Rightarrow k_1 = -i\chi$$

Effettuando la sostituzione, otteniamo:

$$\varphi_{\mathcal{E}}(x) = \begin{cases} c_+^1 e^{\chi x} + c_-^1 e^{-\chi x} & x < -\frac{a}{2} & (1) \\ c_+^2 e^{ik_2 x} + c_-^2 e^{-ik_2 x} & -\frac{a}{2} < x < \frac{a}{2} & (2) \\ c_+^3 e^{\chi x} + c_-^3 e^{-\chi x} & x > \frac{a}{2} & (3) \end{cases}$$

$$\chi = \frac{\sqrt{-2m\mathcal{E}}}{\hbar}; \quad k_2 = \frac{\sqrt{2m(\mathcal{E} - \bar{V})}}{\hbar}$$

Poiché qualsiasi autofunzione (sia nel caso appartenga ad \mathcal{H} che ad \mathcal{S}') non può divergere esponenzialmente, abbiamo immediatamente che $c_-^1 = c_+^3 = 0$.

Notiamo che $\varphi_{\mathcal{E}}(x)$ è la stessa vista per la barriera rettangolare con $\mathcal{E} > \bar{V}$, se non per la sostituzione $k_1 \leftrightarrow -i\chi$. Possiamo allora riadattare tutti i conti già fatti.

Partendo da (3.116) ed effettuando la sostituzione:

$$\begin{pmatrix} c_+^3 \\ c_-^3 \end{pmatrix} = \begin{pmatrix} 0 \\ c_-^1 \end{pmatrix} = (M' M) \begin{pmatrix} c_+^1 \\ c_-^1 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} c_+^1 \\ 0 \end{pmatrix}$$

Sviluppando il prodotto:

$$0 = c_+^3 = \alpha c_+^1 + \underbrace{\beta c_-^1}_0 = \alpha c_+^1 \Rightarrow \alpha = 0 \quad (3.135)$$

Ma abbiamo già calcolato un'espressione esplicita per $\alpha = (M' M)_{11}$ in (3.121), o, nella forma più semplificata, nel complesso coniugato di (3.122):

$$\alpha = e^{-ik_1 a} \left[\cos(k_2 a) + i \frac{k_1^2 + k_2^2}{2k_1 k_2} \sin(k_2 a) \right]$$

Effettuando la sostituzione $k_1 \leftrightarrow -i\chi$ giungiamo a:

$$\alpha = e^{-a\chi} \left[\cos(k_2 a) + \frac{\chi^2 - k_2^2}{2\chi k_2} \sin(k_2 a) \right]$$

Da (3.135) abbiamo quindi un'equazione:

$$\begin{aligned}
& e^{-a\chi} \left[\frac{2\chi k_2 \cos(k_2 a) + (\chi^2 - k_2^2) \sin(k_2 a)}{2\chi k_2} \right] = 0 \\
& \Rightarrow 2\chi k_2 \cos(k_2 a) + (\chi^2 - k_2^2) \sin(k_2 a) = 0 \\
& \Rightarrow 2\chi k_2 \frac{e^{ik_2 a} + e^{-ik_2 a}}{2} + (\chi^2 - k_2^2) \frac{e^{ik_2 a} - e^{-ik_2 a}}{2i} = 0 \\
& \Rightarrow e^{ik_2 a} (2i\chi k_2 + \chi^2 - k_2^2) + e^{-ik_2 a} (2i\chi k_2 - \chi^2 + k_2^2) = 0 \\
& \Rightarrow e^{ik_2 a} (\chi + ik_2)^2 = e^{-ik_2 a} (\chi - ik_2)^2 \\
& \Rightarrow e^{2ik_2 a} = \frac{(\chi - ik_2)^2}{(\chi + ik_2)^2} \Rightarrow \frac{\chi - ik_2}{\chi + ik_2} = \pm e^{ik_2 a} \tag{3.136}
\end{aligned}$$

Da (3.112):

$$\begin{pmatrix} c_+^2 \\ c_-^2 \end{pmatrix} = M \begin{pmatrix} c_+^1 \\ c_-^1 \end{pmatrix} = M \begin{pmatrix} c_+^1 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} c_+^2 = M_{11} c_+^1 \\ c_-^2 = M_{21} c_+^1 \end{cases}$$

E possiamo calcolare M_{11} e M_{21} da (3.119), effettuando la sostituzione $k_1 \leftrightarrow -i\chi$:

$$\begin{aligned}
M_{11} &= \frac{1}{2i} \left(\frac{\chi + ik_2}{k_2} \right) e^{ik_2 \frac{a}{2}} e^{-\chi \frac{a}{2}} \\
M_{21} &= -\frac{1}{2i} \left(\frac{\chi - ik_2}{k_2} \right) e^{-ik_2 \frac{a}{2}} e^{-\chi \frac{a}{2}}
\end{aligned}$$

Dividendo membro a membro:

$$\frac{c_-^2}{c_+^2} = \frac{M_{21}}{M_{11}} \frac{c_+^1}{c_+^1} = -\frac{(\chi - ik_2)}{(\chi + ik_2)} e^{-\cancel{ik_2 \frac{a}{2}}} e^{-\cancel{\chi \frac{a}{2}}} \stackrel{(3.136)}{=} \mp 1 \Rightarrow c_+^2 = \pm c_-^2$$

Abbiamo allora due possibilità:

- Prendendo il +, ossia con $c_+^2 = c_-^2$ otteniamo una funzione d'onda *pari* all'interno della buca finita. Infatti, per $-\frac{a}{2} < x < \frac{a}{2}$:

$$\varphi_{\mathcal{E}}(x) = c_+^2 e^{ik_2 x} + c_-^2 e^{-ik_2 x} = c_+^2 (e^{ik_2 x} + e^{-ik_2 x}) \propto \cos$$

e vale $\varphi_{\mathcal{E}}(x) = \varphi_{\mathcal{E}}(-x)$.

- Prendendo invece il -, con $c_+^2 = -c_-^2$, la $\varphi_{\mathcal{E}}(x)$ all'interno della buca è dispari:

$$\varphi_{\mathcal{E}}(x) = c_+^2 (e^{ik_2 x} - e^{-ik_2 x}) \propto \sin$$

con $\varphi_{\mathcal{E}}(x) = -\varphi_{\mathcal{E}}(-x)$.

Cerchiamo ora un'espressione per l'energia \mathcal{E} . Partiamo da (3.136) ed espandiamo l'esponenziale:

$$\frac{\chi - ik_2}{\chi + ik_2} = \pm e^{ik_2 a} = \pm (\cos(k_2 a) + i \sin(k_2 a))$$

Ricaviamo due equazioni reali prendendo parte *reale* e *immaginaria* di entrambi i membri.

Per la parte reale:

$$\operatorname{Re} \left(\frac{\chi - ik_2}{\chi + ik_2} \right) = \frac{1}{2} \left[\frac{\chi - ik_2}{\chi + ik_2} + \frac{\chi + ik_2}{\chi - ik_2} \right] = \frac{\chi^2 - k_2^2}{\chi^2 + k_2^2} = \pm \cos(k_2 a) \quad (3.137)$$

Mentre per quella immaginaria:

$$\begin{aligned} \operatorname{Im} \left(\frac{\chi - ik_2}{\chi + ik_2} \right) &= \frac{1}{2i} \left[\frac{\chi - ik_2}{\chi + ik_2} - \frac{\chi + ik_2}{\chi - ik_2} \right] = -\frac{2k_2 \chi}{\chi^2 + k_2^2} = \pm \sin(k_2 a) \\ \Rightarrow \frac{2k_2 \chi}{\chi^2 + k_2^2} &= \mp \sin(k_2 a) \end{aligned} \quad (3.138)$$

Ricordando ora i valori delle costanti:

$$\chi = \frac{\sqrt{-2m\mathcal{E}}}{\hbar}; \quad k_2 = \frac{\sqrt{2m(\mathcal{E} - \bar{V})}}{\hbar}$$

Possiamo esplicitare la (3.137):

$$\frac{\chi^2 - k_2^2}{\chi^2 + k_2^2} = \frac{-4m\mathcal{E} + 2m\bar{V}}{-2m\bar{V}} = -1 + 2 \left(\frac{\mathcal{E}}{\bar{V}} \right) = \pm \cos(k_2 a)$$

Possiamo semplificare tale equazione invertendo le formule di bisezione:

$$\begin{aligned} \sin^2 \left(k_2 \frac{a}{2} \right) &= \frac{1 - \cos(k_2 a)}{2} \Rightarrow -\cos(k_2 a) = -1 + 2 \sin^2 \left(k_2 \frac{a}{2} \right) \\ \cos^2 \left(k_2 \frac{a}{2} \right) &= \frac{1 + \cos(k_2 a)}{2} \Rightarrow +\cos(k_2 a) = -1 + 2 \cos^2 \left(k_2 \frac{a}{2} \right) \end{aligned}$$

Otteniamo perciò:

$$-1 + 2 \left(\frac{\mathcal{E}}{\bar{V}} \right) = -1 + 2 \begin{cases} \cos^2 \left(k_2 \frac{a}{2} \right) & (+) \\ \sin^2 \left(k_2 \frac{a}{2} \right) & (-) \end{cases} \quad (3.139)$$

Distinguiamo quindi due casi:

- Prendendo il $+$ in (3.139), a cui corrisponde prendere il $-$ in (3.138), otteniamo:

$$(1) : \begin{cases} \left| \cos \left(k_2 \frac{a}{2} \right) \right| = \sqrt{\frac{\mathcal{E}}{\bar{V}}} \\ \sin(k_2 a) < 0 \end{cases}$$

Con un'opportuna manipolazione, possiamo ricavare al secondo membro di (1) un termine lineare in k_2 . Infatti, elevando al quadrato e sottraendo a 1:

$$1 - \cos^2 \left(k_2 \frac{a}{2} \right) = \sin^2 \left(k_2 \frac{a}{2} \right) = 1 - \frac{\mathcal{E}}{\bar{V}} = \frac{\bar{V} - \mathcal{E}}{\bar{V}}$$

E prendendo di nuovo la radice quadrata:

$$\left| \sin \left(k_2 \frac{a}{2} \right) \right| = \frac{\bar{V} - \mathcal{E}}{\bar{V}} = \frac{k_2}{\bar{k}}$$

dove abbiamo definito \bar{k} :

$$\bar{k} = \frac{\sqrt{-2m\bar{V}}}{\hbar}$$

- Prendendo il $-$ in (3.139) e quindi il $+$ in (3.138):

$$(2) : \begin{cases} \left| \sin \left(k_2 \frac{a}{2} \right) \right| = \sqrt{\frac{\mathcal{E}}{\bar{V}}} \\ \sin(k_2 a) > 0 \end{cases}$$

Analogamente al caso precedente, ricaviamo:

$$\left| \cos \left(k_2 \frac{a}{2} \right) \right| = \frac{k_2}{\bar{k}} \quad (3.140)$$

Le soluzioni (1) e (2) possono essere esplicitate graficamente in figura 3.10, e sono l'intersezione tra la retta e le funzioni seno e coseno in valore assoluto (dove sono accettate solo quelle nelle regioni non proibite). Esse sono le soluzioni dell'equazione agli autovalori per H con $\mathcal{E} < 0$. Notiamo che lo spettro dell'hamiltoniana è unicamente discreto, e quindi le autofunzioni $\varphi_{\mathcal{E}}(x)$ appartengono a L^2 . Inoltre $\sigma(H)|_{\mathcal{E} < 0} = \sigma_P(H)$ è non degenere (degenerazione 1). Se $\bar{V} \rightarrow \infty$ la retta k_2/\bar{k} tende all'asse x ed esaminandone le intersezioni con il coseno e il seno ritroviamo i risultati della buca infinitamente profonda.

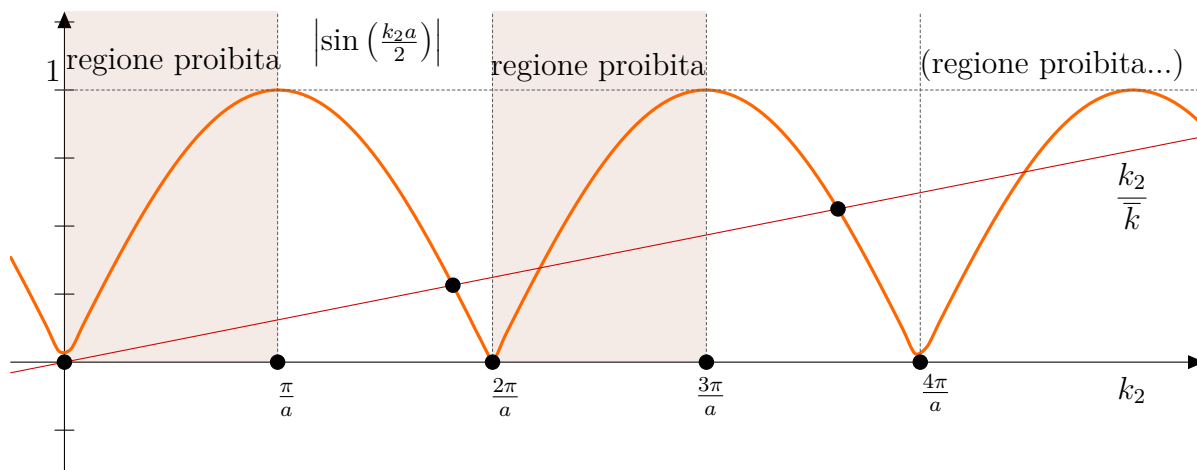
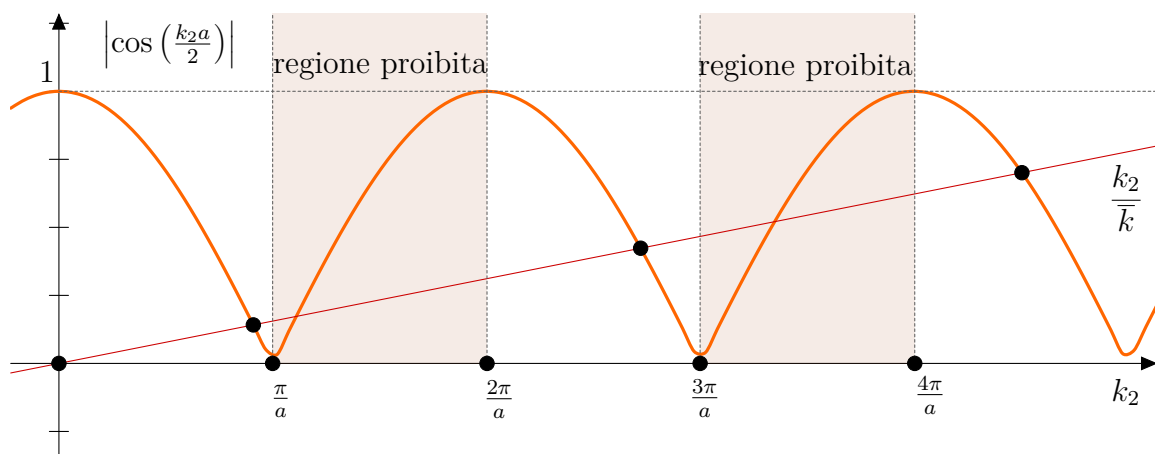


Figura 3.10: (1) e (2) in funzione di k_2 , con il plot del membro di sinistra, e il membro di destra con le intersezioni (cioè le soluzioni).

3.9.7 Regole generali per $\sigma(H)$ per potenziali in 1D (con al più salti)

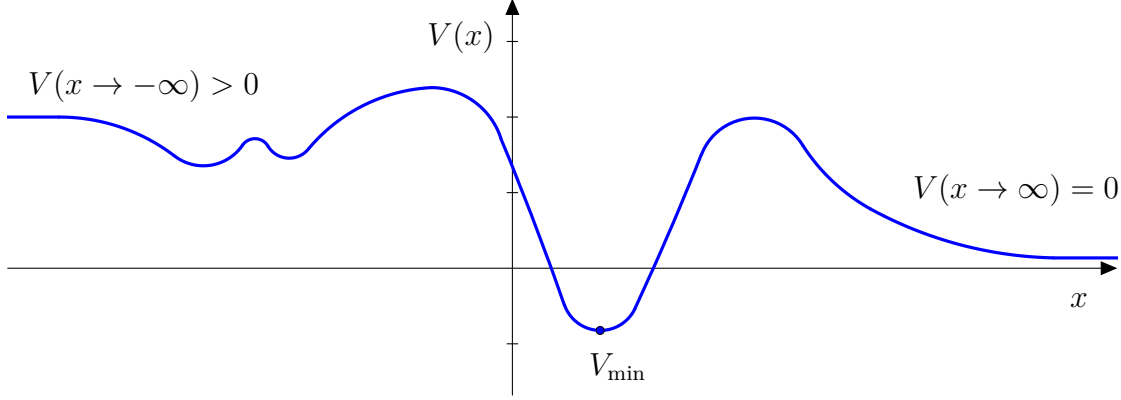


Figura 3.11: Potenziale $V(x)$ generico.

Consideriamo ora il potenziale generico $V(x)$ in figura; richiediamo che V sia regolare con al più discontinuità di salto. Inoltre richiediamo che $V(x \rightarrow +\infty) = 0$, e che $V(x \rightarrow -\infty) > 0$. Supponiamo poi che il minimo assoluto di V sia negativo. Si hanno tre possibili situazioni qualitative per l'energia \mathcal{E} :

Caso $V_{\min} < \mathcal{E} < 0$

Questo è il caso di una particella “confinata in una regione limitata da entrambi i lati”.

Le soluzioni sono del tipo $\varphi_{\mathcal{E}} \sim e^{\mp c \pm x}$. La funzione non può divergere, e in particolare $\varphi_{\mathcal{E}} \in \mathcal{H}$. Lo spettro è inoltre non degenere, cioè ha degenerazione 1. Per dimostrarlo supponiamo che ψ_1 e ψ_2 siano due soluzioni corrispondenti alla stessa energia \mathcal{E} , e dimostriamo che rappresentano lo stesso stato. Si ha che

$$\psi_1'' + \frac{2m}{\hbar^2}(\mathcal{E} - V(x))\psi_1 = 0$$

$$\psi_2'' + \frac{2m}{\hbar^2}(\mathcal{E} - V(x))\psi_2 = 0$$

Dividendo la prima per ψ_1 , la seconda per ψ_2 e sottraendo membro a membro:

$$\begin{aligned} \frac{\psi_1''}{\psi_1} - \frac{\psi_2''}{\psi_2} = 0 &\Rightarrow \psi_1''\psi_2 - \psi_1\psi_2'' = 0 \\ &\Rightarrow \frac{d}{dx}(\psi_1'\psi_2 - \psi_1\psi_2') = 0 \end{aligned}$$

e quindi la funzione $\psi_1'\psi_2 - \psi_1\psi_2'$ deve essere una costante. Tale funzione è regolare perché $\psi_{1,2}$ è derivabile due volte (dato che è soluzione dell'equazione di Schrödinger. Inoltre tale funzione è in \mathcal{H} (perché è combinazione \mathcal{C}^∞ di autofunzioni $\varphi_{\mathcal{E}}$

di autovalori nello spettro discreto $\sigma_P(H)$, ma quindi è quadrato sommabile e necessariamente tende a 0 all'infinito. L'unica funzione costante in \mathcal{H} è quella identicamente nulla. Quindi

$$(\psi'_1\psi_2 - \psi_1\psi'_2)(x) = 0 \quad \Rightarrow \quad \frac{\psi'_1(x)}{\psi_1(x)} = \frac{\psi'_2(x)}{\psi_2(x)}$$

e integrando entrambi i membri in dx si ottiene

$$\ln \psi_1(x) = \ln \psi_2(x) + c \quad \Rightarrow \quad \psi_1(x) = e^c \psi_2(x)$$

dove c è una costante di integrazione. Quindi ψ_1 e ψ_2 differiscono per un fattore moltiplicativo, ma quindi rappresentano lo **stesso stato**. Segue che lo spettro $\sigma(H)|_{V_{\min} < \mathcal{E} < 0}$ è non degenere.

Caso $0 < \mathcal{E} < V(-\infty)$

Questo è il caso di una particella “confinata da un lato”.

La soluzione ha un andamento del tipo:

$$\varphi_{\mathcal{E}}(x) = \begin{cases} c_+ e^{ikx} + c_- e^{-ikx} & x \rightarrow \infty \\ e^{\chi x} & x \rightarrow -\infty \end{cases} \quad \chi = \frac{\sqrt{2m(V(+\infty) - \mathcal{E})}}{\hbar}$$

Poiché la soluzione $\varphi_{\mathcal{E}}$ oscilla infinitamente non può appartenere ad \mathcal{H} , e quindi lo spettro $\sigma(H)|_{0 < \mathcal{E} < V(-\infty)}$ dell'hamiltoniana è necessariamente continuo. Inoltre siccome $\varphi_{\mathcal{E}}(-\infty) = 0$ (dato che $V(-\infty) = 0$) possiamo applicare gli stessi ragionamenti del teorema visto al punto precedente e quindi affermare che $\sigma(H)$ è non degenere.

Caso $\mathcal{E} > V(-\infty)$

Questo è il caso di una particella “non confinata”, come per esempio il caso del potenziale a gradino con $\mathcal{E} > \bar{V}$ visto precedentemente.

Le soluzioni risultano

$$\varphi_{\mathcal{E}}(x) = \begin{cases} c_+^1 e^{ik_1 x} + c_-^1 e^{-ik_1 x} & x \rightarrow +\infty \\ c_+^2 e^{ik_1 x} + c_-^2 e^{-ik_1 x} & x \rightarrow -\infty \end{cases}$$

dove

$$k_1 = \frac{\sqrt{2m(\mathcal{E} - V(-\infty))}}{\hbar}$$

Anche in questo caso le autofunzioni non possono essere elementi di \mathcal{H} , perchè presentano oscillazioni infinite. Quindi lo spettro $\sigma(H)|_{\mathcal{E} > V(-\infty)}$ è continuo. Inoltre lo spettro ha degenerazione 2, perchè sono 4 costanti da determinare a cui vanno sottratte le 2 condizioni di raccordo ($4 - 2 = 2$).

Caso della parità del potenziale

Qualcosa in più sul sistema si può dedurre se si suppone che il potenziale sia una funzione pari, ovvero $V(x) = V(-x)$. In questo caso specifico se $\varphi_{\mathcal{E}}(x)$ è soluzione dell'equazione di Schrödinger stazionaria, allora anche $\varphi_{\mathcal{E}}(-x)$ lo è. Se poi supponiamo anche che lo spettro $\sigma(H)$ sia non degenere allora

$$\varphi_{\mathcal{E}}(x) = c\varphi_{\mathcal{E}}(-x) = c^2\varphi_{\mathcal{E}}(x) \Rightarrow c^2 = 1 \Rightarrow \varphi_{\mathcal{E}}(x) = \pm\varphi_{\mathcal{E}}(-x)$$

Quindi ci sono solo due possibilità: la funzione d'onda può essere pari, dunque $\varphi_{\mathcal{E}}(x) = \varphi_{\mathcal{E}}(-x)$, oppure può essere dispari, ovvero $\varphi_{\mathcal{E}}(x) = -\varphi_{\mathcal{E}}(-x)$. Se poi si definisce l'operatore di parità \mathcal{P} come:

$$(\mathcal{P}\psi)(x) = \psi(-x)$$

allora si ha che gli autostati di H sono anche autostati di \mathcal{P} , con autovalore $+1$ oppure -1 .

3.9.8 Esercizio 2

Una particella di massa m si trova nello stato fondamentale (cioè di energia minima) in una buca di potenziale infinita in $[0, a]$. Al tempo $t = 0$ la parete impenetrabile posta in a viene improvvisamente spostata in $x = 2a$.

1. Calcolare la probabilità che al tempo $\bar{t} > 0$ l'energia della particella sia uguale a quella a $t = 0^-$
2. Che sia invece minore.

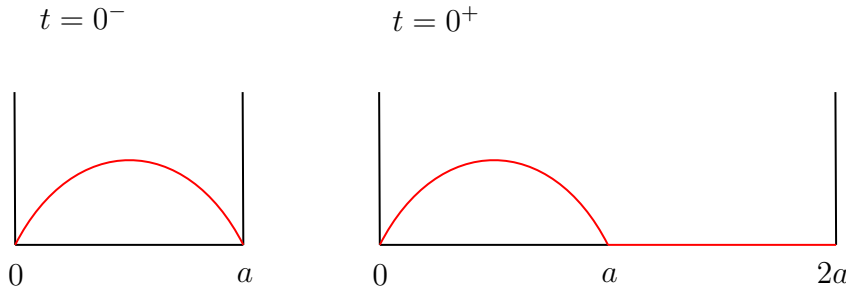


Figura 3.12: Funzione d'onda ψ per $t = 0^-$ e $t = 0^+$

Soluzione

1. Per $t < 0$, $\sigma(H_1)$ è quello della buca in $[0, a]$, e gli autovalori sono:

$$\mathcal{E}_n^1 = \frac{\hbar^2}{2m} \frac{\pi^2 n^2}{a^2}$$

Per $t > 0$, la buca è in $[0, 2a]$ e gli autovalori saranno:

$$\mathcal{E}_n^2 = \frac{\hbar^2}{2m} \frac{\pi^2 n^2}{4a^2}$$

Dopo il cambiamento del sistema, la ψ non è più autofunzione di H . Confrontando le due espressioni per gli autovalori si ha che $\mathcal{E}_1^1 = \mathcal{E}_2^2$, quindi per rispondere alla prima richiesta occorre calcolare $W_{\psi(t>0)}^{H_2}(\mathcal{E}_2^2)$, ossia la probabilità che una misura di energia a $t > 0$ dia come risultato \mathcal{E}_2^2 . In notazione di Dirac:

$$\begin{aligned} W_{\psi(t)}^{H_2}(\mathcal{E}_2^2) &= |\langle \psi(t) | \mathcal{E}_2^2 \rangle|^2 = |\langle \psi(0^+) | e^{i\frac{t}{\hbar}H_2} | \mathcal{E}_2^2 \rangle|^2 = \\ &= |\langle \psi(0^+) | \mathcal{E}_2^2 \rangle e^{i\frac{t}{\hbar}\mathcal{E}_2^2}|^2 = |\langle \psi(0^+) | \mathcal{E}_2^2 \rangle|^2 \end{aligned}$$

dove abbiamo semplificato l'evoluzione temporale riducendola al calcolo della ψ in $t = 0^+$. Abbiamo quindi:

$$\psi(x, t = 0^+) = \psi(x, t = 0^-) = \begin{cases} \sqrt{\frac{2}{a}} \sin \frac{\pi}{a}x & x \in [0, a] \\ 0 & \text{altrove} \end{cases}$$

ed è già normalizzata in $L^2([0, 2a], dx)$ (era normalizzata prima, e abbiamo aggiunto solo degli zeri). Calcoliamo quindi l'autofunzione di H_2 :

$$\langle x | \mathcal{E}_n^2 \rangle = \varphi_n(x) = \sqrt{\frac{1}{a}} \sin \frac{\pi x n}{2a}$$

Vogliamo quindi calcolare il prodotto scalare per esprimere la ψ sulla base delle autofunzioni di H_2 :

$$\begin{aligned} \langle \psi(0^+) | \mathcal{E}_2^2 \rangle &= \int_0^a \sqrt{\frac{2}{a}} \sin \frac{\pi x}{a} \sqrt{\frac{1}{a}} \sin \frac{2\pi x}{2a} dx + \int_a^{2a} 0 \cdot \sqrt{\frac{1}{a}} \sin \frac{2\pi x}{2a} dx = \\ &= \frac{\sqrt{2}}{a} \int_0^a \sin^2 \frac{\pi x}{a} = \frac{\sqrt{2}}{a} \frac{a}{2} = \frac{1}{\sqrt{2}} \end{aligned}$$

Ma quindi:

$$W_{\psi(t>0)}^{H_2}(\mathcal{E}_2) = \frac{1}{2}$$

2. Dato che l'unico autovalore di H_2 con energia minore di $\mathcal{E}_2^2 = \mathcal{E}_1^1$ è quello per $n = 1$, basta calcolare $W_{\psi(t)}^{H_2}(\mathcal{E}_1^2)$. Utilizzando la notazione di Dirac si ha:

$$W_{\psi(t)}^{H_2}(\mathcal{E}_1^2) = |\langle \psi(0^+) | \mathcal{E}_1^2 \rangle|^2$$

Successivamente si calcola il prodotto scalare:

$$\begin{aligned} \langle \psi(0^+) | \mathcal{E}_1^2 \rangle &= \int_0^a \sqrt{\frac{2}{a}} \sin \frac{\pi x}{a} \sqrt{\frac{1}{a}} \sin \frac{\pi x}{2a} dx \stackrel{(a)}{=} \frac{\sqrt{2}}{a} \int_0^a \frac{1}{2} \left[\cos \frac{\pi a}{2a} - \cos \frac{3\pi x}{2a} \right] dx = \\ &\stackrel{(b)}{=} \frac{\sqrt{2}}{a} \frac{1}{2} \left[\frac{2a}{\pi} \sin y \Big|_0^{\frac{\pi}{2}} - \frac{2a}{3\pi} \sin y \Big|_0^{\frac{3\pi}{2}} \right] = \sqrt{2} \left[\frac{1}{\pi} + \frac{1}{3\pi} \right] = \sqrt{2} \frac{4}{3\pi} \end{aligned}$$

dove in (a) si è usata una delle formule di Werner:

$$\sin \alpha \sin \beta = \frac{1}{2} [\cos(\alpha - \beta) - \cos(\alpha + \beta)]$$

mentre in (b) si è effettuato il cambio di variabili $y = \pi x/(2a)$. Arriviamo quindi al risultato finale:

$$W_{\psi(t)}^{H_2}(\mathcal{E}_1^2) = \frac{32}{9\pi^2}$$

m

Capitolo 4

Conseguenze degli assiomi

4.1 Principio di indeterminazione

(Lezione 20 del
12/11/2018)

Supponiamo di voler misurare simultaneamente due osservabili A e B : che vincolo abbiamo sulla precisione della misura simultanea?

Una delle conseguenze degli assiomi è il *principio di indeterminazione* che generalizza a coppie di osservabili arbitrarie quello trovato (euristicamente) da Heisenberg nel 1927 per X e P .

Ricordiamo che la fluttuazione di una osservabile A in uno stato ψ è definita come:

$$(\Delta A)_\psi = \sqrt{\langle (A - \langle A \rangle_\psi)^2 \rangle_\psi} \quad (4.1)$$

Teorema 4.1.1. *Siano A e B autoaggiunti (ossia corrispondenti a osservabili) in $D(A)$ e $D(B)$. Supponiamo che esista un **dominio “comune”** $D \subset \{\psi \in \mathcal{H} | \psi, A\psi, B\psi \in D(A) \cap D(B)\}$ denso in \mathcal{H} (ossia tale che applicare A o B ad un elemento dell’intersezione dei domini $D(A) \cap D(B)$ produce un risultato che resta ancora nell’intersezione). Questa scelta del dominio è tale che tutte le operazioni usate in seguito siano ben definite, e in particolare abbia senso calcolare il **commutatore** $[A, B]$.*

Principio di
indeterminazione

Vale allora, $\forall \psi \in D$:

$$(\Delta A)_\psi (\Delta B)_\psi \geq \left| \frac{\langle [A, B] \rangle_\psi}{2i} \right| = \frac{1}{2} |\langle [A, B] \rangle_\psi|$$

In altre parole, per il *principio di indeterminazione*, il valor medio del commutatore $[A, B]$ di due operatori autoaggiunti costituisce, a meno di un fattore, un *limite inferiore* al prodotto delle incertezze su A e B .

Dimostrazione. Definiamo per semplicità di notazione:

$$\bar{A} = A - \langle A \rangle_\psi \mathbb{I}; \quad \bar{B} = B - \langle B \rangle_\psi \mathbb{I} \quad (4.2)$$

Abbiamo cioè *traslato* i valor medi di A e B a 0.

Per la linearità del commutatore si ha:

$$[\bar{A}, \bar{B}] = [A - \langle A \rangle, B - \langle B \rangle] = [A, B] - [A, \langle B \rangle] - [\langle A \rangle, B] + [\langle A \rangle, \langle B \rangle] = [A, B] \quad (4.3)$$

dato che i valor medi, essendo numeri reali, commutano sempre.
Inoltre, sostituendo (4.2) in (4.1) si ottiene:

$$(\Delta A)_\psi^2 = \langle \bar{A}^2 \rangle_\psi; \quad (\Delta B)_\psi^2 = \langle \bar{B}^2 \rangle_\psi \quad (4.4)$$

Consideriamo il seguente prodotto scalare:

$$\left\langle \left(\frac{\bar{A}}{(\Delta A)_\psi} \pm i \frac{\bar{B}}{(\Delta B)_\psi} \right)^\dagger \left(\frac{\bar{A}}{(\Delta A)_\psi} \pm i \frac{\bar{B}}{(\Delta B)_\psi} \right) \right\rangle \geq 0$$

che è sempre non negativo, dato che, per un generico operatore c , vale, dalla definizione di valor medio:

$$\langle c^\dagger c \rangle_\psi = (\psi, c^\dagger c \psi) = (c\psi, c\psi) = \|c\psi\|^2 \geq 0$$

Sviluppando allora i prodotti in (4.4):

$$\begin{aligned} & \left\langle \left(\frac{\bar{A}^\dagger}{(\Delta A)_\psi} \mp i \frac{\bar{B}^\dagger}{(\Delta B)_\psi} \right) \left(\frac{\bar{A}}{(\Delta A)_\psi} \pm i \frac{\bar{B}}{(\Delta B)_\psi} \right) \right\rangle \geq 0 \\ & \stackrel{(4.4)}{\Rightarrow} \underbrace{\frac{\langle \bar{A}^2 \rangle_\psi}{(\Delta A)_\psi^2}}_{=1} + \underbrace{\frac{\langle \bar{B}^2 \rangle_\psi}{(\Delta B)_\psi^2}}_{=1} \mp i \frac{\overbrace{\langle \bar{B}\bar{A} - \bar{A}\bar{B} \rangle_\psi}^{\langle [\bar{B}, \bar{A}] \rangle_\psi}}{(\Delta A)_\psi (\Delta B)_\psi} \geq 0 \\ & \Rightarrow 2 \geq \mp i \frac{\langle [\bar{A}, \bar{B}] \rangle_\psi}{(\Delta A)_\psi (\Delta B)_\psi} \Rightarrow (\Delta A)_\psi (\Delta B)_\psi \geq \mp \frac{i}{2} \langle [\bar{A}, \bar{B}] \rangle_\psi \end{aligned}$$

Da cui si ha:

$$(\Delta A)_\psi (\Delta B)_\psi \underset{(a)}{\geq} \left| \frac{i}{2} \langle [\bar{A}, \bar{B}] \rangle_\psi \right| \underset{(b)}{=} \left| \frac{\langle [\bar{A}, \bar{B}] \rangle_\psi}{2i} \right| \underset{(4.3)}{=} \left| \frac{\langle [A, B] \rangle_\psi}{2i} \right|$$

Dove in (a) si è usata la relazione:

$$|x| \leq a \Leftrightarrow \begin{cases} x \leq a \\ -x \leq a \end{cases} \Leftrightarrow -a \leq x \leq a$$

e in (b) si è moltiplicato e diviso per i , lasciando assorbire il segno generato dal valore assoluto.

In particolare, se $A = X$ e $B = P$, $D = \mathcal{S}(\mathbb{R})$, $\forall \psi \in D$:

$$(\Delta X)_\psi (\Delta P)_\psi \geq \left| \frac{\langle [X, P] \rangle_\psi}{2i} \right| = \frac{\hbar}{2}$$

Il principio di indeterminazione di Heisenberg è quindi **solo** una conseguenza dell'algebra $[X, P] = i\hbar$.

Dimostrazione alternativa^a

Consideriamo due operatori A e B , e delle funzioni d'onda $|\psi\rangle \in D$, dove D è il “dominio comune” di A e B come definito nelle ipotesi del teorema (4.1.1).

Partiamo dalla definizione di fluttuazione di A , ed espandiamo il valor medio^b:

$$\begin{aligned}(\Delta A)_\psi^2 &= \langle (A - \langle A \rangle_\psi)^2 \rangle_\psi = \langle \psi | (A - \langle A \rangle_\psi)^2 \psi \rangle \stackrel{(a)}{=} \langle (A - \langle A \rangle_\psi) \psi | (A - \langle A \rangle_\psi) \psi \rangle = \\ &= \langle f | f \rangle = \|f\|^2 \quad |f\rangle \equiv (A - \langle A \rangle_\psi) |\psi\rangle\end{aligned}\tag{4.5}$$

Dove in (a) si è usata l'autoaggiuntezza di A .

Analogamente, per l'osservabile B :

$$(\Delta B)_\psi^2 = \langle g | g \rangle = \|g\|^2 \quad |g\rangle \equiv (B - \langle B \rangle_\psi) |\psi\rangle$$

Per il teorema di Cauchy-Schwarz^c si ha che:

$$(\Delta A)_\psi^2 (\Delta B)_\psi^2 = \langle f | f \rangle \langle g | g \rangle \geq |\langle f | g \rangle|^2\tag{4.6}$$

Per ogni numero complesso z vale:

$$|z|^2 = \text{Re}(z)^2 + \text{Im}(z)^2 \geq \text{Im}(z)^2 = \left(\frac{1}{2i}(z - z^*) \right)^2$$

Nel caso di $z = \langle f | g \rangle$, perciò:

$$|\langle f | g \rangle|^2 \geq \left(\frac{1}{2i}[\langle f | g \rangle - \langle g | f \rangle] \right)^2\tag{4.7}$$

(dato che $\langle f | g \rangle^* = \langle g | f \rangle$ per come è definita l'applicazione di un bra a un ket).

Sostituendo (4.7) in (4.6) si giunge a:

$$(\Delta A)_\psi^2 (\Delta B)_\psi^2 \geq \left(\frac{1}{2i}[\langle f | g \rangle - \langle g | f \rangle] \right)^2\tag{4.8}$$

Basta ora sviluppare i braket. Ricordando le definizioni in (4.8) ed espandendo:

$$\begin{aligned}\langle f | g \rangle &= \langle (A - \langle A \rangle_\psi) \psi | (B - \langle B \rangle_\psi) \psi \rangle \stackrel{(a)}{=} \langle \psi | (A - \langle A \rangle_\psi)(B - \langle B \rangle_\psi) \psi \rangle = \\ &= \langle \psi | (AB - A\langle B \rangle_\psi - B\langle A \rangle_\psi + \langle A \rangle_\psi \langle B \rangle_\psi) \psi \rangle = \\ &= \underbrace{\langle \psi | AB \psi \rangle}_{\langle AB \rangle_\psi} - \underbrace{\langle B \rangle_\psi \langle \psi | A \psi \rangle}_{\langle A \rangle_\psi} - \underbrace{\langle A \rangle_\psi \langle \psi | B \psi \rangle}_{\langle B \rangle_\psi} + \langle A \rangle_\psi \langle B \rangle_\psi = \\ &= \langle AB \rangle_\psi - \langle A \rangle_\psi \langle B \rangle_\psi\end{aligned}$$

Analogamente si ottiene:

$$\langle g | f \rangle = \langle BA \rangle_\psi - \langle A \rangle_\psi \langle B \rangle_\psi$$

Da cui:

$$\langle f | g \rangle - \langle g | f \rangle = \langle AB \rangle_\psi - \langle BA \rangle_\psi = \langle [A, B] \rangle_\psi$$

E sostituendo in (4.8) si ottiene la disuguaglianza desiderata:

$$(\Delta A)_\psi^2 (\Delta B)_\psi^2 \geq \left(\frac{\langle [A, B] \rangle_\psi}{2i} \right)^2 \quad (4.9)$$

Nota: la i al denominatore non rende il secondo membro negativo (e quindi ovvia la disuguaglianza), in quanto i valor medi di un commutatore di due operatori autoaggiunti sono *puramente immaginari*. Lo si può dimostrare notando che $[A, B]$ è anti-hermitiano, ossia è pari al suo aggiunto *cambiato di segno*^d:

$$([A, B])^\dagger = (AB - BA)^\dagger = (BA - AB) = -[A, B]$$

Ma allora se λ è il valor medio di $[A, B]$, e λ^* è il coniugato del valor medio di $[A, B]$, ossia il valor medio di $[A, B]^\dagger$, si ha che $\lambda^* = -\lambda$, e tale uguaglianza è valida per $\lambda \in \mathbb{C}$ solamente se λ è puramente immaginario.

Perciò la i al denominatore di (4.9) si semplifica con la i che moltiplica il valor medio del commutatore, e dà un risultato reale. Possiamo allora prendere la radice quadrata di ambo i membri^e e scrivere la relazione finale:

$$(\Delta A)_\psi (\Delta B)_\psi \geq \left| \frac{\langle [A, B] \rangle_\psi}{2i} \right|$$

^a^Sezione 3.4, pag. 108 di [3]

^b^Le $\langle \dots \rangle$ senza il ψ a pedice sono un *braket*, non una media!

^c^Tale teorema esprime il fatto *intuitivo* che la norma di un vettore è sempre maggiore (o al più uguale) alla lunghezza della proiezione di tale vettore su un altro vettore.

^d^L'aggiunto di un prodotto di operatori autoaggiunti segue le *regole notazionali* della trasposta del prodotto di matrici

^e^Facendo tutti i conti, siamo di fronte ad una disequazione del tipo $x^2 \leq a^2$, che ha soluzioni per $x \geq |a|$ se $x \geq 0$, e $x \leq -|a|$ se $x < 0$. Ma ovviamente il prodotto di due incertezze è positivo, e quindi siamo nel primo caso.

4.1.1 Osservazioni sul principio di indeterminazione

- Il principio di indeterminazione vale anche per **stati misti**. Se $A, B \in \mathcal{B}(\mathcal{H})$ (operatori lineari limitati, e quindi estendibili a $D = \mathcal{H}$), ricordando la definizione in (3.37) di valor medio di un operatore in uno stato misto dato dalla matrice densità ρ , si ha che:

$$\langle A \rangle_\rho = \text{Tr}(\rho A); \quad (\Delta A)_\rho^2 = \text{Tr}(\rho A - \text{Tr}(\rho A))^2 \quad (4.10)$$

Vale ancora il principio di indeterminazione:

$$(\Delta A)_\rho (\Delta B)_\rho \geq \left| \frac{\langle [A, B] \rangle_\rho}{2i} \right|$$

Infatti, il passaggio da cui segue il teorema consiste nel notare che il valor medio del prodotto aggiunto-operatore è non negativo, ossia $\langle c^\dagger c \rangle \geq 0$. Ma

Indeterminazione degli stati misti

questo è un risultato generale, che vale anche usando la definizione di valor medio data in (4.10), infatti:

$$\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|; \quad \langle c^\dagger c \rangle_\psi \geq 0 \Rightarrow \text{Tr } \rho(c^\dagger c) = \sum_i p_i \langle \psi_i | c^\dagger c | \psi_i \rangle \geq 0$$

- Talvolta il principio di indeterminazione **sembra violato**, ma ciò è dovuto a *problemi di dominio*, per cui cadono le ipotesi fatte nel teorema. Abbiamo infatti ipotizzato che esista un certo dominio D , per cui è possibile applicare A o B alle $|\psi\rangle \in D$ ottenendo risultati che restano in D . Se ciò non avviene, la conseguenza del teorema potrebbe non valere.

Importanza delle ipotesi del teorema: un controesempio

Per esempio, consideriamo come spazio di Hilbert $\mathcal{H} = (L^2[-\frac{1}{2}, \frac{1}{2}], dx)$. In tale \mathcal{H} si ha che l'operatore *posizione* X è limitato, e quindi $D(X) = \mathcal{H}$.

Consideriamo poi l'operatore *momento* $P = -i\hbar \frac{d}{dx}$ con condizioni periodiche per $D(P)$ (in modo che sia autoaggiunto).

Poiché X è limitato, $(\Delta X)_\psi \leq 1$ (non posso avere un'indeterminazione più grande del dominio!).

Se scegliamo come ψ un **autostato** ψ_n di P , dove $\sigma(P)$ è unicamente discreto (come visto in sezione 3.5.3), sappiamo che (definizione di autostato):

$$(\Delta P)_{\psi_n} = 0$$

Ma allora:

$$(\Delta X)_{\psi_n} (\Delta P)_{\psi_n} = 0$$

in apparente violazione del principio di indeterminazione.

Tuttavia, notiamo che applicando X a $\psi_n(x)$ non otteniamo un elemento nel dominio di P (e perciò certamente non un elemento nell'intersezione tra $D(X) \cap D(P)$ come richiesto dalle ipotesi del teorema).

Infatti, se risolviamo l'equazione agli autovalori per trovare gli autostati $\psi_n(x)$ di P :

$$-i\hbar \frac{d}{dx} \psi_n(x) = \lambda_n \psi_n(x) \Rightarrow \psi_n(x) = \exp\left(-i \frac{\lambda_n}{\hbar} x\right)$$

Imponendo le condizioni (periodiche) al contorno:

$$\psi\left(\frac{1}{2}\right) \stackrel{!}{=} \psi\left(-\frac{1}{2}\right) \Rightarrow \lambda_n = 2\pi n \hbar$$

Perciò gli autostati ψ_n di P sono¹:

$$\psi_n(x) = e^{-i2\pi x n}$$

Tuttavia se applichiamo X a una qualsiasi ψ_n otteniamo un risultato che non è nel dominio di P :

$$X\psi_n(x) = x\psi_n(x) = x e^{-i2\pi x n} \notin D(P)$$

¹^Si noti che, date le condizioni al contorno, è come se stessimo considerando la particella "su un anello". In particolare, queste ψ_n sono normalizzabili, al contrario di quanto accade se definiamo P con dominio tutto \mathbb{R} .

perché manca la periodicità.
Quindi:

$$\psi_n \notin D = \{\psi \in \mathcal{H} | \psi, X\psi, P\psi \in D(X) \cap D(P) = D(P)\}$$

Non potendo trovare un “dominio comune” D per X e P in questo caso, il principio non è applicabile, e infatti la tesi non vale.

Intuitivamente, l'esempio appena portato consiste nell'esaminare una particella che gira su un cerchio, per cui X è più precisamente una coordinata angolare. Si nota allora come sia possibile, in principio, conoscere con precisione arbitraria il momento, senza che l'incertezza sull'angolo possa aumentare a dismisura: sarà al più di 2π .

Nota: tale controesempio non significa che quando le ipotesi non valgono non esista alcun limite inferiore al prodotto delle incertezze di due osservabili: il risultato potrebbe valere lo stesso (ma va dimostrato). Per esempio, nel caso della retta reale, non è possibile trovare controesempi come quello appena visto, anche scegliendo ψ che non sono nel dominio di PX o XP , basta solo richiedere che le incertezze $(\Delta X)_\psi$ e $(\Delta P)_\psi$ siano definite. Per chi fosse interessato, una dimostrazione molto *tecnica* di ciò è data a pag. 246 di “*Quantum theory for Mathematicians*”, Hall, B. C. (2013), Springer. Del resto le considerazioni sul controesempio sono tratte da https://en.wikipedia.org/wiki/Uncertainty_principle#A_counterexample.

4.2 Osservabili compatibili

Il principio di indeterminazione implica che in generale non è possibile effettuare misure simultanee arbitrariamente precise se il commutatore $[A, B] \neq 0$, ma suggerisce che se $[A, B] = 0$ ciò potrebbe essere possibile. Dimostriamo che, in effetti, **simultanea misurabilità** e **commutatività** sono in corrispondenza biunivoca. Iniziamo col caso di spettro discreto e premettiamo una nozione matematica.

Definizione 24. Se \mathcal{H}_1 e \mathcal{H}_2 sono spazi di Hilbert, l'insieme delle **coppie** $\mathcal{H}_1 \oplus \mathcal{H}_2 = \{\psi_1 \in \mathcal{H}_1, \psi_2 \in \mathcal{H}_2\}$ denotate $\psi_1 \oplus \psi_2$ con il prodotto scalare: *Somma diretta*

$$(\psi_1 \oplus \psi_2, \varphi_1 \oplus \varphi_2) \equiv (\psi_1, \varphi_1)_{\mathcal{H}_1} + (\psi_2, \varphi_2)_{\mathcal{H}_2}$$

è uno spazio di Hilbert detto **somma diretta** di \mathcal{H}_1 e \mathcal{H}_2 e denotato con $\mathcal{H}_1 \oplus \mathcal{H}_2$. Nel caso finito dimensionale, $\mathcal{H}_1 \cong \mathbb{C}^N$, $\mathcal{H}_2 \cong \mathbb{C}^M$, si ha:

$$\mathcal{H}_1 \oplus \mathcal{H}_2 \cong \mathbb{C}^{N+M}$$

e i suoi vettori sono vettori colonna del tipo:

$$\begin{pmatrix} \psi_1 \in \mathbb{C}^N \\ \psi_2 \in \mathbb{C}^M \end{pmatrix}$$

Si noti che tra spazi di Hilbert in somma diretta non si ha “interferenza”, cioè tutti gli stati di \mathcal{H}_1 pensati in $\mathcal{H}_1 \oplus \mathcal{H}_2$, ossia come $\psi_1 \oplus 0$, sono ortogonali a quelli di \mathcal{H}_2 pensati in $\mathcal{H}_1 \oplus \mathcal{H}_2$ come $0 \oplus \psi_2$.

Definizione 25. Due osservabili **limitate** a **spettro discreto** si dicono **compatibili** se una misura di prima specie di una osservabile eseguita su un autostato dell'altra la lascia nello stesso autovalore (ossia nello stesso *autospazio*, ma non necessariamente nello stesso *autostato*), cosicché l'ordine delle misure di A e B è irrilevante, e sia A e B sono simultaneamente misurabili con arbitraria precisione. Ciò significa che per entrambe esiste una **base ortonormale di autovettori comuni**: dato che una misura di prima specie proietta il sistema in un *autostato* dell'osservabile misurata, conoscere contemporaneamente due osservabili significa che, dopo le due misure, il sistema si trova in un autostato di entrambe.

Osservabili
compatibili

Per esempio, consideriamo un sistema che inizialmente si trova in uno stato $|\psi\rangle$, e siano A e B due osservabili compatibili: ciò significa che se misuriamo A e otteniamo un valore a , poi misuriamo B e otteniamo un b , rimisurando A riotteniamo a , e così via. In questo senso possiamo dire di conoscere le misure di A e B *contemporaneamente*. Esaminiamo, passo per passo, cosa succede dopo ogni misura.

Partiamo facendo una misura di A sul sistema, da cui otteniamo un risultato $a \in \sigma(A)$. Ciò vuol dire che la funzione d'onda $|\psi\rangle$ iniziale è *collassata* su una $|\psi_a\rangle$ autostato di A .

Misuriamo ora B . Abbiamo due casi:

- $|\psi_a\rangle$ è anche un autostato di B , per esempio di autovalore b . In tal caso la funzione d'onda del sistema resta la stessa (viene *proiettata* su di sé) e otterremo b come risultato della misura. Allora $|\psi_a\rangle$ è un autostato comune a A e B , e se ora rimisuriamo A otterremo con certezza di nuovo a , come voluto dalla richiesta che A e B siano compatibili.
- $|\psi_a\rangle$ non è un autostato di B . In tal caso, se il risultato della misura è un certo $b \in \sigma(B)$, la $|\psi_a\rangle$ è collassata in una funzione d'onda differente, che chiamiamo $|\psi_b\rangle$ e che è un autostato di B . Se ora rimisuriamo A ci aspettiamo (per ipotesi di compatibilità) di riottenere a con certezza: ciò significa che $|\psi_b\rangle$ deve essere un autostato di A . Ma allora $|\psi_b\rangle$ è un autostato comune ad entrambi gli operatori.

Poiché questo ragionamento vale per ogni possibile $|\psi\rangle$ iniziale, la definizione di compatibilità si traduce, geometricamente, nel richiedere che esista una *base di autostati* comune ad A e B , che in particolare possiamo scegliere essere ortonormale. Infatti, le proiezioni descritte sopra funzionano solo se la $|\psi\rangle$ iniziale (generica) è scrivibile come combinazione lineare (eventualmente infinita) di autostati comuni alle due osservabili.

Notiamo ora che tale definizione di compatibilità si traduce in una precisa **proprietà algebrica** degli operatori che descrivono tali osservabili.

Teorema 4.2.1. *Se A e B sono **autoaggiunti** e **limitati** e a **spettro discreto** sono **compatibili** (cioè ammettono una base ortonormale di autovettori comuni) se e solo se **commutano**.*

Compatibilità \leftrightarrow
commutatività
(per operatori
a.a. limitati a
spettro discreto)

Dimostrazione

Osserviamo che, per ipotesi, A e B sono autoaggiunti e **limitati** (da cui continui), e quindi possiamo considerare il loro dominio $D(A) = D(B) = \mathcal{H}$ (basta estendere per continuità $D(A)$ e $D(B)$ densi). In particolare, $A|\psi\rangle \in D(B)$ e $B|\psi\rangle \in D(A)$ per ogni $|\psi\rangle$, e quindi *non si hanno problemi di dominio*.

Partiamo dimostrando che **compatibilità** \Rightarrow **commutatività**

Siano $\sigma(A) = \{\lambda_n\}_{n \in N}$, $\sigma(B) = \{\mu_m\}_{m \in M}$ gli spettri di A e B . Lo spazio degli autovettori comuni di A e B di autovalori λ_n, μ_m è definito da:

$$\mathcal{H}_{(\lambda_n, \mu_m)} = \{|\phi\rangle \in \mathcal{H} \mid A|\phi\rangle = \lambda_n |\phi\rangle, B|\phi\rangle = \mu_m |\phi\rangle\}$$

Se $\mathcal{H}_{(\lambda_n, \mu_m)} \neq 0$ chiameremo i (λ_n, μ_m) **autovalori comuni** di A e B .

Nota: non è necessario che A e B abbiano gli stessi autovalori, e nemmeno lo stesso numero² di autovalori! Per esempio B potrebbe avere un autovalore *degenerare* μ_1 , che ha come autospazio un \mathcal{H}_{μ_1} di dimensione pari alla sua degenerazione, mentre lo stesso \mathcal{H}_{μ_1} *contiene* autovettori di A con autovalori $\lambda_1, \dots, \lambda_n$ (chiaramente, $n < \dim(\mathcal{H}_{\mu_1})$).

L'insieme di tutti gli autovalori comuni è dato dallo **spettro comune**:

$$\sigma(A, B) \equiv \{(\lambda_n, \mu_m) \in \sigma(A) \times \sigma(B) \text{ autovalori comuni} \} \subseteq \sigma(A) \times \sigma(B)$$

Nota: in generale $\sigma(A, B)$ è più piccolo del prodotto cartesiano $\sigma(A) \times \sigma(B)$ dei singoli spettri, poiché può essere che, fissato un λ_n , non tutti i μ_m siano ammessi come autovalori.

Notiamo che $\mathcal{H}_{(\lambda_n, \mu_m)}$ è l'intersezione di autospazi di operatori a spettro discreto, che sono *sottospazi* di \mathcal{H} , e quindi a loro volta spazi di Hilbert. Di conseguenza, anche $\mathcal{H}_{(\lambda_n, \mu_m)}$ è una varietà chiusa di \mathcal{H} , e quindi è di Hilbert. Ammette perciò una **base ortonormale** di autovettori comuni, di dimensione pari alla degenerazione di (λ_n, μ_m) :

$$\{ |(\lambda_n, \mu_m), j\rangle, j = 1, \dots, \dim \mathcal{H}_{(\lambda_n, \mu_m)} \equiv d(\lambda_n, \mu_m) \}$$

(dove j indica la degenerazione dell'autovettore, nel caso alla stessa coppia di autovalori corrispondano più autovettori).

Inoltre, se $(\lambda_n, \mu_m) \neq (\lambda_{n'}, \mu_{m'})$ gli stati in $\mathcal{H}_{(\lambda_n, \mu_m)}$ sono ortogonali agli stati in $\mathcal{H}_{(\lambda_{n'}, \mu_{m'})}$ in quanto autostati di autovalori diversi di operatori autoaggiunti a spettro discreto (ciò deriva dal teorema spettrale).

Se A e B sono compatibili i loro autovettori comuni formano una base di \mathcal{H} : per ottenerla basta *unire* tutti gli autospazi comuni:

$$\mathcal{H} = \bigoplus_{(\lambda_n, \mu_m) \in \sigma(A, B)} \mathcal{H}_{(\lambda_n, \mu_m)}$$

²Parlare di *numero di autovalori* ha senso solo se \mathcal{H} è finito-dimensionale.

e $\{ |(\lambda_n, \mu_m), j\rangle \}, j = 1, \dots, d(\lambda_n, \mu_m), (\lambda_n, \mu_m) \in \sigma(A, B)$ è la base ortonormale in \mathcal{H} di autovettori comuni.

Pertanto possiamo esprimere ogni vettore di \mathcal{H} come combinazione lineare di elementi di questa base. $\forall |\psi\rangle \in \mathcal{H}$ vale:

$$|\psi\rangle = \sum_{\substack{n \in N \\ m \in M}} \sum_{j=1}^{d(\lambda_n, \mu_m)} c_{nmj} |(\lambda_n, \mu_m), j\rangle \quad (4.11)$$

Essendo $|(\lambda_n, \mu_m), j\rangle$ autostati di A e di B (di autovalore λ_n per A e μ_m per B), per definizione di autovettore avremo:

$$\begin{aligned} A |(\lambda_n, \mu_m), j\rangle &= \lambda_n |(\lambda_n, \mu_m), j\rangle \\ B |(\lambda_n, \mu_m), j\rangle &= \mu_m |(\lambda_n, \mu_m), j\rangle \end{aligned}$$

Mettendo tutto insieme:

$$(AB - BA) |(\lambda_n, \mu_m), j\rangle = (A\mu_m - B\lambda_n) |(\lambda_n, \mu_m), j\rangle = (\lambda_n\mu_m - \mu_m\lambda_n) |(\lambda_n, \mu_m), j\rangle = 0$$

Poiché ogni $|\psi\rangle$ generica si può scrivere come combinazione di tali autostati (4.11), allora deve essere:

$$[A, B] |\psi\rangle = 0; \quad \forall |\psi\rangle \in \mathcal{H}$$

Concretizziamo le due **note** appena viste con un esempio. Siano $A, B \in \mathbb{R}^{3 \times 3}$ (matrici 3×3). Supponiamo che A abbia tre autovalori reali $\lambda_1, \lambda_2, \lambda_3$ con autospazi $E_{\lambda_1}, E_{\lambda_2}, E_{\lambda_3}$, mentre B ha due autovalori reali μ_1, μ_2 con autospazi $E_{\mu_1} = \text{Span}(E_{\lambda_1}, E_{\lambda_2})$ e $E_{\mu_2} = E_{\lambda_3}$, da cui $\sigma(A) = \{\lambda_1, \lambda_2, \lambda_3\}$, $\sigma(B) = \{\mu_1, \mu_2\}$.

A e B sono **compatibili**: gli autospazi $E_{\lambda_1}, E_{\lambda_2}, E_{\lambda_3}$ sono generati da 3 autovettori, che costituiscono una base di $\mathbb{R}^{3 \times 3}$, e sono anche autovettori di B .

Avremo allora:

$$\begin{aligned} \mathcal{H}_{(\lambda_1, \mu_1)} &= E_{\lambda_1} \\ \mathcal{H}_{(\lambda_2, \mu_1)} &= E_{\lambda_2} \\ \mathcal{H}_{(\lambda_3, \mu_2)} &= E_{\lambda_3} \\ \mathcal{H}_{(\lambda_1, \mu_2)} &= \mathcal{H}_{(\lambda_2, \mu_2)} = \mathcal{H}_{(\lambda_3, \mu_1)} = \emptyset \end{aligned}$$

Perciò gli autovalori comuni sono:

$$\begin{aligned} \sigma(A, B) &= \{(\lambda_1, \mu_1), (\lambda_2, \mu_1), (\lambda_3, \mu_2)\} \\ &\subset \{(\lambda_1, \mu_1), (\lambda_2, \mu_1), (\lambda_3, \mu_1), (\lambda_1, \mu_2), (\lambda_2, \mu_2), (\lambda_3, \mu_2)\} = \sigma(A) \times \sigma(B) \end{aligned}$$

E notiamo, come ci si aspetta, che:

$$\mathbb{R}^{3 \times 3} = \bigoplus_{(\lambda_n, \mu_m) \in \sigma(A, B)} \mathcal{H}_{(\lambda_n, \mu_m)}$$

Dimostriamo ora il viceversa, ossia che **commutatività** \Rightarrow **compatibilità**.

Abbiamo allora $[A, B] = 0$ per ipotesi. Data una $|\phi_n\rangle$ autostato di A di autovalore λ_n , allora:

$$AB|\phi_n\rangle \underset{(a)}{=} BA|\phi_n\rangle \underset{(b)}{=} \lambda_n B|\phi_n\rangle$$

In (a) si è usata la commutatività, e in (b) il fatto che $|\phi_n\rangle$ sia un autostato di A . Otteniamo quindi che $B|\phi_n\rangle$ è autostato di A di autovalore λ_n , cioè che applicare B ad uno autostato di A non cambia l'autovalore λ_n a cui è associato, ossia B mappa vettori dell'autospazio \mathcal{H}_{λ_n} in altri vettori di \mathcal{H}_{λ_n} .

Ciò suggerisce che vi sia una base comune ad entrambi gli operatori: costruiamola. Iniziamo notando che, poiché l'autospazio \mathcal{H}_{λ_n} è ortogonale a $\mathcal{H}_{\lambda_{n'}}$ perché $\lambda_n \neq \lambda_{n'}$ (autospazi di autovalori distinti), si ha che:

$$\underbrace{\langle \phi_{n'} |}_{\in \mathcal{H}_{\lambda_{n'}}} \underbrace{B | \phi_n \rangle}_{\in \mathcal{H}_{\lambda_n}} = 0 \quad \lambda_n \neq \lambda_{n'}$$

Consideriamo ora una base $|\phi_n\rangle$ di autovettori di A . Tale base esiste per teorema spettrale, dato che:

$$\mathcal{H} = \bigoplus_{n \in N} \mathcal{H}_{\lambda_n}$$

Se ora esaminiamo gli elementi di matrice di B in questa base, otteniamo una *matrice diagonale a blocchi*:

$$B = \begin{pmatrix} \boxed{B|_{\mathcal{H}_{\lambda_1}}} & 0 & 0 \\ 0 & \boxed{B|_{\mathcal{H}_{\lambda_2}}} & 0 \\ 0 & 0 & \ddots \end{pmatrix}$$

Se gli autovalori di B non hanno degenerazione, allora la matrice di sopra è esattamente diagonale, e quindi $|\phi_n\rangle$ è una *autobase* comune ad A e B , e la dimostrazione è finita. Nel caso generale, tuttavia, i blocchi non sono 1×1 , e quindi dobbiamo fare un ulteriore passaggio.

Notiamo che $B|_{\mathcal{H}_{\lambda_n}}$ è autoaggiunto in \mathcal{H}_{λ_n} (lo è in \mathcal{H} , e come appena visto \mathcal{H} si ottiene unendo \mathcal{H}_{λ_n}) e perciò ammette in \mathcal{H}_{λ_n} una base di autovettori, ciascuno corrispondente a una coppia $\lambda_n \in \sigma(A)$, $\mu_m \in \sigma(B)$. Tali autovettori, che denotiamo con $|(\lambda_n, \mu_m), j\rangle$, sono proprio la base di autovettori comuni di A e B che stavamo cercando, e pertanto A e B sono compatibili.

□

Poiché gli operatori autoaggiunti sono in corrispondenza biunivoca con le loro famiglie spettrali e queste famiglie spettrali sono composte di operatori limitati a spettro discreto (sono proiettori, quindi con solo autovalori 0 e 1) è naturale la definizione:

Definizione 26. A e B si dice che **commutano** se commutano le corrispondenti famiglie spettrali.

Commutatività di operatori

Nel caso di operatori limitati a spettro discreto le due cose sono equivalenti. In effetti, se A e B sono limitati e a spettro discreto possiamo scrivere le famiglie spettrali:

$$P^A(\lambda) = \sum_{\lambda_n \in \sigma(A)} \sum_{r=1}^{d(\lambda_n)} H(\lambda - \lambda_n) |\lambda_n, r\rangle \langle \lambda_n, r|$$

$$P^B(\mu) = \sum_{\mu_m \in \sigma(B)} \sum_{s=1}^{d(\mu_m)} H(\mu - \mu_m) |\mu_m, s\rangle \langle \mu_m, s|$$

Se $P^A(\lambda)$ e $P^B(\mu)$ commutano:

$$[P^A(\lambda), P^B(\mu)] = 0 \Leftrightarrow [|\lambda_n, r\rangle \langle \lambda_n, r|, |\mu_m, s\rangle \langle \mu_m, s|] = 0 \quad \forall \lambda_n, \mu_m \quad (4.12)$$

Ma allora, scrivendo A e B tramite decomposizione spettrale:

$$A = \sum_{\lambda_n \in \sigma(A)} \lambda_n \sum_{r=1}^{d(\lambda_n)} |\lambda_n, r\rangle \langle \lambda_n, r|$$

$$B = \sum_{\mu_m \in \sigma(B)} \mu_m \sum_{s=1}^{d(\mu_m)} |\mu_m, s\rangle \langle \mu_m, s|$$

notiamo che $AB - BA$ non è altro che una combinazione lineare di termini nulli per (4.12), e perciò $AB - BA = [A, B] = 0$. Abbiamo allora mostrato che le famiglie spettrali di A e B commutano $\Leftrightarrow [A, B] = 0$ (almeno nel caso di A e B limitati a spettro discreto).

Il caso generale di A e B autoaggiunti (non necessariamente limitati e con spettro generico) non è altrettanto semplice: nella definizione di commutatore entrano problemi di dominio. Tuttavia, si può dimostrare che il caso particolare che abbiamo esaminato ha validità generale, e vale cioè:

Teorema 4.2.2. A e B **autoaggiunti** (con spettro qualsiasi) sono **compatibili** (proprietà fisica) se e solo se **commutano** (proprietà algebrica), ossia se commutano le loro **famiglie spettrali**.

Dimostrazione: omessa.

(Lezione 21 del 15/11/2018)

Chiaramente, due osservabili A e B possono essere (banalmente) compatibili se uno è una funzione dell'altro, $B = f(A)$. Si dicono allora **dipendenti**, e le misure di una *determinano* le misure dell'altra. In particolare si avrà $\sigma(B) = \sigma(f(A)) = f(\sigma(A))$. Se ciò non succede, A e B si dicono **indipendenti**.

Notiamo che la dipendenza di osservabili compatibili si può verificare sperimentalmente (possiamo misurare entrambe le osservabili e vedere se il valore di una coincide con f calcolata nel valore dell'altra).

Definizione 27. Un insieme \mathcal{C} di osservabili compatibili indipendenti si dice **completo** se è massimale, ossia se non esiste alcun insieme di osservabili compatibili indipendenti che lo contenga propriamente (non si possono trovare insiemi strettamente “più grandi”).

Per le osservabili a spettro discreto, il risultato di una misura di I specie è l'individuazione del sottospazio di \mathcal{H} corrispondente all'autovalore trovato (quello *compatibile* con la misura ottenuta).

Ci aspettiamo allora che, una volta misurate tutte le osservabili in \mathcal{C} , si abbia un'informazione *massimale* sul sistema, ossia che la misura di tutte le \mathcal{C} proietti il sistema su uno **stato puro**.

Teorema 4.2.3. Sia $\mathcal{C} = \{A_1, \dots, A_n\}$ un insieme di osservabili compatibili indipendenti a spettro discreto. Allora \mathcal{C} è **completo** se e solo se le osservabili A_1, \dots, A_n hanno un insieme completo (cioè una base tale da verificare la completezza di Dirac) di autovettori comuni **senza degenerazione**.

ICOC (a spettro discreto)

In altre parole:

$$\dim \mathcal{H}_{(a_1, \dots, a_n)} = 1 \quad \forall (a_1, \dots, a_n) \in \sigma(A_1, \dots, A_n)$$

Dim. Partiamo dimostrando che *non degenerare* \Rightarrow *completo*. Consideriamo un'osservabile B compatibile con tutte le osservabili in \mathcal{C} , e verifichiamo che, se la base di autovettori comuni non ha degenerazione, B deve essere per forza **dipendente**, ossia funzione delle osservabili in \mathcal{C} .

Indichiamo con $\{|(a_1, \dots, a_n)\rangle\}$ gli autovettori comuni di \mathcal{C} , dove $(a_1, \dots, a_n) \in \sigma(\mathcal{C})$ sono gli autovalori comuni delle osservabili $A_1, \dots, A_n \in \mathcal{C}$ (e in generale, come visto precedentemente, $\sigma(\mathcal{C}) \neq \sigma(A_1) \times \dots \times \sigma(A_n)$).

Per ipotesi di compatibilità, tali autovettori costituiscono una base completa. Se, in più, non vi è degenerazione, la completezza di Dirac si scrive come:

$$\sum_{(a_1, \dots, a_n) \in \sigma(\mathcal{C})} |(a_1, \dots, a_n)\rangle \langle (a_1, \dots, a_n)| = \mathbb{I} \quad (4.13)$$

Poiché B è compatibile con \mathcal{C} , ogni $|(a_1, \dots, a_n)\rangle$ deve essere anche autovettore di B e quindi anche della famiglia spettrale $P^B(\lambda)$ di B per il teorema (4.2.2).

Vale allora l'equazione agli autovalori per $P^B(\lambda)$:

$$P^B(\lambda) |(a_1, \dots, a_n)\rangle = f_\lambda(a_1, \dots, a_n) |(a_1, \dots, a_n)\rangle \quad (4.14)$$

dove gli *autovalori* sono dati da una qualche funzione $f_\lambda(a_1, \dots, a_n)$, che è iniettiva per la non-degenerazione (cambiare uno degli a_i significa considerare un altro autovettore, che deve quindi avere un altro autovalore poiché tutti gli autospazi hanno dimensione 1. Ma allora f manda punti diversi in punti diversi).

Allora combinando (4.13) con (4.14) si ha:

$$\begin{aligned} P^B(\lambda)\mathbb{I} &= P^B(\lambda) = \sum_{(a_1, \dots, a_n) \in \sigma(\mathcal{C})} f_\lambda(a_1, \dots, a_n) |(a_1, \dots, a_n)\rangle \langle (a_1, \dots, a_n)| \\ &\equiv f_\lambda(A_1, \dots, A_n) \end{aligned}$$

dove si è usata la definizione di rappresentazione spettrale di una funzione delle osservabili.

Quindi B è una funzione di \mathcal{C} individuata da:

$$B = \int_{\mathbb{R}} \lambda dP^B(\lambda) = \int_{\mathbb{R}} \lambda df_{\lambda}(A_1, \dots, A_n)$$

ed è pertanto dipendente da \mathcal{C} , quindi \mathcal{C} è completo.

Viceversa, dimostriamo che se vi è **degenerazione** > 1 allora \mathcal{C} **non è completo**. Supponiamo che per qualche autovalore $(a_1, \dots, a_n) \in \sigma(\mathcal{C})$ l'autospazio $\mathcal{H}_{(a_1, \dots, a_n)}$ abbia dimensione $d(a_1, \dots, a_n) > 1$ (ossia vi è degenerazione).

Allora dato un operatore B autoaggiunto non banale in $\mathcal{H}_{(a_1, \dots, a_n)}$ (ossia una qualsiasi matrice hermitiana in questo autospazio che non sia nulla o multiplo dell'identità), lo estendiamo all'identità sul complemento di $\mathcal{H}_{(a_1, \dots, a_n)}$ in \mathcal{H} (cioè lo estendiamo ad un operatore \tilde{B} definito su tutto \mathcal{H} che è pari a B sull'autospazio $\mathcal{H}_{(a_1, \dots, a_n)}$, e agisce come l'identità su tutto il resto).

Nella base comune di \mathcal{C} , la forma matriciale di B è allora del tipo:

$$\tilde{B} = \begin{pmatrix} 1 & \dots & \dots & \dots & \dots & \dots & 0 \\ \vdots & \ddots & 0 & & & & \vdots \\ \vdots & 0 & 1 & 0 & & & \vdots \\ \vdots & \vdots & 0 & \boxed{B} & 0 & & \vdots \\ \vdots & & & 0 & 1 & 0 & \vdots \\ \vdots & & & & 0 & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix}$$

Dove il blocco centrale ha dimensione pari alla degenerazione $d(a_1, \dots, a_n)$, e quindi è certamente più grande di un 1×1 .

Tale rappresentazione matriciale mostra che B manda vettori dell'autospazio $\mathcal{H}_{(a_1, \dots, a_n)}$ in vettori che rimangono al suo interno, e quindi *non disturba* le misure di tutti gli altri operatori, e infatti commuta³ con A_1, \dots, A_n .

Tuttavia, se misuriamo A_1, \dots, A_n e otteniamo come risultato (a_1, \dots, a_n) lo stato del sistema sarà una combinazione lineare degli autostati di B (data la degenerazione), e scegliendo opportunamente B in modo che abbia più di un autovalore, il risultato di una seguente misura di B non è determinato.

Abbiamo allora trovato un operatore \tilde{B} che è compatibile con tutti i \mathcal{C} e non dipende funzionalmente da essi: perciò \mathcal{C} non è completo.

□

³ Più precisamente, basta seguire passi analoghi alla dimostrazione del teorema di compatibilità. In particolare, B è per ipotesi autoaggiunto, e quindi ammette una base su $\mathcal{H}_{(a_1, \dots, a_n)}$ che diagonalizza il *blocco centrale*. Ma allora ha una base di autovettori comune a tutte le \mathcal{C} , e quindi è compatibile con esse.

Osservazione 5. Poiché una misura su un Insieme Completo di Osservabili Compatibili (ICOC) \mathcal{C} a spettro discreto identifica uno spazio 1-dimensionale in \mathcal{H} , tale misura può essere usata per preparare uno stato puro.

Osservazione 6. Sotto opportune condizioni sugli spazi nucleari ϕ_{A_1, \dots, A_n} comuni a \mathcal{C} (che generalizzano la nozione di *autospazio comune*), il teorema precedente si estende ad un ICOC a spettro arbitrario, includendo gli autovettori generalizzati (cioè possiamo usarlo anche per *spettri continui*).

Questo consente di dare una definizione precisa di **rappresentazione** \mathcal{C} .

Rappresentazioni

Partiamo scrivendo la *completezza generalizzata*⁴ per un ICOC \mathcal{C} :

$$\prod_{i=1}^n \left(\sum_{a_i \in \sigma_P(A_i)} + \int_{\sigma_C(A_i)} da_i \right) |(a_1, \dots, a_n)\rangle \langle (a_1, \dots, a_n)| = \mathbb{I}$$

ove $|(a_1, \dots, a_n)\rangle$ sono autovettori (generalizzati) comuni di \mathcal{C} **nondegeneri**. Per ciò ogni vettore $|(a_1, \dots, a_n)\rangle$ è univocamente individuato dai valori (a_1, \dots, a_n) e ogni $|\psi\rangle \in \mathcal{H}$ si può scrivere:

$$|\psi\rangle = \prod_{i=1}^n \left(\sum_{a_i \in \sigma_P(A_i)} + \int_{\sigma_C(A_i)} da_i \right) |(a_1, \dots, a_n)\rangle \langle (a_1, \dots, a_n) | \psi\rangle$$

Quindi i coefficienti $\{\langle (a_1, \dots, a_n) | \psi\rangle = \psi(a_1, \dots, a_n)\}$, con $(a_1, \dots, a_n) \in \sigma(\mathcal{C})$, identificano univocamente lo stato $|\psi\rangle$ e vengono detti definire lo stato in **rappresentazione** \mathcal{C} .

Per esempio, supponiamo $\mathcal{C} = \{\vec{X}\}$, allora sappiamo che:

$$\langle \vec{x} | \psi\rangle = \psi(\vec{x})$$

Se invece $\mathcal{C} = \{\vec{P}\}$, allora:

$$\langle \vec{p} | \psi\rangle = \tilde{\psi}(\vec{p})$$

Per una particella in \mathbb{R}^3 potremmo scegliere $\mathcal{C} = \{X, P_y, Z\}$, mentre in \mathbb{R} si può avere $\mathcal{C} = \{H, \mathcal{P}\}$, dato che se $[H, \mathcal{P}] = 0$ *rimuove* la degenerazione degli autovalori di H .

Definizione 28. Un insieme di osservabili \mathcal{I} (non compatibili!) si dice **irriducibile** se ogni osservabile che commuta con tutti gli elementi di \mathcal{I} è un multiplo dell'identità \mathbb{I} (ossia il risultato della sua misura è **indipendente dallo stato** - es. come nel caso dell'osservabile che misura la carica elettrica dell'elettrone).

Osservazione 7. Per una particella quantistica elementare *con analogo classico*⁵ $\{\vec{X}, \vec{P}\}$ è un insieme irriducibile.

Infatti sia A un'osservabile che commuta con \vec{X} , poiché $\mathcal{C} = \{\vec{X}\}$ è un ICOC,

⁴La produttoria è data dal fatto che stiamo considerando l'applicazione simultanea di più osservabili. Per esempio, se $A|\psi\rangle = a|\psi\rangle$ e $B|\psi\rangle = b|\psi\rangle$, con A e B compatibili si ha $AB|\psi\rangle = ab|\psi\rangle$. Si noti inoltre l'assenza di sommatorie sugli indici di degenerazione, dato che per un ICOC gli autovettori sono *nondegeneri* per ipotesi.

⁵Ossia senza proprietà unicamente quantistiche, come lo *spin*

necessariamente $A = f(\vec{X})$. Ma se A commuta anche con \vec{P} , vuol dire che (usando la rappresentazione in \vec{x}):

$$\begin{aligned} 0 &= [\vec{P}, \underbrace{f(\vec{X})}_A] \psi(\vec{X}) = \vec{P}(f(\vec{X})\psi(\vec{X})) - f(\vec{X})\vec{P}(\psi(\vec{X})) = \\ &= -i\hbar \vec{\nabla}(f(x)\psi(x)) + i\hbar f(\vec{x})\vec{\nabla}\psi(\vec{x}) = -i\hbar(\vec{\nabla}f)(\vec{x})\psi(\vec{x}) = 0 \quad \forall \psi \end{aligned}$$

Ma allora f è una costante:

$$\vec{\nabla}f(\vec{x}) = 0 = f(\vec{x}) = \text{cost.}$$

4.3 Esercizio 3

Una particella di massa m in 1D è immersa nel potenziale (buca infinita):

$$V(x) = \begin{cases} 0 & |x| < \frac{a}{2} \\ +\infty & \text{altrove} \end{cases}$$

Sia $\mathcal{E} = \frac{2\pi^2\hbar^2}{ma^2}$ l'energia dello stato a $t = -t_0$ ($t_0 > 0$).

*Richieste
dell'esercizio*

1. All'istante $t = 0$ una misura (ideale di I specie) rivela la presenza della particella nella parte destra della buca. Determinare la funzione d'onda dopo la misura.
2. All'istante t_0 si esegue una misura⁶, determinare la probabilità di trovare un'energia più alta di quelle a $t = t_0$.
3. Determinare la probabilità di trovare il valore $+1$ eseguendo una misura di *parità* \mathcal{P} a $t = t_0$
4. Determinare $(\Delta X)_\psi(\Delta P)_\psi$ per tutti gli stati stazionari ψ del sistema.
5. Si dica se P (con condizioni periodiche) e H sono compatibili.
6. Si determini la probabilità di trovare la particella nella metà sinistra della buca per $0 < t < t_0$.

4.3.1 Soluzione

1. Dalla soluzione del potenziale della buca infinita (CFR sezione 3.8.5) sappiamo che gli autovalori di H sono:

$$\mathcal{E}_n = \frac{\hbar^2}{2m} \frac{n^2\pi^2}{a^2} \quad (4.15)$$

⁶^Che, quando non specificato diversamente, è da intendersi come ideale di prima specie.

e le rispettive autofunzioni sono:

$$\varphi_n(x) = \begin{cases} \sqrt{\frac{2}{a}} \cos\left(\frac{n\pi x}{a}\right) & n \in \mathbb{N} \text{ dispari} \\ \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) & n \in \mathbb{N} \text{ pari} \end{cases}$$

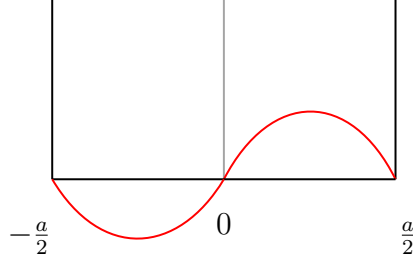


Figura 4.1: ψ iniziale

Ricordiamo che gli autovalori \mathcal{E}_n **non hanno degenerazione**: ogni \mathcal{E}_n si ottiene per due valori, n e $-n$, che però definiscono *lo stesso stato*. Basta infatti sostituirli nelle φ_n : nel caso dispari non cambia nulla, e in quello pari si ottiene un cambio di segno, ossia uno sfasamento che non modifica l'auto-stato.

Notiamo che l'energia iniziale assegnata $\mathcal{E} = \mathcal{E}_2$ (chiaramente il valore assegnato **deve** essere uno di quelli possibili, dato che non ogni valore è possibile come in MC).

Lo stato $|\psi_0\rangle$ a $t = -t_0$ è allora un **autostato** di H (per la precisione $|\varphi_2\rangle$, autofunzione di autovalore \mathcal{E}_2). Si tratta allora di uno **stato stazionario**, ossia di uno stato che rimane inalterato fino a che il sistema resta isolato, ossia fino a quando non si effettua una misura.

Matematicante, gli *stati stazionari* sono le soluzioni dell'equazione di Schrödinger stazionaria, che in effetti è proprio l'equazione agli autovalori di H : $H|\psi\rangle = \mathcal{E}|\psi\rangle$.

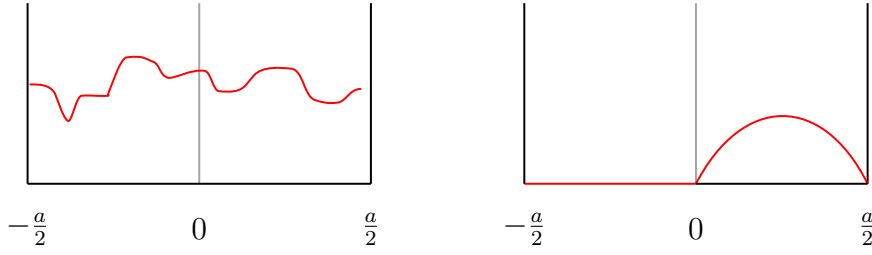
Stati stazionari

Se esaminiamo l'evoluzione di $|\psi_0\rangle$ applicando direttamente la formula (3.41), otteniamo infatti un conto *semplice*:

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}t\mathcal{E}_n} |\mathcal{E}_2\rangle \underbrace{\langle \mathcal{E}_2 | \psi_0 \rangle}_{\langle \varphi_2 | \varphi_2 \rangle = 1} = e^{-\frac{i}{\hbar}\mathcal{E}_2 t} |\psi_0\rangle$$

Dato che tutti gli altri termini della sommatoria sono nulli, essendo $|\psi_0\rangle = |\varphi_2\rangle = |\mathcal{E}_2\rangle$.

Il conto appena fatto mostra come l'evoluzione temporale della $|\psi_0\rangle$ sia un'*oscillazione di fase*. Se esaminiamo la parte reale (o immaginaria) della $|\psi(t)\rangle$ osserviamo infatti un'*onda stazionaria*.



(a) Effetto della misura

(b) ψ dopo la misura

Figura 4.2: Misura della posizione nella buca di potenziale

Calcolando per $t = 0$ giungiamo allora a:

$$\psi(x, t = 0) = e^{-\frac{i}{\hbar} t_0 \mathcal{E}_2} \psi(x, -t_0)$$

Ma la $\psi(x, -t_0)$ di partenza è definita a meno di una fase, e possiamo sceglierla opportunamente in modo che l'esponenziale generato dall'evoluzione temporale scompaia. Giungiamo allora a:

$$\psi(x, 0^-) = \sqrt{\frac{2}{a}} \sin\left(\frac{2\pi x}{a}\right)$$

A $t = 0$ l'atto di misura proietta (assioma di von Neumann) la funzione d'onda sugli autostati compatibili con la misura osservata:

$$\psi(x, 0^+) = P^X\left(\left[0, \frac{a}{2}\right]\right) \psi(x, 0^-) = A \begin{cases} \sqrt{\frac{2}{a}} \sin\left(\frac{2\pi x}{a}\right) & 0 < x < \frac{a}{2} \\ 0 & -\frac{a}{2} < x < 0 \end{cases}$$

Dove A è la nuova costante di normalizzazione per $\psi(x, 0^+)$.

Poiché $\psi(x, 0)$ era già un autostato, la parte di essa che fa parte del “dominio compatibile con la misura”, ossia $[0, +a/2]$, viene *proiettata su se stessa*, e quindi non cambia. Tutto il resto viene invece portato a 0.

Normalizziamo per trovare A :

$$\begin{aligned} 1 &\stackrel{!}{=} \int_{-\frac{a}{2}}^{+\frac{a}{2}} dx |\psi(x, 0^+)|^2 = |A|^2 \int_0^{+\frac{a}{2}} \frac{2}{a} \sin^2\left(\frac{2\pi x}{a}\right) dx = \\ &= |A|^2 \int_0^{+\frac{a}{2}} \frac{2}{a} \frac{1}{2} \left(1 - \cos\left(\frac{4\pi x}{a}\right)\right) dx = \\ &= \frac{|A|^2}{a} \frac{1}{2} \left[x - \frac{a}{4\pi} \sin\left(\frac{4\pi x}{a}\right)\right]_0^{+\frac{a}{2}} = |A|^2 \frac{1}{a} \frac{a}{2} = \frac{|A|^2}{2} = 1 \Rightarrow A = \sqrt{2} \end{aligned}$$

Perciò:

$$\psi(x, 0^+) = \begin{cases} \frac{2}{\sqrt{a}} \sin\left(\frac{2\pi x}{a}\right) & 0 < x < +\frac{a}{2} \\ 0 & -\frac{a}{2} < x < 0 \end{cases}$$

Notiamo che $\psi(x, 0^+)$ **non** è più un autostato di H .

2. Per i calcoli di probabilità è necessaria la funzione d'onda al tempo $t_0 > 0$. Per determinare $\psi(x, t)$ per $t > 0$ (fino a t_0) dobbiamo **espandere** $\psi(x, 0^+)$ negli **autostati** $\varphi_n(x)$ di H :

$$\psi(x, 0^+) = \sum_{n=1}^{\infty} \underbrace{(\varphi_n, \psi(x, 0^+))}_{c_n} \varphi_n(x)$$

Da cui la formula (3.41) conduce a:

$$\psi(x, t) = \sum_{n=1}^{\infty} c_n e^{-\frac{i}{\hbar} \mathcal{E}_n t} \varphi_n(x)$$

Determiniamo quindi i c_n . Per n pari, ed essendo φ_n a valori reali:

$$c_n = \int_0^{+\frac{a}{2}} \underbrace{\sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right)}_{\varphi_n^* = \varphi_n} \underbrace{\frac{2}{\sqrt{a}} \sin\left(\frac{2\pi x}{a}\right)}_{\psi(x, 0^+)} dx$$

Applicando le formule di Werner:

$$\sin(\alpha) \sin(\beta) = \frac{1}{2} (\cos(\alpha - \beta) - \cos(\alpha + \beta))$$

Si ha che:

$$c_n = \frac{2\sqrt{2}}{a} \int_0^{+\frac{a}{2}} \frac{1}{2} \left[\cos \frac{(n-2)\pi x}{a} - \cos \frac{(n+2)\pi x}{a} \right] dx$$

Per $n \neq 2$ si ottiene (ricordando che n pari):

$$\frac{\sqrt{2}}{a} \left[\frac{a}{(n-2)\pi} \sin \frac{(n-2)\pi x}{a} \Big|_0^{+\frac{a}{2}} - \frac{a}{(n+2)\pi} \sin \frac{(n+2)\pi x}{a} \Big|_0^{+\frac{a}{2}} \right] = 0$$

Per $n = 2$ la funzione integranda è \sin^2 , e perciò:

$$\begin{aligned} c_2 &= \frac{2\sqrt{2}}{a} \int_0^{+\frac{a}{2}} \sin^2\left(\frac{2\pi x}{a}\right) dx = \\ &= \frac{2\sqrt{2}}{a} \left[x - \frac{a}{4\pi} \sin\left(\frac{4\pi x}{a}\right) \right]_0^{+\frac{a}{2}} = \frac{\sqrt{2}a}{a} \frac{1}{2} = \frac{1}{\sqrt{2}} \end{aligned}$$

Per n dispari, analogamente:

(Lezione 22 del
16/11/2018)

$$c_n = \int_0^{+\frac{a}{2}} \sqrt{\frac{2}{a}} \cos\left(\frac{n\pi x}{a}\right) \frac{2}{\sqrt{a}} \sin\left(\frac{2\pi x}{a}\right) dx$$

Applicando le formule di Werner:

$$\sin(\alpha) \cos(\beta) = \frac{1}{2} (\sin(\alpha + \beta) + \sin(\alpha - \beta))$$

si giunge a:

$$\begin{aligned}
c_n &= \int_0^{\frac{a}{2}} \frac{2\sqrt{2}}{a} \frac{1}{2} \left[\sin \frac{(2+n)\pi x}{a} + \sin \frac{(2-n)\pi x}{a} \right] dx = \\
&= \frac{\sqrt{2}}{\mathfrak{A}} \left[\frac{\mathfrak{A}}{(2+n)\pi} \left(-\cos \frac{(n+2)\pi x}{a} \right) \Big|_0^{\frac{a}{2}} + \frac{\mathfrak{A}}{(2-n)\pi} \left(-\cos \frac{(2-n)\pi x}{a} \right) \Big|_0^{\frac{a}{2}} \right] = \\
&= \sqrt{2} \left[\frac{1}{(2+n)\pi} + \frac{1}{(2-n)\pi} \right] = \sqrt{2} \frac{4}{4-n^2} \frac{1}{\pi}
\end{aligned}$$

Riepilogando, i valori di $c_n = (\varphi_n, \psi(x, 0^+))$ sono dati da:

$$\begin{aligned}
&\begin{matrix} n \text{ pari} \\ \\ n \text{ dispari} \end{matrix} \quad \begin{cases} c_n = 0 & n \neq 2 \\ c_2 = \frac{1}{\sqrt{2}} \\ c_n = \sqrt{2} \frac{4}{4-n^2} \frac{1}{\pi} \end{cases}
\end{aligned}$$

Possiamo finalmente calcolare la probabilità che la nuova funzione d'onda abbia a t_0 energia più alta di \mathcal{E}_2 . Ciò è equivalente al problema (più semplice) di calcolare la probabilità che $|\psi\rangle$ **non** abbia energia pari a \mathcal{E}_1 o a \mathcal{E}_2 . Applicando allora la formula (3.63) (che è già stata vista proprio nel caso della buca infinita in (3.66)):

$$\begin{aligned}
P_{\psi(t)}^H(\mathcal{E} > \mathcal{E}_2) &= 1 - W_{\psi(t)}^H(\mathcal{E}_1) - W_{\psi(t)}^H(\mathcal{E}_2) \\
W_{\psi(t)}^H(\mathcal{E}_n) &= |(\psi(t), \varphi_n)|^2 = |(\psi(0^+), \varphi_n)|^2 = |c_n|^2
\end{aligned}$$

dato che $\psi(t) = U(t)\psi(0^+)$, e $U(t)$ è unitario, cioè aggiunge solo una fase $\exp(-\frac{i}{\hbar}\mathcal{E}_n t)$ alla $\psi(0^+)$, senza modificare il modulo. Svolgendo allora i conti:

$$P_{\psi(t)}^H(\mathcal{E} > \mathcal{E}_2) = 1 - |c_1|^2 - |c_2|^2 = 1 - \left| \sqrt{2} \left(\frac{1}{\pi} + \frac{1}{3\pi} \right) \right|^2 - \frac{1}{2} \sim 0.14$$

3. Calcoliamo ora la probabilità che la parità applicata a $\psi(t_0)$ dia $+1$, ossia $W_{\psi(t_0)}^{\mathcal{P}}(+1)$. Sappiamo che la parità agisce come:

$$\mathcal{P}\psi(x, t_0) = \psi(-x, t_0)$$

Per calcolare $W_{\psi(t_0)}(+1)$ dobbiamo costruire il proiettore $P^{\mathcal{P}}(\{+1\})$ che proietta un generico stato $|\psi\rangle$ nell'autospazio di \mathcal{P} di autovalore $+1$. Una base di questo autospazio è per esempio data dalle φ_n con n dispari (i cos, che sono pari), e quindi $+1$ è autovalore a degenerazione ∞ .

Infatti, dato che $[H, \mathcal{P}] = 0$, H e \mathcal{P} ammettono una base di autovettori comuni. In particolare, una base comune per H e l'autospazio di \mathcal{P} corrispondente all'autovalore $+1$ è data dai $|n\rangle$ con n dispari, la cui rappresentazione in x è:

$$\langle x | n \rangle = \cos \left(\frac{n\pi x}{a} \right)$$

Il relativo proiettore $P^{\mathcal{P}}(\{+1\})$ si ottiene quindi per somma infinita dei proiettori su tali autostati:

$$P^{\mathcal{P}}(\{+1\}) = \sum_{n \text{ dispari}} |n\rangle \langle n|$$

Analogamente, l'autospazio di \mathcal{P} di autovalore -1 è generato dai $|n\rangle$ con n pari e:

$$\langle x | n \rangle = \sin\left(\frac{n\pi x}{a}\right); \quad P^{\mathcal{P}}(\{-1\}) = \sum_{n \text{ pari}} |n\rangle \langle n|$$

Tuttavia, possiamo evitare di calcolare una somma infinita per trovare il proiettore, notando che $\mathcal{P}^2 = \mathbb{I}$, da cui $(1 + \mathcal{P})/2$ è un proiettore. Infatti:

$$\left(\frac{1 + \mathcal{P}}{2}\right) \left(\frac{1 + \mathcal{P}}{2}\right) = \frac{1}{4} + \frac{1}{2}\mathcal{P} + \frac{1}{4}\mathcal{P}^2 = \frac{1}{4} + \frac{1}{2}\mathcal{P} + \frac{1}{4} = \frac{1 + \mathcal{P}}{2}$$

E tale operatore agisce come:

$$\frac{1 + \mathcal{P}}{2}\psi(x) = \frac{\psi(x) + \psi(-x)}{2} = \begin{cases} \psi(x) & \text{se } \psi(x) \text{ è pari} \\ 0 & \text{se } \psi(x) \text{ è dispari} \end{cases}$$

Quindi $(1 + \mathcal{P})/2$ è il proiettore che proietta nell'autospazio $+1$ di \mathcal{P} che volevamo. La probabilità cercata è allora:

$$\begin{aligned} W_{\psi(t)}^{\mathcal{P}}(+1) &= \left(\psi(t), \frac{\mathbb{I} + \mathcal{P}}{2} \psi(t) \right) \stackrel{(a)}{=} \\ &= \left(\sum_n c_n \exp\left(-\frac{i}{\hbar} \mathcal{E}_n t\right) \varphi_n, \left(\frac{\mathbb{I} + \mathcal{P}}{2}\right) \underbrace{\sum_m c_m \exp\left(-\frac{i}{\hbar} \mathcal{E}_m t\right) \varphi_m}_{m \text{ dispari}} \right) \stackrel{(b)}{=} \\ &= \sum_{n, m \text{ dispari}} c_n^* c_m \exp\left(-\frac{i}{\hbar} (\mathcal{E}_m - \mathcal{E}_n) t\right) \underbrace{(\varphi_n, \varphi_m)}_{\delta_{nm}} = \sum_{n \text{ dispari}} |c_n|^2 \end{aligned}$$

dove in (a) si sono sviluppate le $\psi(t)$ sulla base degli autostati di H come da formula (3.41), e in (b) si è ristretta la somma ai soli indici dispari, dato che, nel secondo membro del prodotto scalare, all'applicazione del proiettore "sopravvivono" solo le φ_m pari, che sono quelle dove l'indice m è dispari.

Di nuovo, non serve calcolare la somma infinita per $W_{\psi(t)}^{\mathcal{P}}(+1)$, poiché sappiamo che, per il caso opposto di $W_{\psi(t)}^{\mathcal{P}}(-1)$, la somma è costituita da un solo termine (in quanto solo $c_2 \neq 0$). Basterà allora calcolare la probabilità che l'operatore parità non dia -1 :

$$W_{\psi(t)}^{\mathcal{P}}(+1) = 1 - W_{\psi(t)}^{\mathcal{P}}(-1) = 1 - |c_2|^2 = 1 - \frac{1}{2} = \frac{1}{2}$$

4. Gli stati stazionari sono gli autostati dell'Hamiltoniana H , ossia le φ_n . Dobbiamo allora calcolare $(\Delta X)_{\varphi_n}(\Delta P)_{\varphi_n}$.

Nota: Nei prossimi passaggi semplificheremo la notazione, abbreviando il valor medio di O su φ_n con $\langle O \rangle_{\varphi_n} \equiv \langle O \rangle_n$ (omettendo la φ).

Partiamo calcolando $(\Delta X)_n$ per n dispari:

$$\begin{aligned}
(\Delta X)_{\varphi_n} &\equiv (\Delta X)_n = \sqrt{\langle X^2 \rangle_n - \langle X \rangle_n^2} \\
\underbrace{\langle X^2 \rangle_n}_{n \text{ dispari}} &= \langle \varphi_n | X^2 \varphi_n \rangle = \langle \varphi_n | x^2 \varphi_n \rangle = \int_{-\frac{a}{2}}^{+\frac{a}{2}} \varphi_n^*(x) x^2 \varphi_n(x) dx \stackrel{\varphi_n(x) \in \mathbb{R}}{=} \int_{-\frac{a}{2}}^{+\frac{a}{2}} \varphi_n(x) x^2 \varphi_n(x) dx \\
&= \int_{-\frac{a}{2}}^{+\frac{a}{2}} \sqrt{\frac{2}{a}} \left(\cos \frac{n\pi x}{a} \right) x^2 \sqrt{\frac{2}{a}} \left(\cos \frac{n\pi x}{a} \right) dx = \\
&= \int_{-\frac{a}{2}}^{+\frac{a}{2}} x^2 \cos^2 \frac{n\pi x}{a} dx = \frac{2}{a} \int_{-\frac{a}{2}}^{+\frac{a}{2}} \frac{x^2}{2} \left(1 + \cos \frac{2n\pi x}{a} \right) dx = \\
&= \frac{1}{a} \frac{x^3}{3} \Big|_{-\frac{a}{2}}^{+\frac{a}{2}} + \frac{1}{a} \frac{a}{2\pi n} \int_{-\frac{a}{2}}^{+\frac{a}{2}} x^2 \frac{d}{dx} \sin \frac{2n\pi x}{a} dx \stackrel{\text{per parti}}{=} \\
&= \frac{a^2}{12} + \frac{1}{2\pi n} x^2 \sin \frac{2n\pi x}{a} \Big|_{-\frac{a}{2}}^{+\frac{a}{2}} - \frac{1}{2\pi n} \int_{-\frac{a}{2}}^{+\frac{a}{2}} 2x \sin \frac{2n\pi x}{a} dx = \\
&= \frac{a^2}{12} - \frac{a}{2(\pi n)^2} x \cos \frac{2\pi n x}{a} \Big|_{-\frac{a}{2}}^{+\frac{a}{2}} + \frac{a}{2(\pi n)^2} \int_{-\frac{a}{2}}^{+\frac{a}{2}} \cos \frac{2\pi n x}{a} dx = \\
&= \frac{a^2}{12} + \frac{a}{2(\pi n)^2} \frac{a}{2} 2(-1)^n + \frac{a}{2(\pi n)^2} \frac{a}{2\pi n} \sin \frac{2\pi n x}{a} \Big|_{-\frac{a}{2}}^{+\frac{a}{2}} = \\
&= a^2 \left(\frac{1}{12} - \frac{1}{2(n\pi)^2} \right) \\
\langle X \rangle_n &= \int_{-\frac{a}{2}}^{+\frac{a}{2}} \sqrt{\frac{2}{a}} \left(\cos \frac{n\pi x}{a} \right) x \sqrt{\frac{2}{a}} \left(\cos \frac{n\pi x}{a} \right) dx = \\
&= \frac{1}{a} \left(\left[\frac{x^2}{2} \right]_{-\frac{a}{2}}^{+\frac{a}{2}} + \int_{-\frac{a}{2}}^{+\frac{a}{2}} x \cos \left(\frac{2n\pi x}{a} \right) dx \right) \stackrel{(a)}{=} 0
\end{aligned}$$

Dato che in (a) stiamo integrando una funzione *dispari* su un dominio simmetrico attorno a 0.

Perciò (n dispari):

$$(\Delta X)_n = a \sqrt{\frac{1}{12} - \frac{1}{2(n\pi)^2}}$$

Calcoliamo ora la fluttuazione del momento:

$$(\Delta P)_n = \sqrt{\langle P^2 \rangle_n - \langle P \rangle_n^2}$$

$\langle P^2 \rangle_n$ lo possiamo ricavare direttamente ricordando che P^2 è funzione di H , e sappiamo già i valori medi di H nei suoi autostati: $\langle H \rangle_n = \mathcal{E}_n$. Ricordando la (4.15) otteniamo:

$$\langle P^2 \rangle_n = \langle 2mH \rangle_n = 2m \frac{\hbar^2}{2m} \frac{\pi^2 n^2}{a^2} = \frac{\hbar^2 \pi^2 n^2}{a^2}$$

Ricordiamo però che il dominio di P su un insieme compatto (come l'intervallo $[-a/2, +a/2]$ che stiamo considerando) richiede, oltre all'integrabilità in modulo quadro e alla regolarità, anche delle *condizioni periodiche* (altrimenti P non è autoaggiunto).

Perciò, quando calcoliamo il valor medio:

$$\langle P \rangle_n = \langle \varphi_n | P \varphi_n \rangle$$

dobbiamo verificare che le $\varphi_n \in D(P)$.

Le *condizioni periodiche* necessarie perché una ψ regolare e modulo quadro integrabile sia in $D(P)$ sono:

$$\psi\left(-\frac{a}{2}\right) = \psi\left(\frac{a}{2}\right)$$

E si verifica immediatamente che le φ_n le rispettano:

$$\varphi_n\left(-\frac{a}{2}\right) = 0 = \varphi_n\left(\frac{a}{2}\right)$$

E allora $\varphi_n \in D(P)$ e possiamo calcolare il valor medio:

$$\begin{aligned} \langle P \rangle_n &= \int_{-\frac{a}{2}}^{+\frac{a}{2}} \sqrt{\frac{2}{a}} \cos\left(\frac{n\pi x}{a}\right) \left(-i\hbar \frac{d}{dx}\right) \sqrt{\frac{2}{a}} \cos\left(\frac{n\pi x}{a}\right) dx = \\ &= \frac{2}{a} i\hbar \frac{n\pi}{a} \int_{-\frac{a}{2}}^{+\frac{a}{2}} \cos\left(\frac{n\pi x}{a}\right) \sin\left(\frac{n\pi x}{a}\right) dx = 0 \end{aligned}$$

(dato che stiamo integrando una funzione dispari su un dominio centrato in 0). Perciò:

$$(\Delta P)_n = \frac{\hbar\pi n}{a}$$

E quindi il prodotto delle incertezze diviene:

$$(\Delta X)_n (\Delta P)_n = a \sqrt{\frac{1}{12} - \frac{1}{2(n\pi)^2}} \frac{\hbar\pi n}{a} = \hbar\pi n \sqrt{\frac{1}{12} - \frac{1}{2(\pi n)^2}} \neq 0 \quad \forall n \text{ dispari}$$

Notiamo che per n pari il calcolo di $(\Delta P)_n$ non cambia. Per il calcolo $(\Delta X)_n$, invece, notiamo che si ha sempre $\langle X \rangle_n = 0$ per un conto analogo a prima. Allo stesso modo si vede che $\langle X^2 \rangle_n$ assume lo stesso valore, ma lo ricalcoleremo illustrando qualche *trucco* per velocizzare il processo. Per n pari:

$$\begin{aligned} \langle X^2 \rangle_n &= \langle \varphi_n | x^2 \varphi_n \rangle = \int_{-\frac{a}{2}}^{+\frac{a}{2}} \frac{2}{a} \sin^2\left(\frac{n\pi x}{a}\right) x^2 dx = \\ &= \frac{2}{a} \frac{1}{2} \int_{-\frac{a}{2}}^{+\frac{a}{2}} x^2 \left(1 - \cos\left(\frac{2n\pi x}{a}\right)\right) dx = \\ &= \frac{1}{a} \left(\left[\frac{x^3}{3} \right]_{-\frac{a}{2}}^{+\frac{a}{2}} - \int_{-\frac{a}{2}}^{+\frac{a}{2}} x^2 \cos\left(\frac{2n\pi x}{a}\right) dx \right) \end{aligned}$$

Come è noto, quando si valuta una primitiva su un dominio simmetrico, se la primitiva è pari allora il risultato è nullo, e se è dispari basta valutarla all'estremo positivo e moltiplicare per 2:

$$f(x) \text{ pari} \Rightarrow [f(x)]_{-a}^{+a} = 0; \quad f(x) \text{ dispari} \Rightarrow [f(x)]_{-a}^{+a} = 2f(a)$$

Perciò immediatamente:

$$\left[\frac{x^3}{3} \right]_{-\frac{a}{2}}^{+\frac{a}{2}} = \frac{2}{3} \left(\frac{a}{2} \right)^3 = \frac{a^3}{12}$$

Il secondo integrale, invece, lo svolgiamo tramite *metodo tabulare*⁷, ossia derivando e integrando ripetutamente e scrivendo i risultati in colonne:

x^2	$\cos\left(\frac{2n\pi x}{a}\right)$
$-2x$	$\frac{a}{2n\pi} \sin\left(\frac{2n\pi x}{a}\right)$
2	$-\left(\frac{a}{2n\pi}\right)^2 \cos\left(\frac{2n\pi x}{a}\right)$
0	$-\left(\frac{a}{2n\pi}\right)^3 \sin\left(\frac{2n\pi x}{a}\right)$

Da cui si ottiene immediatamente:

$$\begin{aligned} \int_{-\frac{a}{2}}^{+\frac{a}{2}} x^2 \cos\left(\frac{2n\pi x}{a}\right) dx = \\ = \left[x^2 \sin\left(\frac{2n\pi x}{a}\right) \frac{a}{2n\pi} + 2x \cos\left(\frac{2n\pi x}{a}\right) \left(\frac{a}{2n\pi}\right)^2 - 2 \sin\left(\frac{2n\pi x}{a}\right) \left(\frac{a}{2n\pi}\right)^3 \right]_{-\frac{a}{2}}^{+\frac{a}{2}} \end{aligned}$$

Notiamo che tutti e tre gli addendi sono funzioni dispari, e quindi basta calcolarle in $+a/2$. Valutando sin e cos

$$\left(\frac{2n\pi x}{a} \right) \Big|_{x=\frac{a}{2}} = n\pi \Rightarrow \sin(n\pi) = 0; \quad \underbrace{\cos(n\pi)}_{n \text{ pari}} = +1$$

Perciò primo e terzo termine si annullano, e basta valutare il secondo in $a/2$ e raddoppiare per simmetria:

$$\int_{-\frac{a}{2}}^{+\frac{a}{2}} x^2 \cos\left(\frac{2n\pi x}{a}\right) dx = 2 \cdot 2 \cdot \frac{a}{2} \cdot (+1) \frac{a^2}{4(n\pi)^2} = \frac{a^3}{2(n\pi)^2}$$

Mettendo tutto insieme si giunge a:

$$\langle X^2 \rangle_n = a^2 \left(\frac{1}{12} - \frac{1}{2(n\pi)^2} \right)$$

e perciò $(\Delta X)_n$ ha lo stesso valore sia per n pari che per n dispari, così come $(\Delta P)_n$. Allora, in generale:

$$(\Delta P)_n (\Delta X)_n = \hbar \pi n \sqrt{\frac{1}{12} - \frac{1}{2(n\pi)^2}} \neq 0 \quad \forall n$$

⁷Si tratta di uno schema per ripetere velocemente multiple integrazioni per parti senza confondersi. Più informazioni qui: bit.ly/2BpkiwQ

5. H e P sono compatibili?

Se H e P sono compatibili, ammettono una base di autovettori comuni. Come ricordato all'inizio, gli autostati φ_n di H sono non degeneri, e quindi sono già una base ortonormale. Ma allora, per la compatibilità, i φ_n dovrebbero essere anche autostati di P : ciò non è possibile perché $(\Delta P)_n \neq 0$ come abbiamo appena visto al punto precedente, e perciò H e P **non** sono compatibili.

Nota che P e P^2 sono compatibili, P^2 e H sono compatibili, ma P e H non lo sono! In generale:

$$[P^2, H] = 0, [P, P^2] = 0 \not\Rightarrow [P, H] = 0$$

La compatibilità è una questione delicata da valutare attentamente sempre *mediante* la definizione, dato che potrebbero sorgere problemi di dominio (per esempio P e H sono compatibili nel caso libero, ma non in questo caso *compatto*).

6. Vogliamo ora trovare la probabilità di trovare la particella nella metà sinistra della buca per tempi t tra 0 e t_0 , ossia, per l'*assioma della probabilità*, $P_{\psi(t)}^X \left(\left[-\frac{a}{2}, 0 \right] \right)$, con $0 < t < t_0$. Essendo $\psi(t)$ già normalizzata, ossia $\|\psi(t)\|^2 = 1$, e ricordando la famiglia spettrale dell'operatore X vista in (3.18), si giunge a:

$$\begin{aligned} P_{\psi(t)}^X \left(\left[-\frac{a}{2}, 0 \right] \right) &= \left(\psi(t), P^X \left(\left[-\frac{a}{2}, 0 \right] \right) \psi(t) \right) = \\ &= \left(\psi(t), \left[P^X(0) - P^X \left(-\frac{a}{2} \right) \right] \psi(t) \right) = \\ &= \int_{\mathbb{R}} \psi^*(t) \left(H(-x) - H \left(-\frac{a}{2} - x \right) \right) \psi(t) dx = \\ &= \int_{-\frac{a}{2}}^0 |\psi(x, t)|^2 dx = \\ &= \int_{-\frac{a}{2}}^0 \sum_n c_n^* \sum_m c_m \exp \left(\frac{i}{\hbar} (\mathcal{E}_n - \mathcal{E}_m) t \right) \varphi_n^*(x) \varphi_m(x) dx = \\ &= \sum_{n,m} c_n^* c_m \gamma_{nm} \exp \left(\frac{i}{\hbar} (\mathcal{E}_n - \mathcal{E}_m) t \right) \quad \gamma_{nm} \equiv \int_{-\frac{a}{2}}^0 \varphi_n^*(x) \varphi_m(x) dx \end{aligned}$$

Per cui, la probabilità che una misura di X trovi la particella in un dato intervallo è semplicemente l'integrale del modulo quadro della funzione d'onda (in rappresentazione in x) in quell'intervallo.

Per m, n dispari, per $m \neq n$ applichiamo le formule di Werner e otteniamo un risultato nullo (le primitive dei cos sono sin che si annullano in 0 e $-a/2$), mentre per $m = n$ si ha l'integrale del \cos^2 , che si scompone con bisezione:

$$\begin{aligned} \gamma_{mn} &= \frac{2}{a} \int_{-\frac{a}{2}}^0 \cos \left(\frac{n\pi x}{a} \right) \cos \left(\frac{m\pi x}{a} \right) dx = \frac{1}{a} \int_{-\frac{a}{2}}^0 \left[\cos \frac{(n+m)\pi x}{a} + \cos \frac{(n-m)\pi x}{a} \right] dx = \\ &= \begin{cases} 0 & n \neq m \\ \frac{1}{a} \int_{-\frac{a}{2}}^0 1 dx = \frac{1}{2} & n = m \end{cases} \end{aligned}$$

Negli altri casi, dobbiamo esaminare solo per $n = 2, m = 2$, oppure uno dei due dispari e l'altro 2 (tutti gli altri sono nulli perché $c_n = 0$ per $n \neq 2$ pari). Si ha allora:

$$\gamma_{22} = \frac{2}{a} \int_{-\frac{a}{2}}^0 \sin^2 \left(\frac{2\pi x}{a} \right) dx = \frac{1}{2}$$

E nell'altro caso, con n dispari:

$$\begin{aligned} \gamma_{2n} = \gamma_{n2} &= \frac{2}{a} \int_{-\frac{a}{2}}^0 \sin \frac{2\pi x}{a} \cos \frac{n\pi x}{a} dx = \frac{1}{a} \int_{-\frac{a}{2}}^0 \left[\sin \frac{(n+2)\pi x}{a} + \sin \frac{(2-n)\pi x}{a} \right] dx = \\ &= - \left[\frac{1}{\pi(n+2)} + \frac{1}{\pi(2-n)} \right] = \frac{1}{\pi} \frac{4}{n^2 - 4} \end{aligned}$$

Allora, combinando tutti i casi per calcolare la sommatoria (notando che nei termini dove $n = m$ l'esponenziale dà 1):

$$\begin{aligned} P_{\psi(t)}^X \left(\left[-\frac{a}{2}, 0 \right] \right) &= \sum_{n,m} c_n c_m \exp \left(\frac{i}{\hbar} (\mathcal{E}_n - \mathcal{E}_m) t \right) \gamma_{nm} = \\ &= \frac{|c_2|^2}{2} + \sum_{n \text{ dispari}} \frac{|c_n|^2}{2} + \\ &+ \sum_{n \text{ dispari}} \frac{1}{\pi} \frac{4}{n^2 - 4} c_n c_2 \left(\exp \left(-\frac{i}{\hbar} (\mathcal{E}_n - \mathcal{E}_2) t \right) + \exp \left(\frac{i}{\hbar} (\mathcal{E}_n - \mathcal{E}_2) t \right) \right) = \\ &\stackrel{(a)}{=} \frac{1}{4} + \frac{1}{4} + \sum_{n \text{ dispari}} \underbrace{\frac{4}{\pi(n^2 - 4)}}_{\gamma_{2n}} \underbrace{\frac{1}{\sqrt{2}}}_{c_2} \underbrace{\sqrt{2} \frac{4}{(4 - n^2)} \frac{1}{\pi}}_{c_n} \underbrace{\frac{e^{ik} + e^{-ik}}{2}}_{\cos(k)} 2 \\ &k = \frac{t}{\hbar} (\mathcal{E}_n - \mathcal{E}_2) \end{aligned}$$

dove in (a) si è usato il fatto che, come visto al punto 2, le $|c_n|^2$ sono le probabilità di trovare la particella con una certa energia \mathcal{E}_n , e quindi la loro somma è 1, da cui:

$$\sum_n |c_n|^2 = 1 = |c_2|^2 + \sum_{n \text{ dispari}} |c_n|^2 \Rightarrow \sum_{n \text{ dispari}} |c_n|^2 = 1 - |c_2|^2 = \frac{1}{2}$$

Arriviamo perciò a:

$$P_{\psi(t)}^X \left(\left[-\frac{a}{2}, 0 \right] \right) = \frac{1}{2} - \sum_{n \text{ dispari}} \left(\frac{1}{\pi} \frac{4}{n^2 - 4} \right)^2 2 \cos \left(\frac{t}{\hbar} (\mathcal{E}_n - \mathcal{E}_2) \right)$$

4.4 Sistemi quantistici composti

Consideriamo due sistemi quantistici S^1 e S^2 con spazio di Hilbert \mathcal{H}_1 e \mathcal{H}_2 . Come si descrive il sistema $S^1 \cup S^2$, e in particolare i suoi stati puri?

In MC, se Ω_1 e Ω_2 sono gli spazi delle fasi di S^1 e S^2 , lo spazio delle fasi del sistema composto $S^1 \cup S^2$ è il **prodotto cartesiano** $\Omega_1 \times \Omega_2$ e quindi gli stati puri sono dati da un prodotto di delta:

$$\delta^{(1)}(q^{(1)} - q_0^{(1)})\delta(p^{(1)} - p_0^{(1)})\delta^{(2)}(q^{(2)} - q_0^{(2)})\delta(p^{(2)} - p_0^{(2)}) \quad (q^{(1)}, p^{(1)}) \in \Omega_1, (q^{(2)}, p^{(2)}) \in \Omega_2$$

Quindi gli stati puri di $S^1 \cup S^2$ sono il prodotto degli stati puri dei sottosistemi.

Proviamo ora a costruire l'analogo in MQ, in cui considereremo all'inizio il caso di due particelle *distinguibili*. Partiamo dall'equazione di Schrödinger della particella singola:

$$i\hbar \frac{\partial \psi(\vec{x}, t)}{\partial t} = H(\vec{x}, -i\hbar \vec{\nabla}) \psi(\vec{x}, t) \quad (4.16)$$

in cui l' H quantistica è ottenuta da quella classica con la sostituzione $\vec{p} \rightarrow -i\hbar \vec{\nabla}$. Poiché nel caso di due particelle l'hamiltoniana classica è della forma:

$$H(\vec{x}^{(1)}, \vec{p}^{(1)}, \vec{x}^{(2)}, \vec{p}^{(2)})$$

Potremmo pensare di estendere (4.16) nel seguente modo:

$$i\hbar \frac{\partial \psi(\vec{x}^{(1)}, \vec{x}^{(2)}, t)}{\partial t} = H(\vec{x}^{(1)}, -i\hbar \vec{\nabla}^{(1)}, \vec{x}^{(2)}, -i\hbar \vec{\nabla}^{(2)}) \psi(\vec{x}^{(1)}, \vec{x}^{(2)}, t)$$

Quindi la funzione d'onda a tempo fissato diviene:

$$\psi(\vec{x}^{(1)}, \vec{x}^{(2)}) \in L^2(\mathbb{R}^6, d^3x^{(1)} d^3x^{(2)})$$

Per capire come si descrivono i sistemi composti dobbiamo capire la relazione tra lo spazio $L^2(\mathbb{R}^6, d^3x^{(1)} d^3x^{(2)})$ (suggerito dall'estensione del caso classico) e gli spazi di Hilbert "iniziali" delle due particelle *considerate separatamente*:

$$\mathcal{H}^{(1)} = L^2(\mathbb{R}^3, d^3x^{(1)}); \quad \mathcal{H}^{(2)} = L^2(\mathbb{R}^3, d^3x^{(2)})$$

Prima di farlo, è necessario sviluppare una nozione di algebra lineare.

(Lezione 23 del
19/11/2018)

Definizione 29. Dati due spazi di Hilbert \mathcal{H}_1 , \mathcal{H}_2 , siano $\varphi_1 \in \mathcal{H}_1$, $\varphi_2 \in \mathcal{H}_2'$ funzionali lineari⁸, e si denoti con $\varphi_1 \otimes \varphi_2$ la forma bilineare sul prodotto cartesiano $\mathcal{H}_1 \times \mathcal{H}_2$ definita da:

Prodotto tensore

$$\varphi_1 \otimes \varphi_2(\underbrace{\{\psi_1, \psi_2\}}_{\mathcal{H}_1 \times \mathcal{H}_2}) = (\varphi_1, \psi_1)_{\mathcal{H}_1} (\varphi_2, \psi_2)_{\mathcal{H}_2} \quad (4.17)$$

In altre parole, *combiniamo* due funzionali lineari definiti su due spazi diversi \mathcal{H}_1 e \mathcal{H}_2 per definire un nuovo funzionale, che stavolta agisce su *coppie* di vettori del tipo $\{\psi_1 \in \mathcal{H}_1, \psi_2 \in \mathcal{H}_2\}$, applicando su ciascun membro il *rispettivo funzionale* (ossia

⁸ Dato che sussiste un isomorfismo tra \mathcal{H} e il suo duale \mathcal{H}' , si ha che agli effetti $\mathcal{H} = \mathcal{H}'$, e quindi non è indispensabile indicare l'apice.

φ_1 su ψ_1 , e φ_2 su ψ_2) per poi fare il prodotto dei risultati. Dato che l'azione su entrambi gli elementi della coppia è lineare, tale nuovo funzionale - che chiamiamo $\varphi_1 \otimes \varphi_2$, si dice **bilineare**. Valgono cioè:

$$\begin{aligned}\varphi_1 \otimes \varphi_2(\{a\psi_a + b\psi_b, \psi_2\}) &= a\varphi_1 \otimes \varphi_2(\{\psi_a, \psi_2\}) + b\varphi_1 \otimes \varphi_2(\{\psi_b, \psi_2\}) \\ \varphi_1 \otimes \varphi_2(\{\psi_1, a\psi_a + b\psi_b\}) &= a\varphi_1 \otimes \varphi_2(\{\psi_1, \psi_a\}) + b\varphi_1 \otimes \varphi_2(\{\psi_1, \psi_b\})\end{aligned}$$

e la relazione analoga in cui si scambiano i due termini della coppia.

Vorremmo ora che lo spazio di questi nuovi funzionali sia di Hilbert. Costruiamolo per passi.

Denotiamo con \mathcal{E} l'insieme delle combinazioni lineari *finite* di forme bilineari definite come sopra. In tal modo su \mathcal{E} è sempre definita la somma di funzionali e il loro prodotto per uno scalare, e quindi \mathcal{E} è (ovviamente) uno spazio vettoriale, la cui base sarà data dai funzionali ottenuti *moltiplicando tensorialmente* i funzionali delle basi di \mathcal{H}_1 e \mathcal{H}_2 . Esplicitamente, se $\{e_i\}$ è una base di \mathcal{H}'_1 e $\{v_j\}$ di \mathcal{H}'_2 , una base di \mathcal{E} è data da $\{e_i \otimes v_j\} \forall i, j$.

Rendiamo ora \mathcal{E} pre-hilbertiano definendo su di esso un *prodotto scalare* come:

$$(\varphi_1 \otimes \varphi_2, \chi_1 \otimes \chi_2) \equiv (\varphi_1, \chi_1)_{\mathcal{H}_1} (\varphi_2, \chi_2)_{\mathcal{H}_2} \quad (4.18)$$

In effetti basta definirlo sulla base di \mathcal{E} e poi estenderlo *per linearità*: precisamente, dato che ogni *vettore* di \mathcal{E} è combinazione lineare di elementi della base, il prodotto scalare tra due generici vettori può essere definito, grazie alla sua linearità, come combinazione dei prodotti scalari degli elementi di base.

Nota: qui stiamo definendo un prodotto scalare *tra forme bilineari* (che sono i “vettori” di \mathcal{E}), e quindi tale formula non va confusa con la definizione data in (4.17).

Il prodotto scalare che abbiamo introdotto definisce una *norma*, e quindi una nozione di *convergenza*. Possiamo allora *completare*⁹ \mathcal{E} aggiungendo ad esso anche i vettori ψ che si ottengono come combinazioni lineari *infinite*:

$$\psi = \sum_{nm} c_{nm} \psi_n^{(1)} \otimes \psi_m^{(2)}$$

se, ovviamente, tali somme hanno senso, ossia se *convergono* secondo la norma indotta dal prodotto scalare in (4.18).

Una volta fatto questo, lo spazio \mathcal{E} è **completo**, e quindi di Hilbert, e lo denotiamo con $\mathcal{H}_1 \otimes \mathcal{H}_2$. Tale spazio si dice **prodotto tensore** di \mathcal{H}_1 e \mathcal{H}_2 .

Possiamo rendere questa definizione “più semplice” concretizzandola utilizzando le basi degli spazi di Hilbert.

Teorema 4.4.1. *Se $\{e_n^{(1)}\}_{n \in N}$ è una base per \mathcal{H}_1 , e $\{e_m^{(2)}\}_{m \in M}$ per \mathcal{H}_2 , allora $\{e_n^{(1)} \otimes e_m^{(2)}\}_{n \in N, m \in M}$ è una base per $\mathcal{H}_1 \otimes \mathcal{H}_2$.*

Base dello spazio prodotto

*Valgono allora i seguenti **isomorfismi** (che esaminiamo senza dimostrare):*

⁹ Possiamo farlo per il *teorema del completamento*, dato che \mathcal{E} è unitario. CFR pag. 10 di [1]

- $L^2(\mathbb{R}^n, d^n x) \otimes L^2(\mathbb{R}^m, d^m y) \cong L^2(\mathbb{R}^{n+m}, d^n x d^m y)$
Cioè, prendendo un $\phi_n(x) \in L^2(\mathbb{R}^n, d^n x)$ e un $\psi_m(y) \in L^2(\mathbb{R}^m, d^m y)$, costruiamo un isomorfismo tra gli spazi sopra considerati tramite:

$$U(x, y) \in L^2(\mathbb{R}^{n+m}, d^n x d^m y) \quad (4.19)$$

$$U : \phi_n(x) \otimes \psi_m(y) \mapsto \phi_n(x) \psi_m(y) \quad (4.20)$$

dove x è un vettore n -dimensionale e y un vettore m -dimensionale, e ϕ_n, ψ_m sono funzioni (modulo quadro integrabili) di questi vettori, estendendo poi per linearità a combinazioni lineari finite e poi per continuità agli interi spazi.

Alternativamente, per definire U basta specificare cosa fa se applicata ai vettori $\{v_i\}$ e $\{w_j\}$ delle basi di $L^2(\mathbb{R}^n)$ e $L^2(\mathbb{R}^m)$, e poi **estendere per linearità e continuità**.

Nello specifico, fissati i valori di $U(v_i \otimes w_j)$, U è una funzione di vettori $a \in \mathbb{R}^{n+m}$. Spezzando $a = (a', a'')$ (prime n componenti e restanti m), dato che a' e a'' sono combinazioni lineari delle basi:

$$a' = \sum_i c_i v_i; \quad a'' = \sum_j d_j w_j$$

Possiamo scrivere:

$$U(a) = U((a', a'')) = U\left(\sum_i c_i v_i, \sum_j d_j w_j\right) = \sum_{ij} c_i d_j U(v_i \otimes w_j)$$

che ha un valore definito dato che abbiamo fissato gli $U(v_i \otimes w_j)$.

- Un altro isomorfismo interessante è:

$$L^2(\mathbb{R}^n, d^n x) \otimes \mathbb{C}^l \cong \underbrace{L^2(\mathbb{R}^n, d^n x) \oplus \cdots \oplus L^2(\mathbb{R}^n, d^n x)}_{l\text{-volte}} \quad (4.21)$$

In generale, se V e W sono spazi vettoriali, $V^* \otimes W$ contiene le funzioni lineari che vanno da $V \rightarrow W$. Più precisamente esiste un isomorfismo^a F :

$$F : V^* \otimes W \rightarrow \text{Hom}(V, W)$$

$$F\left(\underbrace{\varphi \otimes w}_{\in V^* \otimes W}\right)(v) = \underbrace{\varphi(v)w}_{\in W}$$

dove $\varphi \in V^*$ è un funzionale, e $w \in W$ è un vettore.

Finora abbiamo parlato però di prodotto tensore tra funzionali, mentre nell'espressione soprastante w è chiaramente un vettore. Cosa significa, perciò, $\varphi \otimes w$?

L'idea sta nel notare che se lo spazio duale W^* è quello delle funzioni lineari $W \rightarrow \mathbb{R}$ (indicato come $\text{Hom}(W, \mathbb{R})$), il duale del duale W^{**} contiene le funzioni $W^* \rightarrow \mathbb{R}$, e per uno spazio finito-dimensionale coincide (cioè, è canonicamente isomorfo) con W stesso. Possiamo cioè vedere i vettori di W come “funzioni di funzionali”, ossia un vettore di W applicato ad un funzio-

nale di W^* produce un numero.

Preso allora $\varphi \in V^*$, e $w \in W$, il loro prodotto tensore sarà una funzione bilineare che si applica contemporaneamente agli elementi a cui si applicherebbero singolarmente φ e w , ossia, rispettivamente, vettori $v \in V$ e funzionali $\phi \in W^*$: $(\varphi \otimes w) : V \times W^* \rightarrow \mathbb{R}$. Esplicitamente sarà proprio il prodotto delle singole “applicazioni”:

$$(\varphi \otimes w)(v, \phi) = \varphi(v)\phi(w)$$

^aUna dimostrazione di ciò è data a pag. 5 della Lezione 9 disponibile sul Moodle del corso di Geometria Differenziale tenuto dal prof. Bottacin.

Si ha quindi che $L^2(\mathbb{R}^n) \otimes \mathbb{C}^l$ contiene le funzioni (modulo quadro integrabili) da $\mathbb{R}^n \rightarrow \mathbb{C}^l$. Ma tali funzioni non sono altro che un vettore di “normali” funzioni $L^2(\mathbb{R}^n)$, che possiamo scrivere come:

$$\begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \vdots \\ \psi_l(x) \end{pmatrix} \quad \psi_i(x) \in L^2(\mathbb{R}^n, d^n x) \quad i = 1, \dots, l$$

- Ricordando che $\mathbb{R}^3 \cong \mathbb{R}_+ \times S^2$ (cioè possiamo rappresentare un punto dello spazio in coordinate sferiche, scrivendo $d^3x = r^2 dr \cdot d\Omega$, con $d\Omega = \sin \theta d\theta d\varphi$), nella notazione dei prodotti tensori avremo che:

$$L^2(\mathbb{R}^3, d^3x) \cong L^2(\mathbb{R}_+, r^2 dr) \otimes L^2(S^2, d\Omega)$$

Operatori in $\mathcal{H}_1 \otimes \mathcal{H}_2$ della forma $A^{(1)} \otimes A^{(2)}$ con $A^{(1)}$ che agisce su \mathcal{H}_1 e $A^{(2)}$ su \mathcal{H}_2 sono definiti per estensione (lineare e continua) da:

Prodotto tensore di operatori

$$(A^{(1)} \otimes A^{(2)})(\varphi^{(1)} \otimes \varphi^{(2)}) = A^{(1)}\varphi^{(1)} \otimes A^{(2)}\varphi^{(2)} \quad (4.22)$$

Attenzione: tuttavia gli operatori in $\mathcal{H}_1 \otimes \mathcal{H}_2$ sono in generale della forma di combinazioni lineari:

$$\sum_{m,n} c_{mn} A_m^{(1)} \otimes A_n^{(2)}$$

(dove la somma è potenzialmente infinita). Perciò non tutti gli operatori sullo “spazio grande” sono ottenibili come prodotti tensori degli operatori sugli “spazi piccoli” che lo compongono.

Ragionando in dimensione finita, un esempio *banale* di ciò è dato qui di seguito. Siano $\vec{v}, \vec{w} \in \mathbb{R}^2$, e A una forma bilineare, definita come:

$$\vec{v} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \vec{w} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad A(\vec{v}, \vec{w}) = 3x_1 y_1$$

In questo caso è facile scomporre A nel prodotto di due funzioni che agiscono

rispettivamente su \vec{v} e su \vec{w} :

$$A(\vec{v}, \vec{w}) = f(\vec{v})g(\vec{w}) \quad f(\vec{v}) = 3x_1 \quad g(\vec{w}) = y_1$$

E in questo senso si potrebbe scrivere:

$$f, g : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad A = f \otimes g : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}, \quad A \in \mathbb{R}'^2 \otimes \mathbb{R}'^2$$

Ma una *scomposizione* analoga non si può più fare per la forma bilineare B :

$$B(\vec{v}, \vec{w}) = 3x_1y_1 + x_2y_2$$

Notiamo però che i singoli addendi nella definizione di B sono scomponibili come prodotti tensori di opportune funzioni. In questo senso B è una *combinazione lineare* di prodotti tensori.

Se matematicamente ciò è ovvio, vedremo che le conseguenze fisiche sono radicali. Uno stato nello spazio *prodotto-tensore* $\mathcal{H}_1 \otimes \mathcal{H}_2$ non è detto che sia il prodotto tensore di uno stato in \mathcal{H}_1 e uno stato in \mathcal{H}_2 : potrebbe benissimo essere una combinazione lineare di prodotti del genere.

Un caso particolare di operatori definiti sullo spazio prodotto tensore è dato da $\underbrace{A^{(1)} \otimes \mathbb{I}^{(2)}}_a$ e $\underbrace{\mathbb{I}^{(1)} \otimes A^{(2)}}_b$ (per cui ciascuno dei due “agisce” solo sulla rispettiva parte

dello spazio, e agisce come identità su tutto il resto).

Essi commutano tra loro, infatti:

$$\begin{aligned} & \underbrace{(A^{(1)} \otimes \mathbb{I}^{(2)})}_a \underbrace{(\mathbb{I}^{(1)} \otimes A^{(2)})}_b (\varphi_1^{(1)} \otimes \varphi^{(2)}) = (A^{(1)} \otimes \mathbb{I}^{(2)}) (\varphi^{(1)} \otimes A^{(2)} \varphi^{(2)}) = \\ & = (A^{(1)} \varphi^{(1)} \otimes A^{(2)} \varphi^{(2)}) = (\mathbb{I}^{(1)} \otimes A^{(2)}) (A^{(1)} \varphi^{(1)} \otimes \varphi^{(2)}) = \\ & = \underbrace{(\mathbb{I}^{(1)} \otimes A^{(2)})}_b \underbrace{(A^{(1)} \otimes \mathbb{I}^{(2)})}_a (\varphi^{(1)} \otimes \varphi^{(2)}) \end{aligned}$$

Partendo allora da:

$$\underbrace{L^2(\mathbb{R}^3, d^3x)}_{\ni \psi(\vec{x})} \otimes \underbrace{L^2(\mathbb{R}^3, d^3y)}_{\ni \psi(\vec{y})} \cong \underbrace{L^2(\mathbb{R}^6, d^3x d^3y)}_{\ni \psi(\vec{x}, \vec{y})}$$

Abbiamo un’idea per estendere l’equazione di Schrödinger al caso dei sistemi composti. Enunciamo quindi l’assioma:

Assioma 6. *Quando un sistema quantistico consiste in N sistemi **distinguibili** con spazi di Hilbert \mathcal{H}_i , $i = 1, \dots, N$, il sistema totale ha come spazio di Hilbert il prodotto tensore degli spazi di Hilbert dei costituenti, cioè:*

Assioma dei sistemi composti

$$\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_N = \bigotimes_{i=1}^N \mathcal{H}_i$$

e le osservabili $A^{(i)}$ del sottosistema i -esimo nel sistema totale sono date da:

$$\mathbb{I}^{(1)} \otimes \dots \otimes \mathbb{I}^{(i-1)} \otimes A^{(i)} \otimes \mathbb{I}^{(i+1)} \otimes \dots \otimes \mathbb{I}^{(N)}$$

In notazione di Dirac, se $|\psi^{(1)}\rangle \in \mathcal{H}_1$ e $|\phi^{(2)}\rangle \in \mathcal{H}_2$ lo stato prodotto tensore è denotato “implicitamente” con:

$$|\psi^{(1)}\rangle |\phi^{(2)}\rangle \equiv |\psi^{(1)}\rangle \otimes |\phi^{(2)}\rangle$$

Poiché, come abbiamo visto, le osservabili che agiscono su ciascun sistema commutano con le osservabili che agiscono su tutti gli altri sottosistemi, se \mathcal{C}_i è un ICOC per il sistema S_i , allora un ICOC per il sistema totale $S = \bigcup_{i=1}^N S_i$ è l'insieme degli ICOC dei sottosistemi $\{\mathcal{C}_i\}_{i=1,\dots,N}$.

Fermiamoci un attimo a riflettere su tali risultati. Se utilizziamo l'interpretazione ondulatoria, per cui le particelle sono rappresentate da *onde di probabilità*, quello che abbiamo ottenuto è sconvolgente. Ci aspettiamo che la funzione d'onda del sistema composto *appartenga ancora allo spazio tridimensionale*. Abbiamo invece trovato che, per 2 particelle in 3D, la $\psi(\vec{x}, \vec{y})$ è definita sullo spazio dei prodotti cartesiani, che è a 6 dimensioni!

La ragione di ciò è che abbiamo descritto lo spazio degli stati puri come **lineare**, e perciò bisogna poter sommare i prodotti delle funzioni dei due sottosistemi, e quindi lo spazio deve avere per forza una dimensione più grande.

Per esempio, consideriamo una misura di *polarizzazione* dei fotoni, che in 2D sono scritte come combinazione lineare di vettori $|x\rangle$ e $|y\rangle$.

Facendo una misura (ideale di prima specie) di x sul primo fotone e di x sul secondo, lo stato finale è indicato da $|x^{(1)}\rangle |x^{(2)}\rangle$. Analogamente per y : $|y^{(1)}\rangle |y^{(2)}\rangle$. Ma lo stato più generale non sarà uno di questi, ma una combinazione lineare del tipo:

$$|x^{(1)}\rangle |x^{(2)}\rangle + |y^{(1)}\rangle |y^{(2)}\rangle$$

che non si può scrivere come prodotto degli stati dei due sistemi “separati” (è una combinazione lineare di tali prodotti).

Osservazione 8. **Importante!**

Per la **linearità** degli spazi vettoriali, nel nostro caso di Hilbert, e per la definizione di **prodotto tensore**, esistono stati del sistema composto che **non sono scrivibili** come prodotto tensore di stati dei sottosistemi (basta che siano una combinazione lineare di prodotti di stati singoli “non separabile” in un unico prodotto).

Stati di questo tipo si dicono **entangled** (=allacciati). Essi descrivono una situazione in cui le proprietà dei sottosistemi non sono indipendenti, forzando una “realtà allacciata” dei sottosistemi.

Questo è un fenomeno puramente quantistico, perché è necessaria la natura vettoriale per definire \otimes ! Ciò è alla base di alcuni dei fenomeni più sconcertanti e delle tecnologie quantistiche, come il paradosso EPR, il teletrasporto quantistico, computer quantistico, crittografia quantistica.

Osservazione 9. Riprendiamo lo stato entangled visto prima:

$$|\psi\rangle \sim |x^{(1)}\rangle |x^{(2)}\rangle + |y^{(1)}\rangle |y^{(2)}\rangle$$

Normalizziamolo e riscriviamolo in una notazione “più suggestiva”:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\rightarrow_1\rangle |\rightarrow_2\rangle + |\uparrow_1\rangle |\uparrow_2\rangle)$$

Cosa succede se facciamo una misura su di esso? Proviamo, per esempio, a misurare la polarizzazione, lungo \rightarrow , del fotone 1. Poiché lo stato del fotone 1 è una combinazione di $|\rightarrow\rangle$ e $|\uparrow\rangle$ *in parti uguali*, la probabilità di trovare il fotone orientato lungo \rightarrow o meno è del 50%:

$$p(\rightarrow_1) = \left| \left\langle \rightarrow \left| \frac{1}{\sqrt{2}}(|\rightarrow_1\rangle + |\uparrow_1\rangle) \right. \right\rangle \right|^2 = \frac{1}{2}$$

A seguito della misura (ideale di prima specie), lo stato *collasserà* nell’autospazio *compatibile* con l’autovalore misurato. Per esempio, se rileviamo il fotone 1 con polarizzazione \rightarrow (ossia se *attraversa* un filtro orientato lungo quell’asse), il nuovo stato del sistema sarà ottenuto applicando il *proiettore* $|\rightarrow_1\rangle \langle\rightarrow_1| \otimes \mathbb{I}$ a $|\psi\rangle$, da cui si ottiene:

$$|\psi_f\rangle = (|\rightarrow_1\rangle \langle\rightarrow_1| \otimes \mathbb{I}) \frac{1}{\sqrt{2}} (|\rightarrow_1\rangle |\rightarrow_2\rangle + |\uparrow_1\rangle |\uparrow_2\rangle) = |\rightarrow_1\rangle |\rightarrow_2\rangle$$

dove ciascuna parte del proiettore agisce sui termini evidenziati con lo stesso colore.

Notiamo ora che se misuriamo la polarizzazione del fotone 2 nel nuovo stato $|\psi_f\rangle$ otterremo *sempre* il risultato \rightarrow , cosa che non era vera per $|\psi\rangle$ iniziale. Ciò significa che la misura di polarizzazione del fotone 1 ci ha dato informazione anche sulla polarizzazione del fotone 2.

Ciò inizialmente potrebbe non sembrare strano, dato che “più misure in una volta sola” si possono fare anche in fisica classica. Per esempio, si considerino due sfere di masse e velocità iniziali conosciute, che urtano elasticamente ad un certo istante t . A t^+ basta misurare la velocità di una sola delle due per sapere, *per conservazione del momento*, la velocità dell’altra.

In MQ, tuttavia, la faccenda è complicata dal fatto che, a differenza del caso classico, le osservabili non sono *sempre definite*: nello stato $|\psi\rangle$, il primo fotone *non* ha una polarizzazione “predeterminata” prima che essa venga misurata. Perciò la misura sul primo fotone ritorna un valore *scelto a caso*, e quella sul secondo avrà lo *stesso valore della prima*. Tuttavia i due fotoni del sistema possono essere separati a grande distanze, e le due misure fatte contemporaneamente, in modo che *non vi sia possibilità* che un qualsiasi segnale possa trasmettersi dalla prima alla seconda, se non a velocità superiori della luce, il che è paradossale.

Sarebbe come lanciare una moneta e sapere *con certezza* che un’altra moneta, dall’altro capo del mondo, è atterrata con la stessa faccia. Ciò avrebbe senso, classicamente, solo se le due monete fossero *truccate*, ossia avessero una sola faccia (uguale in entrambe), e il lancio della prima moneta sia dovuto all’ignoranza dello sperimentatore che non sa di avere a che fare con una moneta del genere.

Una teoria del genere è detta *a variabili nascoste*. Come vedremo in un prossimo capitolo, la supposizione che il risultato di una misura sia in un qualche modo determinato da “variabili inaccessibili” che si comportino rispettando la località è incompatibile con le previsioni matematiche, e sperimentalmente verificabili, della MQ.

(Lezione 24 del
22/11/2018)

4.5 Simmetrie e Algebre di Lie

Nella sezione 3.8 abbiamo esaminato il caso della simmetria per traslazioni temporali. Cosa succede, invece, per le **traslazioni spaziali** o le **rotazioni**?

4.5.1 Teoria dei gruppi e delle algebre di Lie

Introduciamo qualche altro utile concetto matematico. Partiamo esaminando un particolare caso di *simmetria fisica* nella seguente definizione:

Definizione 30. Se una **simmetria fisica** lascia **invariata l’Hamiltoniana**, allora è detta **simmetria dinamica**. Matematicamente, sia U la rappresentazione del gruppo che descrive tale simmetria, che *agisce* su un ket come $|\psi'\rangle = U|\psi\rangle$, e supponiamo che H sia ben definita in $D(H)$ denso. Se U descrive una simmetria fisica, allora vale:

*Simmetria
dinamica*

$$\langle x'|H'|x'\rangle \stackrel{!}{=} \langle x|H|x\rangle \Rightarrow \langle Ux|H'|Ux\rangle = \langle x|U^\dagger H' U|x\rangle \stackrel{!}{=} \langle x|H|x\rangle$$

Perché U sia una simmetria dinamica, deve essere anche $H' = H$, e quindi:

$$H' \equiv U H U^\dagger = H$$

Intuitivamente, se U è una traslazione, per trasformare $H \rightarrow H'$, con H operatore che prende vettori in \mathcal{H} e li trasforma in \mathcal{H} bisogna “trasformare” *input* e *output* (esattamente come nella formula del cambio di base vista in algebra lineare). Perciò H' si ottiene applicando H al vettore “trasformato all’indietro” da U^\dagger (nell’esempio, traslato di $-t$), e poi traslando in avanti con U il risultato del conto.

L’**analogo classico** delle *simmetrie dinamiche* è dato dalle *trasformazioni* che mantengono invariate le *parentesi di Poisson*.

Nota: una generica simmetria *fisica* in MQ ha, come unica richiesta, la preservazione delle probabilità di transizione (non c’è alcun cenno sull’Hamiltoniana, che potrebbe anche non esistere). Per una simmetria dinamica la richiesta è nettamente più forte, ossia che l’Hamiltoniana proprio non cambi, e che quindi *le soluzioni del moto* restino esattamente le stesse.

Se le simmetrie fisiche sono descritte da un gruppo continuo G allora U è necessariamente unitario (\bar{U} è escluso, come dimostrato nell’equazione (3.39)). In particolare, se $g \in G$, gli $U(g)$ formano un *gruppo a un parametro di operatori*

unitari. Ciò significa che, prendendo per esempio $G = (\mathbb{R}, +)$, per cui allora un elemento del gruppo è un $\lambda \in \mathbb{R}$, le $U(\lambda)$ soddisfano le seguenti relazioni:

$$U(0) = \mathbb{I}; \quad U(\lambda_1)U(\lambda_2) = U(\lambda_1 + \lambda_2)$$

E il *teorema di Stone* ci assicura che \exists un operatore autoaggiunto definito da:

$$A \equiv \frac{1}{i} \frac{dU(\lambda)}{d\lambda} \Big|_{\lambda=0} \quad (4.23)$$

con dominio $D(A)$ denso in \mathcal{H} . A è detto il **generatore (infinitesimo)** di $U(\lambda)$.

Se $U(\lambda)$ è una *simmetria dinamica* sappiamo anche che deve valere:

$$U(\lambda)HU(\lambda)^\dagger = H \quad (4.24)$$

Da ciò segue che H *non dipende da* λ , e perciò la sua derivata rispetto a λ deve essere nulla.

Ma allora, derivando ambo i membri di (4.24) rispetto a λ otteniamo:

$$\begin{aligned} 0 &= \frac{d}{d\lambda} U(\lambda)HU(\lambda)^\dagger \Big|_{\lambda=0} = \frac{dU}{d\lambda} HU(\lambda)^\dagger + U(\lambda)H \frac{dU(\lambda)^\dagger}{d\lambda} \Big|_{\lambda=0} = \\ &\stackrel{(a)}{=} \frac{[H, A]}{i} \end{aligned} \quad (4.25)$$

dove in (a) si è usato il teorema di Stone, per cui:

$$A = \frac{1}{i} \frac{dU(\lambda)}{d\lambda} \Big|_{\lambda=0} \Rightarrow \frac{dU(\lambda)}{d\lambda} \Big|_{\lambda=0} = iA; \quad \frac{dU(\lambda)^\dagger}{d\lambda} \Big|_{\lambda=0} = -iA$$

e ricordando che $U(\lambda)$ e $U(\lambda)^\dagger$ calcolati per $\lambda = 0$ coincidono con l'identità \mathbb{I} .

Da (4.25) abbiamo allora ottenuto che il **generatore** di un gruppo a un parametro derivato da una simmetria dinamica **commuta** con l'Hamiltoniana.

Ma allora, in visuale di Heisenberg:

$$\frac{dA^H(t)}{dt} = \frac{[A^H(t), H]}{i\hbar} \stackrel{(a)}{=} \frac{U^\dagger(t)[A, H]U(t)}{i\hbar} \stackrel{(4.25)}{=} 0 \quad (4.26)$$

dove in (a) si è usato $A^H(t) = U^\dagger(t)AU(t)$, e si sono “portati fuori” gli U grazie al fatto che H è invariante rispetto alla loro azione.

Abbiamo allora trovato che $A^H(t) \equiv A$, ovvero il generatore **non evolve** nel tempo. Questa è la versione quantistica delle “**costanti del moto**”, e la relazione (4.26) associa a simmetrie continue leggi di conservazione¹⁰.

Finora ci siamo concentrati sulle proprietà “di gruppo”: abbiamo definito una simmetria, e fatto uso di risultati di *teoria delle rappresentazioni* per concretizzare

¹⁰^Questo è anche il collegamento con l'analogo classico visto a Fisica Matematica con il *teorema di Noether*

Generatore di una simmetria dinamica commuta con H

Il generatore di una simmetria dinamica è costante

l'azione di un gruppo sui sistemi che ci interessano.

Sappiamo però fin dall'inizio che i gruppi “continui” con cui abbiamo a che fare hanno una struttura molto più grande: sono **gruppi di Lie**, ossia sono anche *varietà differenziabili*. Esaminiamo qualche dettaglio in più di questo aspetto.

In breve, una varietà differenziale M è un insieme (di punti) che *localmente* “assomiglia” ad un aperto di \mathbb{R}^n . Matematicamente ciò significa che, preso un intorno U (sufficientemente piccolo) di un generico punto di M , è possibile “rappresentarlo” in \mathbb{R}^n tramite una funzione $\varphi : U \cap M \rightarrow V \subset \mathbb{R}^n$, detta *carta locale*, che ha la proprietà di essere un **omeomorfismo** (ossia è continua e invertibile con inversa continua)¹¹.

Nella pratica useremo sempre l'inversa $\varphi^{-1} : \mathbb{R}^n \supset V \rightarrow U \subset M$, detta *parametrizzazione* (locale), che associa punti di \mathbb{R}^n a elementi della varietà. Dato che $G = M$ è un gruppo (di Lie), un suo generico elemento $g \in G$ sarà parametrizzato da n parametri come $g(x_1, \dots, x_n)$.

Consideriamo tale funzione **analitica**, ossia di classe \mathcal{C}^∞ e tale che il suo sviluppo in serie di Taylor converga ad essa per qualsiasi punto.

G con la parametrizzazione adottata è detto **gruppo di Lie** a n parametri reali (o, eventualmente, complessi).

Ad esempio, le *traslazioni* in \mathbb{R}^3 sono un gruppo di Lie a 3 parametri, così come le *rotazioni*. Il *gruppo di Lorentz* è a 6 parametri (3 rotazioni + 3 boost), e quello di *Poincaré* è a 10 parametri (6 di Lorentz + 3 traslazioni spaziali + 1 traslazione temporale).

Scegliamo la parametrizzazione in modo che $g(0, \dots, 0) = e$, dove e è l'elemento neutro del gruppo G .

22/11/2018

Se ora calcoliamo le derivate in e lungo le direzioni date dalla base canonica, troviamo i vettori e_α che fungono da base dello *spazio tangente* $T_e G$ al gruppo presso il suo elemento neutro:

$$\left. \frac{d}{dx_\alpha} g(0, 0, \dots, x_\alpha, 0, \dots, 0) \right|_{x_\alpha=0} \equiv e_\alpha \quad \alpha = 1, \dots, n \quad (4.27)$$

$T_e G$ ha di per sé la struttura di uno spazio vettoriale: possiamo cioè sommare vettori tangenti o moltiplicarli per uno scalare. In realtà, grazie alla struttura di gruppo, possiamo trovare anche una *terza operazione* che associa due vettori in $T_e G$ e ne produce un terzo. Tale operazione (bilineare) è detta *parentesi di Lie*, e si indica con $[\cdot, \cdot]$. Con questa operazione, l'insieme degli $\{e_\alpha\}_{\alpha=1, \dots, n}$ costituisce la base di un'**algebra**¹², detta di Lie, denotata con $\text{Lie } G$ o \mathfrak{g} in cui la moltiplicazione si denota con il simbolo $[\dots, \dots]$ (come il commutatore).

¹¹Generalmente saranno necessarie più φ_i , ciascuna associata a un U_i , con l'insieme degli U_i che “ricopre” la varietà M . Può allora succedere che, per certi punti di M , siano definite *più di una carta locale*: vogliamo che tali funzioni siano “compatibili”, ossia che si possa passare da una all'altra componendo con una *funzione di transizione* η_{ij} , che deve essere \mathcal{C}^∞

¹²In inglese il termine tecnico è *algebra over a field*: https://en.wikipedia.org/wiki/Algebra_over_a_field

In effetti, se G è un gruppo moltiplicativo di matrici, ossia se i suoi elementi $g(\vec{x}) \in \mathcal{M}_{n \times n}$ sono matrici, l'esistenza di e_α è garantita dall'analiticità della parametrizzazione (per cui le derivate soprastanti esistono), e dal fatto che le matrici formano un'algebra, e perciò possiamo sommarle e moltiplicarle tra loro. In particolare, anche gli elementi dello spazio tangente saranno *matrici* (le cui componenti sono le derivate, lungo le direzioni della base canonica, delle componenti delle matrici $g(\vec{x})$ di partenza). In tal caso la parentesi di Lie è proprio il commutatore:

$$[e_\alpha, e_\beta] \equiv e_\alpha e_\beta - e_\beta e_\alpha$$

e da qui si comprende il perché di identificare le due notazioni.

$\forall a, b, c \in \mathfrak{g}$ la parentesi di Lie soddisfa le seguenti proprietà:

- $[a, b] = -[b, a]$
- $[[a, b], c] + [[b, c], a] + [[c, a], b] = 0$ (identità di **Jacobi**)
- $\forall \alpha, \beta \in \mathbb{R}$ (o a \mathbb{C} se G è complesso) vale la **linearità**:

$$[a, \alpha b + \beta c] = \alpha[a, b] + \beta[a, c]$$

(che sono proprio le stesse proprietà del commutatore).

Se invece consideriamo G come gruppo di Lie “generico” (ossia non necessariamente di matrici), la cosa si fa più complicata, dato che non possiamo, in linea di principio, “moltiplicare tra loro” elementi di $T_e G$, e quindi non si può usare immediatamente la definizione del *braket di Lie* come commutatore.

Dato che tal caso non è di nostro interesse in fisica, diamo qui solo qualche cenno e fonte per i più curiosi.

Per prima cosa, è possibile identificare ogni *vettore tangente* in $T_g G$ con un operatore che, se applicato ad una funzione differenziabile, ne calcola la *derivata direzionale* nella medesima direzione, e valuta tale derivata nel punto g di tangenza.

Un campo vettoriale definito su M è allora una *sezione del fibrato tangente*, ossia una funzione che associa punti di M a *vettori tangenti*, ossia *operatori di derivata direzionale* calcolata in quei punti. Un campo X è perciò, in generale, una funzione che prende in input una f differenziabile e determina una $X(f)$ che è ancora una funzione differenziabile, ed è pari alla derivata della f lungo le direzioni dei vettori tangenti che formano il campo X stesso. Cioè, valutare $X(f)_g$ in un punto g , è come prendere il vettore tangente $x_g \in T_g G$ appartenente al campo e applicarlo a f .

In questo modo, dati due campi vettoriali definiti sulla varietà $X, Y \in \mathcal{T}(G)$, potremmo pensare di comporli - visto che sono derivazioni, e quindi funzioni. Tuttavia $X \circ Y$ *non* è una derivazione (per esempio non rispetta la regola di Leibniz), mentre lo è $X \circ Y - Y \circ X \equiv [X, Y]$ (essenzialmente: svolgendo i conti con carte locali si nota che i termini che prima impedivano Leibniz ora si cancellano nella differenza). Ecco perciò la definizione della *parentesi di Lie*.

Nel paragrafo di sopra, tuttavia, abbiamo parlato di parentesi di Lie di elementi

dello spazio tangente $T_e G$, e non di campi vettoriali! In realtà le due cose sono le stesse - si può infatti dimostrare che un vettore in $T_e G$ (o, equivalentemente, in un qualsiasi $T_g G$ con $g \in G$ generico, si sceglie l'elemento neutro solo per comodità e convenzione) *determina univocamente* un campo vettoriale sull'intero G , che è detto *left invariant vector field*. Intuitivamente, stiamo usando la struttura di gruppo per *traslare* il vettore inizialmente scelto in $T_e G$ e trovarne i corrispettivi tangenti ad ogni punto di G (graficamente, sfruttiamo la *simmetria* data dal gruppo per costruire tutti gli altri vettori del campo).

Tramite queste costruzioni otteniamo campi vettoriali X e Y associati univocamente a vettori $x, y \in T_e G$. Per X e Y sappiamo calcolare il bracket di Lie, e perciò definiamo $[x, y] \equiv [X, Y]$. Si può dimostrare, con tanti e lunghi calcoli, che tale costruzione è coerente, ossia è compatibile con le operazioni definite su $T_e G$.

A cosa serve tutto ciò? Lavorare su $T_e G$, che è un *normale* spazio vettoriale, è molto *più comodo* che lavorare con gli elementi di una varietà G . Inoltre si dimostra che è possibile ricavare gli elementi di G (se si trovano entro un intorno *connesso* sufficientemente piccolo di e) a partire da quelli di \mathfrak{g} , ossia dai vettori di $T_e G$.

L'idea è la seguente. Abbiamo visto che un vettore su $T_e G$ può essere esteso, mediante la struttura di gruppo, a un campo di vettori su tutta G , che indichiamo con K . Tale K non è altro che una *sezione del fibrato tangente*, ossia una funzione che associa un punto di g a un vettore tangente in quel punto, ossia contenuto in $T_e G$. K definisce delle *curve integrali*: sono le curve $\gamma : \mathbb{R} \supset [a, b] \rightarrow G, t \mapsto \gamma(t)$ tali che se $\gamma(t) = g \in G$, $\gamma'(t)$ è proprio il vettore $K(g)$ “indicato” dal campo in quel punto. Queste curve integrali si trovano, in pratica, risolvendo un sistema di equazioni differenziali. Se partiamo da $K(e) \in \mathfrak{g} = T_e G$, avremo un problema di Cauchy del tipo:

$$\begin{cases} \gamma'(t) = K(\gamma(t)) \\ \gamma(0) = e \end{cases}$$

Se cambiamo l'elemento di \mathfrak{g} di partenza, avremo un altro campo vettoriale K , e di conseguenza altre curve integrali. In che modo? Proviamo ad esplorarlo definendo la seguente mappa:

$$\begin{aligned} \exp : \mathfrak{g} &\rightarrow G \\ \mathfrak{g} \ni v &\mapsto \gamma(1) \in G \end{aligned}$$

Pittorescamente, l'idea è questa: fissare un elemento di \mathfrak{g} consiste nel fissare una “direzione” verso cui spostarsi su G , partendo dal punto e , e la mappa \exp dà proprio il punto di G raggiunto *dopo una sola “unità di tempo”*.

Tale mappa è detta *mappa esponenziale* non a caso: se G è un gruppo di matrici, il sistema di equazioni differenziali sarà lineare, e risolubile con le tecniche viste in Analisi 3 tramite l'*esponenziale di matrice*. Dato un G generico l'immagine si complica - e in genere conviene lavorare in una rappresentazione lineare di tale G , in cui i suoi elementi sono identificati da matrici.

Si può dimostrare che questa mappa è “bella”, nel senso che preserva relazioni di commutazione, è compatibile con le rappresentazioni, etc.

Perciò, in conclusione, ricavata una base di \mathfrak{g} , data dai *generatori* della simmetria, e studiati i loro commutatori, possiamo ricavare il comportamento di tutti^a gli altri punti del gruppo! Questa è *un'intuizione* dell'importanza delle algebre di Lie.

Tale introduzione *discorsiva* è lungi dall'essere formale ed esaustiva (cosa che richiederebbe ben più pagine). Lasciamo perciò alcune **fonti** per chi volesse approfondire:

- L'associazione tra vettori tangenti e derivate direzionali valutate nel punto è ben spiegata nella Lezione4_2 delle dispense del prof. Francesco Bottacin (<https://www.math.unipd.it/~bottacin/geomdiff.htm>) per il corso di *Geometria Differenziale*.
Sempre nelle stesse dispense, l'associazione tra campi vettoriali e derivazioni è esplicitata nella Lezione11, e la definizione di parentesi di Lie si trova alla Lezione12.
- Per l'identificazione tra campi e vettori tangenti all'elemento neutro, lasciamo un'*intuizione grafica* presso bit.ly/2C66xDy (risposte di *ACuriousMind* e *Vectornaut*, danno anche un'intuizione sui risultati del prossimo paragrafo), e un punto di partenza per ulteriori approfondimenti è https://en.wikipedia.org/wiki/Lie_group#The_Lie_algebra_associated_with_a_Lie_group.

^aQueste frasi vanno intese con 5-6 asterischi: ci sono tantissimi dettagli matematici su cui stiamo glossando per semplicità.

Se G è un gruppo moltiplicativo di matrici, allora si dimostra che ogni elemento di G in un intorno dell'identità I può essere scritto come:

$$g(x_1, \dots, x_n) = \exp \left(\sum_{\alpha=1}^n x_{\alpha} e_{\alpha} \right) \quad (4.28)$$

e e_{α} è detto *generatore* del gruppo a un parametro $e^{x_{\alpha} e_{\alpha}}$, $x_{\alpha} \in \mathbb{R}$, e avremo n gruppi del genere.

Poiché $\{e_{\alpha}\}_{\alpha=1, \dots, n}$ è una base di \mathfrak{g} , che è uno spazio vettoriale, devono esistere delle costanti $f_{\alpha\beta\gamma}$ che consentano di rappresentare il commutatore $[e_{\alpha}, e_{\beta}]$ (che è ancora un elemento di \mathfrak{g} per come è definito il *braket di Lie*) come combinazione lineare degli elementi della base. Cioè, esplicitamente:

$$[e_{\alpha}, e_{\beta}] = \sum_{\gamma=1}^n f_{\alpha\beta\gamma} e_{\gamma}$$

Tali costanti $f_{\alpha\beta\gamma}$ sono dette **costanti di struttura** di \mathfrak{g} .

Si dimostra che l'algebra di Lie di G e del suo gruppo di ricoprimento universale \tilde{G} sono isomorfe, come suggerito dal fatto che l'algebra coinvolge il gruppo solo in un intorno dell'identità e (dato che la base che la genera è fatta di *derivate* calcolate

nell'identità), e localmente G e \tilde{G} sono omeomorfi.

Come per i gruppi, anche per le algebre di Lie si possono definire le rappresentazioni.

Una rappresentazione \mathcal{D} di un'algebra di Lie \mathfrak{g} in uno spazio vettoriale V è una mappa da \mathfrak{g} nello spazio degli operatori lineari su V , $\mathcal{L}(V)$, che preserva la struttura di algebra, ossia è *compatibile* con il *braket di Lie*:

$$\begin{aligned}\mathcal{D} : a \in \mathfrak{g} &\mapsto \mathcal{D}(a) \in \mathcal{L}(V) \\ \mathcal{D}([a, b]) &= [\mathcal{D}(a), \mathcal{D}(b)] \equiv \mathcal{D}(a)\mathcal{D}(b) - \mathcal{D}(b)\mathcal{D}(a) \quad \forall a, b \in \mathfrak{g}\end{aligned}$$

Nota: Se G non è un gruppo di matrici, il commutatore $[a, b]$ non ha senso immediato (cosa vuol dire applicare un vettore tangente ad un altro?), ma va definito nel modo specificato nel precedente box di approfondimento. Tuttavia, anche in tal caso, il membro a destra resta ben definito, dato che $\mathcal{D}(a)$, $\mathcal{D}(b)$ sono operatori lineari (matrici) e quindi è definito il loro commutatore nel senso usuale (poiché per le matrici è definita un'algebra, ossia si possono sommare/moltiplicare tra loro).

La rappresentazione preserva anche la struttura vettoriale di \mathfrak{g} , per cui la rappresentazione di una combinazione lineare è la combinazione lineare delle rappresentazioni:

$$\mathcal{D}(\alpha a + \beta b) = \alpha \mathcal{D}(a) + \beta \mathcal{D}(b)$$

In particolare:

$$[\mathcal{D}(e_\alpha), \mathcal{D}(e_\beta)] = \mathcal{D}([e_\alpha, e_\beta]) = \mathcal{D}\left(\sum_{\gamma} f_{\alpha\beta\gamma} e_\gamma\right) = \sum_{\gamma} f_{\alpha\beta\gamma} \mathcal{D}(e_\gamma)$$

Per cui basta sapere la rappresentazione degli elementi di base, e le *costanti di struttura*, ed è possibile calcolare qualsiasi commutatore nella rappresentazione \mathcal{D} .

Una rappresentazione \mathcal{D} dell'algebra di Lie \mathfrak{g} è *indotta* da una rappresentazione D stavolta del **gruppo** di Lie G su V . Ricordando che $D(g) \in \mathcal{L}(V)$ si ha:

$$\left. \frac{d}{dx_\alpha} D(g(0, \dots, 0, x_\alpha, 0, \dots, 0)) \right|_{x_\alpha=0} = \mathcal{D}(e_\alpha) \quad \alpha = 1, \dots, n \quad (4.29)$$

E tali $\mathcal{D}(e_\alpha)$ formano una *rappresentazione* di \mathfrak{g} in V (o in un dominio $\subseteq V$ in cui è ben definita, per evitare i problemi di dominio che potrebbero sorgere in spazi ∞ -dimensionali).

Concretizziamo tutto ciò nel caso delle **traslazioni temporali**, dove $G = (\mathbb{R}, +)$, e $U(t)$ operatore unitario è una *rappresentazione unitaria* in \mathcal{H} di G , ossia, utilizzando la notazione appena vista, $U(t) = D(g(t))$.

Allora utilizzando (4.29) possiamo ricavare una rappresentazione della relativa algebra di Lie. Notiamo però che l'unica derivata, rispetto a t , possiamo trovarla

immediatamente applicando il teorema di Stone:

$$\underbrace{A}_{-\frac{H}{\hbar}} = \frac{1}{i} \frac{dU(t)}{dt} \Big|_{t=0} \Rightarrow \frac{d}{dt} D(g(x_1)) \Big|_{x_1=0} = \frac{d}{dt} U(t) \Big|_{t=0} = \frac{H}{i\hbar} = \mathcal{D}(e_1)$$

Dato che G è isomorfo a \mathbb{R} , possiamo parametrizzarlo *canonicamente* con se stesso, ossia tramite la parametrizzazione $g(t) = t$. Possiamo allora calcolare esplicitamente il generatore del gruppo, secondo l'espressione in (4.27):

$$\frac{dg(t)}{dt} \Big|_{t=0} \Rightarrow \frac{dt}{dt} \Big|_{t=0} = \mathbf{1} = e_1$$

E perciò $H/(i\hbar)$ è una rappresentazione $\mathcal{D}(1)$ del generatore $\mathbf{1}$ dell'algebra di Lie delle traslazioni in $D(H) \subseteq \mathcal{H}$ (non è in generale tutto \mathcal{H} , dato che non è detto che l'Hamiltoniana sia *ovunque ben definita*).

4.6 Le traslazioni spaziali

Possiamo ora facilmente discutere la rappresentazione delle *traslazioni spaziali* seguendo lo stesso schema delle traslazioni temporali.

Partiamo considerando dall'**omogeneità dello spazio**, per cui i risultati di un esperimento non dipendono dal luogo in cui viene svolto. In MQ ciò significa che, ogni traslazione delle coordinate $\vec{x} \rightarrow \vec{x} + \vec{a}$, con $\vec{a} \in \mathbb{R}^3$ non modifica le probabilità di transizione $|\langle \phi | \psi \rangle|^2$.

Consideriamo quindi il gruppo *additivo* $G = (\mathbb{R}^3, +)$ delle *traslazioni spaziali*. Per teorema di Wigner esiste allora una rappresentazione proiettiva $G \ni \vec{a} \mapsto \hat{U}(\vec{a})$, che agisce su *raggi vettori* ($\hat{U} : \mathcal{PH} \rightarrow \mathcal{PH}$) ed è compatibile con l'operazione di gruppo (addizione):

$$\hat{U}(\vec{a})\hat{U}(\vec{b}) = \hat{U}(\vec{a} + \vec{b}) \quad (4.30)$$

Per il teorema di Bargmann, dato che il gruppo di ricoprimento universale di $(\mathbb{R}^3, +)$ è se stesso (dato che è semplicemente connesso), possiamo passare da una rappresentazione proiettiva a una unitaria $\vec{a} \mapsto U(\vec{a})$ sul ricoprimento.

Dato che gli $U(\vec{a})$ formano un gruppo di matrici, per quanto visto in (4.28), possiamo scrivere ognuno di essi come esponenziale dei generatori del gruppo, che indichiamo con e_1, e_2, e_3 . Si ha quindi:

$$U(\vec{a}) = U(a_1, a_2, a_3) = \exp(a_1 e_1 + a_2 e_2 + a_3 e_3) \quad (4.31)$$

Procediamo calcolando i generatori. Spostiamoci in rappresentazione in $\{\vec{x}\}$. Detti $|\vec{x}\rangle$ gli autovettori di autovalore x dell'operatore posizione \vec{X} . $U(\vec{a})$ *agisce* su di essi per traslazione:

$$U(\vec{a}) |\vec{x}\rangle = |\vec{x} + \vec{a}\rangle$$

Prendendo il coniugato di entrambi i membri si giunge a:

$$\langle \vec{x} | U(\vec{a})^\dagger = \langle \vec{x} + \vec{a} | \quad (4.32)$$

Si ha che $U(\vec{a})^\dagger$ corrisponde alla *traslazione inversa* $U(\vec{a})^{-1} = U(-\vec{a})$: lo si vede della proprietà (4.30) scrivendo $\mathbb{I} = U(\vec{0}) = U(\vec{a} - \vec{a}) = U(\vec{a})U(-\vec{a})$. Ma allora sostituendo nell'equazione (4.32) e poi rinominando $-\vec{a} \rightarrow \vec{a}$:

$$\langle \vec{x} | U(-\vec{a}) = \langle \vec{x} + \vec{a} | \Rightarrow \langle \vec{x} | U(\vec{a}) = \langle \vec{x} - \vec{a} | \quad (4.33)$$

Possiamo ora usare questo risultato per calcolare l'azione di $U(\vec{a})$ su $|\psi\rangle$ in rappresentazione $\{\vec{x}\}$:

$$(U(\vec{a})\psi)(\vec{x}) \equiv \langle \vec{x} | U(\vec{a}) | \psi \rangle \stackrel{(4.33)}{=} \langle \vec{x} - \vec{a} | \psi \rangle = \psi(\vec{x} - \vec{a})$$

In questo modo abbiamo scritto gli stati come *funzioni* (o *vettori* di $L^2(\mathbb{R}^3, d^3x)$) che possiamo derivare facilmente. Calcoliamo perciò i generatori dall'espressione (4.27):

$$e_j \equiv \frac{\partial}{\partial a_j} (U(\vec{a})\psi)(\vec{x}) \Big|_{\vec{a}=0} = \frac{\partial}{\partial a_j} \psi(\vec{x} - \vec{a}) \Big|_{\vec{a}=0} = -\frac{\partial}{\partial x_j} \psi(\vec{x}) = \left(\frac{P_j}{i\hbar} \psi \right) (\vec{x})_j$$

dove $P_j = -i\hbar \partial_{x_j}$ è la j -esima componente del momento in rappresentazione $\{\vec{x}\}$. Sostituendo allora quanto trovato in (4.31), e sintetizzando la somma con un prodotto scalare, giungiamo a:

$$U(\vec{a}) = \exp \left(-\frac{i}{\hbar} \vec{a} \cdot \vec{P} \right)$$

4.7 Momento angolare e rotazioni

Una osservabile che in MQ si discosta fortemente dal comportamento classico è il **momento angolare**.

In MC, utilizzando le coordinate cartesiane in \mathbb{R}^3 , le componenti del momento angolare sono (in convenzione di Einstein, con gli indici in lettere latine che vanno intesi da 1 a 3):

$$l_i = (\vec{x} \times \vec{p})_i = \epsilon_{ijk} x_j p_k \quad \vec{l} = \{l_1, l_2, l_3\}$$

dove ϵ_{ijk} è il tensore di Levi-Civita così definito:

$$\epsilon_{ijk} = \begin{cases} 1 & \text{se } i, j, k = 1, 2, 3 \text{ o permutazioni pari} \\ -1 & \text{se } i, j, k \text{ si ottiene da } 1, 2, 3 \text{ per permutazioni dispari} \\ 0 & \text{altrimenti} \end{cases}$$

Il momento angolare ha spettro continuo in MC.

In MQ, nonostante \vec{X} e \vec{P} abbiano spettro continuo in \mathbb{R}^3 , \vec{L} risulterà avere spettro discreto!

Per semplicità non discuteremo eventuali problemi di dominio: basti sapere che tutte le operazioni che faremo sono lecite in:

$$D = \mathcal{S}(\mathbb{R}^3)$$

Volendo si potrebbe definire un dominio più ampio, ma farlo risulta complicato e oltre gli obiettivi di questo corso.

Calcoliamo l'algebra di \vec{L} , ossia tutti i possibili commutatori tra L_1 , L_2 e L_3 . Risulta:

$$[L_i, L_j] = i\hbar\epsilon_{ijk}L_k \quad L_i = \epsilon_{ijk}X_jP_k$$

Tale risultato si potrebbe derivare applicando la *quantizzazione canonica* di Dirac delle parentesi di Poisson $[\cdot, \cdot] = i\hbar\{\cdot, \cdot\}$, e utilizzando il risultato in MC:

$$\{l_i, l_j\} = \epsilon_{ijk}l_k$$

Alternativamente, si può procedere per calcolo diretto. Per esempio, calcoliamo $[L_1, L_2]$. Risulta comodo sfruttare le seguenti due proprietà del commutatore:

$$\begin{aligned} [A, BC] &= [A, B]C + B[A, C] \\ [AB, C] &= [A, C]B + A[B, C] \end{aligned}$$

che si verificano immediatamente espandendo i commutatori da definizione.

Allora:

$$\begin{aligned} [L_1, L_2] &= [X_2P_3 - X_3P_2, X_3P_1 - X_1P_3] = \\ &= X_2[P_3, X_3]P_1 + X_1[X_3, P_3]P_2 = \\ &= -i\hbar X_2P_1 + i\hbar X_1P_2 = i\hbar L_3 \end{aligned}$$

(al primo conto, prova a calcolare *tutti* i commutatori)

Definiamo ora $\vec{L}^2 = L_1^2 + L_2^2 + L_3^2$. Notiamo che \vec{L}^2 commuta con ogni L_j :

$$\begin{aligned} [\vec{L}^2, L_j] &= [L_i L_i, L_j] = L_i [L_i, L_j] + [L_i, L_j] L_i = \\ &= i\hbar \underbrace{\epsilon_{ijk}}_{ik \text{ antisimmetrici}} \underbrace{(L_i L_k + L_k L_i)}_{ik \text{ simmetrici}} = 0 \end{aligned}$$

dato che si tratta della *contrazione* di un tensore A_{ik} antisimmetrico con uno S_{ik} simmetrico, che è nulla. Infatti, scambiando l'ordine degli indici, il tensore simmetrico non cambia, mentre quello antisimmetrico cambia di segno:

$$A_{ik}S_{ik} = -A_{ki}S_{ki} \Rightarrow A_{ik}S_{ik} + A_{ki}S_{ki} = 0$$

Ora basta rinominare gli indici nel secondo termine (possiamo farlo dato che gli indici su cui si somma sono *muti*), e raccogliere. Si ottiene perciò $2A_{ik}S_{ik} = 0$, da

cui segue il risultato utilizzato.

Perciò, riepilogando, mentre L_i e L_j , per $i \neq j$ **non** sono compatibili, \vec{L}^2 e L_j sono compatibili $\forall j$.

Vediamo la relazione di \vec{L} con le rotazioni. Calcoliamo prima:

$$\begin{aligned} [L_i, X_l] &= [\epsilon_{ijk} X_j P_k, X_l] = \epsilon_{ijk} X_j [P_k, X_l] = \\ &= \epsilon_{ijk} X_j (-i\hbar \delta_{kl}) = -i\hbar \epsilon_{ijk} X_j = i\hbar \epsilon_{ilj} X_j \\ [L_i, P_l] &= i\hbar \epsilon_{ilj} P_j \end{aligned}$$

Analogamente a quanto trovato per le traslazioni spaziali, i cui generatori erano dati dal momento (lineare), ci aspettiamo che il momento angolare funga da generatore delle rotazioni. Procediamo, allora, *a ritroso*, partendo da $e_3 = L_3/\hbar$ come generatore di un gruppo di matrici, e verifichiamo che da esso si ottengano effettivamente le rotazioni attorno all'asse \hat{z} .

Come già visto, e_3 è un elemento dell'algebra di Lie \mathfrak{g} del gruppo di operatori unitari $U(\varphi)$ che, come vogliamo verificare, rappresenta su \mathcal{H} le rotazioni. Il passaggio da e_3 a $U(\varphi)$ è dato, come già visto, dalla mappa esponenziale:

$$U(\varphi) = \exp\left(-i\varphi \frac{L_3}{\hbar}\right)$$

Come agisce $U(\varphi)$ su un'osservabile, per esempio sulla posizione \vec{X} ? Come già visto nel caso di H , la formula è analoga a quella del *cambio di base* vista in algebra lineare:

$$X_i(\varphi) = U(\varphi) X_i U(\varphi)^\dagger = \exp\left(-i\varphi \frac{L_3}{\hbar}\right) X_i \exp\left(i\varphi \frac{L_3}{\hbar}\right)$$

Procediamo per gradi, partendo dal calcolo dell'esponenziale. In generale, l'esponenziale di una matrice A è:

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$$

Detta B un'altra matrice, con cui A potrebbe o meno commutare, proviamo ad espandere l'espressione $e^{-A} B e^A$ nella speranza di trovare un qualche *schema* che

velocizzi il conto:

$$\begin{aligned}
e^A &= \mathbb{I} + A + \frac{1}{2!}A^2 + \dots \\
e^{-A} &= \mathbb{I} - A + \frac{1}{2!}A^2 + \dots \\
e^{-A}Be^A &= \left(\mathbb{I} - A + \frac{1}{2!}A^2 + \dots\right) \left(B + BA + \frac{1}{2!}BA^2 + \dots\right) = \\
&= B + BA + \frac{1}{2!}BA^2 + \dots - AB - ABA - \frac{1}{2}ABA^2 + \dots \\
&+ \frac{1}{2}A^2B + \frac{1}{2}A^2BA + \frac{1}{4}A^2BA^2 + \dots = \\
&= B + (BA - AB) + \frac{1}{2!}(BA^2 - 2ABA + A^2B) + \dots = \\
&= B + [B, A] + \frac{1}{2!}[[B, A], A] + \dots
\end{aligned}$$

Con un po' di formalismo in più si giunge allora alla *formula di Hadamard*:

$$e^{-A}Be^A = \sum_{n=0}^{\infty} \frac{1}{n!} \underbrace{[[[\dots [B, A], A], A], \dots, A]}_{n \text{ volte}}$$

dove la differenza di segno è puramente convenzionale, e serve per considerare una rotazione di φ positivo come *antioraria*.

In questi conti stiamo ignorando ogni problema di dominio: per esempio non è detto a priori che le potenze n -esime di A siano definite. Il tutto, in ogni caso, funziona all'interno del formalismo (ben più avanzato) degli spazi di Hilbert equipaggiati, che però non esploriamo.

Svolgiamo lo stesso conto per gli operatori che ci interessano, ossia poniamo nella formula di Hadamard:

$$A = i\frac{\varphi}{\hbar}L_3 \quad B = X_1$$

Calcolando i vari termini, ci accorgiamo della comparsa delle espansioni di sin e cos:

$$\begin{aligned}
\exp\left(i\frac{\varphi}{\hbar}L_3\right)X_1\exp\left(-i\frac{\varphi}{\hbar}L_3\right) &= X_1 - \frac{i\varphi}{\hbar}[X_1, L_3] + \frac{1}{2}\left(\frac{-i\varphi}{\hbar}\right)^2[[X_1, L_3], L_3] + \dots = \\
&= X_1 - i\frac{\varphi}{\hbar}(-i\hbar X_2) + \frac{1}{2}\left(\frac{i\varphi}{\hbar}\right)^2 - i\hbar \underbrace{[X_2, L_3]}_{i\hbar X_1} + \dots = \\
&= X_1 - \varphi X_2 - \frac{1}{2}\varphi^2 X_1 + \dots \\
&= X_1 \left(1 - \frac{1}{2}\varphi^2 \dots\right) - X_2 (\varphi + \dots) = \\
&= X_1 \cos \varphi - X_2 \sin \varphi
\end{aligned}$$

Conti analoghi portano alle *trasformazioni* delle altre componenti dell'operatore posizione \vec{X} :

$$\begin{aligned}\exp\left(i\frac{\varphi}{\hbar}L_3\right)X_2\exp\left(-i\frac{\varphi}{\hbar}L_3\right) &= X_2\cos\varphi + X_1\sin\varphi \\ \exp\left(i\frac{\varphi}{\hbar}L_3\right)X_3\exp\left(-i\frac{\varphi}{\hbar}L_3\right) &= X_3\end{aligned}$$

Mettendo tutto insieme otteniamo perciò:

$$\exp\left(i\frac{\varphi}{\hbar}L_3\right)\vec{X}\exp\left(-i\frac{\varphi}{\hbar}L_3\right) = R(\varphi, \hat{z})\vec{X}$$

con:

$$R(\varphi, \hat{z}) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \vec{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$$

Dove \hat{z} è il versore lungo z , e $R(\varphi, \hat{z})$ è la matrice di rotazione (antioraria) di un angolo φ intorno all'asse z .

Più in generale, se \vec{n} è un versore di \mathbb{R}^3 :

$$U(\varphi, \vec{n}) \equiv \exp\left(-i\frac{\varphi}{\hbar}\vec{L} \cdot \vec{n}\right)$$

è l'operatore (unitario) che descrive le rotazioni di un angolo φ attorno al versore \vec{n} .

Ripetendo gli stessi conti per l'operatore momento \vec{P} si ottiene la stessa cosa:

$$\exp\left(i\frac{\varphi}{\hbar}L_3\right)\vec{P}\exp\left(-i\frac{\varphi}{\hbar}L_3\right) = R(\varphi, \hat{x})\vec{P}$$

Deriviamo allora la trasformazione di una generica funzione d'onda $\psi(\vec{x})$ in rappresentazione $\{\vec{x}\}$. Procediamo alla stessa maniera utilizzata per le traslazioni spaziali, notando che $U(\varphi, \vec{n})$ *agisce* “da destra” su un bra come una *rotazione inversa*:

$$\langle \vec{x} | U(\varphi, \vec{n}) = \langle R^{-1}(\varphi, \vec{n})\vec{x} |$$

E applicando ciò ad una $|\psi\rangle$ generica otteniamo:

$$\langle \vec{x} | U(\varphi, \vec{n}) |\psi\rangle = \langle R^{-1}(\varphi, \vec{n})\vec{x} | \psi\rangle = \psi(R^{-1}(\varphi, \vec{n})\vec{x})$$

Confrontando con le rappresentazioni delle algebre di Lie dei gruppi riconosciamo che:

$$\frac{L_j}{i\hbar} = \frac{\partial(U(\varphi, \vec{n})\psi)}{\partial(\varphi n_j)}(\vec{x}) \Big|_{\varphi=0}$$

è la rappresentazione in \mathcal{H} (o meglio in un dominio denso) dei generatori dell'algebra di Lie delle rotazioni.

Perché una rotazione di 2π lascia \vec{x} invariante:

$$U(2\pi, \vec{n})\psi(\vec{x}) = \psi(\vec{x})$$

ossia $U(2\pi, \vec{n}) = \mathbb{I} = \exp\left(i\frac{2\pi}{\hbar}\vec{L} \cdot \vec{n}\right)$.

$$1 = \sigma(\mathbb{I}) = \sigma\left(\exp\left(i\frac{2\pi}{\hbar}\vec{L} \cdot \vec{n}\right)\right) = \exp\left(i\frac{2\pi}{\hbar}\sigma(\vec{L} \cdot \vec{n})\right) \Rightarrow \sigma(\vec{L} \cdot \vec{n}) \subseteq \hbar\mathbb{Z}$$

4.8 Momento angolare generale

Nei paragrafi precedenti abbiamo ricavato una rappresentazione unitaria $U(\varphi, \vec{n})$ delle rotazioni procedendo in analogia con la MC. Ci chiediamo: è questa *la rappresentazione più generale*, o nella stranezza della MQ è possibile trovare altre possibilità?

L'unico modo di capir ciò è ripartire, in analogia alle traslazioni temporali o spaziali, dai concetti di teoria dei gruppi e delle rappresentazioni che abbiamo introdotto negli scorsi capitoli.

Per l'**isotropia dello spazio** sappiamo che le *rotazioni* devono lasciare *invariate* le *probabilità di transizione*. Rappresentiamo allora le rotazioni con operatori che agiscono sugli stati, ossia in \mathcal{PH} , associando a ogni elemento del gruppo delle rotazioni uno di questi operatori: $G \ni g \mapsto \hat{U}(g)$.

Per il *teorema di Wigner*, tale rappresentazione *induce* una rappresentazione *proiettiva unitaria* su \mathcal{H} , ossia possiamo vedere gli $\hat{U}(g)$ (che lavorano nello spazio proiettivo \mathcal{PH} degli stati puri), come $U(g)$ che invece agiscono sul “normale” spazio di Hilbert \mathcal{H} definiti *a meno di un fattore*.

Per rimuovere questa ambiguità potremmo usare il *teorema di Bargmann*.

Le rotazioni sono descritte dal gruppo $SO(3)$, la cui algebra di Lie è indicata con lettere minuscole come $\mathfrak{so}(3)$. Detto un elemento $R \in SO(3)$:

- R è una matrice 3×3 , come indicato dal (3).
- R è una matrice **Ortogonale**, ossia la sua inversa coincide con la trasposta:
 $R^T R = \mathbb{I}$
- R è una matrice **Speciale**, ossia con $\det R = +1$ (le matrici ortogonali, di per sé, hanno $\det = \pm 1$).

Il problema è che mentre per le traslazioni temporali/spaziali partivamo da $G = (\mathbb{R}, +)$ o $G = (\mathbb{R}^3, +)$, entrambi già semplicemente connessi, qui abbiamo $G = SO(3)$ che *non* è semplicemente connesso!

Per vedere il perché, esaminiamone la sua struttura di varietà differenziale, su cui fissiamo una parametrizzazione.

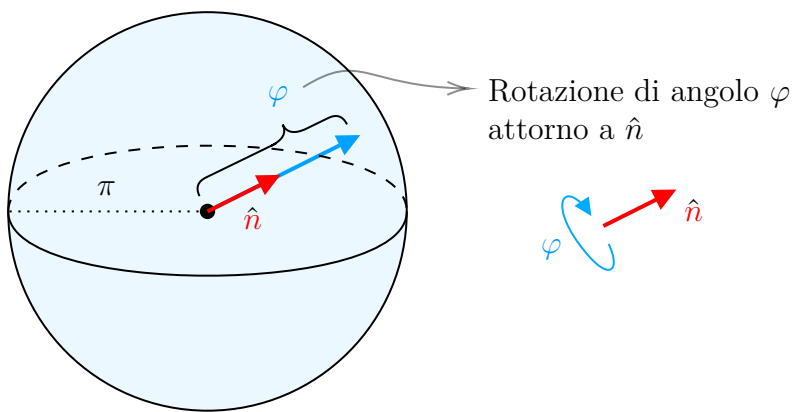


Figura 4.3: Parametrizzazione di $SO(3) \cong B^3 / \sim$

Consideriamo la *palla* B^3 di raggio π , ossia la *sfera* 3D considerata con bordo e punti interni, che sappiamo parametrizzare, e mostriamo che è *equivalente* a $SO(3)$. Detto O il centro di B^3 , un punto $P \in B^3$ definisce una direzione OP , che interpretiamo come l'*asse* di una rotazione. La distanza di P da O (che sarà al più π), indica invece l'*angolo di rotazione*, che sarà (convenzionalmente) in senso orario osservando “da O verso la punta di OP ”¹³.

In questo modo, usando $\varphi \in [0, \pi]$ e $\hat{n} \in \mathbb{R}^3$ possiamo individuare ogni punto in B^3 , e quindi ogni rotazione in $SO(3)$.

In realtà c'è una correzione importante da effettuare. Notiamo che, per la nostra convenzione, ruotare di π attorno a \hat{n} produce lo stesso risultato di ruotare di π attorno a $-\hat{n}$ ¹⁴, dato che differiscono di una rotazione di 2π equivalente alla rotazione nulla. Ciò significa che, geometricamente, dobbiamo considerare i punti antipodali della *superficie esterna* di B^3 come *equivalenti*, visto che corrispondono a rotazioni di π con \hat{n} opposti. Se \sim è tale relazione di equivalenza, indichiamo la palla con i punti antipodali identificati con B^3 / \sim , che è finalmente *isomorfa* al gruppo $SO(3)$ delle rotazioni.

Possiamo allora notare che B^3 / \sim non è semplicemente connessa: dato che i punti antipodali sono *lo stesso punto*, un diametro di B^3 / \sim è un *percorso chiuso*, che però non si può contrarre con continuità a un punto senza spezzarlo, dato che i suoi estremi sono vincolati a muoversi in modo *speculare* sulla superficie di B^3 .

Ciò significa che il gruppo di ricoprimento di $SO(3)$ non può essere $SO(3)$ stesso:

$$\widetilde{SO(3)} \neq SO(3)$$

Si può dimostrare¹⁵ che il ricoprimento $\widetilde{SO(3)}$ è quello delle matrici 2×2 Unitarie (complesse) a determinante $+1$ (**S**peciali), che chiamiamo $SU(2)$ (e la relativa

¹³Il che equivale ad applicare la *regola della mano destra* puntando il pollice verso P e osservando la direzione in cui si piegano le altre quattro dita.

¹⁴Se ciò non è immediatamente chiaro, basta disegnare una freccia verso l'alto su un foglio, e ruotare il tutto di 180° in senso antiorario: la freccia punterà in basso. Stesso risultato si ottiene riportandosi alla situazione iniziale e ruotando di 180° in senso orario.

¹⁵L'idea sta nell'utilizzare i *quaternioni*, che possiamo considerare come un'estensione dei numeri complessi con “3 unità immaginarie invece che una”, che sono in corrispondenza biunivoca con i punti di $S^3 \cong SU(2)$ - e in effetti la loro esistenza deriva proprio dal fatto che su S^3 si

algebra di Lie è indicata con $\mathfrak{su}(2)$). Gli elementi T di $SU(2)$ hanno la forma:

$$T = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \in SU(2)$$

Con la condizione $|a|^2 + |b|^2 = \det(T) = +1$. Possiamo riscrivere tale condizione separando parti reali e immaginarie:

$$(\operatorname{Re} a)^2 + (\operatorname{Im} a)^2 + (\operatorname{Re} b)^2 + (\operatorname{Im} b)^2 = 1$$

Otteniamo l'equazione di una sfera¹⁶ S^3 in 4 dimensioni, che sappiamo essere semplicemente connessa (lo sono tutte le S^n per $n > 1$).

Si dimostra che ogni punto di S^3 parametrizzata *cartesianamente* con $(\operatorname{Re} a, \operatorname{Im} a, \operatorname{Re} b, \operatorname{Im} b)$ corrisponde, *a meno del segno*, a un'unica rotazione in $SO(3)$. Matematicamente, si ha una mappa suriettiva $\tilde{g} : SU(2) \cong \tilde{SO}(3) \rightarrow SO(3)$ (che nello specifico è un *ricoprimento a due fogli*) data da:

$$\pm(\operatorname{Re} a, \operatorname{Im} a, \operatorname{Re} b, \operatorname{Im} b) \mapsto \tilde{g}(\operatorname{Re} a, \operatorname{Im} a, \operatorname{Re} b, \operatorname{Im} b)$$

Geometricamente, \tilde{g} collega punti antipodali su S^3 alla stessa rotazione in $SO(3)$.

Cerchiamo di dare un'intuizione *pittorresca* della situazione. Spostiamoci in dimensione inferiore, e partiamo da B^2 , ossia da un *cerchio* (inteso come circonferenza + punti interni), i cui punti rappresentano le *trasformazioni* che stiamo esaminando¹⁷. Come prima, identifichiamo gli antipodi di B^2 , dato che nella nostra analogia corrisponderebbero alle *stesse trasformazioni*. Abbiamo quindi che B^2 non è semplicemente connesso (un diametro di B^2 è un percorso chiuso che non si può contrarre a un punto con continuità).

Vogliamo trovare un'altra varietà \tilde{G} , che sia semplicemente connessa, tale che esista una mappa suriettiva $\tilde{g} : \tilde{G} \rightarrow B^2$, ossia per cui ogni punto di B^2 abbia *almeno una controimmagine* in \tilde{G} .

Proviamo con S^2 (visto che nel caso a dimensione maggiore abbiamo usato S^3), che è la superficie di una sfera tridimensionale - ed è semplicemente connessa.

Allora, pittorrescamente, si consideri una sfera unitaria centrata all'origine di un piano cartesiano, e il piano $\hat{x}\hat{y}$ che la *taglia* a metà. \tilde{g} si può interpretare come la

può definire una struttura di gruppo (di Lie) - cosa che per esempio non si può fare su S^2 . Si procede dimostrando che un quaternioni *a meno del segno* individua una rotazione tridimensionale: perciò esiste una mappa suriettiva tra $SU(2)$ e $SO(3)$, e dato che S^3 è semplicemente connessa, $SU(2)$ è il ricoprimento universale di $SO(3)$. I dettagli della dimostrazione sono disponibile nella lezione 4 – 1 del corso di Geometria Differenziale tenuto dal professor Francesco Bottacin

¹⁶Con sfera intendiamo sempre la *superficie sferica*. S^n è quindi la superficie n -dimensionale di una sfera in $n + 1$ dimensioni: S^2 è la superficie sferica di una normale sfera 3D, mentre S^3 è la “superficie” tridimensionale di una ipersfera, che consideriamo immersa in \mathbb{R}^4 .

¹⁷*Approfondimento*: In realtà, i punti di B^2 non rappresentano, come ci si potrebbe aspettare, le rotazioni in 2D, che sono invece individuate dai punti di B^1 - con antipodi identificati - ossia dalla circonferenza unitaria $S^1 \cong U(1)$. In effetti il ricoprimento S^2 , inteso come varietà, è *incompatibile* con una struttura di gruppo.

proiezione dei punti della sfera sul piano. In particolare, notiamo che ogni punto del cerchio unitario B^2 sul piano $\hat{x}\hat{y}$ si ottiene dalla proiezione di almeno un punto di S^2 , e quindi \tilde{g} è una mappa suriettiva..

Graficamente, notiamo che ci sono sempre due punti di S^2 che *si proiettano* nello stesso punto di B^2 : uno “dall’alto” e uno “dal basso”. In realtà dobbiamo considerare che i punti di S^2 all’intersezione con il piano $z = 0$ sono *proiettati su se stessi*, ma dobbiamo ricordarci che in B^2 abbiamo identificato i punti antipodali, e quindi anche qui è come avere *due proiezioni* per lo stesso punto.

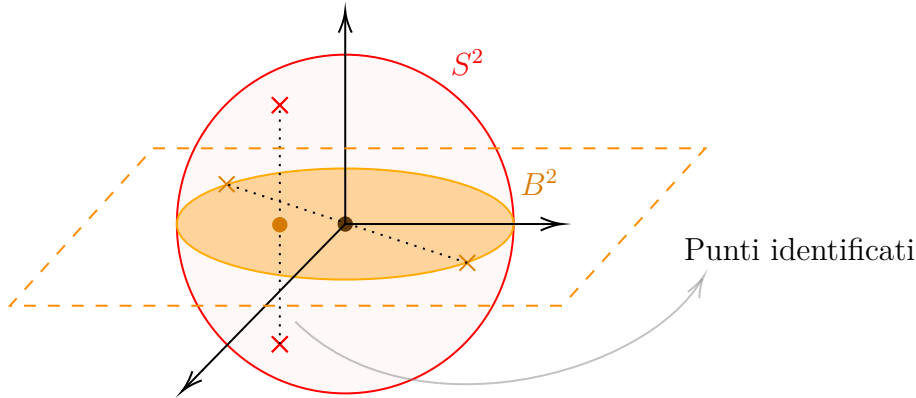


Figura 4.4: Ogni punto di B^2/\sim è *originato* dalla proiezione di esattamente *due* punti di S^2

Possiamo allora comprendere la definizione di \tilde{g} come *a meno del segno*: così facendo stiamo “eliminando tutti i doppioni”. Visivamente, ciò significa considerare un solo emisfero (per esempio S^2_+), con *mezzo bordo* (dato che anche per i punti su $z = 0$ vi sono *due proiezioni*).

Notiamo altre due cose:

- Un diametro di B^2 , che è il cammino chiuso e non contraibile a un punto che prima dava problemi, risulta come la proiezione di una *circonferenza massimale* su S^2 , che stavolta è contraibile a un punto.
- Una rotazione di 2π corrispondeva in B^3 (e quindi anche in B^2 , per analogia) a passare da un punto sul *bordo* al suo punto antipodale (e infatti i due coincidono su B^2). La stessa cosa non si ha più su S^2 , dato che tale punto è proiezione di *due punti* non coincidenti. Solo ripetendo un’ulteriore rotazione di 2π anche in S^2 i due punti coincideranno.

Stessa cosa si può dimostrare, perdendo però ogni visualizzazione, nel caso a dimensione 3 che ci interessa, ossia per $SO(3) \cong B^3/\sim$ e $S^3 \cong SU(2) \cong \widetilde{SO}(3)$.

Ciò ha conseguenze fondamentali: se decidiamo di utilizzare una rappresentazione unitaria senza ambiguità di fasi (e quindi *non proiettiva*) tramite il teorema di Bargmann, dovremo spostarci sul gruppo di ricoprimento, che *non* segue le stesse “regole” del gruppo di partenza: su B^3/\sim la rotazione di 2π è nulla, su S^3 è quella di 4π a essere nulla.

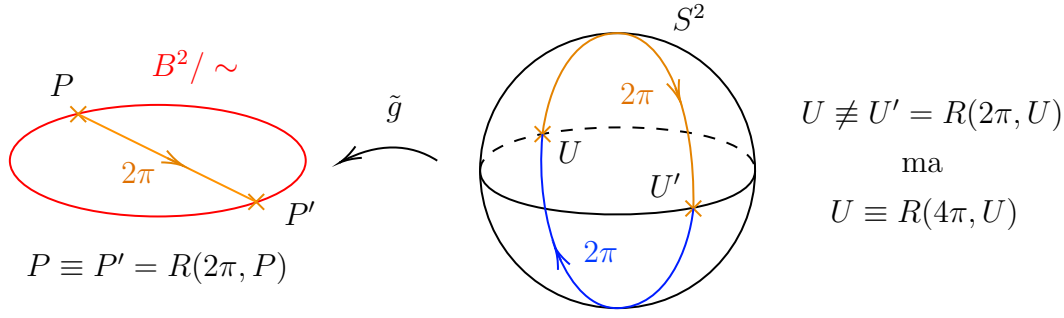


Figura 4.5: Una rotazione di 2π su B^2/\sim “non sposta il punto” P , dato che P e P' sono identificati da \sim , ma lo sposta sul ricoprimento S^2 , dove partendo da U si ritorna a U solo dopo una rotazione di 4π

Applicando allora il teorema di Bargmann, troviamo che una rappresentazione unitaria in \mathcal{H} di una rotazione $\tilde{R} \in SU(2)$ sul gruppo di ricoprimento universale di $SO(3)$ sarà una certa $\tilde{U}(\varphi, \vec{n})$, che per teorema di Stone si ottiene dalla mappa esponenziale applicata ad un elemento (o meglio ad una sua rappresentazione) dell'algebra di Lie $\mathfrak{su}(2)$. Detta \vec{J} una rappresentazione in un insieme (denso) di \mathcal{H} dei generatori (base) dell'algebra di Lie di $SU(2)$ si ha:

$$\tilde{U}(\varphi, \vec{n}) = \exp\left(-i\frac{\varphi}{\hbar}\vec{n} \cdot \vec{J}\right)$$

Come abbiamo visto nella teoria, l'algebra di Lie dipende unicamente dalle caratteristiche *locali* del gruppo da cui deriva. Poiché un gruppo e ricoprimento universale sono *localmente identificati*, da $\tilde{SO}(3) \cong SU(2)$, si ha $\mathfrak{su}(2) \cong \mathfrak{so}(3)$. In particolare, l'algebra di \vec{J} è la stessa dei generatori delle rotazioni, quindi la stessa di \vec{L} . In pratica, possiamo riutilizzare per \vec{J} i commutatori già calcolati per \vec{L} :

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k \quad (4.34)$$

Poiché tuttavia una rotazione di 2π in $\tilde{SO}(3)$ non è equivalente alla rotazione nulla, per \vec{J} **non** abbiamo più la condizione sugli autovalori $\sigma(\vec{J} \cdot \vec{n}) \subseteq \hbar\mathbb{Z}$.

\vec{J} sarà chiamato **momento angolare** se soddisfa (4.34) e $\vec{L} = \vec{X} \wedge \vec{P}$ verrà detto **momento angolare orbitale**.

Da (4.34) sappiamo che, come ricavato per \vec{L} , \vec{J}^2, J_3 sono compatibili, e quindi possiamo cercare *autovettori comuni*.

Una procedura efficiente per far ciò consiste nel notare che, trovato un autovettore comune, possiamo “costruire” tutti gli altri, tramite l'applicazione di opportuni operatori che “alzano” e “abbassano” gli autovalori.

Definiamo allora $J_{\pm} \equiv J_1 \pm iJ_2$, ove $J_- = (J_+)^{\dagger}$. Ricaviamone subito le relazioni di commutazione. Ricordando (4.34) e sfruttando la bilinearità del commutatore

si ha:

$$\begin{aligned} [J_3, J_{\pm}] &= [J_3, J_1 \pm iJ_2] = i\hbar(\overbrace{\epsilon_{312}}^1 J_2 \pm i\overbrace{\epsilon_{321}}^{-1} J_1) = \\ &= i\hbar(J_2 \mp iJ_1) = \hbar(iJ_2 \pm J_1) = \pm\hbar(J_1 \pm iJ_2) = \pm\hbar J_{\pm} \end{aligned} \quad (4.35)$$

$$\begin{aligned} [J_+, J_-] &= [J_1 + iJ_2, J_1 - iJ_2] = i[J_2, J_1] - i[J_1, J_2] = \\ &= i(i\hbar\overbrace{\epsilon_{123}}^1 J_3) - i(i\hbar\overbrace{\epsilon_{213}}^{-1} J_3) = 2\hbar J_3 \end{aligned} \quad (4.36)$$

Sia ora $|\lambda, m\rangle$ un autovettore comune di \vec{J}^2 e J_3 , espresso in notazione di Dirac, di autovalore $\hbar^2\lambda$ per \vec{J}^2 e $\hbar m$ per J_3 , ossia tale che:

$$\vec{J}^2 |\lambda, m\rangle = \hbar^2\lambda |\lambda, m\rangle \quad (4.37)$$

$$J_3 |\lambda, m\rangle = \hbar m |\lambda, m\rangle \quad (4.38)$$

In tal modo abbiamo “adimensionalizzato” la notazione dell’autovettore, dato che $\lambda, m \in \mathbb{R}$.

Mostriamo ora che J_+ e J_- , se applicati a J_3 rispettivamente ne “alzano” e “abbassano” l’autovalore $\hbar m$, e quindi ci permettono di trovare gli altri autovalori. Infatti, applicandoli a J_3 otteniamo:

$$\begin{aligned} J_3 J_{\pm} |\lambda, m\rangle &= ([J_3, J_{\pm}] + J_{\pm} J_3) |\lambda, m\rangle \stackrel{(4.35)}{=} J_{\pm} (J_3 \pm \hbar) |\lambda, m\rangle = \\ &\stackrel{(4.38)}{=} J_{\pm} (m\hbar \pm \hbar) |\lambda, m\rangle = \hbar(m \pm 1) J_{\pm} |\lambda, m\rangle \end{aligned} \quad (4.39)$$

ovvero $J_{\pm} |\lambda, m\rangle$ è autovettore di J_3 con autovalore $\hbar(m \pm 1)$ (solo se $|\lambda, m\rangle$ è autovettore e $J_{\pm} |\lambda, m\rangle \neq 0$).

Il passo successivo sta nel riscrivere \vec{J}^2 utilizzando gli operatori J_{\pm} e J_3 . Partiamo esaminando un classico risultato dell’algebra alla luce di operatori che non commutano. La scomposizione:

$$a^2 + b^2 = (a - ib)(a + ib)$$

deve essere adattata per la non commutatività. Infatti, se calcoliamo il termine a destra per A e B che non commutano:

$$\begin{aligned} (A - iB)(A + iB) &= A^2 + iAB - iBA + B^2 = A^2 + B^2 + i[A, B] \\ \Rightarrow A^2 + B^2 &= (A - iB)(A + iB) - i[A, B] \end{aligned}$$

Possiamo applicare ciò al nostro caso, dato che \vec{J}^2 è una somma di quadrati:

$$\begin{aligned} \vec{J}^2 &= J_1^2 + J_2^2 + J_3^2 = (J_1 \mp iJ_2)(J_1 \pm iJ_2) \mp i[J_1, J_2] + J_3^2 = \\ &\stackrel{(4.35)}{=} J_{\mp} J_{\pm} \pm \hbar J_3 + J_3^2 \end{aligned} \quad (4.40)$$

Riarrangiando, troviamo che:

$$J_{\mp} J_{\pm} = \vec{J}^2 \mp \hbar J_3 - J_3^2 \quad (4.41)$$

Ciò che abbiamo appena calcolato ci è utile per esaminare il *modulo* degli autovettori di J_3 “aumentati” o “abbassati” rispettivamente da J_+ e J_- .

$$\begin{aligned} 0 \leq \|J_+ |\lambda, m\rangle\|^2 &= \langle \lambda, m | J_+^\dagger J_+ |\lambda, m\rangle = \langle \lambda, m | J_- J_+ |\lambda, m\rangle = \\ &\stackrel{(4.41)}{=} \langle \lambda, m | \vec{J}^2 - \hbar J_3 - J_3^2 |\lambda, m\rangle = \hbar^2(\lambda - m - m^2) = \hbar^2(\lambda - m(m+1)) \end{aligned} \quad (4.42)$$

$$\begin{aligned} 0 \leq \|J_- |\lambda, m\rangle\|^2 &= \langle \lambda, m | J_+ J_- |\lambda, m\rangle = \\ &\stackrel{(4.41)}{=} \langle \lambda, m | \vec{J}^2 + \hbar J_3 - J_3^2 |\lambda, m\rangle = \hbar^2(\lambda - m(m-1)) \end{aligned} \quad (4.43)$$

Poiché λ è fissato, notiamo allora che il modulo degli autovettori di autovalore maggiore o minore rispettivamente sale o scende. Ci aspettiamo che ciò non possa andare avanti all’infinito, dato che il modulo non può scendere sotto zero, né salire infinitamente. In effetti, nei prossimi passaggi mostreremo che esistono un m_{\min} e un m_{\max} .

Partiamo sommando le disuguaglianze (4.42) e (4.43):

$$\lambda - m^2 \geq 0 \Rightarrow \lambda \geq m^2 \geq 0 \Rightarrow \exists \sqrt{\lambda}$$

E quindi calcolando la radice quadrata di ambo i membri giungiamo a:

$$|m| \leq \sqrt{\lambda} \Rightarrow -\sqrt{\lambda} \leq m \leq +\sqrt{\lambda} \quad (4.44)$$

Sappiamo da (4.39) che, applicando J_+ e J_- a $|\lambda, m\rangle$, possiamo alzare o abbassare arbitrariamente l’autovalore di J_3 senza modificare λ , ma ciò avviene solo a meno che l’applicazione di J_+ o J_- non dia il vettore nullo.

Perché allora sia la (4.39) che la (4.44) siano contemporaneamente vere, si ha che fissato λ deve esistere un m_{\min} tale che $J_- |\lambda, m_{\min}\rangle = 0$, e simmetricamente anche un m_{\max} , per cui $J_+ |\lambda, m_{\max}\rangle = 0$, ossia due autovalori che “interrompono la catena” e oltre i quali non si può più né scendere né salire¹⁸.

Cosa succede applicando \vec{J}^2 agli autoket di m max o min? Vediamolo:

$$\begin{aligned} \vec{J}^2 |\lambda, m_{\max}\rangle &= \hbar^2 \lambda |\lambda, m_{\max}\rangle \stackrel{(4.40)}{=} \underbrace{(J_- J_+ + \hbar J_3 + J_3^2)}_0 |\lambda, m_{\max}\rangle = \\ &= \hbar^2(m_{\max} + m_{\max}^2) |\lambda, m_{\max}\rangle \Rightarrow \lambda = m_{\max} + m_{\max}^2 \\ \vec{J}^2 |\lambda, m_{\min}\rangle &= \hbar^2 \lambda |\lambda, m_{\min}\rangle \stackrel{(4.40)}{=} \underbrace{(J_+ J_- - \hbar J_3 + J_3^2)}_0 |\lambda, m_{\min}\rangle = \\ &= (-m_{\min} + m_{\min}^2) |\lambda, m_{\min}\rangle \Rightarrow \lambda = -m_{\min} + m_{\min}^2 \end{aligned}$$

dato che J_+ applicato all’autoket con m_{\max} non può alzarlo ulteriormente, e quindi ritorna il vettore nullo, e stessa cosa per J_- applicato all’autoket con m_{\min} .

¹⁸Il riferimento è voluto.

Dato che in tutto ciò λ non cambia, possiamo scrivere:

$$\lambda = m_{\max} + m_{\max}^2 = -m_{\min} + m_{\min}^2$$

che è verificata se $m_{\max} = -m_{\min} \equiv j > 0$, da cui $\lambda = j(j+1)$.

Poiché applicando $N+1$ volte J_- all'autoket $|j, j = m_{\max}\rangle$ dobbiamo ottenere 0 per un qualche N abbiamo che:

$$\underbrace{J_-^N |j, j\rangle}_{\propto |j, j-N\rangle} \propto |j, m_{\min} = -j\rangle \Rightarrow j - N = -j \Rightarrow 2j = N \in \mathbb{N} \Rightarrow j \in \frac{\mathbb{N}}{2} \Rightarrow m \in \frac{\mathbb{Z}}{2}$$

L'idea è che, come abbiamo visto, i valori possibili per m vanno da $m_{\min} = -j$ a $m_{\max} = +j$, sono a distanza unitaria (a meno di un \hbar) uno dall'altro e sono in *numero finito*. Possiamo pensarli come una successione di $2N+1$ *puntini* “centrata” attorno allo 0 della retta reale. Se N è pari, avremo un autovalore a 0 e gli altri a $\pm i$, ma se N è dispari allora gli autovalori saranno per forza del tipo $\pm i/2$.

Esplicitamente, fissato allora $j \in \mathbb{N}/2$ i possibili valori di m sono:

$$\{j, j-1, j-2, \dots, -j+1, -j\}$$

che coprono $2j+1$ possibilità.

Se imponiamo a $|j, m\rangle$ di avere norma 1:

$$J_{\pm} |j, m\rangle = c_{\pm}(j, m) |j, m \pm 1\rangle$$

$$\|J_{\pm} |j, m\rangle\|^2 = |c_{\pm}(j, m)|^2$$

*Coefficienti di
normalizzazione*

Da cui:

$$\|J_{\pm} |\lambda, m\rangle\|^2 = \hbar^2(\lambda - m(m \pm 1)) = \hbar^2(j(j+1) - m(m \pm 1)) = |c_{\pm}(j, m)|^2$$

Con una scelta opportuna di fase possiamo fissare il segno (positivo) della radice, e quindi scrivere:

$$J_{\pm} |j, m\rangle = \sqrt{j(j+1) - m(m \pm 1)} |j, m \pm 1\rangle$$

Abbiamo allora dimostrato che, se esiste un autoket $|j, j\rangle$ “da cui partire” vale:

$$\sigma_P(\vec{J}^2; J_3) = \underbrace{\{\hbar^2(j(j+1)), j \in \frac{\mathbb{N}}{2}\}}_{\sigma(\vec{J}^2)} \underbrace{\{\hbar m, m = -j, -j+1, \dots, j\}}_{\sigma(J_3)}$$

Notiamo che lo spettro è puramente **discreto**, dato che nei ragionamenti abbiamo usato delle *norme*, e quindi gli autostati devono essere normalizzabili. Inoltre partendo da un punto qualsiasi e *costruendo* gli altri autovalori otteniamo *tutti* gli autovalori, dato che, come vedremo, i loro autovettori formano una base di L^2 .

Il vero aspetto importante, tuttavia, è dato dal fatto che per ricavare tale spettro comune siamo partiti *unicamente* dalle relazioni di commutazione, e in particolare da:

$$[J_i, J_k] = i\hbar \epsilon_{ijk} J_k$$

Ciò significa che le conclusioni raggiunte valgono anche per tutte le altre osservabili che soddisfano tali relazioni!

*(Lezione 26 del
26/11/2018)*

4.8.1 Gli autostati del momento angolare

Concentriamoci sul caso particolare del momento angolare “analogo al caso classico” $\vec{J} = \vec{L} = \vec{X} \wedge \vec{P}$, che chiamiamo *momento angolare orbitale*.

Vogliamo trovare la forma esplicita degli autoket $|l, m\rangle$ comuni a \vec{L}^2 e L_3 , con $L \in \mathbb{N}$ (e non $\mathbb{N}/2$, dato che per \vec{L} “classico” una rotazione di 2π è *equivalente* a una rotazione nulla), e $|m| \leq l$. In effetti, se riusciamo a dimostrare l’esistenza di $|l, l\rangle$ (trovandolo esplicitamente) abbiamo automaticamente l’esistenza di tutti i altri autovettori, data la possibilità di *costruirli* tramite L_{\pm} , e se l’insieme di tutti i tali autovettori è una base di L^2 abbiamo mostrato anche che lo spettro $\sigma(\vec{L}^2, L_z)$ è puramente *discreto*.

Conviene allora partire esprimendo gli operatori che ci servono, ossia L_3 e L_{\pm} (da cui si ricava anche \vec{L}^2), in coordinate “compatibili” con la simmetria rotazionale che stiamo studiando, ossia le **coordinate polari** $\{r, \theta, \varphi\}$.

Ricordiamo il passaggio da coordinate sferiche a cartesiane e viceversa:

1. L_3, L_{\pm}, L^2 in coordinate polari

$$\begin{cases} x = r \sin \theta \cos \varphi \\ y = r \sin \theta \sin \varphi \\ z = r \cos \theta \end{cases} \Leftrightarrow \begin{cases} r = \sqrt{x^2 + y^2 + z^2} \\ \cos \theta = \frac{z}{\sqrt{x^2 + y^2 + z^2}} \\ \tan \varphi = \frac{y}{x} \end{cases}$$

Dato che L_3 e L_{\pm} in rappresentazione $\{\vec{x}\}$ contengono le derivate $\partial_x, \partial_y, \partial_z$, conviene ricavarsi anche la trasformazione delle basi degli spazi *tangenti* dalle coordinate cartesiane a quelle polari e viceversa¹⁹. Procedendo direttamente dalla *chain-rule*:

$$\begin{aligned} \frac{\partial}{\partial \varphi} &= \frac{\partial x}{\partial \varphi} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \varphi} \frac{\partial}{\partial y} + \frac{\partial z}{\partial \varphi} \frac{\partial}{\partial z} = \\ &= -r \sin \theta \sin \varphi \frac{\partial}{\partial x} + r \sin \theta \cos \varphi \frac{\partial}{\partial y} = \\ &= -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} \end{aligned} \tag{4.45}$$

E tale derivazione ci basta per esprimere L_3 in coordinate polari:

$$L_3 = -i\hbar x \frac{\partial}{\partial y} + i\hbar y \frac{\partial}{\partial x} = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) = \frac{\hbar}{i} \frac{\partial}{\partial \varphi}$$

Per L_{\pm} è invece necessario un po’ più di lavoro²⁰. Partendo alla *chain-rule* vista in (4.45), notiamo che possiamo scrivere la i -esima derivata rispetto alle coordinate sferiche come prodotto scalare tra il vettore delle derivazioni rispetto alle

¹⁹ΛL’idea è che il cambio di coordinate oltre a trasformare anche i *punti* di \mathbb{R}^3 , trasforma anche i *vettori*, che precisamente vanno intesi come appartenenti allo *spazio tangente* a \mathbb{R}^3 in un dato punto (che nel caso di \mathbb{R}^n , agli effetti, coincide - cioè è *canonicamente isomorfo* - con \mathbb{R}^n stesso). Tali vettori sono *combinazioni lineari* degli elementi della base dello spazio tangente, che è data (nel caso canonico in coordinate locali) proprio dalle *derivazioni* $\partial/\partial_x, \partial/\partial_y, \partial/\partial_z$

²⁰ΛSi tratta di quel genere di conti che è bene fare almeno una volta nella vita!

coordinate cartesiane e la i -esima riga della *trasposta* della matrice jacobiana del cambiamento di coordinate. Esplicitamente:

$$\begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \\ \frac{\partial}{\partial \varphi} \end{pmatrix} = \left(\frac{\partial(x, y, z)}{\partial(r, \theta, \varphi)} \right)^T \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} = \begin{pmatrix} \sin \theta \cos \varphi & \sin \theta \sin \varphi & \cos \theta \\ r \cos \theta \cos \varphi & r \cos \theta \sin \varphi & -r \sin \theta \\ -r \sin \theta \sin \varphi & r \sin \theta \cos \varphi & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}$$

E per la relazione inversa è necessario invertire la matrice:

$$\begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} = \left[\left(\frac{\partial(x, y, z)}{\partial(r, \theta, \varphi)} \right)^T \right]^{-1} \begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \\ \frac{\partial}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \sin \theta \cos \varphi & \frac{1}{r} \cos \theta \cos \varphi & -\frac{\sin \varphi}{r \sin \theta} \\ \sin \theta \sin \varphi & \frac{1}{r} \cos \theta \sin \varphi & \frac{\cos \varphi}{r \sin \theta} \\ \cos \theta & -\frac{\sin \theta}{r} & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \\ \frac{\partial}{\partial \varphi} \end{pmatrix} \quad (4.46)$$

Ricaviamo allora:

$$\begin{aligned} L_{\pm} &\equiv L_1 \pm iL_2 = (X_2P_3 - X_3P_2) \pm i(X_3P_1 - X_1P_3) = \\ &= \left(-i\hbar y \frac{\partial}{\partial z} + i\hbar z \frac{\partial}{\partial y} \right) \pm i \left(-i\hbar z \frac{\partial}{\partial x} + i\hbar x \frac{\partial}{\partial x} \right) = \\ &= \hbar \left[-iy \frac{\partial}{\partial z} + i z \frac{\partial}{\partial y} \pm z \frac{\partial}{\partial x} \mp x \frac{\partial}{\partial z} \right] = \\ &= \hbar \left[(\mp x - iy) \frac{\partial}{\partial z} + z \left(\pm \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \right] = \\ &= \hbar \left[\mp (x \pm iy) \frac{\partial}{\partial z} \pm z \left(\frac{\partial}{\partial x} \pm i \frac{\partial}{\partial y} \right) \right] \end{aligned}$$

Dalle equazioni del cambiamento di coordinate abbiamo le sostituzioni:

$$\begin{aligned} x \pm iy &= r(\sin \theta \cos \varphi \pm i \sin \theta \sin \varphi) = \\ &= r \sin \theta (\cos \varphi \pm i \sin \varphi) = r \sin \theta e^{\pm i\varphi} \\ z &= r \cos \theta \end{aligned}$$

Dopo le quali giungiamo a:

$$L_{\pm} = \hbar \left[\mp r \sin \theta e^{\pm i\varphi} \frac{\partial}{\partial z} \pm r \cos \theta \left(\frac{\partial}{\partial x} \pm i \frac{\partial}{\partial y} \right) \right]$$

Otteniamo le ultime sostituzioni da (4.46):

$$\begin{aligned} \frac{\partial}{\partial z} &= \frac{1}{r} \left(r \cos \theta \frac{\partial}{\partial r} - \sin \theta \frac{\partial}{\partial \theta} \right) \\ \frac{\partial}{\partial x} \pm i \frac{\partial}{\partial y} &= \frac{1}{r} \left[r \sin \theta (\cos \varphi \pm i \sin \varphi) \frac{\partial}{\partial r} + \right. \\ &\quad \left. + \cos \theta (\cos \varphi \pm i \sin \varphi) \frac{\partial}{\partial \theta} + \right. \\ &\quad \left. + \frac{1}{\sin \theta} (-\sin \varphi \pm i \cos \varphi) \frac{\partial}{\partial \varphi} \right] = \\ &= \frac{1}{r} e^{\pm i\varphi} \left(r \sin \theta \frac{\partial}{\partial r} + \cos \theta \frac{\partial}{\partial \theta} \pm \frac{i}{\sin \theta} \frac{\partial}{\partial \varphi} \right) \end{aligned}$$

Da cui, infine, giungiamo a:

$$\begin{aligned}
 L_{\pm} &= \hbar e^{\pm i\varphi} \left[\mp \frac{r' \sin \theta}{r'} \left(r \cos \theta \frac{\partial}{\partial r} - \sin \theta \frac{\partial}{\partial \theta} \right) \pm \frac{r' \cos \theta}{r'} \left(r \sin \theta \frac{\partial}{\partial r} + \cos \theta \frac{\partial}{\partial \theta} \pm \frac{i}{\sin \theta} \frac{\partial}{\partial \varphi} \right) \right] = \\
 &= \hbar e^{\pm i\varphi} \left[\pm \sin^2 \theta \frac{\partial}{\partial \theta} \pm \cos^2 \theta \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \varphi} \right] = \\
 &= \hbar e^{\pm i\varphi} \left(i \cot \theta \frac{\partial}{\partial \varphi} \pm \frac{\partial}{\partial \theta} \right)
 \end{aligned}$$

Poiché anche \vec{L}^2 soddisfa la relazione vista per \vec{J}^2 , possiamo scrivere:

$$\begin{aligned}
 \vec{L}^2 &= L_+ L_- + L_3^2 - \hbar L_3 = \\
 &= \hbar^2 \left(e^{i\varphi} \left(i \cot \theta \frac{\partial}{\partial \varphi} + \frac{\partial}{\partial \theta} \right) \left[e^{-i\varphi} \left(i \cot \theta \frac{\partial}{\partial \varphi} - \frac{\partial}{\partial \theta} \right) \right] - \frac{\partial^2}{\partial \varphi^2} + i \frac{\partial}{\partial \varphi} \right)
 \end{aligned}$$

Attenzione! Le derivate evidenziate si applicano a tutto ciò che è alla loro destra, e perciò purtroppo non vale usare il prodotto notevole:

$$e^{i\varphi} \left(i \cot \theta \frac{\partial}{\partial \varphi} + \frac{\partial}{\partial \theta} \right) \left[e^{-i\varphi} \left(i \cot \theta \frac{\partial}{\partial \varphi} - \frac{\partial}{\partial \theta} \right) \right] \neq \left(i \cot \theta \frac{\partial}{\partial \varphi} \right)^2 - \frac{\partial^2}{\partial \theta^2}$$

Per tale prodotto otteniamo invece:

$$\begin{aligned}
 &= i e^{i\varphi} \cot \theta \frac{\partial}{\partial \varphi} \left(e^{-i\varphi} \left(i \cot \theta \frac{\partial}{\partial \varphi} - \frac{\partial}{\partial \theta} \right) \right) + e^{i\varphi} \frac{\partial}{\partial \theta} \left(e^{-i\varphi} \left(i \cot \theta \frac{\partial}{\partial \varphi} - \frac{\partial}{\partial \theta} \right) \right) = \\
 &= i e^{i\varphi} \cot \theta \left(-i e^{-i\varphi} \left(i \cot \theta \frac{\partial}{\partial \varphi} - \frac{\partial}{\partial \theta} \right) + e^{-i\varphi} \left(i \cot \theta \frac{\partial^2}{\partial \varphi^2} - \frac{\partial}{\partial \theta \partial \varphi} \right) \right) + \\
 &+ e^{i\varphi} e^{-i\varphi} \left(-\frac{\partial^2}{\partial \theta^2} + i \cot \theta \frac{\partial}{\partial \varphi \partial \theta} - \frac{i}{\sin^2 \theta} \frac{\partial}{\partial \varphi} \right) = \\
 &= i \cot \theta \left(\cot \theta \frac{\partial}{\partial \varphi} + i \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial^2}{\partial \varphi^2} - \frac{\partial}{\partial \theta \partial \varphi} \right) - \frac{\partial^2}{\partial \theta^2} + i \cot \theta \frac{\partial}{\partial \varphi \partial \theta} - \frac{i}{\sin^2 \theta} \frac{\partial}{\partial \varphi} = \\
 &= i \cot^2 \theta \frac{\partial}{\partial \varphi} - \cot \theta \frac{\partial}{\partial \theta} - \cot^2 \theta \frac{\partial^2}{\partial \varphi^2} - \frac{\partial^2}{\partial \theta^2} - \frac{i}{\sin^2 \theta} \frac{\partial}{\partial \varphi}
 \end{aligned}$$

Sostituendo nell'espressione per \vec{L}^2 e raccogliendo i termini simili otteniamo:

$$\vec{L}^2 = \hbar^2 \left[-\frac{\partial^2}{\partial \theta^2} + i \left(\cot^2 \theta - \frac{1}{\sin^2 \theta} + 1 \right) \frac{\partial}{\partial \varphi} - (\cot^2 \theta + 1) \frac{\partial^2}{\partial \varphi^2} - \cot \theta \frac{\partial}{\partial \theta} \right]$$

Ricordando la relazione goniometrica:

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

possiamo semplificare le due parentesi tonde. Giungiamo infine a:

$$\vec{L}^2 = -\hbar^2 \left[\frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} \right]$$

Che possiamo “comprimere” ulteriormente applicando Leibniz *al contrario*:

$$\frac{\partial^2}{\partial \theta^2} + \frac{\cos \theta}{\sin \theta} \frac{\partial}{\partial \theta} = \frac{1}{\sin \theta} \left(\cos \theta \frac{\partial}{\partial \theta} + \frac{\partial^2}{\partial \theta^2} \right) = \frac{1}{\sin \theta} \left(\frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} \right)$$

E finalmente otteniamo:

$$\vec{L}^2 = -\hbar^2 \left[\frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} \right]$$

Vediamo che L_3 , L_{\pm} e \vec{L}^2 agiscono solo su θ e φ e non su r . Possiamo allora usare l'isomorfismo “delle coordinate sferiche”:

$$L^2(\mathbb{R}^3, d^3x) \cong L^2(\mathbb{R}_+, r^2 dr) \otimes L^2(S^2, d\Omega = \sin \theta d\theta d\varphi)$$

e cerchiamo le autofunzioni $f_{lm}(\theta, \varphi)$ di \vec{L}^2 e L_3 in $L^2(S^2, d\Omega)$ di autovalori rispettivamente $\hbar^2 l(l+1)$ e $\hbar m$, risolvendo le equazioni (differenziali) agli autovalori date da:

$$L_3 f_{lm}(\theta, \varphi) = \hbar m f_{lm}(\theta, \varphi) \quad (4.47)$$

$$\vec{L}^2 f_{lm}(\theta, \varphi) = \hbar^2 l(l+1) f_{lm}(\theta, \varphi) \quad (4.48)$$

Utilizziamo una tecnica comune nella risoluzione di equazioni differenziali alle derivate parziali, che sta nel cercare soluzioni *fattorizzabili* del tipo:

2. *Autofunzioni fattorizzate*

$$f_{lm}(\theta, \varphi) = \Theta_{lm}(\theta) \Phi_{lm}(\varphi)$$

Sostituendo nella (4.47) otteniamo (sopprimendo temporaneamente gli indici lm per alleggerire la notazione):

$$L_3 \Theta(\theta) \Phi(\varphi) = \frac{\hbar}{i} \Theta(\theta) \frac{\partial}{\partial \varphi} \Phi(\varphi) = \hbar m \Theta(\theta) \Phi(\varphi)$$

che è a variabili separabili, e conduce a:

$\Phi(\varphi)$

$$\frac{\Phi'(\varphi)}{\Phi(\varphi)} = im \Rightarrow \Phi(\varphi) \sim e^{im\varphi}$$

Ricavare $\Theta_{lm}(\theta)$ richiede invece un po' più di conti. Partiamo notando che, poiché in \vec{L}^2 l'unica derivata rispetto a φ è al quadrato, si ha che le $\Theta_{lm}(\theta)$ *non* dipendono dal segno di m , ossia:

$$\Theta_{lm}(\theta) = \Theta_{l,-m}(\theta) = \Theta_{l,|m|}(\theta)$$

Ciò si verifica applicando \vec{L}^2 ad un'autofunzione:

$$\begin{aligned} \vec{L}^2 \Theta(\theta) \Phi(\varphi) &= -\hbar^2 \left[\frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} \right] \Theta(\theta) e^{im\varphi} = \hbar^2 l(l+1) \Theta(\theta) e^{im\varphi} = \\ &= -\hbar^2 e^{im\varphi} \left(-\frac{m^2}{\sin^2 \theta} + \frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} \right) \Theta(\theta) = l(l+1) \Theta(\theta) e^{im\varphi} \end{aligned}$$

E perciò $\Theta(\theta)$ dipende solo da m^2 , e perciò non “avverte” cambiamenti di segno di m .

L'equazione differenziale appena ricavata è difficile da risolvere direttamente²¹, ma fortunatamente possiamo utilizzare i risultati ottenuti esaminando l'algebra di \vec{L} . Supponiamo infatti che esista un autoket $|l, l\rangle$ “massimo”, nel senso che L_+ non possa “alzarlo” ulteriormente, e tale che applicando ad esso L_- iterativamente sia possibile ottenere *tutti* gli altri autoket:

$\Theta(\theta)$ tramite L_+

$$L_+ |l, l\rangle = 0 \quad (4.49)$$

$$|l, m\rangle \propto L_-^{l-m} |l, l\rangle \quad (m > 0) \quad (4.50)$$

Riscrivendo (4.49) in rappresentazione in *coordinate sferiche* $\{r, \theta, \varphi\}$ giungiamo a un'equazione differenziale a variabili separabili:

$$\begin{aligned} \hbar e^{i\varphi} \left(i \cot \theta \frac{\partial}{\partial \varphi} + \frac{\partial}{\partial \theta} \right) (\Theta_u(\theta) e^{il\varphi}) &= 0 \\ \Rightarrow i \cot \theta \Theta_u(\theta) i l e^{il\varphi} + e^{il\varphi} \frac{\partial}{\partial \theta} \Theta_u(\theta) &= 0 \\ \Rightarrow \left(-l \cot \theta + \frac{\partial}{\partial \theta} \right) \Theta_u(\theta) &= 0 \\ \Rightarrow \frac{\partial}{\partial \theta} \Theta_u(\theta) &= l \frac{\cos \theta}{\sin \theta} \Theta_u(\theta) \\ \Rightarrow \frac{1}{\cos \theta} \frac{\partial}{\partial \theta} \Theta_u(\theta) &= \frac{l}{\sin \theta} \Theta_u(\theta) \\ \Rightarrow \frac{d}{d \sin \theta} \Theta_u(\theta) &= \frac{l}{\sin \theta} \Theta_u(\theta) \Rightarrow \Theta_u(\theta) \sim (\sin \theta)^l \end{aligned}$$

(Nell'ultimo passaggio si è effettuato un *cambio di variabili* “sul posto”, ponendo $t = \sin \theta$, da cui $dt = \cos \theta d\theta$, e sostituendo senza introdurre t nell'equazione.

Le $\Theta(\theta)$ e $\Phi(\varphi)$ sono state ricavate *a meno della costante di integrazione*, e perciò abbiamo:

$$f_u(\theta, \varphi) \propto (\sin \theta)^l e^{il\varphi}$$

La costante è fissata *normalizzando*, ossia imponendo:

$$\int |f_u(\theta, \varphi)|^2 d\Omega \stackrel{!}{=} 1$$

ma per ora non ci serve calcolarla²².

Una volta ottenuto $|l, l\rangle$ possiamo ottenere un generico $|l, m\rangle$ “abbassandolo” tramite (4.50). In coordinate sferiche ciò porta a:

$$f_{l,|m|}(\theta, \varphi) \propto \left(\frac{L_-}{\hbar} \right)^{l-|m|} ((\sin(\theta))^l e^{il\varphi}) = \left[e^{-i\varphi} \left(i \cot \theta \frac{\partial}{\partial \varphi} - \frac{\partial}{\partial \theta} \right) \right]^{l-|m|} ((\sin(\theta))^l e^{il\varphi})$$

²¹ Lo si può comunque fare, per esempio tramite metodi con *serie di potenze*, come accennato qui: http://www.eng.fsu.edu/~dommelen/quantum/style_a/nt_soll2.html

²² La procedura è comunque disponibile qui <http://hitoshi.berkeley.edu/221a/sphericalharmonics.pdf>

Cerchiamo ora di riscrivere tale espressione in una forma migliore. Sostituendo:

$$L_3 = \frac{\hbar}{i} \frac{\partial}{\partial \varphi} \Rightarrow i \frac{\partial}{\partial \varphi} = -\frac{L_3}{\hbar}$$

calcoliamo la prima *applicazione* di L_i/\hbar all'autostato:

$$\begin{aligned} f_{l,l-1}(\theta, \varphi) &\propto -e^{-i\varphi} \left(\frac{L_3}{\hbar} + \frac{\partial}{\partial \theta} \right) (\sin \theta)^l e^{im\varphi} \stackrel{(4.47)}{=} -e^{i\varphi(l-1)} \left(\frac{\hbar l}{\hbar} \cot \theta + \frac{\partial}{\partial \theta} \right) (\sin \theta)^l = \\ &= -e^{i\varphi(l-1)} \left(l \cot \theta + \frac{\partial}{\partial \theta} \right) (\sin \theta)^l \end{aligned}$$

dove l'autostato $e^{il\varphi}$ viene *abbassato* dall' $e^{-i\varphi}$ iniziale. Se riapplichiamo L_-/\hbar , facendolo agire solo sull'esponentiale iniziale e lasciando il resto così com'è, otteniamo:

$$\begin{aligned} f_{l,l-2}(\theta, \varphi) &\propto +e^{-i\varphi} \left(\frac{L_3}{\hbar} + \frac{\partial}{\partial \theta} \right) e^{i\varphi(l-1)} \left(l \cot \theta + \frac{\partial}{\partial \theta} \right) (\sin \theta)^l = \\ &= e^{i\varphi(l-2)} \left((l-1) \cot \theta + \frac{\partial}{\partial \theta} \right) \left(l \cot \theta + \frac{\partial}{\partial \theta} \right) (\sin \theta)^l \end{aligned}$$

Ignorando l'alternanza del segno (dato che, in ogni caso, gli autostati sono definiti *a meno di una fase*), possiamo generalizzare il *pattern* appena trovato:

$$\begin{aligned} f_{l,|m|}(\theta, \varphi) &\propto e^{i\varphi|m|} \left(\frac{\partial}{\partial \theta} + (|m|+1) \cot \theta \right) \left(\frac{\partial}{\partial \theta} + (|m|+2) \cot \theta \right) \cdots \quad (4.51) \\ &\cdots \left(\frac{\partial}{\partial \theta} + (l-1) \cot \theta \right) \left(\frac{\partial}{\partial \theta} + l \cot \theta \right) (\sin \theta)^l \end{aligned}$$

L'ultimo trucco sta nell'usare un'*identità furba* per semplificare maggiormente l'espressione:

$$\frac{\partial}{\partial \theta} + k \cot \theta = \frac{1}{\sin^k \theta} \frac{\partial}{\partial \theta} \sin^k \theta$$

che si verifica immediatamente applicando il secondo membro ad una generica $f(\theta)$:

$$\begin{aligned} \frac{1}{\sin^k \theta} \frac{d}{d\theta} (\sin^k \theta f(\theta)) &= \frac{1}{\sin^k \theta} \left(k \sin^{k-1}(\theta) \cos(\theta) f(\theta) + \sin^k(\theta) \frac{\partial}{\partial \theta} \right) = \\ &= \left[k \cot \theta + \frac{\partial}{\partial \theta} \right] f(\theta) \end{aligned}$$

Sostituendola in (4.51) otteniamo:

$$\begin{aligned}
 f_{l,|m|}(\theta, \varphi) &\propto e^{i|m|\varphi} \left(\frac{1}{\sin^{|m|+1}(\theta)} \frac{\partial}{\partial \theta} \sin^{|m|+1}(\theta) \right) \left(\frac{1}{\sin^{|m|+2}(\theta)} \frac{\partial}{\partial \theta} \sin^{|m|+2}(\theta) \right) \dots \\
 &\dots \left(\frac{1}{\sin^{l-1}(\theta)} \frac{\partial}{\partial \theta} \sin^{l-1}(\theta) \right) \left(\frac{1}{\sin^l(\theta)} \frac{\partial}{\partial \theta} \sin^l(\theta) \right) \sin^l(\theta) = \\
 &= e^{i|m|\varphi} \sin^{-|m|}(\theta) \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \right) \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \right) \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \right) \dots \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \right) \sin^{2l}(\theta) = \\
 &= e^{i|m|\varphi} \sin^{-|m|}(\theta) \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \right)^{l-|m|} \sin^{2l}(\theta) = \\
 &= e^{i|m|\varphi} \frac{1}{\sin^{|m|}(\theta)} \left(\frac{d}{d \cos \theta} \right)^{l-|m|} \sin^{2l}(\theta)
 \end{aligned}$$

Le $f_{lm}(\theta, \varphi)$ *normalizzate* sono dette **armoniche sferiche**, e assumono la seguente forma:

3. Armoniche sferiche $Y_l^m(\theta, \varphi)$

$$Y_l^m(\theta, \varphi) \equiv f_{lm}(\theta, \varphi) = \sqrt{\frac{(2l+1)}{4\pi}} \sqrt{\frac{(l+|m|)!}{(l-|m|)!}} \frac{1}{2^l l!} \frac{1}{(\sin \theta)^{|m|}} \frac{\partial^{l-|m|}}{\partial \cos \theta^{l-|m|}} (\sin \theta)^{2l} e^{im\varphi}$$

Spesso vengono riscritte in termini dei *polinomi di Legendre* $P_l^m(x)$:

$$\begin{aligned}
 Y_l^m &= \sqrt{\frac{2l+1}{4\pi}} \sqrt{\frac{(l+m)!}{(l-m)!}} e^{im\varphi} P_l^m(\cos \theta) \\
 P_l^m &= \frac{(-1)^m}{2^l l!} (1-x^2)^{\frac{m}{2}} \frac{d^{l+m}}{dx^{l+m}} (x^2-1)^l
 \end{aligned}$$

Calcoliamo le prime:

$$\begin{aligned}
 Y_0^0(\theta, \varphi) &= \frac{1}{\sqrt{4\pi}} \quad \int |Y_0^0|^2 d\Omega = 1 \\
 Y_1^0(\theta, \varphi) &= \sqrt{\frac{3}{4\pi}} \cos \theta \\
 Y_1^{\pm 1}(\theta, \varphi) &= \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\varphi}
 \end{aligned}$$

(Intuitivamente, se applichiamo un'evoluzione temporale, la prima è costante, la seconda è data da una *rotazione* di θ , la terza da una variazione sia di θ che di φ - che corrisponde in senso semiclassico “alla Bohr” ad un'orbita obliqua).

Esaminiamo cosa succede se trasformiamo le armoniche sferiche tramite la parità, ossia se effettuiamo le sostituzioni che “cambiano di segno”:

Parità delle armoniche sferiche

$$\begin{aligned}
 \vec{x} &\rightarrow -\vec{x} \\
 \theta &\rightarrow \pi - \theta \\
 \varphi &\rightarrow \varphi + \pi
 \end{aligned}$$

otteniamo:

$$\begin{aligned} Y_l^m(\pi - \theta, \varphi + \pi) &= \frac{1}{(-1)^{l-|m|}} (-1)^{2l} (-1)^m Y_l^m(\theta, \varphi) = \\ &= (-1)^l Y_l^m(\theta, \varphi) \end{aligned}$$

Perciò per l pari l'armonica è una funzione pari, e per l dispari è dispari.

Si può dimostrare che le funzioni $\{Y_l^m = \langle \theta, \varphi | l, m \rangle, l \in \mathbb{N}, m \in \mathbb{Z} \cap [-l, l]\}$ costituiscono una **base ortonormale** in $L^2(S^2, d\Omega)$ e pertanto:

*Armoniche
sferiche come
base ON*

- Esiste un autoket $|l, l\rangle \in L^2$ comune a \vec{L}^2 e L_3 , da cui, tramite L_- , si possono costruire *tutti* gli autovalori di \vec{L}^2 e L_3 , i cui rispettivi autoket appartengono *sempre* a L^2 .

Si ha allora che lo spettro di \vec{L}^2 è *puramente discreto*, ed è dato da:

*\vec{L}^2 è a spettro
discreto*

$$\sigma(\vec{L}^2) = \sigma_P(\vec{L}^2) = \{\hbar^2 l(l+1), l \in \mathbb{N}\}$$

Notiamo in particolare che *non vale* la **congettura di Bohr**, che suppose che gli autovalori di \vec{L}^2 assumessero la forma $\hbar^2 l^2$.

Se denotiamo con \mathcal{H}_l il sottospazio di $L^2(S^2, d\Omega)$ corrispondente all'autovalore $\hbar^2 l(l+1)$ di \vec{L}^2 possiamo trovare lo spettro di L_3 , anch'esso *puramente discreto*:

$$\sigma\left(L_3 \Big|_{\mathcal{H}_l}\right) = \sigma_P\left(L_3 \Big|_{\mathcal{H}_l}\right) = \hbar([-l, l] \cap \mathbb{Z})$$

- Essendo gli autovettori comuni $|l, m\rangle$ non degeneri in $L^2(S^2, d\Omega)$, per il teorema sulle osservabili compatibili, $\{\vec{L}^2, L_3\}$ costituisce un **ICOC** in $L^2(S^2, d\Omega)$
- Poiché $\mathcal{H}_l \perp \mathcal{H}_{l'}$ per $l \neq l'$ (in quanto autospazi associati ad autovalori differenti di uno spettro discreto), possiamo decomporre:

$$L^2(S^2, d\Omega) \cong \bigoplus_{l \in \mathbb{N}} \mathcal{H}_l \quad \dim(\mathcal{H}_l) = 2l + 1$$

- Due componenti diverse del momento angolare L_i, L_j con $i \neq j$ **non** sono **compatibili**, dato che $[L_i, L_j] = i\hbar \epsilon_{ijk} L_k$.

Applicando allora il principio di indeterminazione (con $D = \mathcal{S}(\mathbb{R}^3)$) vale:

$$(\Delta L_i)_\psi (\Delta L_j)_\psi \geq \frac{\hbar}{2} |\epsilon_{ijk} \langle L_k \rangle_\psi|$$

Più in generale, se \hat{n}, \hat{m} sono versori in \mathbb{R}^3 :

$$(\Delta(\vec{L} \cdot \hat{n}))_\psi (\Delta(\vec{L} \cdot \hat{m}))_\psi \geq \frac{\hbar}{2} |\langle \hat{n} \times \hat{m} \cdot \vec{L} \rangle_\psi|$$

Quindi i momenti angolare \vec{L} lungo le direzioni \hat{n} e $\hat{m} \neq \hat{n}$ si possono misurare simultaneamente con precisione arbitraria solo sullo stato $l = 0$, dove infatti vale anche $m = 0$, e quindi:

$$\vec{L}|0, 0\rangle = 0 \quad L_+, L_-, L_3|0, 0\rangle = 0$$

La non misurabilità simultanea del momento angolare in direzioni diverse spiega come sia possibile che \vec{L} sia quantizzato *lungo ogni direzione*. Ciò è incomprensibile nella teoria *semi-classica* di Bohr: infatti, se intendiamo \vec{L} come un vettore in senso classico, che ha una “direzione ben definita ad ogni istante”, è impossibile che ogni sua proiezione lungo un qualsiasi asse sia un multiplo intero di una certa costante \hbar .

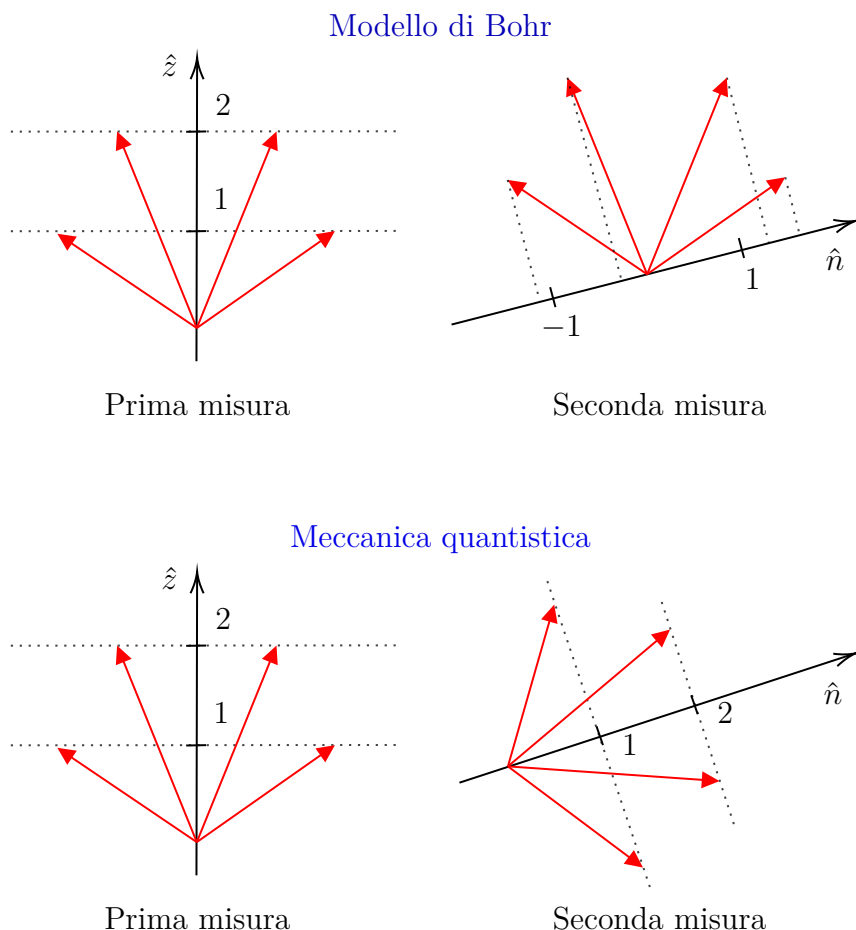


Figura 4.6: Confronto tra modello semiclassico di Bohr e teoria quantistica. Nel primo caso si introduce “ad hoc” la quantizzazione del momento angolare, ma misure lungo assi diversi sono ancora considerate compatibili, e ciò crea problemi perché è impossibile che lo *stesso vettore* abbia proiezioni *quantizzate* lungo un qualsiasi asse (a meno che non sia il vettore nullo). La situazione è risolta in MQ, dove misure di componenti diverse di \vec{L} sono incompatibili, e fare una *modifica* il risultato dell’altra.

Ipotizziamo infatti di misurare $\vec{L} \cdot \hat{z}$: otterremo allora un multiplo di \hbar , che per esempio sarà scelto tra $-2, -1, 0, +1, +2$. Se ora misuriamo lungo \hat{n} generico, se il momento ha la stessa direzione che abbiamo determinato prima, chiaramente possiamo scegliere \hat{n} in modo che il risultato *non possa essere un multiplo intero di \hbar* , ma ciò è in contrasto con gli esperimenti, dove comunque si osserva un valore multiplo intero di \hbar .

Classicamente non ha senso, e neanche nell'idea di Bohr, dove le “orbite” sono ancora “reali”, nel senso che “esistono” e non sono *collassate*, ossia determinate, dall'atto di misura.

In MQ, invece, fare la nuova misura lungo \vec{n} è come se “ridefinisse” i vettori del momento in modo che le loro proiezioni lungo l'asse scelto siano multipli interi di \hbar . Ma questo si può fare perché *non* si può mai misurare simultaneamente lungo due assi, dato che una misura irrimediabilmente modifica lo stato del sistema.

4.8.2 Lo spin

Nelle sezioni precedenti, abbiamo osservato che in MQ \vec{L} non è il più generale momento angolare, che invece è dato da \vec{J} , per cui valgono solo $[J_i, J_j] = i\hbar\epsilon_{ijk}J_k$, senza alcuna condizione sulle rotazioni di 2π che devono essere per forza equivalenti alla rotazione nulla.

Abbiamo però determinato che gli autovettori comuni di \vec{L}^2 e L_3 dati da $\{|l, m\rangle, l \in \mathbb{N}, m \in \mathbb{Z} \cap [-l, l]\}$ costituiscono una **base** per la parte angolare di $L^2(\mathbb{R}^3, d^3x)$ su cui agiscono le rotazioni generate da \vec{L} .

Ma allora se $\vec{J} \neq \vec{L}$, si ha che \vec{J} non può agire più solo sullo spazio di Hilbert $L^2(\mathbb{R}^3, d^3x)$, ma su uno spazio più grande, che sarà del tipo $L^2(\mathbb{R}^3, d^3x) \otimes \mathcal{H}_s$.

Vediamo il perché. Supponiamo che \vec{L} e \vec{J} agiscano entrambi su $L^2(\mathbb{R}^3, d^3x)$. Le relazioni di commutazione per \vec{L} sono date da:

$$[L_i, X_l] = i\hbar\epsilon_{ilk}X_k; \quad [L_i, P_l] = i\hbar\epsilon_{ilk}P_k$$

E relazioni analoghe valgono anche per \vec{J} , dato che è anch'esso generatore delle rotazioni:

$$[J_i, X_l] = i\hbar\epsilon_{ilk}X_k; \quad [J_i, P_l] = i\hbar\epsilon_{ilk}P_k$$

Sottraendo i corrispettivi commutatori a membro giungiamo a:

$$[J_i - L_i, \vec{X}] = 0 \quad [J_i - L_i, \vec{P}] = 0$$

In $L^2(\mathbb{R}^3)$ sappiamo però che $\{\vec{X}, \vec{P}\}$ è un sistema *irriducibile* di osservabili, e ogni osservabile che commuta con tutti gli elementi di tale sistema deve assumere autovalori che non dipendono in alcun modo dallo stato in cui è calcolata, e quindi è un multiplo dell'identità. Vale quindi:

$$J_i - L_i = c_i \mathbb{I} \Rightarrow J_i = L_i + c_i \mathbb{I}$$

Ma allora otteniamo:

$$[J_i, J_j] = [L_i + c_i \mathbb{I}, L_j + c_j \mathbb{I}] = i\hbar\epsilon_{ijk}L_k \neq i\hbar\epsilon_{ijk}J_k$$

a meno che le $c_j = 0$, ma allora si avrebbe $\vec{J} = \vec{L}$, che è il contrario dell'ipotesi da cui siamo partiti.

\vec{J} agisce su uno spazio di Hilbert “più grande”

Deve esistere allora una nuova osservabile \vec{S} che agisce su \mathcal{H}_s come momento angolare (ossia come *generatore* delle rotazioni) in modo che:

$$\vec{J} = \vec{L} \otimes \mathbb{I} + \mathbb{I} \otimes \vec{S} \rightsquigarrow \vec{J} = \vec{L} + \vec{S}$$

Tale \vec{S} è detto **momento angolare di spin** o semplicemente spin, e valgono le analoghe proprietà del momento angolare:

$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k$$

$$\sigma(\vec{S}^2; S_3) = \{\hbar^2 s(s+1), s \in \mathbb{N}/2; \hbar(s, s-1, \dots, -s)\}$$

In particolare s non è più vincolato a stare in \mathbb{N} , poiché non c'è più il vincolo fisico che una rotazione di 2π in \mathcal{H}_s che non è il “comune spazio classico” $L^2(\mathbb{R}^3)$ sia banale, ossia equivalente alla rotazione nulla.

Notiamo che poiché \vec{S} agisce su uno spazio \mathcal{H}_s che è *distinto* da L^2 , commuta automaticamente con qualsiasi operatore \vec{A} che agisce *solo* su L^2 , dato che:

$$[\vec{A} \otimes \mathbb{I}, \mathbb{I} \otimes \vec{S}] = 0$$

In particolare, \vec{S} commuta con \vec{X} , \vec{P} e \vec{L} .

Inoltre $\{\vec{X}, \vec{P}, \vec{S}\}$ formano un *insieme irriducibile*, stavolta per tutte le particelle elementari quantistiche (mentre $\{\vec{X}, \vec{P}\}$ sono un sistema irriducibile solo per le particelle quantistiche “con analogo classico”, ossia senza spin).

Notiamo ora che \vec{S}^2 commuta con \vec{X} , \vec{P} , \vec{S} , ossia con tutti gli elementi del sistema irriducibile. Perciò deve essere un multiplo dell'identità, e in particolare deve assumere un valore specifico, indipendente dallo stato.

Ogni particella, perciò, è caratterizzata da un *autovalore fissato* di \vec{S}^2 , della forma $\hbar^2 s(s+1)$, con $s \in \mathbb{N}/2$.

Notiamo ora che \mathcal{H}_s è generato dagli autoket dello spin, e dato che le singole componenti di \vec{S} non commutano tra di loro, possiamo prendere una di esse, come S_3 , come ICOC in \mathcal{H}_s .

Per quanto ricavato discutendo il momento angolare, lo spettro di S_3 è discreto, con $2s+1$ possibili autovalori (che vanno a incrementi di \hbar da $-\hbar s$ a $+\hbar s$). Per ciascuno di essi avremo un autoket $|s_z\rangle$, e tali autoket costituiscono una base di \mathcal{H}_s . Si ha perciò che $\dim \mathcal{H}_s = 2s+1$ e quindi $\mathcal{H}_s \cong \mathbb{C}^{2s+1}$.

Perciò, lo spazio “totale” per gli stati di una singola particella *dotata di spin* è dato dal prodotto tensore tra lo spazio “fisico reale” L^2 e lo “spazio di spin” \mathcal{H}_s . Usando l'isomorfismo introdotto in (4.21) possiamo allora scrivere:

$$\mathcal{H} = L^2(\mathbb{R}^3, d^3x) \otimes \mathcal{H}_s \cong L^2(\mathbb{R}^3, d^3x) \otimes \mathbb{C}^{2s+1} \cong \underbrace{L^2(\mathbb{R}^3, d^3x) \oplus \dots \oplus L^2(\mathbb{R}^3, d^3x)}_{2s+1 \text{ volte}}$$

(Lezione 27 del
28/11/2018)

$\{\vec{X}, \vec{P}, \vec{S}\}$ come
ICOC “generale”

\vec{S}^2 ha un valore
fissato

In altre parole, un vettore di \mathcal{H} in rappresentazione $\{\vec{X}, S_3\}$ si può vedere come un vettore colonna a $2s + 1$ componenti:

$$\begin{pmatrix} \psi_1(\vec{x}) \\ \psi_2(\vec{x}) \\ \vdots \\ \psi_{2s+1}(\vec{x}) \end{pmatrix} \in L^2(\mathbb{R}^3) \oplus \cdots \oplus L^2(\mathbb{R}^3)$$

In natura, gran parte delle particelle hanno spin non nullo, e non sono quindi descrivibili in uno spazio *puramente* “classico”. In effetti, notiamo che tutte le particelle “di materia” - come elettroni, protoni e neutroni - hanno spin $s = \frac{1}{2}$.

Riepilogando, i gradi di libertà aggiuntivi dati dallo spin derivano dal fatto che le *simmetrie* offerte dall’isotropia dello spazio agiscono su raggi vettori in \mathcal{PH} , e non direttamente sui vettori di \mathcal{H} , dato che non possono “percepire” la “fase assoluta” di uno stato, ma solo una differenza tra fasi: questo è un fatto verificato sperimentalmente.

Per teorema di Bargmann possiamo realizzare tali simmetrie come rappresentazioni unitarie che agiscono su \mathcal{H} , al costo di passare *sul gruppo di ricoprimento* delle rotazioni, che non è $SO(3)$ stesso, ma $SU(2)$. Localmente, i due sono indistinguibili (e in effetti generano la stessa algebra di Lie), ma globalmente no, per cui ad esempio rotazioni equivalenti in $SO(3)$ che differiscono di 2π “sono proiettate” in punti (ossia rotazioni) distinti in $SU(2)$, dove è necessaria un’ulteriore differenza di 2π (per un totale di 4π) per ritornare effettivamente al punto di partenza. Ciò offre agli stati di \mathcal{H} , per cui usiamo quest’ultima rappresentazione, una maggiore “libertà”, che si traduce proprio nell’esistenza dello spin - da cui, come vedremo, derivano le proprietà di tutta la *materia* come la conosciamo!

Concentriamoci ora sul caso $s = \frac{1}{2}$. Sappiamo che:

$$\mathcal{H}_{s=\frac{1}{2}} \cong \mathbb{C}^2 \qquad \sigma(S_3) = \left\{ -\frac{\hbar}{2}, +\frac{\hbar}{2} \right\}$$

Perciò le singole componenti S_i di \vec{S} sono matrici hermitiane 2×2 . Potremmo allora pensare a \vec{S} come ad un “vettore di matrici” (o meglio, un tensore).

Cerchiamo ora una base di $\mathcal{H}_{s=\frac{1}{2}}$, che ci aspettiamo essere formata da tre componenti, dato che una matrice 2×2 hermitiana ha 3 “gradi di libertà”: 2 sulla diagonale e un altro qualsiasi (di cui il quarto elemento sarà il complesso coniugato).

Sia $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z) = (\sigma_1, \sigma_2, \sigma_3)$ definito da:

$$\vec{S} = \frac{\hbar}{2} \vec{\sigma}$$

In tal modo le matrici (hermitiane) σ_i sono 2×2 con entrate adimensionali e, come vedremo, di modulo 1. Dalle relazioni di commutazione delle componenti di \vec{S} si

ricavano relazioni per le σ_i :

$$\begin{aligned} [S_i, S_j] &= i\hbar\epsilon_{ijk}S_k \Rightarrow \left[\frac{\hbar}{2}\sigma_i, \frac{\hbar}{2}\sigma_j \right] = i\frac{\hbar^2}{2}\epsilon_{ijk}\sigma_k \\ &\Rightarrow [\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \end{aligned} \quad (4.52)$$

Per semplicità e convenzione, scegliamo S_3 come diagonale. Dato che gli autovalori di S_3 sono $\pm\hbar/2$, e abbiamo estratto $\hbar/2$ dalla definizione delle σ_i , si ha che σ_3 avrà ± 1 sulla diagonale:

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Notiamo che $\sigma_3^2 = \mathbb{I}$. Dato che la scelta di partire proprio da σ_3 (associata a \hat{z}) è puramente convenzionale, dato che per l'isotropia dello spazio non esiste una direzione preferenziale, si ha che anche $\sigma_1^2 = \sigma_2^2 = \mathbb{I}$. Matematicamente, se P è la rotazione degli assi che scambia \hat{z} con \hat{x} (o, equivalentemente, con \hat{y}), per isotropia abbiamo che, detto $\sigma' = \vec{\sigma} \cdot \hat{n}$, con \hat{n} versore generico:

$$(\sigma')^2 = (P^\dagger \sigma_z P)^2 = P^{-1} \sigma_z P P^{-1} \sigma_z P = P^\dagger \sigma_z^2 P = P^\dagger \mathbb{I} P = \mathbb{I}$$

Dato che $\sigma_y^2 = \mathbb{I}$, allora commuta con ogni componente:

$$\begin{aligned} 0 &= [\sigma_y^2, \sigma_z] = \sigma_y^2 \sigma_z - \sigma_z \sigma_y^2 = \sigma_y (\sigma_y \sigma_z - \sigma_z \sigma_y) + (\sigma_y \sigma_z - \sigma_z \sigma_y) \sigma_y = \\ &= \sigma_y [\sigma_y, \sigma_z] - [\sigma_z, \sigma_y] \sigma_y = \sigma_y [\sigma_y, \sigma_z] + [\sigma_y, \sigma_z] \sigma_y = \\ &\stackrel{(4.52)}{=} \sigma_y 2i\sigma_x + 2i\sigma_x \sigma_y = 2i(\sigma_y \sigma_x + \sigma_x \sigma_y) \end{aligned}$$

Relazioni analoghe valgono anche per le altre coppie di σ_i, σ_j . Possiamo riscriverle in forma compatta tramite la nozione di **anticommutatore**:

Relazioni di anti-commutazione

$$\{\sigma_i, \sigma_j\} \equiv \sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij} \quad (4.53)$$

In altre parole, σ_i con diversi indici *anticommutano* tra loro, cioè:

$$\{\sigma_i, \sigma_j\} = 0 \Rightarrow \sigma_i \sigma_j = -\sigma_j \sigma_i \quad \forall i \neq j$$

Le σ_i sono matrici hermitiane (dato che \vec{S} è un'osservabile), ossia $\sigma_i = \sigma_i^\dagger$. Hanno quindi la forma:

$$\begin{pmatrix} a & b \\ b^* & c \end{pmatrix} \quad a, c \in \mathbb{R}, \quad b \in \mathbb{C}$$

Imponiamo la condizione di anticommutazione con $\sigma_3 = \sigma_z$ che abbiamo già fissato:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ b^* & c \end{pmatrix} = \begin{pmatrix} a & b \\ -b^* & -c \end{pmatrix} = - \begin{pmatrix} a & b \\ b^* & c \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -a & b \\ -b^* & c \end{pmatrix} \Rightarrow c = 0 = a$$

Perciò σ_x e σ_y hanno la forma:

$$\sigma_x, \sigma_y = \begin{pmatrix} 0 & b \\ b^* & 0 \end{pmatrix}$$

Per σ_x consideriamo la scelta più semplice, che sia compatibile con $\sigma_x^2 = \mathbb{I}$:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Dalla relazione di anticommutazione $\sigma_x \sigma_y = -\sigma_y \sigma_x$, otteniamo che per σ_y deve essere $b = -b^*$, ossia b deve essere puramente immaginario. Da $\sigma_y^2 = \mathbb{I}$ avremo poi che $b = i$.

Riepilogando, otteniamo la seguente scelta per le matrici $\sigma_x, \sigma_y, \sigma_z$:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Tali matrici sono dette **matrici di Pauli**²³, e costituiscono la versione “adimensionalizzata” delle osservabili associate alle componenti cartesiane dello spin \vec{S} , dato che siamo partiti da $\vec{S} = \frac{\hbar}{2} \vec{\sigma}$.

Per trovare invece la componente di \vec{S} lungo un versore generico \hat{n} di \mathbb{R}^3 , data da $\vec{S} \cdot \hat{n}$, passiamo in coordinate sferiche, indicando $\hat{n} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$ - dove, come al solito, $\theta \in [0, \pi]$ è l'angolo rispetto alla verticale \hat{z} .

$$\begin{aligned} \vec{S} \cdot \vec{n} &= \frac{\hbar}{2} \vec{\sigma} \cdot \hat{n} = \frac{\hbar}{2} \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \sin \theta \cos \varphi + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \sin \theta \sin \varphi + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cos \theta \right] = \\ &= \frac{\hbar}{2} \begin{pmatrix} \cos \theta & \sin \theta (\cos \varphi - i \sin \varphi) \\ \sin \theta (\cos \varphi + i \sin \varphi) & -\cos \theta \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \cos \theta & \sin \theta e^{-i\varphi} \\ \sin \theta e^{i\varphi} & -\cos \theta \end{pmatrix} \end{aligned}$$

Sappiamo già che gli autovalori di tale matrice saranno $\pm \hbar/2$. Cerchiamone i relativi autovettori:

$$\left[\vec{S} \cdot \hat{n} - \left(\pm \frac{\hbar}{2} \right) \right] \begin{pmatrix} a \\ b \end{pmatrix} \stackrel{!}{=} 0 \Rightarrow \begin{pmatrix} \cos \theta \mp 1 & \sin \theta e^{i\varphi} \\ \sin \theta e^{i\varphi} & -\cos \theta \mp 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0$$

Da cui si ottengono due equazioni linearmente dipendenti, per cui basta scrivere la prima:

$$a(\cos \theta \mp 1) + b \sin \theta e^{-i\varphi} = 0$$

Riscrivendo ogni termine, tramite identità goniometriche, in funzione di $\theta/2$:

$$\begin{aligned} a \left(\cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2} \mp \left(\cos^2 \frac{\theta}{2} + \sin^2 \frac{\theta}{2} \right) \right) + b 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{-i\varphi} &= 0 \\ +\frac{\hbar}{2} : \quad \frac{b}{a} &= \frac{\sin \theta/2}{\cos \theta/2} e^{i\varphi} \quad -\frac{\hbar}{2} : \quad \frac{b}{a} = -\frac{\cos \theta/2}{\sin \theta/2} e^{i\varphi} \end{aligned}$$

²³ Si può anche dimostrare che $i\sigma_x, i\sigma_y, i\sigma_z$ formano una base per l'algebra di Lie $\mathfrak{su}(2)$, il che ha intuitivamente senso, dato che lo spin nasce proprio dalla discussione dei generatori delle rotazioni in \mathbb{R}^3 .

Scegliamo ora $a = \cos \frac{\theta}{2} e^{-i\varphi/2}$ per il primo caso, e $a = \sin \frac{\theta}{2} e^{-i\varphi/2}$ per il secondo. Otteniamo allora, per gli autostati dello spin in una generica direzione \hat{n} inclinata di θ rispetto a \hat{z} e *ruotata* di φ sul piano $\perp \hat{z}$ (che per $\psi > 0$ è antioraria se vista “dalla punta” di $+\hat{z}$):

$$\begin{aligned} \left| s_n = \frac{\hbar}{2} \right\rangle &\equiv \left| +\frac{1}{2} \right\rangle_n = \begin{pmatrix} \cos \frac{\theta}{2} e^{-i\varphi/2} \\ \sin \frac{\theta}{2} e^{i\varphi/2} \end{pmatrix} \\ \left| s_n = -\frac{\hbar}{2} \right\rangle &\equiv \left| -\frac{1}{2} \right\rangle_n = \begin{pmatrix} \sin \frac{\theta}{2} e^{-i\varphi/2} \\ \cos \frac{\theta}{2} e^{i\varphi/2} \end{pmatrix} \end{aligned} \quad (4.54)$$

La scelta di a così effettuata è tale che per $\theta = 0$, ossia considerando come direzione $\hat{n} = +\hat{z}$, gli autostati sono scritti come:

$$\left| +\frac{1}{2} \right\rangle \equiv |+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad \left| -\frac{1}{2} \right\rangle \equiv |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Nota: ciò vale $\forall \varphi$. Infatti, in questo caso, per $\varphi \neq 0$ si otterrebbe, per esempio, $e^{-i\varphi} |+\rangle$, che è equivalente a $|+\rangle$ una volta normalizzato. Ciò è ragionevole, dato che una rotazione attorno a \hat{z} lascia invariati i vettori diretti come \hat{z} .

Tale osservazione ci permette, perciò, di scrivere gli autostati dello spin misurato in una direzione generica (che denoteremo con $|+\rangle_n$ e $|-\rangle_n$) come combinazione lineare degli autoket dello spin misurato lungo \hat{z} (denotati come $|+\rangle$ e $|-\rangle$, senza alcun pedice). Esplicitando le (4.54):

$$\begin{aligned} |+\rangle_n &= \cos \frac{\theta}{2} e^{-i\varphi/2} |+\rangle + \sin \frac{\theta}{2} e^{+i\varphi/2} |-\rangle \\ |-\rangle_n &= -\sin \frac{\theta}{2} e^{-i\varphi/2} |+\rangle + \cos \frac{\theta}{2} e^{i\varphi/2} |-\rangle \end{aligned} \quad (4.55)$$

Per esempio, per $\theta = \pi/2$ e $\varphi = 0$, ossia per $\hat{n} = +\hat{x}$, otteniamo:

$$|+\rangle_x = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle); \quad |-\rangle_x = \frac{1}{\sqrt{2}}(|-\rangle - |+\rangle)$$

Tali formule sono utili nel caso di più misure di spin lungo assi diversi. Per esempio, se un sistema è inizialmente nello stato $|+\rangle_x$, dalla relazione appena vista avremo che una misura di spin lungo \hat{z} darà con probabilità uguali ($p = 1/2$) uno dei due risultati $\pm\hbar/2$ possibili.

4.8.3 Composizione dei momenti angolari

Ci poniamo ora il problema di *comporre* più momenti angolari²⁴, al fine di calcolare il *momento angolare totale* di un sistema.

Conti del genere compaiono, per esempio, quando si considerano sia il momento angolare orbitale \vec{L} di un elettrone, che quello *intrinseco* di spin \vec{S} , e si cerca di

²⁴ΛFonti a [6]

determinare le loro relazioni con $\vec{J} = \vec{L} + \vec{S}$.

Occupiamoci fin da subito del caso più generale. Consideriamo due operatori di momento angolare $\vec{J}^{(1)}$ e $\vec{J}^{(2)}$ che agiscono rispettivamente su spazi *distinti* \mathcal{H}_{j_1} e \mathcal{H}_{j_2} di un sistema composto dato da $\mathcal{H} = \mathcal{H}_{j_1} \otimes \mathcal{H}_{j_2}$, ossia corrispondono a “gradi di libertà distinti”:

$$\vec{J}^{(1)} = \vec{J}^{(1)} \otimes \mathbb{I}^{(2)} \quad \vec{J}^{(2)} = \mathbb{I}^{(1)} \otimes \vec{J}^{(2)}$$

Poiché $\vec{J}^{(1)}$ e $\vec{J}^{(2)}$ commutano (dato che agiscono su *spazi distinti*), possiamo realizzare un ICOC per il sistema unendo i due ICOC dei sistemi separati, ossia considerando $\{(\vec{J}^{(1)})^2, \vec{J}_3^{(1)}; (\vec{J}^{(2)})^2, \vec{J}_3^{(2)}\}$. Indichiamo un autoket di questo ICOC con $|j_1, j_2; m_1, m_2\rangle$, e tale autoket verifica le seguenti equazioni agli autovalori:

1. ICOC
 $\{(\vec{J}^{(1)})^2, (\vec{J}^{(2)})^2, \vec{J}_3^{(1)}, \vec{J}_3^{(2)}\}$

$$(\vec{J}^{(1)})^2 |j_1, j_2; m_1, m_2\rangle = j_1(j_1 + 1)\hbar^2 |j_1, j_2; m_1, m_2\rangle$$

$$(\vec{J}^{(2)})^2 |j_1, j_2; m_1, m_2\rangle = j_2(j_2 + 1)\hbar^2 |j_1, j_2; m_1, m_2\rangle$$

$$\vec{J}_3^{(1)} |j_1, j_2; m_1, m_2\rangle = m_1\hbar |j_1, j_2; m_1, m_2\rangle$$

$$\vec{J}_3^{(2)} |j_1, j_2; m_1, m_2\rangle = m_2\hbar |j_1, j_2; m_1, m_2\rangle$$

dove $m_1, m_2 \in \mathbb{Z}$ con $|m_1| \leq j_1$ e $|m_2| \leq j_2$.

Consideriamo ora il momento angolare totale \vec{J} in \mathcal{H} . Si ha:

$$\vec{J} = \vec{J}^{(1)} \otimes \mathbb{I}^{(2)} + \mathbb{I}^{(1)} \otimes \vec{J}^{(2)} \equiv \vec{J}^{(1)} + \vec{J}^{(2)}$$

Notiamo che:

$$\vec{J}^2 = (\vec{J}^{(1)})^2 + (\vec{J}^{(2)})^2 + 2\vec{J}^{(1)} \cdot \vec{J}^{(2)}$$

Poiché $(\vec{J}^{(1)})^2$ commuta con tutti i termini che compaiono in \vec{J}^2 , e la stessa cosa succede per $(\vec{J}^{(2)})^2$, si ha che:

$$[(\vec{J}^{(1)})^2, \vec{J}^2] = 0 \quad [(\vec{J}^{(2)})^2, \vec{J}^2] = 0$$

ossia possiamo misurare contemporaneamente il modulo del momento di ciascun sottosistema e del sistema totale - che è individuato dai numeri j_1 , j_2 e j . Tuttavia $\vec{J}_3^{(1)}$ non commuta con le altre $\vec{J}_i^{(1)}$, e quindi nemmeno con \vec{J}^2 , dato che quest'ultimo contiene un prodotto scalare che coinvolge tutte le $\vec{J}_i^{(1)}$. Stessa cosa succede, simmetricamente, per $\vec{J}_3^{(2)}$. Perciò ci aspettiamo di non poter misurare contemporaneamente m_1 , m_2 e j .

Da queste considerazioni, possiamo formare un altro ICOC, dato da $\{(\vec{J}^{(1)})^2, (\vec{J}^{(2)})^2, \vec{J}^2, \vec{J}_3\}$, i cui autoket sono denotati con $|j_1, j_2; j, m\rangle$, oppure, quando è chiaro quali siano j_1 e j_2 , semplicemente con $|j, m\rangle$.

Tali autoket verificano le seguenti equazioni agli autovalori:

2. ICOC
 $\{(\vec{J}^{(1)})^2, (\vec{J}^{(2)})^2, \vec{J}^2, \vec{J}_3\}$

$$(\vec{J}^{(1)})^2 |j_1, j_2; j, m\rangle = j_1(j_1 + 1)\hbar^2 |j_1, j_2; j, m\rangle$$

$$(\vec{J}^{(2)})^2 |j_1, j_2; j, m\rangle = j_2(j_2 + 1)\hbar^2 |j_1, j_2; j, m\rangle$$

$$\vec{J}^2 |j_1, j_2; j, m\rangle = j(j + 1)\hbar^2 |j_1, j_2; j, m\rangle$$

$$\vec{J}_3 |j_1, j_2; j, m\rangle = m\hbar |j_1, j_2; j, m\rangle$$

Le due collezioni di osservabili sono *incompatibili* tra loro, ossia un autoket per una non corrisponde a un unico autoket dell'altra, ma a una certa combinazione lineare. Fisicamente, ciò significa che partendo da stati con j_1, j_2, j, m fissati, misure ripetute produrranno valori di m_1 e m_2 non unici, ma all'interno di un *range*. Vogliamo ora determinare quali siano questi *range*, e cioè trovare un modo per passare da una base all'altra.

Partiamo notando che \vec{J}_3 commuta con $\vec{J}_3^{(1)}$ e $\vec{J}_3^{(2)}$, e perciò se questi ultimi due hanno autovalori $\hbar m_1$ e $\hbar m_2$, $J_3 = J_3^{(1)} + J_3^{(2)}$ avrà autovalore $\hbar m = \hbar(m_1 + m_2)$. Poiché m_1 ed m_2 variano all'interno di range determinati da j_1 e j_2 , ci chiediamo quale sia il range di m a partire da j_1 e j_2 . Da questo, ricordando che m varia a sua volta tra j e $-j$, vogliamo trovare anche il range possibile di valori di j , ossia degli autovalori di \vec{J}^2 .

Raccogliamo in tabella tutti i possibili valori di m_1 e m_2 , con i rispettivi valori assunti da m , ordinandoli per m decrescenti:

m	m_1	m_2
$j_1 + j_2$	j_1	j_2
$j_1 + j_2 - 1$	$j_1 - 1$	j_2
	j_1	$j_2 - 1$
$j_1 + j_2 - 2$	$j_1 - 2$	j_2
	$j_1 - 1$	$j_2 - 1$
	j_1	$j_2 - 2$

Innanzitutto, il numero di possibilità per m_1 ed m_2 (ossia di righe nella tabella) è dato dal numero di autostati comuni di $J_3^{(1)}$ e $J_3^{(2)}$ nell'autospazio di $\{(\vec{J}^{(1)})^2, (\vec{J}^{(2)})^2\}$ di autovalore (j_1, j_2) , che ammonta a $(2j_1 + 1)(2j_2 + 1)$. Il range di m si trova subito esaminando i casi limite: $m_{\max} = j_1 + j_2$ ed $m_{\min} = -(j_1 + j_2)$. Determinare i valori possibili di j , invece, richiede qualche passaggio in più. Partiamo notando che, poiché j è definito come m_{\max} , il valore massimo j_{\max} è proprio $j_1 + j_2$. Per le regole del momento angolare, possiamo pensare che tale j_{\max} “generi” i $2j_{\max} + 1$ valori di m ad esso associati, che denominiamo “multipletto discendente”, e sono dati da:

$$m_{j_{\max}} = j_{\max}, j_{\max} - 1, j_{\max} - 2, \dots, -j_{\max}$$

Ciascuno di questi m sarà a sua volta “generato” da opportuni m_1 ed m_2 tali che $m = m_1 + m_2$. Perciò, nella tabella di sopra, evidenziamo in giallo le righe che sono “generate” da j_{\max} .

Notiamo, come ci si aspetta, che così facendo non abbiamo evidenziato tutte le righe: in altre parole, non basta un unico j totale per “generare” tutti i possibili valori di m_1 e m_2 .

Notiamo che il valore massimo di m rimasto “scoperto”, ossia con una riga non evidenziata, è $j_1 + j_2 - 1$. Tale m *massimo* corrisponderà perciò ad un secondo

*Multipletti
discendenti*

valore per j , da cui costruiamo, allo stesso modo di prima, un secondo multipletto, stavolta evidenziato in azzurro. Di nuovo, notiamo che $j_1 + j_2 - 2$ ha ancora una riga non evidenziata - e quindi anch'esso corrisponderà ad un (terzo) valore di j , e ad un altro multipletto.

Con questo ragionamento abbiamo associato ogni valore possibile di j a un *multipletto discendente* di valori di m , che conterrà $2j + 1$ possibilità. Se troviamo il numero N di multipletti sapremo immediatamente il range di j : sarà proprio

$$j_{\min} = j_{\max} - N + 1 = j_1 + j_2 - N + 1$$

dove il $+1$ è necessario perché dobbiamo contare anche il j_{\max} di partenza nel totale dei multipletti.

Come possiamo determinare N ? Sappiamo che $|j_1, j_2; j, m\rangle$ e $|j_1, j_2; m_1, m_2\rangle$ sono basi (diverse) dello stesso spazio di Hilbert, e perciò devono avere lo stesso numero di elementi, che è dato da $(2j_1 + 1)(2j_2 + 1)$ come visto precedentemente. Dato che ad ogni j sono associate $2j + 1$ possibilità, e j varia tra j_{\max} (che conosciamo) a j_{\min} (che dipende da N), possiamo costruire la seguente equazione:

$$\begin{aligned} (2j_1 + 1)(2j_2 + 1) &= [2j_{\max} + 1] + [2(j_{\max} - 1) + 1] + \cdots + [2j_{\min} + 1] = \\ &= [2(j_1 + j_2) + 1] + [2(j_1 + j_2 - 1) + 1] + \cdots + [2(j_1 + j_2 - (N - 1)) + 1] = \\ &= N(2(j_1 + j_2) + 1) - 2 \sum_{n=0}^{N-1} n = N(2(j_1 + j_2) + 1) - \frac{2N(N - 1)}{2} = \\ &= N(2(j_1 + j_2) + 1 - (N - 1)) = N(2j_1 + 1 + 2j_2 + 1 - N) \end{aligned}$$

L'unica soluzione si ha allora per $N = 2j_2 + 1$.

Perciò, i possibili valori di j (da cui ricaviamo gli autovalori $\hbar^2 j(j + 1)$ di \vec{J}^2), con $j_1 > j_2$, variano tra $j_{\max} = j_1 + j_2$ e $j_{\min} = j_1 + j_2 - N + 1 = j_1 - j_2$:

$$j = j_1 + j_2, j_1 + j_2 - 1, \dots, j_1 - j_2$$

Per ottenere il caso generale, per $j_1 < j_2$ o $j_1 > j_2$ basta inserire un valore assoluto:

$$j = j_1 + j_2, \dots, |j_1 - j_2|$$

Possiamo allora scomporre lo spazio \mathcal{H} come somma diretta degli autospazi degli autovalori di \vec{J}^2 , individuati da j , ciascuno di dimensione $2j + 1$:

$$\mathcal{H}_{j_1} \otimes \mathcal{H}_{j_2} = \bigoplus_{j=|j_1-j_2|}^{j_1+j_2} \mathcal{H}_j$$

Per esempio, nel caso di due particelle con spin $\frac{1}{2}$ (es. 2 elettroni):

$$\underbrace{\mathcal{H}_{\frac{1}{2}}}_{\dim 2} \otimes \underbrace{\mathcal{H}_{\frac{1}{2}}}_{\dim 2} = \underbrace{\mathcal{H}_0}_{\dim 1} \oplus \underbrace{\mathcal{H}_1}_{\dim 3}$$

Occupiamoci ora di determinare come le basi canoniche $|j_1, m_1\rangle \otimes |j_2, m_2\rangle \equiv |j_1, j_2, m_1, m_2\rangle$ di $\mathcal{H}_{j_1} \otimes \mathcal{H}_{j_2}$ e quella del “momento totale” $|j_1, j_2; j, m\rangle \equiv |j, m\rangle$ di \mathcal{H}_j si rappresentano una in termini dell'altra.

Partiamo usando la completezza di Dirac:

$$\mathbb{I}_{\mathcal{H}_{j_1} \otimes \mathcal{H}_{j_2}} = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |j_1, j_2, m_1, m_2\rangle \langle j_1, j_2, m_1, m_2|$$

Da cui deriva:

$$|j, m\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \langle j_1, j_2; m_1, m_2 | j, m\rangle |j_1, j_2; m_1, m_2\rangle \quad (4.56)$$

I termini $\langle j_1, j_2, m_1, m_2 | j, m\rangle$, con la convenzione $\langle j_1, j_2, j_1, j - j_1 | j, j\rangle \geq 0$, sono detti coefficienti di **Clebsch-Gordan** e denotati:

*Coefficienti di
Clebsch-Gordan*

$$\langle j_1, j_2, m_1 m_2 | j, m\rangle \equiv C_{m_1, m_2, m}^{j_1, j_2, j}$$

Nella discussione precedente abbiamo notato che vale sempre $m = m_1 + m_2$, e perciò:

$$C_{m_1, m_2, m}^{j_1, j_2, j} = 0 \text{ se } m \neq m_1 + m_2$$

Inoltre j può assumere valori solo in $\{|j_1 - j_2|, \dots, j_1 + j_2\}$, e perciò:

$$C_{m_1, m_2, m}^{j_1, j_2, j} = 0 \text{ se } j \notin \{|j_1 - j_2|, \dots, j_1 + j_2\}$$

Notiamo poi che, poiché gli autoket delle due basi sono normalizzati, la somma delle *proiezioni* di un certo $|j_1, j_2; j, m\rangle$ su tutti i $|j_1, j_2; m_1, m_2\rangle$ (con $|m_1| \leq j_1$ e $|m_2| \leq j_2$) deve fare 1:

$$\sum_{m_1=-j_1}^{+j_1} \sum_{m_2=-j_2}^{+j_2} |\langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m\rangle|^2 = 1$$

Ciò deriva direttamente dalla completezza:

$$\begin{aligned} 1 &= \langle j_1, j_2; j, m | j_1, j_2; j, m\rangle = \sum_{m_1=-j_1}^{+j_1} \sum_{m_2=-j_2}^{+j_2} \langle j_1, j_2; j, m | j_1, j_2; m_1, m_2\rangle \langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m\rangle = \\ &= \sum_{m_1} \sum_{m_2} |\langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m\rangle|^2 \end{aligned} \quad (4.57)$$

Vale una cosa analoga se partiamo da m_1 e m_2 fissati e sommiamo su j e m .

Ci serve ora un metodo generale per ricavare tutti i coefficienti non nulli. Un'idea è partire dal fatto che $|j_{\max}, m_{\max}\rangle = |j_1 + j_2, j_1 + j_2\rangle$ corrisponde, a meno di un fattore, solamente a $|j_1, j_2; j_1, j_2\rangle$, dato che è l'unico autoket che verifica $m = m_1 + m_2$. Fissiamo tale fattore a 1, seguendo la *convenzione di Condon-Shortley*, e scriviamo:

$$|j_1 + j_2, j_1 + j_2\rangle = |j_1, j_2; j_1, j_2\rangle$$

Possiamo ora *abbassare* entrambi i membri applicando l'operatore $J_- = J_-^{(1)} + J_-^{(2)}$:

$$\begin{aligned} J_- |j_1 + j_2, j_1 + j_2\rangle &= \sqrt{(j_1 + j_2)(j_1 + j_2 + 1) - (j_1 + j_2)(j_1 + j_2 - 1)} |j_1 + j_2, j_1 + j_2 - 1\rangle \\ (J_-^{(1)} + J_-^{(2)}) |j_1, j_2; j_1, j_2\rangle &= \sqrt{j_1(j_1 + 1) - j_1(j_1 - 1)} |j_1, j_2; j_1 - 1, j_2\rangle + \\ &+ \sqrt{j_2(j_2 + 1) - j_2(j_2 - 1)} |j_1, j_2; j_1, j_2 - 1\rangle \end{aligned} \quad (4.58)$$

Uguagliando i termini evidenziati, abbiamo trovato un'espressione che relazione $|j_1 + j_2, j_1 + j_2 - 1\rangle$ a $|j_1, j_2; j_1 - 1, j_2\rangle$ e $|j_1, j_2; j_1, j_2 - 1\rangle$:

$$\sqrt{2(j_1 + j_2)} |j_1 + j_2, j_1 + j_2 - 1\rangle = \sqrt{2j_1} |j_1, j_2; j_1 - 1, j_2\rangle + \sqrt{2j_2} |j_1, j_2; j_1, j_2 - 1\rangle \quad (4.59)$$

Possiamo reiterare la procedura per abbassare ulteriormente l'autovalore m , ottenendo $|j_1 + j_2, j_1 + j_2 - 2\rangle$, $|j_1 + j_2, j_1 + j_2 - 3\rangle$ e così via: in generale potremo trovare $|j_1 + j_2, m\rangle$ per ogni valore di $|m| \leq j_1 + j_2$.

Per abbassare j , invece, sfruttiamo l'ortonormalità dei $|j, m\rangle$, imponendo:

$$\langle j_1 + j_2, j_1 + j_2 - 1 | j_1 + j_2 - 1, j_1 + j_2 - 1 \rangle = 0$$

Il termine a sinistra è conosciuto (è combinazione di $|j_1, j_2; j_1 - 1, j_2\rangle$ e $|j_1, j_2; j_1, j_2 - 1\rangle$ come visto in (4.58)), e perciò possiamo risolvere per il termine a destra. Una volta ottenuto, possiamo poi reiterare questi due passaggi per ottenere tutti gli altri autoket.

In particolare, se vogliamo trovare delle relazioni esplicite per ricavare i coefficienti di Clebsch-Gordan, abbassiamo/alziamo $|j, m\rangle$ dato da (4.56):

$$\begin{aligned} J_\pm |j, m\rangle &= \sqrt{j(j+1) - m(m \pm 1)} |j, m \pm 1\rangle = \\ &= \sum_{m'_1} \sum_{m'_2} \langle j_1, j_2; m'_1, m'_2 | j, m \rangle \cdot \end{aligned} \quad (4.60)$$

$$\begin{aligned} &\cdot \left(\sqrt{j_1(j_1 + 1) - m'_1(m'_1 \pm 1)} |j_1, j_2; m'_1 \pm 1, m'_2\rangle + \right. \\ &\quad \left. + \sqrt{j_2(j_2 + 1) - m'_2(m'_2 \pm 1)} |j_1, j_2; m'_1, m'_2 \pm 1\rangle \right) \end{aligned} \quad (4.61)$$

E prendendo il prodotto scalare con $\langle j_1, j_2; m_1, m_2 |$ si *sopprimono* (per l'ortonormalità) tutti i termini per cui $m'_1 - 1 \neq m_1$ e $m'_2 - 1 \neq m_2$ e perciò si giunge alla relazione:

$$\begin{aligned} &\sqrt{j(j+1) - m(m \pm 1)} \langle j_1, j_2; m_1, m_2 | j, m \pm 1 \rangle = \\ &= \sqrt{j_1(j_1 + 1) - m_1(m_1 \mp 1)} \langle j_1, j_2; m_1 \mp 1, m_2 | j, m \rangle + \\ &+ \sqrt{j_2(j_2 + 1) - m_2(m_2 \mp 1)} \langle j_1, j_2; m_1, m_2 \mp 1 | j, m \rangle \end{aligned} \quad (4.62)$$

Dato che nelle precedenti relazioni che *collegano* i coefficienti di Clebsch-Gordan *non* compaiono termini complessi, si ha che tali coefficienti sono puramente reali. Perciò si può passare da una base all'altra e *viceversa* con gli stessi coefficienti:

$$\langle j_1, j_2; j, m | j_1, j_2; m_1, m_2 \rangle = \langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m \rangle$$

Quanto visto basta per determinare tutti i coefficienti, *a meno di una fase*, che viene fissata dalla convenzione:

$$\langle j_1, j_2, j_1, j - j_1 | j, j \rangle \geq 0$$

4.8.4 Esempio di calcolo dei coefficienti di Clebsh-Gordan

Applichiamo quanto appena visto al caso, semplice ma interessante, di due particelle (es. elettroni) con spin pari a $\frac{1}{2}$. Avremo allora $j_1 = j_2 = \frac{1}{2}$, e $|m_1| \leq \frac{1}{2}$, $|m_2| \leq \frac{1}{2}$.

Il momento totale j varia tra $|j_1 - j_2|$ e $j_1 + j_2$, e quindi si ha $0 \leq j \leq 1$. In altre parole, il sistema dei due elettroni avrà o spin pari a 1, o spin nullo.

Determiniamo tutti i coefficienti di Clebsh-Gordan per questo sistema. Per farlo, aiutiamoci con una tabella:

m_1	m_2				
1/2	1/2	x			
1/2	-1/2				
-1/2	1/2				
-1/2	-1/2				
		1	0	-1	0
		m			
		1		0	
		j			

Tabella 4.1: Tabella dei coefficienti di Clebsh-Gordan per due elettroni: struttura

La x nella tabella, per esempio, rappresenta:

$$\langle j_1, j_2; m_1, m_2 | j, m \rangle = C_{m_1, m_2, m}^{j_1, j_2, j} = C_{\frac{1}{2}, -\frac{1}{2}, 0}^{\frac{1}{2}, \frac{1}{2}, 1} = \left\langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \middle| 1, 0 \right\rangle$$

Partiamo subito ponendo a 0 tutti i coefficienti per cui $m \neq m_1 + m_2$:

m_1	m_2				
1/2	1/2	0	0	0	
1/2	-1/2	0		0	
-1/2	1/2	0		0	
-1/2	-1/2	0	0		0
		1	0	-1	0
		m			
		1		0	
		j			

La condizione di normalizzazione (4.57), tradotta nella rappresentazione tabulare che stiamo considerando, ci dice che la somma dei quadrati degli elementi di una

riga (o di una colonna) *deve* fare 1. Ciò significa che possiamo inserire un ± 1 nelle righe o colonne che hanno un solo elemento non nullo:

m_1	m_2					
1/2	1/2	± 1	0	0	0	
1/2	-1/2	0		0		
-1/2	1/2	0		0		
-1/2	-1/2	0	0	± 1	0	
		1	0	-1	0	m
			1		0	j

Usiamo quindi la relazione di recursione (4.62) per *abbassare* $|1, 1\rangle$, con $j_1 = j_2 = \frac{1}{2}, j = 1, m = 0, m_1 = \pm \frac{1}{2}, m_2 = \mp \frac{1}{2}$, o, in maniera equivalente, la (4.59), dato che stiamo abbassando l'autoket “più alto”. In ogni caso, si giunge a:

$$\sqrt{2}|1, 0\rangle = \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \right\rangle + \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle$$

E prendendo il prodotto scalare con $\langle 1/2, 1/2; -1/2, 1/2|$ e poi con $\langle 1/2, 1/2; -1/2, 1/2|$, otteniamo che:

$$\left\langle \frac{1}{2}, \frac{1}{2}; \pm \frac{1}{2}, \mp \frac{1}{2} \middle| 1, 0 \right\rangle = \frac{1}{\sqrt{2}}$$

Giungiamo quindi a:

m_1	m_2					
1/2	1/2	± 1	0	0	0	
1/2	-1/2	0	$\pm \frac{1}{\sqrt{2}}$	0		
-1/2	1/2	0	$\pm \frac{1}{\sqrt{2}}$	0		
-1/2	-1/2	0	0	± 1	0	
		1	0	-1	0	m
			1		0	j

Per trovare i coefficienti della quarta colonna, potremmo utilizzare ancora la relazione di ricursione, oppure notare che, poiché le entrate di ogni riga/colonna rappresentano le espansioni di basi ortonormali, ogni riga/colonna deve essere *ortogonale* alle altre. Scegliamo allora le entrate della quarta colonna in modo che il prodotto scalare con la seconda dia 0, e che la somma dei loro quadrati dia 1. Vi è un'unica possibilità:

m_1	m_2					
1/2	1/2	± 1	0	0	0	
1/2	-1/2	0	$\pm \frac{1}{\sqrt{2}}$	0	$\pm \frac{1}{\sqrt{2}}$	
-1/2	1/2	0	$\pm \frac{1}{\sqrt{2}}$	0	$\mp \frac{1}{\sqrt{2}}$	
-1/2	-1/2	0	0	± 1	0	
		1	0	-1	0	\mathbf{m}
			1		0	\mathbf{j}

La scelta dei segni è arbitraria, basta che siano mantenute le *relazioni* tra un termine e l'altro. Per convenzione si prende:

$$\langle j_1, j_2; j_1, j - j_1 | j, j \rangle \geq 0 \Rightarrow \left\langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2} \middle| 1, 1 \right\rangle \geq 0$$

e tutti gli altri segni di conseguenza. Perciò:

m_1	m_2					
1/2	1/2	1	0	0	0	
1/2	-1/2	0	$\frac{1}{\sqrt{2}}$	0	$\frac{1}{\sqrt{2}}$	
-1/2	1/2	0	$\frac{1}{\sqrt{2}}$	0	$-\frac{1}{\sqrt{2}}$	
-1/2	-1/2	0	0	1	0	
		1	0	-1	0	\mathbf{m}
			1		0	\mathbf{j}

Tabella 4.2: Coefficienti di Clebsh-Gordan per un sistema di due particelle a spin $\frac{1}{2}$

Grazie ai coefficienti appena calcolati possiamo finalmente esprimere una base in termine dell'altra.

Poniamo, per semplicità di notazione, gli autostati dello spin (calcolati in (4.54)) con:

$$\left| +\frac{1}{2} \right\rangle = |+\rangle \quad \left| -\frac{1}{2} \right\rangle = |-\rangle$$

Allora, una base di $\mathcal{H}_{j_1} \otimes \mathcal{H}_{j_2}$ è data dal prodotto tensore delle basi di ciascuno:

$$\{|+\rangle, |-\rangle\} \otimes \{|+\rangle, |-\rangle\} = \{|+\rangle |+\rangle, |-\rangle |-\rangle, |+\rangle |-\rangle, |-\rangle |+\rangle\}$$

Dove si usa la notazione di Dirac che *omette* il prodotto tra i ket: $|+\rangle |+\rangle \equiv |+\rangle \otimes |+\rangle$.

Sappiamo che:

$$\mathcal{H}_{s=\frac{1}{2}} \otimes \mathcal{H}_{s=\frac{1}{2}} = \mathcal{H}_{j=0} \oplus \mathcal{H}_{j=1}$$

E nello spazio “di destra”, la base canonica è:

$$\mathcal{H}_0 \rightarrow \{|0, 0\rangle\} \quad \mathcal{H}_1 \rightarrow \{|1, -1\rangle, |1, 0\rangle, |1, 1\rangle\}$$

Osservando la tabella 4.2 possiamo scrivere questi ultimi autoket in termini dei primi.

Per \mathcal{H}_0 :

$$|0, 0\rangle = \frac{|+\rangle |-\rangle - |-\rangle |+\rangle}{\sqrt{2}}$$

Mentre per \mathcal{H}_1 :

$$|1, 1\rangle = |+\rangle |+\rangle \quad |1, 0\rangle = \frac{|+\rangle |-\rangle + |-\rangle |+\rangle}{\sqrt{2}} \quad |1, -1\rangle = |-\rangle |-\rangle$$

Nota: In questo caso si procede velocemente anche procedendo per “abbassamento diretto” tramite $S_- = S_-^{(1)} + S_-^{(2)}$, e notando che:

$$\begin{aligned} S_- |+\rangle |+\rangle &= S_-^{(1)} |+\rangle |+\rangle + S_-^{(2)} |+\rangle |+\rangle = |-\rangle |+\rangle + |+\rangle |-\rangle \\ S_- (|+\rangle |-\rangle - |-\rangle |+\rangle) &= (S_-^{(1)} + S_-^{(2)}) (|+\rangle |-\rangle - |-\rangle |+\rangle) = |-\rangle |-\rangle - |-\rangle |-\rangle = 0 \end{aligned}$$

(Lezione 28 del
29/11/2018)

4.9 Esercizio 4

Una particella con spin $s = \frac{1}{2}$ è descritta, per la componente spaziale in \mathbb{R}^3 , dalla funzione d'onda $\psi(\vec{r})$:

$$\psi(\vec{r}) = x e^{-a|\vec{r}|} \quad a > 0 \text{ cost.}$$

1. Si determinino i possibili risultati e le relative probabilità di una misura di L_3
2. Si assuma che la componente di spin lungo \hat{z} sia $m_s = \hbar/2$. Si calcolino le probabilità che una misura di $J_3 = L_3 + S_3$ dia come risultato 0 o $3\hbar/2$.
3. Si calcoli la probabilità che una misura congiunta di \vec{J}^2 e J_3 dia come risultato $(3\hbar/4, -\hbar/2)$.
4. Supponendo che la misura del punto 3 sia stata eseguita con i risultati indicati, una successiva misura di S_3 a $t = 0$ dà il valore $m_s = -\hbar/2$. Se l'Hamiltoniana è:

$$H = c \vec{L} \cdot \vec{S} \quad c \neq 0 \text{ cost.}$$

si determini l'evoluto temporale dello stato a $t > 0$.

4.9.1 Soluzione

1. Poiché abbiamo a che fare con gli operatori del momento angolare \vec{L}^2 e L_3 , conviene riscrivere la $\psi(\vec{r})$ come combinazione lineare degli autoket comuni a \vec{L}^2 e L_3 , ossia come combinazione delle *armoniche sferiche*, che hanno la forma:

$$Y_l^m(\theta, \varphi) = \sqrt{\frac{2l+1}{4\pi}} \sqrt{\frac{(l+|m|)!}{(l-|m|)!}} \frac{1}{2^l l!} \frac{1}{\sin \theta^{|m|}} \frac{\partial^{l-|m|}}{\partial \cos \theta^{l-|m|}} (\sin \theta)^{2l} e^{im\varphi} (\text{sgn } m)^m (-1)^l$$

Partiamo passando in coordinate sferiche, notando che $e^{-a|\vec{r}|}$ ha già simmetria sferica, e infatti diviene e^{-ar} , mentre per il resto usiamo $x = r \sin \theta \cos \varphi$, giungendo a:

$$\psi(r, \theta, \varphi) = r \sin \theta \cos \varphi e^{-ar}$$

Notiamo poi che possiamo riscrivere:

$$x = r \sin \theta \cos \varphi = r \sin \theta \frac{e^{i\varphi} + e^{-i\varphi}}{2}$$

come combinazione delle armoniche con $l = 1$ e $m = \pm 1$:

$$Y_1^1 = \sqrt{\frac{3}{4\pi}} \sqrt{2!} \frac{1}{2} \sin \theta e^{i\varphi} (-1) = -\sqrt{\frac{3}{8\pi}} \sin \theta e^{i\varphi}$$

$$Y_1^{-1} = \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\varphi}$$

Perciò possiamo riscrivere la funzione d'onda in termini di armoniche sferiche:

$$\psi(\vec{x}) = x e^{-a|\vec{x}|} \propto -r e^{-ar} (Y_1^1 - Y_1^{-1})$$

Poiché $m = \pm 1$, i possibili valori ottenuti da una misura di L_3 sono $\pm \hbar$. Il fatto che abbiano coefficienti uguali significa che i due sono equiprobabili, e quindi:

$$W_\psi^{L_3}(\pm \hbar) = \frac{1}{2}$$

2. Per determinare le probabilità di una misura di \vec{J} o J_3 , è necessario spostarsi dalla base $|l, s; m, m_s\rangle$ a quella $|J, M\rangle$:

$$\underbrace{\mathcal{H}_l \otimes \mathcal{H}_s}_{|l, s, m, m_s\rangle} = \bigoplus_{\substack{J=|l-s| \\ |J, M\rangle}}^{l+s} \mathcal{H}_J$$

Partiamo riscrivendo lo stato $|\psi\rangle$ iniziale in notazione di Dirac, ricordando che dai dati del problema $l = 1$, $s = \frac{1}{2}$, $m = \pm 1$, $m_s = \frac{1}{2}$. Trascuriamo,

inoltre, la componente radiale, dato che non modifica in alcun modo le misure di \vec{J} e J_3 , e normalizziamo opportunamente lo stato così ottenuto²⁵:

$$|\psi\rangle = \frac{|1, \frac{1}{2}, 1, \frac{1}{2}\rangle - |1, \frac{1}{2}, -1, \frac{1}{2}\rangle}{\sqrt{2}}$$

Sappiamo che $M = J, J-1, \dots, -J$, e che J può assumere valori da $l-s = \frac{1}{2}$ a $l+s = \frac{3}{2}$, ossia solo $\{\frac{1}{2}, \frac{3}{2}\}$. I valori di M sono allora ottenuti da questi sottraendo un intero, e sono perciò $\{-\frac{3}{2}, -\frac{1}{2}, +\frac{1}{2}, +\frac{3}{2}\}$. Notiamo che tra di essi non compare 0, che perciò non potrà mai essere il risultato di una misura. Perciò:

$$W_\psi^{J_3}(0) = 0$$

D'altro canto, per la probabilità di ottenere $J_3 = \frac{3}{2}$ dobbiamo calcolare:

$$W_\psi^{J_3}\left(\frac{3}{2}\hbar\right) = \left|\left\langle\frac{3}{2}, \frac{3}{2}\right|\psi\right\rangle\right|^2$$

Ricordando che $J_3 = M = m + m_s$, se scriviamo le prime combinazioni:

M	m	m_s
$\frac{3}{2}$	1	$\frac{1}{2}$
$\frac{1}{2}$	0	$\frac{1}{2}$
	$\frac{1}{2}$	0
\vdots	\vdots	\vdots

notiamo che $\frac{3}{2}$ è il *massimo* valore di J_3 , e quindi l'unico modo di ottenerlo è scegliendo l e s massimi, ossia rispettivamente pari a 1 e 1/2 (evidenziati in giallo). Scriviamo quindi:

$$\begin{aligned} W_\psi^{J_3}\left(\frac{3}{2}\hbar\right) &= \left|\left\langle\frac{3}{2}, \frac{3}{2}\right|\psi\right\rangle\right|^2 = \left|\left\langle 1, \frac{1}{2}; 1, \frac{1}{2}\right|\psi\right\rangle\right|^2 = \\ &= \left|\left\langle 1, \frac{1}{2}; 1, \frac{1}{2}\right|\frac{1}{\sqrt{2}}\left(|1, \frac{1}{2}; 1, \frac{1}{2}\rangle - |1, \frac{1}{2}; 1, -\frac{1}{2}\rangle\right)\right\rangle\right|^2 = \frac{1}{2} \end{aligned}$$

Nel svolgere i conti del braket abbiamo utilizzato l'ortonormalità della base $|l, s; m, m_s\rangle$.

3. Vogliamo ora calcolare:

$$W_\psi^{\vec{J}^2, J_3}\left(\frac{3}{4}\hbar^2, -\frac{\hbar}{2}\right) = \left|\left\langle \underbrace{\frac{1}{2}}_J, \underbrace{-\frac{1}{2}}_M \right|\psi\right\rangle\right|^2 \quad (4.63)$$

²⁵Matematicamente, stiamo utilizzando l'isomorfismo $L^2(\mathbb{R}^3, d^3x) \cong L^2(\mathbb{R}_+, r^2 dr) \otimes L^2(S^2, d\Omega = \sin\theta d\theta d\varphi)$ per lavorare solo in $L^2(S^2)$, dove trascuriamo perciò la componente radiale.

Dato che \vec{J}^2 ha autovalori della forma $\hbar^2 J(J+1)$, da cui abbiamo ricavato:

$$\hbar^2 J(J+1) = \frac{3}{4} \hbar^2 \Rightarrow J = \frac{1}{2}$$

Poiché $|\psi\rangle$ è scritta nella base $|l, s; m, m_s\rangle$, per svolgere il braket conviene riscrivere anche $\left|\frac{1}{2}, -\frac{1}{2}\right\rangle$ in questa base. Possiamo farlo per “abbassamenti” successivi, partendo da:

$$\left|\frac{3}{2}, \frac{3}{2}\right\rangle = \left|1, \frac{1}{2}; 1, \frac{1}{2}\right\rangle$$

Applicando J_- e utilizzando l'identità (4.59), che è la forma già semplificata per il caso comune di abbassare un autoket *massimo*:

$$\underbrace{J_- \left|\frac{3}{2}, \frac{3}{2}\right\rangle}_{=(J_-^{(1)} + J_-^{(2)})|1, \frac{1}{2}, 1, \frac{1}{2}\rangle} = \sqrt{2 \left(1 + \frac{1}{2}\right)} \left|\frac{3}{2}, \frac{1}{2}\right\rangle = \sqrt{3} \left|\frac{3}{2}, \frac{1}{2}\right\rangle$$

$$\begin{aligned} (J_-^{(1)} + J_-^{(2)}) \left|1, \frac{1}{2}; 1, \frac{1}{2}\right\rangle &= \sqrt{2 \cdot 1} \left|1, \frac{1}{2}; 0, \frac{1}{2}\right\rangle + \sqrt{2 \cdot \frac{1}{2}} \left|1, \frac{1}{2}; 1, -\frac{1}{2}\right\rangle = \\ &= \sqrt{2} \left|1, \frac{1}{2}; 0, \frac{1}{2}\right\rangle + \left|1, \frac{1}{2}; 1, -\frac{1}{2}\right\rangle \end{aligned}$$

Dall'uguaglianza dei due termini azzurri ricaviamo:

$$\left|\frac{3}{2}, \frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}} \left|1, \frac{1}{2}; 0, \frac{1}{2}\right\rangle + \frac{1}{\sqrt{3}} \left|1, \frac{1}{2}; 1, -\frac{1}{2}\right\rangle \quad (4.64)$$

Vogliamo ora scendere ulteriormente, per arrivare a $\left|\frac{1}{2}, -\frac{1}{2}\right\rangle$. Per farlo, dovremo anche abbassare l'autovalore di J , per cui usiamo l'ortogonalità:

$$\langle j_1 + j_2, j_1 + j_2 - 1 | j_1 + j_2 - 1, j_1 + j_2 - 1 \rangle = 0 \Rightarrow \left|\frac{3}{2}, \frac{1}{2}\right\rangle \perp \left|\frac{1}{2}, \frac{1}{2}\right\rangle$$

Notiamo che quest'ultimo autoket può essere combinazione solo di due $|l, s; m, m_s\rangle$, come si può notare dalle prime combinazioni di $M = m + m_s$ scritte sopra, di cui stavolta ci interessano solo i termini evidenziati in azzurro.

Scriviamo allora il nostro obiettivo $\left|\frac{1}{2}, \frac{1}{2}\right\rangle$ come combinazione lineare di questi autostati:

$$\left|\frac{1}{2}, \frac{1}{2}\right\rangle = c_1 \left|1, \frac{1}{2}, 0, \frac{1}{2}\right\rangle + c_2 \left|1, \frac{1}{2}, 1, -\frac{1}{2}\right\rangle$$

Calcolando allora il prodotto scalare con (4.64) (sfruttando l'ortonormalità dei $|l, s; m, m_s\rangle$) e imponendo che il risultato sia nullo per l'ortogonalità:

$$0 \stackrel{!}{=} \left\langle \frac{1}{2}, \frac{1}{2} \left| \frac{3}{2}, \frac{1}{2} \right\rangle = c_1 \sqrt{\frac{2}{3}} + \frac{c_2}{\sqrt{3}}$$

E per la normalizzazione si deve avere $c_1^2 + c_2^2 = 1$, con $c_2 > 0$ per la convenzione:

$$\langle j_1, j_2, j_1, j - j_1 | j, j \rangle \geq 0 \Rightarrow c_2 = \left\langle \frac{1}{2}, \frac{1}{2} \left| 1, \frac{1}{2}, 1, -\frac{1}{2} \right\rangle > 0$$

In questo caso il modo più rapido per trovare c_1 e c_2 è utilizzare il fatto di algebra lineare per cui dato un vettore (a, b) , i vettori perpendicolari *con la stessa norma* sono $(-b, a)$ o $(b, -a)$. Possiamo allora partire da (4.64) e da $(\sqrt{\frac{2}{3}}, \frac{1}{\sqrt{3}})$ e trovare (c_1, c_2) seguendo lo schema $(-b, a)$, che rispetta la convenzione di $c_2 > 0$:

$$c_1 = -\sqrt{\frac{1}{3}}; \quad c_2 = \sqrt{\frac{2}{3}}$$

Riepilogando, abbiamo ottenuto:

$$\begin{aligned} \left| \frac{3}{2}, \frac{3}{2} \right\rangle &= \left| 1, \frac{1}{2}; 1, \frac{1}{2} \right\rangle \\ \left| \frac{3}{2}, \frac{1}{2} \right\rangle &= \sqrt{\frac{2}{3}} \left| 1, \frac{1}{2}; 0, \frac{1}{2} \right\rangle + \frac{1}{\sqrt{3}} \left| 1, \frac{1}{2}; 1, -\frac{1}{2} \right\rangle \\ \left| \frac{1}{2}, \frac{1}{2} \right\rangle &= -\frac{1}{\sqrt{3}} \left| 1, \frac{1}{2}; 0, \frac{1}{2} \right\rangle + \sqrt{\frac{2}{3}} \left| 1, \frac{1}{2}; 1, -\frac{1}{2} \right\rangle \end{aligned} \quad (4.65)$$

Per ottenere $\left| \frac{1}{2}, -\frac{1}{2} \right\rangle$ basta abbassare $\left| \frac{1}{2}, \frac{1}{2} \right\rangle$ mediante J_- . Utilizziamo allora la formula (4.61), con $j = \frac{1}{2}$ e $m = \frac{1}{2}$. Il membro di sinistra diviene:

$$\begin{aligned} J_- \left| \frac{1}{2}, \frac{1}{2} \right\rangle &= \sqrt{j(j+1) - m(m-1)} |j, m-1\rangle = \\ &= \sqrt{\frac{1}{2} \left(\frac{1}{2} + 1 \right) - \frac{1}{2} \left(\frac{1}{2} - 1 \right)} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \end{aligned}$$

Mentre quello di destra è dato da:

$$\begin{aligned} \sum_{m'_1 = -j_1}^{+j_1} \sum_{m'_2 = -j_2}^{+j_2} \langle j_1, j_2; m'_1, m'_2 | j, m \rangle \left(\sqrt{j_1(j_1+1) - m'_1(m'_1-1)} |j_1, j_2; m'_1-1, m'_2\rangle + \right. \\ \left. + \sqrt{j_2(j_2+1) - m'_2(m'_2-1)} |j_1, j_2; m'_1, m'_2-1\rangle \right) \end{aligned}$$

dove $j_1 = l = 1$, $j_2 = s = \frac{1}{2}$, $j = m = \frac{1}{2}$, e $\langle j_1, j_2; m'_1, m'_2 | j, m \rangle$ sono i coefficienti evidenziati in giallo in (4.65), e sono non nulli solo per $(m'_1, m'_2) = (0, \frac{1}{2}), (1, -\frac{1}{2})$. Svolgendo allora i conti otteniamo:

$$\begin{aligned} &\left(-\frac{1}{\sqrt{3}} \right) \left[\sqrt{1(1+1) - 0 \left(0 - \frac{1}{2} \right)} \left| 1, \frac{1}{2}; -1, \frac{1}{2} \right\rangle + \sqrt{\frac{1}{2} \left(\frac{1}{2} + 1 \right) - \frac{1}{2} \left(\frac{1}{2} - 1 \right)} \left| 1, \frac{1}{2}; 0, -\frac{1}{2} \right\rangle \right] + \\ &+ \left(\sqrt{\frac{2}{3}} \right) \left[\sqrt{1(1+1) - 1(1-1)} \left| 1, \frac{1}{2}; 0, -\frac{1}{2} \right\rangle + \sqrt{\frac{1}{2} \left(\frac{1}{2} + 1 \right) - \left(-\frac{1}{2} \right) \left(-\frac{1}{2} - 1 \right)} \left| 1, \frac{1}{2}; 1, -\frac{3}{2} \right\rangle \right] = \\ &= -\sqrt{\frac{2}{3}} \left| 1, \frac{1}{2}; -1, \frac{1}{2} \right\rangle - \frac{1}{\sqrt{3}} \left| 1, \frac{1}{2}; 0, -\frac{1}{2} \right\rangle + \frac{2}{\sqrt{3}} \left| 1, \frac{1}{2}; 0, -\frac{1}{2} \right\rangle = \\ &= -\sqrt{\frac{2}{3}} \left| 1, \frac{1}{2}; -1, \frac{1}{2} \right\rangle + \frac{1}{\sqrt{3}} \left| 1, \frac{1}{2}; 0, -\frac{1}{2} \right\rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \end{aligned} \quad (4.66)$$

Nota: dato che $|1, \frac{1}{2}; 1, -\frac{1}{2}\rangle$ ha già lo spin minimo, non è possibile abbassarlo ulteriormente, e infatti il coefficiente di $|1, \frac{1}{2}; 1, -\frac{3}{2}\rangle$ risulta nullo.

Riepilogando, per ottenere $|\frac{1}{2}, -\frac{1}{2}\rangle$ abbiamo usato questa strategia:

$$|\frac{3}{2}, \frac{3}{2}\rangle \xrightarrow{J_-} |\frac{3}{2}, \frac{1}{2}\rangle \perp |\frac{1}{2}, \frac{1}{2}\rangle \xrightarrow{J_-} |\frac{1}{2}, -\frac{1}{2}\rangle$$

Usando il risultato ottenuto in (4.66), possiamo allora finalmente calcolare il braket in (4.63), e quindi la probabilità cercata:

$$W_{\psi}^{\vec{J}^2, J_3} \left(\frac{3}{4}\hbar^2, -\frac{\hbar}{2} \right) = \left| \left\langle \frac{1}{2}, -\frac{1}{2} \left| \frac{1}{\sqrt{2}} \left(\left| 1, \frac{1}{2}, -1, \frac{1}{2} \right\rangle - \left| 1, \frac{1}{2}, 1, \frac{1}{2} \right\rangle \right) \right. \right|^2 = \left| \sqrt{\frac{2}{3}} \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{3}$$

4. La misura di \vec{J}^2 e J_3 , di risultato $(\frac{3\hbar^2}{4}, -\frac{\hbar}{2})$, proietta lo stato del sistema sull'autoket $|\frac{1}{2}, -\frac{1}{2}\rangle$, che abbiamo già determinato al punto precedente:

$$|\frac{1}{2}, -\frac{1}{2}\rangle = \underbrace{\frac{1}{\sqrt{3}} |1, \frac{1}{2}, 0, -\frac{1}{2}\rangle}_{m_s = -1/2} - \underbrace{\sqrt{\frac{2}{3}} |1, \frac{1}{2}, -1, \frac{1}{2}\rangle}_{m_s \neq -1/2} \quad (4.67)$$

In questo stato misuriamo S_3 , ottenendo $-\frac{\hbar}{2}$, ossia $m_s = -\frac{1}{2}$. Ciò fa sì che tutti i termini di (4.67) che non sono compatibili con tale risultato svaniscano, e perciò il nuovo stato $|\psi(0^+)\rangle$ conterrà solo i rimanenti:

$$|\psi(0^+)\rangle = |1, \frac{1}{2}; 0, -\frac{1}{2}\rangle$$

Vogliamo ora calcolare l'evoluto temporale $|\psi(t)\rangle$:

$$|\psi(t)\rangle = \exp\left(-\frac{i}{\hbar}tH\right) |\psi(0^+)\rangle \quad (4.68)$$

Partiamo allora dall'espressione dell'Hamiltoniana:

$$H = c\vec{L} \cdot \vec{S}$$

In questa forma, non sappiamo ricavare direttamente lo spettro di $\sigma(\vec{L} \cdot \vec{S})$, per cui risulterebbe difficile riscrivere $|\psi(0^+)\rangle$ come combinazione di autostati di H . Notiamo però che possiamo riscrivere tale H in funzione di operatori “che sappiamo gestire”.

Sappiamo infatti che:

$$\vec{J} = \vec{L} + \vec{S} = \vec{L} \otimes \mathbb{I} + \mathbb{I} \otimes \vec{S}$$

e $[\vec{L}, \vec{S}] = 0$. Allora possiamo elevare al quadrato “senza problemi”:

$$\vec{J}^2 = \vec{L}^2 + \vec{S}^2 + 2\vec{L} \cdot \vec{S} \Rightarrow \vec{L} \cdot \vec{S} = \frac{1}{2} (\vec{J}^2 - \vec{L}^2 - \vec{S}^2)$$

E quindi otteniamo:

$$H = \frac{c}{2}(\vec{J}^2 - \vec{L}^2 - \vec{S}^2)$$

Vogliamo inoltre che $|\psi(0^+)\rangle$ sia scritta come combinazione di autoket di H , ossia di autoket comuni a \vec{J}^2 , \vec{L}^2 , \vec{S}^2 . Attualmente $|\psi(0^+)\rangle$ è però scritta nella base $|l, s; m, m_s\rangle$, che è fatta di autoket di \vec{L}^2 e \vec{S}^2 , ma non di \vec{J}^2 . Dobbiamo allora passare in base $|l, s; J, M\rangle$ riscrivendo $|\psi(0^+)\rangle$ come:

$$\left|1, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle = \sum_{J=\frac{1}{2}}^{\frac{3}{2}} \sum_{M=-J}^J \langle J, M | 1, \frac{1}{2}, 0, -\frac{1}{2} \rangle |J, M\rangle \quad (4.69)$$

I coefficienti di Clebsh-Gordan sono non nulli solo per $M = m + m_s = 0 + (-\frac{1}{2}) = -\frac{1}{2}$ e perciò gli unici termini che contribuiscono alla sommatoria sono $|\frac{3}{2}, -\frac{1}{2}\rangle$ e $|\frac{1}{2}, -\frac{1}{2}\rangle$. Quest'ultimo l'abbiamo già ricavato:

$$\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = -\sqrt{\frac{2}{3}} \left|1, \frac{1}{2}, -1, \frac{1}{2}\right\rangle + \frac{1}{\sqrt{3}} \left|1, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle$$

E ci manca solo ricavare il primo. Ragionando come prima, possiamo per esempio abbassare $|\frac{3}{2}, \frac{1}{2}\rangle$ che già conosciamo:

$$\begin{aligned} J_- \left|\frac{3}{2}, \frac{1}{2}\right\rangle &= \sqrt{\frac{3}{2} \left(\frac{3}{2} + 1\right) - \frac{1}{2} \left(-\frac{1}{2}\right)} \left|\frac{3}{2}, -\frac{1}{2}\right\rangle = \\ &= (L_- + S_-) \left(\sqrt{\frac{2}{3}} \left|1, \frac{1}{2}, 0, \frac{1}{2}\right\rangle + \frac{1}{\sqrt{3}} \left|1, \frac{1}{2}, 1, -\frac{1}{2}\right\rangle \right) = \\ &= \sqrt{\frac{2}{3}} \left[\sqrt{1(1+1)+0} \left|1, \frac{1}{2}, -1, \frac{1}{2}\right\rangle + \sqrt{\frac{1}{2} \left(\frac{1}{2} + 1\right) - \frac{1}{2} \left(-\frac{1}{2}\right)} \left|1, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle \right] + \\ &\quad + \frac{1}{\sqrt{3}} \left[\sqrt{1(1+1) - 1(1-1)} \left|1, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle + 0 \right] \\ \Rightarrow \left|\frac{3}{2}, -\frac{1}{2}\right\rangle &= \frac{1}{\sqrt{3}} \left|1, \frac{1}{2}, -1, \frac{1}{2}\right\rangle + \sqrt{\frac{2}{3}} \left|1, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle \end{aligned}$$

Sostituendo in (4.69) e calcolando, otteniamo:

$$\begin{aligned} \left|1, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle &= \left\langle \frac{3}{2}, -\frac{1}{2} \left| 1, \frac{1}{2}, 0, -\frac{1}{2} \right\rangle \right| \left|\frac{3}{2}, -\frac{1}{2}\right\rangle + \left\langle \frac{1}{2}, -\frac{1}{2} \left| 1, \frac{1}{2}, 0, -\frac{1}{2} \right\rangle \right| \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \\ &= \sqrt{\frac{2}{3}} \left|\frac{3}{2}, -\frac{1}{2}\right\rangle + \frac{1}{\sqrt{3}} \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \end{aligned}$$

Detti $|\varphi_n\rangle = |J, M\rangle$ gli autostati di H , possiamo espandere la (4.68) come:

$$|\psi(t)\rangle = \sum_n \exp\left(-\frac{i}{\hbar} \mathcal{E}_n(t - t_0)\right) |\varphi_n\rangle \langle \varphi_n | \psi(0^+) \rangle$$

Calcoliamo gli \mathcal{E}_n dalle equazioni agli autovalori:

$$\begin{aligned} H \left| \frac{3}{2}, -\frac{1}{2} \right\rangle &= \frac{c}{2} (\overbrace{J^2}^1 - \overbrace{\tilde{L}^2}^{1/2} - \overbrace{\tilde{S}^2}^{1/2}) \left| \frac{3}{2}, -\frac{1}{2} \right\rangle = \\ &= \frac{c}{2} \hbar^2 \left[\frac{3}{2} \left(\frac{3}{2} + 1 \right) - 1(1+1) - \frac{1}{2} \left(\frac{1}{2} + 1 \right) \right] \left| \frac{3}{2}, -\frac{1}{2} \right\rangle = \\ &= \frac{c}{2} \hbar^2 \left| \frac{3}{2}, -\frac{1}{2} \right\rangle \Rightarrow \mathcal{E}_1 = \frac{c}{2} \hbar^2 \end{aligned}$$

$$\begin{aligned} H \left| \frac{1}{2}, -\frac{1}{2} \right\rangle &= \frac{c}{2} \hbar^2 \left[\frac{1}{2} \left(\frac{1}{2} + 1 \right) - 1(1+1) - \frac{1}{2} \left(\frac{1}{2} + 1 \right) \right] \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \\ &= -c \hbar^2 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \Rightarrow \mathcal{E}_2 = -c \hbar^2 \end{aligned}$$

E infine:

$$\begin{aligned} |\psi(t)\rangle &= \exp \left(-i \frac{t}{\hbar} H \right) \left| 1, \frac{1}{2}, 0, -\frac{1}{2} \right\rangle = \\ &= \exp \left(-\frac{i}{\hbar} \mathcal{E}_1 t \right) \sqrt{\frac{2}{3}} \left| \frac{3}{2}, -\frac{1}{2} \right\rangle + \exp \left(-\frac{i}{\hbar} \mathcal{E}_2 t \right) \frac{1}{\sqrt{3}} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \end{aligned}$$

Inserendo allora i valori di \mathcal{E}_1 e \mathcal{E}_2 e reintroducendo il termine radiale, otteniamo infine:

$$|\psi(t)\rangle = -r e^{-ar} \left[\exp \left(-i \frac{c}{2} \hbar t \right) \sqrt{\frac{2}{3}} \left| \frac{3}{2}, -\frac{1}{2} \right\rangle + e^{i \hbar t} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \right]$$

(Lezione 29 del
30/11/2018)

4.10 Oscillatore armonico

Occupiamoci ora di studiare il comportamento di un **oscillatore armonico**, che ha applicazioni per lo studio di svariati fenomeni, come le vibrazioni di cristalli, o per l'espansione dei campi in modi normali: si ha infatti che ogni moto periodico nei pressi di un *minimo* dell'energia potenziale (ossia un punto di equilibrio) è ben approssimato da un oscillatore armonico²⁶.

In questa sezione ci limiteremo al caso **unidimensionale**. In MC, dette q e p rispettivamente la posizione e il momento di un punto di massa m , legato ad una molla di costante elastica k centrata all'origine, si ottiene la seguente Hamiltoniana:

Oscillatore
classico

$$H = \frac{p^2}{2m} + \frac{1}{2} m \omega^2 q^2 \quad \omega = \sqrt{\frac{k}{m}}$$

²⁶ΛCFR [3] pag. 47.

dove ω è la *pulsazione* del moto armonico, che è esplicitamente dato da:

$$q(t) = A \sin(\omega t) + B \cos(\omega t)$$

Poiché p e q possono potenzialmente assumere qualsiasi valore ($\sigma(p) = \sigma(q) = \mathbb{R}$), e in MQ lo spettro di una somma è la somma degli spettri, si ha che $\sigma(H) = \mathbb{R}$, ed è quindi *puramente continuo*.

Scopriremo che, invece, in MQ $\sigma(H)$ è *discreto*, e un ICOC per il sistema è dato dal solo H .

Oscillatore
quantistico

Lavoriamo in rappresentazione $\{x\}$, fissando $\mathcal{H} = L^2(\mathbb{R}, dx)$. L'Hamiltoniana quantistica H_0 si ottiene sostituendo le “versioni quantistiche” di posizione e momento, ossia P e X , che in rappresentazione $\{x\}$ sono dati da:

$$X = x \quad P = -i\hbar \frac{d}{dx}$$

Giungiamo allora a:

$$H_0 = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + m \frac{\omega^2}{2} x^2; \quad D(H_0) = \mathcal{S}(\mathbb{R})$$

Si dimostra che $H_0^\dagger \equiv H$ è autoaggiunto (cioè $H_0^\dagger = (H_0^\dagger)^\dagger$) e questa è l'hamiltoniana dell'oscillatore armonico, se vogliamo essere matematicamente precisi. Ma per semplicità consideriamo in seguito le operazioni definite in $D(H_0)$.

L'equazione di Schrödinger stazionaria è allora:

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m \omega^2 x^2 \right] \psi(x) = \mathcal{E} \psi(x)$$

Per risolverla si può procedere in modo *diretto* e *analitico* tramite le serie di potenze, oppure - come faremo - sfruttando delle ingegnose proprietà algebriche, adottando tecniche simili a quelle usate nel caso del momento angolare.

Partiamo riscrivendo H_0 in una forma più “suggestiva”:

$$H_0 = \frac{1}{2m} \left[\left(\frac{\hbar}{i} \frac{d}{dx} \right)^2 + (m\omega x)^2 \right] = \frac{1}{2m} [P^2 + (m\omega X)^2]$$

L'idea è di *fattorizzare* H_0 , usando l'identità dell'algebra:

$$u^2 + v^2 = (u - iv)(u + iv)$$

Dato che stiamo lavorando con operatori che in genere *non commutano*, la fattorizzazione non sarà perfetta - ma ci consentirà comunque di ottenere una soluzione. Prima di fattorizzare, raccogliamo un fattore $m\omega$ per “simmetrizzare” i due quadrati, e anche un \hbar per “adimensionalizzarli”. Tale passaggio serve unicamente per velocizzare i conti successivi, togliendo di torno delle fastidiose costanti:

$$H_0 = \frac{\hbar\omega}{2} \left[\underbrace{\frac{P^2}{\hbar m \omega}}_{B^2} + \underbrace{\frac{m\omega X^2}{\hbar}}_{A^2} \right] \rightarrow \hbar\omega \left(\frac{A - iB}{\sqrt{2}} \right) \left(\frac{A + iB}{\sqrt{2}} \right)$$

A e B non sono altro che “riscalamenti” di momento P e posizione X , e quindi li chiamiamo in accordo:

$$A \equiv X' \equiv \sqrt{\frac{m\omega}{\hbar}} X \quad B \equiv P' \equiv \frac{1}{\sqrt{m\hbar\omega}} P$$

I due termini ottenuti dal raccoglimento sono invece denotati con a e a^\dagger :

$$a = \frac{X' + iP'}{\sqrt{2}} \quad a^\dagger = \frac{X' - iP'}{\sqrt{2}}$$

Notiamo che possiamo riscrivere X e P in termini di a e a^\dagger :

$$X = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger) \quad P = i\sqrt{\frac{\hbar m\omega}{2}}(a^\dagger - a)$$

Come ci si può aspettare, a e a^\dagger **non** commutano, dato che sono combinazione di X' e P' , che a loro volta non commutano:

$$\begin{aligned} [X, P] = i\hbar &\Rightarrow [X', P'] = \left[\sqrt{\frac{m\omega}{\hbar}} X, \frac{1}{\sqrt{m\omega\hbar}} P \right] = \frac{[X, P]}{\hbar} = i \\ [a, a^\dagger] &= \frac{1}{2} [X' + iP', X' - iP'] = \frac{i}{2} [P', X'] - \frac{i}{2} [X', P'] = -\frac{i^2}{2} - \frac{i^2}{2} = 1 \end{aligned}$$

Ci aspettiamo allora che la “fattorizzazione” fatta sopra non sia *perfetta*, ma introduca una differenza (evidenziata in giallo) data proprio dal commutatore. Svolgendo infatti i conti:

$$\begin{aligned} a^\dagger a &= \left(\frac{X' - iP'}{\sqrt{2}} \right) \left(\frac{X' + iP'}{\sqrt{2}} \right) = \frac{X'^2}{2} + \frac{P'^2}{2} + \frac{i[X', P']}{2} = \\ &= \frac{m\omega}{2\hbar} + \frac{p^2}{2m\omega\hbar} - \frac{1}{2} = H' - \frac{1}{2} \\ &\Rightarrow H = \hbar\omega H' = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right) \end{aligned}$$

Dove con H' denotiamo l'Hamiltoniana da cui abbiamo rimosso le costanti $\hbar\omega$ precedentemente raccolte - ossia tale che $H = \hbar\omega H'$.

Definiamo allora l'operatore $N = a^\dagger a$. Poiché H è autoaggiunto in $D(H)$, anche N lo è, e in effetti uno è funzione dell'altro:

$$H = \hbar\omega \left(N + \frac{1}{2} \right) \quad \sigma(H) = \hbar\omega \left(\sigma(N) + \frac{1}{2} \right) \quad (4.70)$$

Perciò, se troviamo lo spettro $\sigma(N)$, automaticamente abbiamo lo spettro $\sigma(H)$ che ci interessa.

Mimando quanto fatto nel caso del momento angolare, esaminiamo allora l'algebra di N , a e a^\dagger , studiandone i commutatori:

$$\begin{aligned}[N, a] &= [a^\dagger a, a] = [a^\dagger, a]a = -a \\ [N, a^\dagger] &= [a^\dagger a, a^\dagger] = a^\dagger[a, a^\dagger] = a^\dagger\end{aligned}$$

Denotando $a^\dagger \equiv a_+$, $a \equiv a_-$, ritroviamo allora una relazione del tipo $[N, a_\pm] = \pm a_\pm$, che è del tutto simile alla relazione $[J_3, J_\pm] = \pm \hbar J_\pm$ precedentemente ottenuta. In un certo senso, perciò, comparando i commutatori con quelli ottenuti studiando il momento angolare, avremmo $a \sim J_-$, $a^\dagger \sim J_+$, $N \sim J_3$.

Ciò ci suggerisce come continuare. Supponiamo infatti che esista un certo ψ_λ autovettore di N di autovalore λ , per cui $N\psi_\lambda = \lambda\psi_\lambda$.

Tale autovalore deve essere positivo²⁷, ossia $\lambda \geq 0$.

Dimostriamolo. Applicando l'equazione agli autovalori al calcolo del valor medio, si ha:

$$(\psi_\lambda, N\psi_\lambda) = (\psi_\lambda, \lambda\psi_\lambda) = \lambda \|\psi_\lambda\|^2$$

Ma per la definizione di $N = a^\dagger a$, vale anche:

$$(\psi_\lambda, N\psi_\lambda) = (\psi_\lambda, a^\dagger a\psi_\lambda) = (a\psi_\lambda, a\psi_\lambda) = \|a\psi_\lambda\|^2 \geq 0$$

E perciò $\lambda \geq 0$, e anche $\lambda = 0 \Leftrightarrow a\psi_\lambda = 0$.

Partendo da un qualsiasi autovettore ψ_λ di N , notiamo che $a\psi_\lambda$ è ancora autovettore di N , ma di autovalore $\lambda - 1$ (ammesso che $a\psi_\lambda \neq 0$), e analogamente lo è $a^\dagger\psi_\lambda$, con autovalore $\lambda + 1$. In altre parole, a e a^\dagger rispettivamente *abbassano* e *alzano* gli autovalori di N .

Dimostriamolo:

$$\begin{aligned}Na\psi_\lambda &= ([N, a] + aN)\psi_\lambda = (-a + a\lambda)\psi_\lambda = (\lambda - 1)a\psi_\lambda \\ Na^\dagger\psi_\lambda &= ([N, a^\dagger] + a^\dagger N)\psi_\lambda = (a^\dagger + a^\dagger\lambda)\psi_\lambda = (\lambda + 1)a^\dagger\psi_\lambda\end{aligned}$$

Per questa ragione, a è chiamato **operatore di distruzione** (o annichilazione), dato che *abbassa* un autovalore, e a^\dagger **operatore di creazione**²⁸. Ma allora iterando n -volte l'applicazione di a si ha che anche $\lambda - n$ è un autovalore corrispondente all'autovettore $a^n\psi_\lambda$ (se è diverso da 0).

Operatori di annichilazione (a) e creazione (a^\dagger)

²⁷ Ciò deriva da un fatto generale, per cui l'energia "associata" ad una qualsiasi soluzione normalizzabile dell'equazione di Schrödinger stazionaria deve essere sempre $\geq V_{min}$. In questo caso, N determina l'energia associata ad un autoket, dato che si può calcolare H da esso, e $V_{min} = 0$ per l'Hamiltoniana H_0 di partenza. L'idea, esaminata nel problema 2.2 a pag. 39 di [3], è che da $\frac{d^2\psi}{dx^2} = \frac{2m}{\hbar^2}[V(x) - \mathcal{E}]\psi$, se $V(x) > \mathcal{E}$ la derivata seconda di ψ e ψ stessa hanno sempre lo stesso segno. Ma se ψ è normalizzabile, deve "partire da 0" per $x \rightarrow \pm\infty$, e non appena vi si allontana non può tornare indietro, data l'assenza di punti di flesso - non può perciò essere $\neq 0$ e normalizzabile allo stesso tempo.

²⁸ Una motivazione migliore di tali nomi deriva dal fatto che i campi sono sviluppati come modi normali di oscillatori armonici, e in teoria dei campi le particelle sono *eccitazioni* di questi campi, si ha che tali operatori effettivamente *creano* e *distruggono* particelle.

Tuttavia, abbiamo dimostrato che $\lambda \geq 0$, ossia che tutti gli autovalori di N sono positivi. Ma scegliendo un n sufficientemente grande, si ha che $\lambda - n < 0$ - e ciò non può essere. L'unico modo è allora che λ sia un intero positivo, che chiamiamo m , cosicché dopo m iterazioni $a^m \psi_\lambda = (\lambda - m) \psi_\lambda = 0$ e la catena si interrompe. Avremo quindi una successione di autovalori di N dati da $m, m-1, \dots, 0$.

Ma allora applicando a ψ_0 (che esiste se esiste ψ_λ) potenze di a^\dagger , possiamo alzare di quanto vogliamo gli autovalori, dato che non esiste in generale un "tetto" per l'energia. Partendo da 0 e salendo *a passi interi*, otteniamo che tutti i numeri naturali fanno parte dello spettro di N .

$$\sigma(N) = \mathbb{N}$$

Naturalmente, tutto ciò funziona ammesso che esista un ψ_λ iniziale da cui partire. Poiché partendo da un *qualsiasi* ψ_λ consente di generare tutti gli autovalori, conviene, per semplicità, cercare di dimostrare l'esistenza per $\lambda = 0$, ossia risolvere l'equazione agli autovalori:

$$a |\psi_0\rangle = 0 \Rightarrow \frac{X' + iP'}{\sqrt{2}} |\psi_0\rangle = 0$$

Passando in rappresentazione in $\{x\}$ otteniamo un'equazione differenziale a variabili separabili:

$$\left(\sqrt{\frac{m\omega}{\hbar}} x + i \left(-\frac{i\hbar}{\sqrt{m\hbar\omega}} \frac{d}{dx} \right) \right) \psi_0(x) = 0$$

Con la sostituzione:

$$x' = \sqrt{\frac{m\omega}{\hbar}} x \Rightarrow \frac{d}{dx} = \frac{dx'}{dx} \frac{d}{dx'} = \sqrt{\frac{m\omega}{\hbar}} \frac{d}{dx'}$$

Il problema si semplifica notevolmente. Separando le variabili e integrando in x' :

$$\left(x' + \frac{d}{dx'} \right) \psi_0(x') = 0 \Rightarrow \frac{\psi_0'(x')}{\psi_0(x')} = -x' \Rightarrow \ln |\psi_0(x')| = -\frac{x'^2}{2} + c$$

Dopodiché basta esponenziare per trovare la soluzione cercata, che ha la forma di una *gaussiana*:

$$\psi_0(x') = A e^{-\frac{x'^2}{2}} = A \exp\left(-\frac{m\omega}{2\hbar} x^2\right) \in L^2$$

Dove A è fissata dalla normalizzazione:

$$A = \left(\frac{m\omega}{\pi\hbar} \right)^{1/4}$$

Riconosciamo in $\psi_0(x')$ la funzione di Hermite $h_0(x')$. Possiamo ora iterare a^\dagger per ottenere tutti gli altri autostati, che corrispondono ancora a funzioni di Hermite $h_n(x')$ di ordine n -esimo:

$$(a^\dagger)^n \psi_0(x) = \left(x' - \frac{d}{dx'} \right)^n h_0(x') \sim h_n(x')$$

Più precisamente le funzioni di Hermite $h_n(x)$ sono definite da:

$$h_n(x) = (2^n n! \sqrt{\pi})^{-\frac{1}{2}} \left(x - \frac{d}{dx} \right)^n \exp \left(-\frac{x^2}{2} \right)$$

Si dimostra che le $\{h_n(x)\}_{n \in \mathbb{N}}$ costituiscono una base ON²⁹ in $L^2(\mathbb{R}, dx)$. Ne segue allora che gli autovalori $n \in \mathbb{N}$ esauriscono $\sigma(N)$, quindi scrivendo $h_n(x) = \langle x | n \rangle$ vale la completezza di Dirac $\sum_{n \in \mathbb{N}} |n\rangle \langle n| = \mathbb{I}$. Perciò non c'è $\sigma_C(N)$ e lo spettro è non degenere.

Notiamo ora che:

1. Dalla relazione tra gli spettri in (4.70) ricaviamo $\sigma(H)$:

$$\sigma(H) = \sigma_P(H) = \hbar\omega \left(\mathbb{N} + \frac{1}{2} \right)$$

Perciò, come conseguenza diretta dell'**algebra** di X e P ($\Leftrightarrow a, a^\dagger$), si ha che lo spettro dell'energia dell'oscillatore armonico è **quantizzato**³⁰.

2. Perché lo spettro è puramente **discreto** e **non degenere** dal teorema sulle osservabili compatibili segue che H è un ICOC per l'oscillatore armonico quantistico.

Facciamo, infine, qualche osservazione sui risultati ottenuti:

1. A differenza di quanto accade in MC, lo spettro della somma di due operatori *non* è la somma degli spettri:

$$\sigma(H) \sim \sigma(X^2 + P^2) \neq \sigma(X^2) + \sigma(P^2)$$

dato che $\sigma(X^2)$ e $\sigma(P^2)$ sono continui, mentre $\sigma(H)$ è discreto.

2. La famiglia spettrale di H si ottiene, in analogia a quanto visto per X , sommando funzioni a gradino di Heaviside $H(\lambda)$ centrate sui singoli autovalori:

$$P^H(\lambda) = \sum_{n \in \mathbb{N}} H \left(\lambda - \hbar\omega \left(n + \frac{1}{2} \right) \right) |n\rangle \langle n|$$

Tramite questa definizione, possiamo dare esplicitamente il dominio di autoaggiuntezza di H come:

$$\begin{aligned} D(H) &= \{ \psi \in \mathcal{H} \mid \int \lambda^2 d(\psi, P^H(\lambda)\psi) < \infty \} \\ &= \{ \psi \in \mathcal{H} \mid |\psi\rangle = \sum_{n \in \mathbb{N}} c_n |n\rangle, \quad \sum_{n \in \mathbb{N}} (\hbar\omega)^2 \left(n + \frac{1}{2} \right)^2 |c_n|^2 < \infty \} \end{aligned}$$

²⁹^CFR pag. 17 di [1]

³⁰^In realtà abbiamo utilizzato anche l'esistenza della soluzione di un'equazione differenziale agli autovalori. In realtà si può dimostrare - in maniera più lunga e *difficile* senza questa ipotesi

3. Alzare o abbassare autostati tramite a^\dagger e a non produce autostati normalizzati. In generale, la procedura introduce un fattore $c(n)$, per cui:

$$a^\dagger |n\rangle = c_+(n) |n+1\rangle \quad a |n\rangle = c_-(n) |n-1\rangle$$

Fissiamo allora:

$$\langle n | n \rangle \stackrel{!}{=} 1 \quad \forall n$$

e cerchiamo di determinare le costanti c_+ e c_- .

Partiamo considerando a^\dagger , e applicandolo a sinistra e a destra di $\langle n | n \rangle$.

Otteniamo due espansioni equivalenti:

$$\langle a^\dagger n | a^\dagger n \rangle = |c_+(n)|^2 \langle n+1 | n+1 \rangle$$

$$\langle a^\dagger n | a^\dagger n \rangle = \langle n | a a^\dagger | n \rangle = \langle n | (a, a^\dagger) + a^\dagger a | n \rangle = \langle n | (1 + N) | n \rangle \stackrel{(a)}{=} (n+1) \langle n | n \rangle$$

dove in (a) si è usato il fatto che $|n\rangle$ è autovalore di autovalore n per N .

Uguagliando le due espressioni si ottiene:

$$|c_+(n)|^2 = (n+1) \Rightarrow c_+(n) = \sqrt{n+1}$$

D'altro canto, ripetendo gli stessi ragionamenti nel caso di a si ottiene:

$$\langle a n | a n \rangle = |c_-(n)|^2 \langle n-1 | n-1 \rangle$$

$$\langle a n | a n \rangle = \langle n | a^\dagger a | n \rangle = \langle n | N | n \rangle \stackrel{(a)}{=} n \langle n | n \rangle$$

Da cui:

$$|c_-(n)|^2 = n \Rightarrow c_-(n) = \sqrt{n}$$

Riepilogando:

$$a^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle \quad a |n\rangle = \sqrt{n} |n-1\rangle \quad (4.71)$$

4. Nella *old quantum theory* Planck era giunto a $\mathcal{E}_n \cong \hbar\omega n$, da cui però manca il termine $\hbar\omega/2$, che dà l'energia di punto zero $\hbar\omega(n+1/2)$, necessaria per il principio di indeterminazione.

Dimostriamo, infatti, che l'energia dell'oscillatore armonico ha un minimo $\neq 0$. Data una $\psi \in D(H)$ generica, il valor medio di H è dato da:

$$\begin{aligned} (\psi, H\psi) &= \left(\psi, \frac{P^2}{2m} \psi \right) + m \frac{\omega^2}{2} (\psi, X^2 \psi) = \\ &= \frac{1}{2m} \langle P^2 \rangle_\psi + \frac{m\omega^2}{2} \langle X^2 \rangle_\psi = \\ &\stackrel{(a)}{=} \frac{1}{2m} [(\Delta P)_\psi^2 + \langle P \rangle_\psi^2] + \frac{m\omega^2}{2} [(\Delta X)_\psi^2 + \langle X \rangle_\psi^2] \\ &\stackrel{(b)}{\geq} \frac{(\Delta P)_\psi^2}{2m} + \frac{m\omega^2}{2} (\Delta X)_\psi^2 \stackrel{(c)}{\geq} \frac{(\Delta P)_\psi^2}{2m} + \frac{m\omega^2}{2} \frac{\hbar^2}{4(\Delta P)_\psi^2} \\ &\stackrel{(d)}{\geq} \frac{m\omega\hbar}{2 \cdot 2m} + \frac{m\omega^2}{2} \frac{\hbar^2 2}{4m\omega\hbar} = \frac{\hbar\omega}{2} \end{aligned}$$

Dove in (a) abbiamo usato la formula per la fluttuazione:

$$(\Delta A)_\psi^2 = \langle A^2 \rangle_\psi - \langle A \rangle_\psi^2 \Rightarrow \langle A^2 \rangle_\psi = (\Delta A)_\psi^2 + \langle A \rangle_\psi^2$$

In (b) si è maggiorata la somma rimuovendo i termini (positivi) $\langle P \rangle^2$ e $\langle X \rangle^2$, mentre in (c) si è applicato il principio di indeterminazione di Heisenberg, in modo da ottenere un'espressione *in una sola variabile* facile da minimizzare:

$$(\Delta X)_\psi^2 (\Delta P)_\psi^2 \geq \left(\frac{\hbar}{2}\right)^2 \Rightarrow (\Delta X)_\psi^2 \geq \frac{\hbar^2}{4(\Delta P)_\psi^2}$$

In (d), infine, si è maggiorata la somma con il suo *minimo*, per cui $(\Delta P)^2 = m\omega\hbar/2$. Infatti, posto $y \equiv (\Delta P)_\psi^2$ per semplicità di notazione, il minimo è dato da:

$$\begin{aligned} (\Delta P)_\psi^2 \equiv y &\Rightarrow \frac{d}{dy} \left[\frac{y}{2m} + \frac{m\omega^2}{2} \frac{\hbar^2}{4y} \right] \stackrel{!}{=} 0 \\ &\Rightarrow \frac{1}{2m} + \frac{m\omega^2\hbar^2}{8} \left(-\frac{1}{y^2} \right) = 0 \\ &\Rightarrow y^2 = \frac{m^2\omega^2\hbar^2}{4} \Rightarrow y = \frac{m\omega\hbar}{2} \end{aligned}$$

Notiamo che lo stato fondamentale dell'oscillatore armonico, per cui $H = \hbar\omega/2$ soddisfa il principio di indeterminazione con *l'uguaglianza*. In altre parole, in un sistema costituito da un oscillatore armonico nello stato fondamentale, si possono determinare posizione e momento con la *minima incertezza*.

4.10.1 Esercizio 5

L'Hamiltoniana di una particella di massa m in 1D è data da:

$$H = \frac{P^2}{2m} + AX^2 + BX + C; \quad A, B, C \in \mathbb{R}, \quad A > 0$$

Fisicamente ciò corrisponde ad un oscillatore armonico in un campo elettrico costante o in presenza di forza gravitazionale.

1. Mostrare che il sistema con una opportuna traslazione di parametro α diventa equivalente a un oscillatore armonico, con hamiltoniana H' , determinando α .
2. Si calcolino spettro e autofunzioni di H , denotando con $\psi_n(x)$ gli autostati dell'Hamiltoniana dell'oscillatore armonico H' ottenuta in 1, in rappresentazione $\{x\}$.
3. All'istante $t = 0$ lo stato della particella sia descritto da:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\phi_0\rangle + |\phi_1\rangle)$$

ove le $|\phi_n\rangle$ sono i ket corrispondenti al livello energetico n -esimo (ordinati in modo crescente) di H .

Si determini il valor medio di X per $t > 0$.

4. Si dimostri che l'operatore \mathcal{P}_1 così definito:

$$\mathcal{P}_1 \equiv \exp\left(i\frac{B}{2A\hbar}P\right) \mathcal{P} \exp\left(-i\frac{B}{2A\hbar}P\right)$$

con \mathcal{P} parità e P momento, commuta con H .

5. Si esprima \mathcal{P}_1 in funzione di H (dato che H costituisce un ICOC, e i due operatori sono compatibili per il punto precedente, i due devono essere funzionalmente dipendenti).

Soluzione

1. Una generica *traslazione spaziale* $U(\alpha)$ ha la forma:

$$U(\alpha) = \exp\left(-i\frac{\alpha}{\hbar}P\right)$$

In visuale di Heisenberg, gli operatori X e P sono *trasformati* da $U(\alpha)$ secondo le relazioni:

$$U^\dagger(\alpha)XU(\alpha) = X + \alpha \quad U^\dagger(\alpha)PU(\alpha) = P$$

Come ci si aspetta, una traslazione spaziale altera solo le posizioni e non il momento.

Scriviamo la relativa trasformazione dell'Hamiltoniana:

$$\begin{aligned} H' = U^\dagger(\alpha)HU(\alpha) &= \frac{P^2}{2m} + A(X + \alpha)^2 + B(x + \alpha) + C = \\ &= \frac{P^2}{2m} + AX^2 + (B + 2\alpha A)X + \alpha^2 A + B\alpha + C \end{aligned}$$

Vogliamo che la traslazione conduca all'Hamiltoniana di un oscillatore armonico, che nel caso generale ha la forma:

$$H = \frac{P^2}{2m} + \frac{m^2\omega^2}{2}X^2 + V_0$$

Dobbiamo allora imporre che il termine lineare in X di H' scompaia, ossia che $B + 2\alpha A = 0$. Ciò significa che abbiamo fissato α :

$$\alpha = -\frac{B}{2A}$$

E allora H' è data da:

$$H' = \frac{P^2}{2m} + A X^2 + \underbrace{\left(C - \frac{B^2}{4A}\right)}_{\equiv V_0}$$

2. Cerchiamo ora spettro e autofunzioni di H . $U(\alpha)$ è una trasformazione unitaria, che perciò non intacca gli spettri, e quindi $\sigma(H) = \sigma(H')$. Ma $\sigma(H')$ è lo spettro di un oscillatore armonico, che è unicamente determinato una volta che sappiamo la sua ω . Basta allora invertire:

$$A = \frac{m\omega^2}{2} \Rightarrow \omega = \sqrt{\frac{2A}{m}}$$

E quindi:

$$\sigma(H) = \sigma(H') = \hbar\omega \left(\mathbb{N} + \frac{1}{2} \right) + V_0$$

Occupiamoci ora delle autofunzioni. Siano $|\psi_n\rangle$ gli autostati di H' , che soddisfano l'equazione agli autovalori per H' :

$$H' |\psi_n\rangle = \left[\hbar\omega \left(n + \frac{1}{2} \right) + V_0 \right] |\psi_n\rangle$$

Ma $H' = U^\dagger(\alpha) H U(\alpha)$. Moltiplicando a sinistra per $U(\alpha)$ entrambi i membri, otteniamo:

$$\cancel{U(\alpha)} U^\dagger(\alpha) H U(\alpha) |\psi_n\rangle = \left[\hbar\omega \left(n + \frac{1}{2} \right) + V_0 \right] U(\alpha) |\psi_n\rangle$$

Perciò $|\phi_n\rangle \equiv U(\alpha) |\psi_n\rangle$ è soluzione dell'equazione agli autovalori per H . Passando allora in rappresentazione $\{x\}$, in notazione di Dirac si ha:

$$\psi_n(x) = \langle x | \psi_n \rangle \quad \phi_n(x) = \langle x | \phi_n \rangle$$

E gli autostati di H , che denotiamo con $\phi_n(x)$, si ottengono traslando le $\psi_n(x)$ tramite $U(\alpha)$:

$$\phi_n(x) = \langle x | U(\alpha) | \psi_n \rangle = \psi_n(x - \alpha)$$

ossia gli autostati di H sono le funzioni di Hermite traslate di $\alpha = -B/(2A)$.

3. Vogliamo ora calcolare l'evoluto temporale del valor medio di X , ossia:

$$\langle X \rangle_{\psi(t)} = \langle \psi | \exp \left(\frac{itH}{\hbar} \right) X \exp \left(-\frac{itH}{\hbar} \right) | \psi \rangle \quad |\psi\rangle = \frac{1}{\sqrt{2}} (|\phi_0\rangle + |\phi_1\rangle) \quad (4.72)$$

Per fare i conti, conviene passare ad H' , che è collegata ad H da una traslazione. Partiamo notando che traslazione temporale $U(\alpha)$ ed evoluzione temporale $T(t)$ sono *indipendenti*, e commutano tra loro: $[U(\alpha), T(t)] = 0$. Perciò possiamo applicare le due trasformazioni nell'ordine che vogliamo. Per esempio, in visuale di Schrödinger, lo stato *trasformato* $|\psi'\rangle$ si ottiene:

$$|\psi'\rangle = T(t)U(\alpha) |\psi\rangle = U(\alpha)T(t) |\psi\rangle$$

Ma allora, passando in visuale di Heisenberg, possiamo traslare l'operatore di evoluzione temporale, oppure evolvere l'operatore di traslazione. Scegliamo la prima via (più semplice). Avremo allora:

$$\exp\left(\pm \frac{i}{\hbar} t H'\right) = U(\alpha)^\dagger \exp\left(\pm \frac{i}{\hbar} t H\right) U(\alpha)$$

Invertendo, in modo da ottenere la sostituzione da fare in (4.72), giungiamo a:

$$\exp\left(\pm \frac{itH}{\hbar}\right) = U(\alpha) \exp\left(\pm \frac{itH'}{\hbar}\right) U(\alpha)^\dagger$$

Sostituendo allora in (4.72):

$$\begin{aligned} (4.72) &= \langle \psi | U(\alpha) \exp\left(\frac{itH'}{\hbar}\right) \underbrace{U(\alpha)^\dagger X U(\alpha)}_{X+\alpha} \exp\left(-\frac{itH'}{\hbar}\right) U^\dagger(\alpha) | \psi \rangle = \\ &\stackrel{(a)}{=} \alpha + \langle \psi | U(\alpha) \exp\left(\frac{itH'}{\hbar}\right) X \exp\left(-\frac{itH'}{\hbar}\right) U^\dagger(\alpha) | \psi \rangle = \end{aligned} \quad (4.73)$$

dove in (a) abbiamo spostato fuori il parametro α , che è lasciato invariato dalle esponenziali (dato che quelle a sinistra annullano quelle a destra). Cominciamo a calcolare l'espressione "a ritroso", partendo da destra, ossia da $U^\dagger(\alpha) | \psi \rangle$.

Sappiamo che:

$$| \phi_n \rangle = U(\alpha) | \psi_n \rangle \Rightarrow U(\alpha)^\dagger | \phi_n \rangle = | \psi_n \rangle$$

E allora:

$$U(\alpha)^\dagger | \psi \rangle = U(\alpha)^\dagger \left(\frac{| \psi_0 \rangle + | \psi_1 \rangle}{2} \right) = \frac{| \psi_0 \rangle + | \psi_1 \rangle}{2}$$

Per $\langle \psi | U(\alpha)$ basta prendere il coniugato di quanto appena trovato:

$$\langle \psi | U(\alpha) = \frac{\langle \psi_0 | + \langle \psi_1 |}{2}$$

Riportando il tutto in (4.73):

$$(4.73) = \alpha + \left(\frac{\langle \psi_0 | + \langle \psi_1 |}{\sqrt{2}} \right) \exp\left(\frac{itH'}{\hbar}\right) X \exp\left(-\frac{itH'}{\hbar}\right) \left(\frac{| \psi_0 \rangle + | \psi_1 \rangle}{\sqrt{2}} \right) \quad (4.74)$$

Occupiamoci ora del termine evidenziato. Le $| \psi_0 \rangle$ e $| \psi_1 \rangle$ sono già autostati di H' , e perciò possiamo subito espandere l'esponenziale:

$$\begin{aligned} \exp\left(-\frac{itH'}{\hbar}\right) \left(\frac{| \psi_0 \rangle + | \psi_1 \rangle}{\sqrt{2}} \right) &= \frac{1}{\sqrt{2}} \left[\exp\left(-\frac{i}{\hbar} t \mathcal{E}_0\right) | \psi_0 \rangle + \exp\left(-\frac{i}{\hbar} t \mathcal{E}_1\right) | \psi_1 \rangle \right] = \\ &= \frac{1}{\sqrt{2}} \left[\exp\left(-\frac{i}{\hbar} \left(\hbar \frac{\omega}{2} + V_0 \right) t\right) | \psi_0 \rangle + \exp\left(-\frac{i}{\hbar} \left(\frac{3}{2} \hbar \omega + V_0 \right) t\right) | \psi_1 \rangle \right] \end{aligned}$$

Dato che:

$$\mathcal{E}_n = \hbar\omega \left(n + \frac{1}{2}\right) + V_0$$

Tutto ciò si può scrivere in maniera più semplice estraendo dagli esponenziali un termine $\exp\left(-i\left(\frac{1}{2}\omega t + \frac{V_0}{\hbar}\right)\right)$:

$$\frac{1}{\sqrt{2}} \exp\left(-i\left(\frac{1}{2}\omega t + \frac{V_0}{\hbar}\right)\right) [|\psi_0\rangle + e^{-i\omega t} |\psi_1\rangle]$$

Prendendo il complesso coniugato, abbiamo calcolato anche la rispettiva parte a sinistra. Sostituendo nell'espressione generale, e semplificando gli esponenziali comuni raccolti nel bra e nel ket (che sono solamente fattori moltiplicativi, e quindi li possiamo spostare):

$$(4.74) = \alpha + \frac{1}{2} [\langle\psi_0| + e^{i\omega t} \langle\psi_1|] X [|\psi_0\rangle + e^{-i\omega t} |\psi_1\rangle] \quad (4.75)$$

Le $|\psi_0\rangle$ e $|\psi_1\rangle$ non sono però autostati di X . Tuttavia, possiamo riscrivere X come combinazione di a e a^\dagger , di cui conosciamo l'azione sulle $|\psi_n\rangle$. Usiamo allora l'identità:

$$X = \left(\frac{\hbar}{2m\omega}\right)^{\frac{1}{2}} (a + a^\dagger)$$

Gli operatori a e a^\dagger rispettivamente *abbassano* e *alzano* gli autostati. Perciò, dei 4 braket che compaiono espandendo la (4.75), ne restano solo 2, come si dimostra usando le relazioni di normalizzazione ricavate dalla formula in (4.71) e l'ortonormalità delle $|\psi_n\rangle$:

$$\begin{aligned} \langle\psi_0| a |\psi_1\rangle &= \langle\psi_0| \sqrt{1}\psi_0\rangle = 1 & \langle\psi_0| a^\dagger |\psi_1\rangle &= \langle\psi_0| \sqrt{2}\psi_2\rangle = 0 \\ \langle\psi_1| a |\psi_0\rangle &= \langle\psi_1| 0\rangle = 0 & \langle\psi_1| a^\dagger |\psi_0\rangle &= \langle\psi_1| \sqrt{1}\psi_1\rangle = 1 \end{aligned}$$

E perciò arriviamo a:

$$\begin{aligned} (4.75) &= \alpha + \frac{1}{2} \left(\frac{\hbar}{2m\omega}\right)^{\frac{1}{2}} [\langle\psi_0| + e^{i\omega t} \langle\psi_1|] (a + a^\dagger) [|\psi_0\rangle + e^{-i\omega t} |\psi_1\rangle] = \\ &= \alpha + \frac{1}{2} \left(\frac{\hbar}{2m\omega}\right)^{\frac{1}{2}} [\langle\psi_0| a |\psi_1\rangle e^{-i\omega t} + \langle\psi_1| a^\dagger |\psi_0\rangle e^{i\omega t}] = \\ &= \alpha + \left(\frac{\hbar}{2m\omega}\right)^{\frac{1}{2}} \left[\frac{e^{i\omega t} + e^{-i\omega t}}{2}\right] = \alpha + \left(\frac{\hbar}{2m\omega}\right)^{\frac{1}{2}} \cos(\omega t) \end{aligned}$$

Come ci si potrebbe aspettare, il valor medio della misura *oscilla* di moto armonico attorno ad $x = \alpha$, in analogia con il caso classico.

4. Notiamo che l'operatore \mathcal{P}_1 così definito non è altro che la *parità* (*anti*)*traslata*³¹ di $\alpha = -B/(2A)$:

$$\mathcal{P}_1 \equiv \exp\left(\frac{i}{\hbar} \frac{B}{2A} P\right) \mathcal{P} \exp\left(-\frac{i}{\hbar} \frac{B}{2A} P\right) = U(\alpha) \mathcal{P} U(\alpha)^\dagger$$

³¹^Si noti che la posizione degli aggiunti è invertita rispetto alla traslazione *diretta*. Intuitivamente, per applicare \mathcal{P} a H , che non è "centrato" in 0 ma in α , dobbiamo prima antitraslarlo, poi calcolare \mathcal{P} e poi ritraslare il risultato.

Vogliamo verificare che $[\mathcal{P}_1, H]$, e che quindi il valor medio di \mathcal{P}_1 sia indipendente dall'evoluzione temporale del sistema. Procediamo per calcolo diretto:

$$\begin{aligned} [\mathcal{P}_1, H] &= [U(\alpha)\mathcal{P}U(\alpha)^\dagger, H] = U(\alpha)\mathcal{P}U(\alpha)^\dagger H - HU(\alpha)\mathcal{P}U(\alpha)^\dagger = \\ &\stackrel{(a)}{=} U(\alpha)\mathcal{P}U(\alpha)^\dagger \underbrace{HU(\alpha)U^\dagger(\alpha)}_{H'} - U(\alpha)\underbrace{U^\dagger(\alpha)HU(\alpha)}_{H'}\mathcal{P}U(\alpha)^\dagger = \\ &= U(\alpha)\mathcal{P}H'U(\alpha)^\dagger - U(\alpha)H'\mathcal{P}U(\alpha)^\dagger = U(\alpha)[\mathcal{P}, H']U(\alpha)^\dagger \stackrel{(b)}{=} 0 \end{aligned}$$

dove in (a) abbiamo moltiplicato opportunamente per $U(\alpha)U(\alpha)^\dagger = \mathbb{I}$ in modo da passare a H' ed evidenziare poi un commutatore, e in (b) abbiamo usato il fatto che:

$$[\mathcal{P}, H'] = 0$$

in quanto H' è l'hamiltoniana di un oscillatore armonico, il cui potenziale è *pari*.

5. Dato che H è un ICOC, e \mathcal{P}_1 è compatibile con H per il punto precedente, allora $\mathcal{P}_1 = f(H)$.

Partiamo scrivendo l'operatore parità, che possiamo poi *controtraslare* per ottenere $\mathcal{P}_1 = U(\alpha)\mathcal{P}U(\alpha)^\dagger$. Dall'espressione degli autostati dell'oscillatore armonico, dati dalle funzioni di Hermite:

$$\psi_n(x) \propto \left(x - \frac{d}{dx}\right)^n e^{-\alpha x^2}$$

Notiamo che le $\psi_n(x)$ con n pari sono funzioni pari, e quelle con n dispari sono dispari. Possiamo allora ottenere la forma operatoriale di \mathcal{P} :

$$\mathcal{P}|\psi_n\rangle = (-1)^n |\psi_n\rangle = (-1)^N |\psi_n\rangle$$

Basta quindi ricavare N in funzione di H . Basta allora invertire la relazione:

$$H' = \hbar\omega \left(N + \frac{1}{2}\right) + V_0 \Rightarrow N = \frac{H' - V_0}{\hbar\omega} - \frac{1}{2}$$

E perciò otteniamo:

$$\mathcal{P} = (-1)^{\left[\frac{H' - V_0}{\hbar\omega} - \frac{1}{2}\right]}$$

Troviamo infine \mathcal{P}_1 *controtraslando*:

$$\begin{aligned} \mathcal{P}_1 &= U(\alpha)\mathcal{P}U(\alpha)^\dagger = (-1)^{\left[\frac{U(\alpha)H'U^\dagger(\alpha) - V_0}{\hbar\omega} - \frac{1}{2}\right]} = \\ &= (-1)^{\left[\frac{H - V_0}{\hbar\omega} - \frac{1}{2}\right]} \end{aligned}$$

4.11 Potenziale centrale

Occupiamoci ora dello studio di potenziali $U(r)$ a simmetria centrale³², ossia che dipendono solo dalla distanza r da un *centro* fissato. Tale caso è infatti di grande rilevanza fisica, dato che ci permetterà di creare un modello dettagliato dell'*atomo di idrogeno*.

Partiamo considerando, in generale, un sistema di due particelle distinguibili in \mathbb{R}^3 (che saranno protone ed elettrone nel caso dell'atomo di idrogeno) di massa rispettivamente m_1 e m_2 , che interagiscono tra loro mediante un potenziale centrale $U(r)$, dove $r = |\vec{x}_1 - \vec{x}_2|$ è la distanza che le separa.

L'Hamiltoniana **classica** è data da:

$$H = \frac{\vec{p}_1^2}{2m_1} + \frac{\vec{p}_2^2}{2m_2} + U(r) \quad r = |\vec{x}_1 - \vec{x}_2|$$

Quantisticamente, il sistema composto ha stati nel prodotto tensore degli spazi di Hilbert dei costituenti, ossia $\mathcal{H} = L^2(\mathbb{R}^3, d^3x_1) \otimes L^2(\mathbb{R}^3, d^3x_2)$, e l'Hamiltoniana si ottiene sostituendo a \vec{p} il relativo operatore quantistico $-i\hbar\nabla$ (in rappresentazione $\{\vec{x}\}$):

$$H_{12} = -\frac{\hbar^2}{2m_1}\Delta_1 - \frac{\hbar^2}{2m_2}\Delta_2 + U(r) \quad r = |\vec{x}_1 - \vec{x}_2|$$

Per poter sfruttare la simmetria sferica del sistema, conviene riscrivere il tutto in funzione delle variabili \vec{R} della posizione del Centro di Massa (CM), e del vettore \vec{x} della distanza relativa tra le particelle:

$$\vec{R} = \frac{m_1\vec{x}_1 + m_2\vec{x}_2}{m_1 + m_2}; \quad \vec{x} = \vec{x}_1 - \vec{x}_2; \quad \mathcal{H} \cong L^2(\mathbb{R}^3, d^3R) \otimes L^2(\mathbb{R}^3, d^3x)$$

Da cui otteniamo:

$$H_{12} = \underbrace{-\frac{\hbar^2}{2M}\Delta_R}_{H_R} - \underbrace{\frac{\hbar^2}{2m}\Delta_x}_{H} + U(r) \quad M = m_1 + m_2; \quad m = \frac{m_1m_2}{m_1 + m_2} \quad (4.76)$$

dove riconosciamo in m la *massa ridotta* del sistema.

Esaminiamo i domini di autoaggiuntezza. H_R è l'Hamiltoniana di una particella libera, di massa M e centrata al CM, ed è definita in:

$$D(H_R) = \{\psi \in L^2(\mathbb{R}^3, d^3R) \mid \vec{p}_R^2 \tilde{\psi}(\vec{p}_R) \in L^2(\mathbb{R}^3, d^3p_R)\}$$

Per H il caso è più delicato, ed è necessario fare ipotesi sulla natura del potenziale $U(r)$. Si può dimostrare (per il teorema di *Kato-Rellich*), che se $U(r) \underset{r \rightarrow 0}{\sim} 1/r^\alpha$ per $\alpha < 3/2$ e $U(r) \underset{r \rightarrow \infty}{\sim} 0$, con $U(r) \in L^2([0, 1], r^2 dr)$ allora H è autoaggiunto sullo stesso dominio del caso della particella libera, ossia:

$$D(H) = \{\psi \in L^2(\mathbb{R}^3, d^3x) \mid \vec{p}^2 \tilde{\psi}(\vec{p}) \in L^2(\mathbb{R}^3, d^3p)\}$$

*Moto di due
particelle nel sdr
del centro di
massa*

*Teorema di
Kato-Rellich*

³²^Alcune fonti a riguardo sono disponibili in [5]

Scriviamo allora l'equazione di Schrödinger stazionaria per il sistema:

$$\left[-\frac{\hbar^2}{2M}\Delta_R - \frac{\hbar^2}{2m}\Delta_x + U(r) \right] \Psi(\vec{R}, \vec{x}) = \mathcal{E}_{tot} \Psi(\vec{R}, \vec{x}) \quad (4.77)$$

Dato che ogni termine dell'Hamiltoniana dipende rispettivamente solo da \vec{R} o da \vec{x} , e mai da entrambi contemporaneamente, possiamo cercare una soluzione $\Psi(\vec{R}, \vec{x})$ *fattorizzata*:

$$\Psi(\vec{R}, \vec{x}) = \varphi(\vec{R})\psi(\vec{x})$$

Sostituendo in (4.77) otteniamo:

$$-\frac{\hbar^2}{2M}(\Delta_R\varphi)\psi + \varphi \left[-\frac{\hbar^2}{2m}\Delta_x\psi + U(r)\psi \right] = \mathcal{E}\varphi\psi$$

Dividendo per $\varphi\psi$:

$$\underbrace{-\frac{\hbar^2}{2M} \frac{\Delta_R\varphi}{\varphi}}_{\mathcal{E}_{CM}} + \underbrace{\left[-\frac{\hbar^2}{2m} \frac{\Delta\psi}{\psi} + U(r) \right]}_{\mathcal{E}} = \mathcal{E}_{tot}$$

Spezzando l'energia in $\mathcal{E}_{tot} = \mathcal{E} + \mathcal{E}_{CM}$, otteniamo due equazioni distinte. In particolare, la parte che deriva da H_R genera l'equazione di Schrödinger di una particella libera, che sappiamo trattare:

$$-\frac{\hbar^2}{2M}\Delta_R\varphi = \mathcal{E}_{CM}\varphi$$

Occupiamoci allora della parte relativa ad H , che è ben più interessante data la presenza del potenziale centrale $U(r)$.

Poiché $U(r)$ ha simmetria sferica, notiamo che H è invariante per rotazioni:

*Simmetria sferica
di H*

$$\exp\left(-i\frac{\varphi}{\hbar}\vec{L} \cdot \vec{n}\right) H \exp\left(i\frac{\varphi}{\hbar}\vec{L} \cdot \vec{n}\right) = H \Rightarrow [H, \vec{L}] = 0; \quad [H, \vec{L}^2] = 0$$

Pertanto H , \vec{L}^2 , L_3 sono tre osservabili compatibili tra loro, e perciò ammettono una base di autoket comuni $|\mathcal{E}, l, m\rangle$. Spezziamo allora il problema in parte *radiale* e *angolare*, tramite l'isomorfismo “delle coordinate sferiche” $L^2(\mathbb{R}^3, d^3x) \cong L^2(\mathbb{R}_+, r^2 dr) \otimes L^2(S^2, d\Omega)$.

Sappiamo che in $L^2(S^2, d\Omega)$, $\{\vec{L}^2, L_z\}$ costituiscono un ICOC. Sapremo allora che $\{H, \vec{L}^2, L_3\}$ è un ICOC per il sistema *ridotto* con hamiltoniana H se, una volta fissati i valori di \vec{L}^2 e L_z , tramite i loro autovalori l e m , l'hamiltoniana $H(l, m)$ in $L^2(\mathbb{R}_+, r^2 dr)$ risulta avere uno **spettro non degenero**. In altre parole, vogliamo verificare che fissati l e m , per un dato \mathcal{E} ci sia un *solo* autoket comune $|\mathcal{E}, l, m\rangle$.

*Autoket comuni
 $|\mathcal{E}, l, m\rangle$*

Dimostriamolo. Partiamo riscrivendo \vec{L}^2 in termini di \vec{P}^2 (che compare in H), in modo da poter cercare agilmente gli autostati comuni $|\mathcal{E}, l, m\rangle$. Con la convenzione di Einstein, sfruttiamo l'identità:

$$\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl} \quad (4.78)$$

Per cui:

$$\begin{aligned}
\vec{L}^2 &= L_i L_i = (\epsilon_{ijk} X_j P_k)(\epsilon_{ilm} X_l P_m) \stackrel{(4.78)}{=} (\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) X_j P_k X_l P_m = \\
&\stackrel{(a)}{=} \delta_{jl} \delta_{km} X_j (X_l P_k + \underbrace{[P_k, X_l]}_{-i\hbar\delta_{kl}} P_m) - \delta_{jm}\delta_{kl} X_j P_k (P_m X_l + \underbrace{[X_l, P_k]}_{i\hbar\delta_{lm}}) = \\
&\stackrel{(b)}{=} \vec{X}^2 \vec{P}^2 - i\hbar(\vec{X} \cdot \vec{P}) - (\vec{X} \cdot \vec{P}) \underbrace{(\vec{P} \cdot \vec{X})}_{\vec{X} \cdot \vec{P} - 3i\hbar} - i\hbar(\vec{X} \cdot \vec{P}) = \\
&\stackrel{(c)}{=} \vec{X}^2 \vec{P}^2 - i\hbar\vec{X} \cdot \vec{P} - (\vec{X} \cdot \vec{P})^2 + 3i\hbar\vec{X} \cdot \vec{P} - i\hbar\vec{X} \cdot \vec{P} = \\
&= \vec{X}^2 \vec{P}^2 - (\vec{X} \cdot \vec{P})^2 + i\hbar\vec{X} \cdot \vec{P} \tag{4.79}
\end{aligned}$$

In (a) abbiamo usato i commutatori per portare le X_j e X_l vicine, e analogamente P_k e P_m , in modo da ottenere i successivi quadrati e prodotti scalari.

In (b), espandendo i prodotti, notiamo che³³:

$$\delta_{jl}\delta_{km}\delta_{kl} = \delta_{jl}\delta_{lm} = \delta_{jm}$$

E quindi, per esempio:

$$\delta_{jl}\delta_{km}\delta_{kl}X_jP_m = \delta_{jm}X_jP_m = X_jP_j = \vec{X} \cdot \vec{P}$$

Analogamente vale:

$$\delta_{jm}\delta_{kl}\delta_{lm} = \delta_{jm}\delta_{km} = \delta_{jk}$$

In (c) “invertiamo” il prodotto scalare tramite i commutatori. In notazione di Einstein, e calcolando alla fine le somme:

$$\vec{X} \cdot \vec{P} = X_i P_i = [X_i, P_i] + P_i X_i = 3i\hbar + \vec{P} \cdot \vec{X} \Rightarrow \vec{P} \cdot \vec{X} = \vec{X} \cdot \vec{P} - 3i\hbar$$

Invertendo la (4.79) otteniamo \vec{P}^2 in termini di \vec{L}^2 :

\vec{P}^2 in funzione di \vec{L}^2

$$\vec{P}^2 = \vec{X}^{-2}[(\vec{X} \cdot \vec{P})^2 + \vec{L}^2 - i\hbar\vec{X} \cdot \vec{P}] \tag{4.80}$$

che è pronto per essere sostituito in H . Prima di farlo, però, conviene passare in *coordinate sferiche*. Facciamolo un termine alla volta, ricordando il passaggio da $\vec{\nabla} = (\partial_x, \partial_y, \partial_z)^t$ a $(\partial_r, \partial_\theta, \partial_\varphi)^t$ visto in (4.46).

$$\vec{X} \cdot \vec{P} = -i\hbar\vec{x} \cdot \vec{\nabla} = -i\hbar r \frac{\partial}{\partial r}$$

³³Si può pensare a una δ_{jm} come tale da “identificare” gli indici. Il prodotto di più delta identifica tra loro più indici contemporaneamente. Un trucco per risolvere le uguaglianze è la “semplificazione” data dalla contrazione di indici uguali confinanti $\delta_{kl}\delta_{lm} = \delta_{km}$. Si ricorda inoltre che la delta di Kronecker è *simmetrica*: $\delta_{km} = \delta_{mk}$.

Dato che:

$$\begin{aligned}
\vec{x} \cdot \vec{\nabla} &= r \sin \theta \cos \varphi \left(\sin \theta \cos \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \cos \varphi \frac{\partial}{\partial \theta} - \frac{\sin \varphi}{r \sin \theta} \frac{\partial}{\partial \varphi} \right) + \\
&+ r \sin \theta \sin \varphi \left(\sin \theta \sin \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \sin \varphi \frac{\partial}{\partial \theta} + \frac{\cos \varphi}{r \sin \theta} \frac{\partial}{\partial \varphi} \right) + \\
&+ r \cos \theta \left(\cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \right) = \\
&= r [\sin^2 \theta \cos^2 \varphi + \sin^2 \theta \sin^2 \varphi + \cos^2 \theta] \frac{\partial}{\partial r} = r \frac{\partial}{\partial r}
\end{aligned}$$

Ricaviamo poi:

$$(\vec{X} \cdot \vec{P})^2 = \left(-i\hbar r \frac{\partial}{\partial r} \right) \left(-i\hbar r \frac{\partial}{\partial r} \right) = -\hbar^2 r \frac{\partial}{\partial r} r \frac{\partial}{\partial r}$$

E sostituendo tutto ciò in (4.80) otteniamo:

$$\vec{P}^2 = \frac{1}{r^2} \left[-\hbar^2 r \frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \vec{L}^2 - \hbar^2 r \frac{\partial}{\partial r} \right] = \frac{1}{r^2} \left[-\hbar^2 r \left(\frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{\partial}{\partial r} \right) + \vec{L}^2 \right] \quad (4.81)$$

Possiamo semplificare ulteriormente questa espressione, notando che:

$$r \frac{\partial}{\partial r} = \frac{\partial}{\partial r} r - 1$$

Come tutte le uguaglianze di operatori, il modo corretto per dimostrarla sta nell'applicarla ad una generica funzione $f(r)$:

$$\begin{aligned}
r \frac{\partial}{\partial r} f(r) &= r f'(r) \\
\left(\frac{\partial}{\partial r} r - 1 \right) f(r) &= \frac{\partial}{\partial r} (r f(r)) - f(r) = f(r) + r f'(r) - f(r) = r f'(r)
\end{aligned}$$

Il termine evidenziato in (4.81) diviene perciò:

$$\frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{\partial}{\partial r} = \frac{\partial}{\partial r} \left(\frac{\partial}{\partial r} r - 1 \right) + \frac{\partial}{\partial r} = \frac{\partial^2}{\partial r^2} r - \frac{\partial}{\partial r} + \frac{\partial}{\partial r} = \frac{\partial^2}{\partial r^2} r$$

E infine otteniamo:

$$\vec{P}^2 = \frac{1}{r^2} \left[-\hbar^2 r \frac{\partial^2}{\partial r^2} r + \vec{L}^2 \right] \quad (4.82)$$

Sostituendo nell'Hamiltoniana H (4.76), giungiamo a:

$$H = \frac{1}{2m} \left[-\frac{\hbar^2}{r} \frac{\partial^2}{\partial r^2} r + \frac{\vec{L}^2}{r^2} \right] + U(r) \quad (4.83)$$

Per ottenere una qualche intuizione sulla natura *fisica* del termine evidenziato confrontiamo quanto abbiamo ottenuto con il caso classico.

Confronto con il caso classico

Come è noto dai corsi di Fisica, un punto di massa m posto in un potenziale centrale $U(r)$ (che ipotizziamo centrato nell'origine) sperimenta una forza $\vec{F} = -\vec{\nabla}U(r)$ diretta verso O . Il momento torcente rispetto al polo O è allora nullo: $\vec{\tau} = \vec{r} \times \vec{F} = 0$, e perciò il momento angolare è costante:

$$\frac{d\vec{L}}{dt} = \vec{\tau} = 0$$

Scomponendo allora l'energia \mathcal{E} della particella nei contributi dovuti al moto $v_r \hat{r}$ verso O , e quello $v_\perp \hat{\theta}$ ad esso perpendicolare, si ha:

$$\mathcal{E} = \frac{1}{2}mv_r^2 + \frac{1}{2}mv_\perp^2 + U(r) = \frac{p^2}{2m} + U(r)$$

dove $\vec{p} = m\vec{v}$ è il momento della particella. Possiamo riscrivere \mathcal{E} utilizzando il momento angolare $|\vec{L}| = mrv_\perp$, e introducendo il *momento radiale* $p_r = mv_r$:

$$\mathcal{E} = \frac{1}{2m} \left(p_r^2 + \frac{L^2}{r^2} \right) \Rightarrow p^2 = p_r^2 + \frac{1}{r^2}L^2 \quad (4.84)$$

Confrontando (4.84) con (4.82) scopriamo che il termine evidenziato è proprio il quadrato del *momento radiale* P_r^2 .

Nota: il momento radiale P_r non si ottiene semplicemente dalla “radice quadrata” di P_r^2 , ma è dato da:

$$P_r = -i\hbar \frac{1}{r} \frac{\partial}{\partial r} r$$

Infatti:

$$P_r^2 = \left(-i\hbar \frac{1}{r} \frac{\partial}{\partial r} r \right) \left(-i\hbar \frac{1}{r} \frac{\partial}{\partial r} r \right) = -\hbar^2 \frac{1}{r} \frac{\partial^2}{\partial r^2} r$$

Definiamo allora il **momento radiale** $P_r = -i\hbar \frac{1}{r} \frac{\partial}{\partial r} r$ in $L^2(\mathbb{R}_+, r^2 dr)$, e mostriamo che il suo quadrato $P_r^2 = -\hbar^2 \frac{1}{r^2} \frac{\partial^2}{\partial r^2} r$ è autoaggiunto nel dominio:

Momento radiale quadro P_r^2 e sua autoaggiuntezza

$$D(P_r^2) = \{ \psi \in L^2(\mathbb{R}_+, r^2 dr), \psi \text{ regolari con } \lim_{r \rightarrow 0} r\psi(r) = 0 \}$$

Infatti, prendendo:

$$\varphi \in D((P_r^2)^\dagger), \psi \in D(P_r^2)$$

e calcolando gli elementi di matrice (dividendo per le costanti in modo da alleggerire la notazione), si ha che:

$$\left(\varphi, \frac{P_r^2 \psi}{-\hbar^2} \right) = \int_0^{+\infty} r^2 dr \varphi^* \frac{1}{r} \frac{\partial^2}{\partial r^2} (r\psi(r)) \quad (4.85)$$

dove l'integrazione su $r^2 dr$ deriva dalla componente radiale della misura delle coordinate sferiche:

$$dV = r^2 dr \wedge \sin \theta d\theta \wedge d\varphi$$

Integriamo per parti due volte l'espressione (4.85), per esempio tramite il *metodo tabulare*:

$r\varphi^*$	$\partial_{r^2}(r\psi)$
$-\partial_r(r\varphi^*)$	$\partial_r(r\psi)$
$\partial_{r^2}(r\varphi^*)$	$r\psi$

Tabella 4.3: Metodo tabulare per l'integrazione per parti

Da cui:

$$(4.85) = (r\varphi^*) \frac{\partial}{\partial r}(r\psi) \Big|_0^{+\infty} - \frac{\partial}{\partial r}(r\varphi^*) r\psi \Big|_0^{+\infty} + \int_0^{+\infty} dr (r\psi) \frac{\partial^2}{\partial r^2}(r\varphi^*) \quad (4.86)$$

Riconosciamo nell'ultimo termine il prodotto scalare tra $P_r^2/(-\hbar^2)$ applicato a φ^* e ψ , come desiderato. Inoltre, essendo $\varphi, \psi \in L^2$, sappiamo che $\varphi^*(+\infty) \rightarrow 0$, e $\psi(+\infty) \rightarrow 0$, e perciò:

$$(4.86) = - \left(r\varphi^* \frac{\partial}{\partial r}(r\psi) \right) (0) + \left(\frac{\partial}{\partial r}(r\varphi^*) r\psi \right) (0) + \left(\frac{P_r^2}{-\hbar^2} \varphi, \psi \right)$$

Poiché $\psi \in D(P_r^2)$, vale $r\psi \xrightarrow{r \rightarrow 0} 0$, e perciò il termine giallo è nullo. Se vogliamo che P_r^2 coincida con il suo aggiunto, dobbiamo imporre che anche il termine evidenziato si annulli. Ma non abbiamo alcuna condizione su $\partial_r(r\psi)$, e quindi dobbiamo imporre per forza:

$$\lim_{r \rightarrow 0} r\varphi(0) = 0$$

Osserviamo che in questo modo si impone la stessa condizione sia per il dominio di P_r^2 che per quello del suo aggiunto $(P_r^2)^\dagger$ e aggiunto, da cui:

$$D(P_r^2) = D((P_r^2)^\dagger)$$

Ciò, assieme all'identità appena verificata per gli elementi di matrice:

$$\left(\varphi, \frac{P_r^2 \psi}{-\hbar^2} \right) = \left(\frac{P_r^2 \varphi}{-\hbar^2}, \psi \right)$$

dimostra l'autoaggiuntezza di P_r^2 .

L'Hamiltoniana *ridotta* H del sistema è allora:

$$H = \frac{1}{2m} \left[P_r^2 + \frac{\vec{L}}{r^2}(\theta, \varphi) \right] + U(r) \quad \lim_{r \rightarrow 0} r\psi(r) = 0$$

Scriviamo quindi l'equazione di Schrödinger stazionaria:

$$\left[\frac{1}{2m} \left(\vec{P}_r^2 + \frac{\vec{L}^2}{r^2}(\theta, \varphi) \right) + U(r) \right] \psi(r, \theta, \varphi) = \mathcal{E} \psi(r, \theta, \varphi) \quad (4.87)$$

Vogliamo ora trovare esplicitamente gli autoket comuni $|\epsilon, l, m\rangle$. Notiamo che termini radiali e angolari sono *separati*, e ciò suggerisce di cercare soluzioni *fattorizzabili*. Di queste, abbiamo già ricavato la forma degli autoket angolari $|l, m\rangle$, che sono proprio le armoniche sferiche: $Y_l^m(\theta, \varphi)$. Perciò fattorizziamo le autofunzioni comuni ψ come:

$$\psi_{\epsilon lm}(r, \theta, \varphi) = h_{\epsilon lm}(r) Y_l^m(\theta, \varphi)$$

Data la simmetria sferica del sistema, ci aspettiamo che l'autofunzione radiale $h(r)$ *non* dipenda dall'autovalore m - se così fosse avremmo una direzione (\hat{z}) distinguibile dalle altre. In effetti, H dipende solo da operatori radiali e da \vec{L}^2 , e non da L_3 . Esplicitamente, se proviamo ad abbassare/alzare gli autovalori di m in ψ tramite L_{\pm} , $h_{\epsilon lm}$ resta invariata:

$$\begin{aligned} L_{\pm} \psi_{\epsilon lm}(\vec{r}) &= \hbar \sqrt{l(l+1) - m(m \pm 1)} \psi_{\epsilon l, m \pm 1}(\vec{r}) \\ L_{\pm} \psi_{\epsilon lm}(\vec{r}) &= h_{\epsilon lm}(r) L_{\pm} Y_l^m(\theta, \varphi) = \hbar \sqrt{l(l+1) - m(m \pm 1)} h_{\epsilon lm}(r) Y_l^{m \pm 1}(\theta, \varphi) \end{aligned}$$

E perciò $h_{\epsilon lm}(r) = h_{\epsilon l, m \pm 1}(r)$. In altre parole, h è indipendente dall'indice m , che perciò possiamo rimuovere. Giungiamo allora a:

$$\psi_{\epsilon lm}(r, \theta, \varphi) = h_{\epsilon m}(r) Y_l^m(\theta, \varphi) \quad (4.88)$$

Per cui valgono le equazioni agli autovalori:

$$H |\epsilon, l, m\rangle = \mathcal{E}_{\epsilon l} |\epsilon, l, m\rangle \quad \vec{L}^2 |\epsilon, l, m\rangle = \hbar^2 l(l+1) |\epsilon, l, m\rangle \quad L_3 |\epsilon, l, m\rangle = \hbar m |\epsilon, l, m\rangle$$

Ci rimane solo da trovare un'espressione esplicita per $h_{\epsilon l}(r)$, per cui è però necessario definire la forma di $U(r)$. Occupiamoci per il momento di trovare alcune proprietà generali. Sostituendo (4.88) in (4.87) e svolgendo le equazioni agli autovalori, giungiamo a:

$$\left[-\frac{\hbar^2}{r} \frac{d^2}{dr^2} r + \frac{\hbar^2 l(l+1)}{r^2} - 2m(\mathcal{E} - U(r)) \right] h_{\epsilon l}(r) = 0 \quad (4.89)$$

Possiamo semplificare tale equazione includendo l' r evidenziato nella funzione³⁴, e cioè definendo:

$$\chi_{\epsilon l}(r) = r h_{\epsilon l}(r)$$

Dalla condizione $\lim_{r \rightarrow 0} r \psi(r) = 0$ abbiamo:

$$\lim_{r \rightarrow 0} \chi_{\epsilon l}(r) = 0$$

³⁴ΛE, per gli altri termini, usiamo l' r al denominatore del primo termine

Perciò la (4.89) diviene:

$$\frac{d^2 \chi_{\ell l}(r)}{dr^2} + \left[\frac{2m}{\hbar^2} (\mathcal{E} - U(r)) - \frac{l(l+1)}{r^2} \right] \chi_{\ell l}(r) = 0 \quad (4.90)$$

La condizione di normalizzazione si riduce a:

$$\infty > \int_0^\infty |h_{\ell l}(r)|^2 r^2 dr = \int_0^\infty |\chi_{\ell l}(r)|^2 dr$$

Notiamo allora che introdurre $\chi_{\ell l}(r)$ permette di trattare un problema *radiale* (in coordinate sferiche) come uno unidimensionale *cartesiano* in $L^2(\mathbb{R}_+, d\mathbf{r})$. Lo possiamo vedere esplicitamente “inglobando” il termine evidenziato in verde nella (4.90) in un potenziale *effettivo* a l fissato:

$$U_l^{\text{eff}}(r) \equiv U(r) + \frac{\hbar^2}{2m} \frac{l(l+1)}{r^2}$$

dove il termine aggiuntivo è analogo a quello dell'*energia centrifuga* del caso classico, che funge da “barriera” che impedisce a una particella con momento angolare non nullo (e quindi $l \neq 0$) di precipitare sull'origine, dove violerebbe la conservazione del momento angolare.

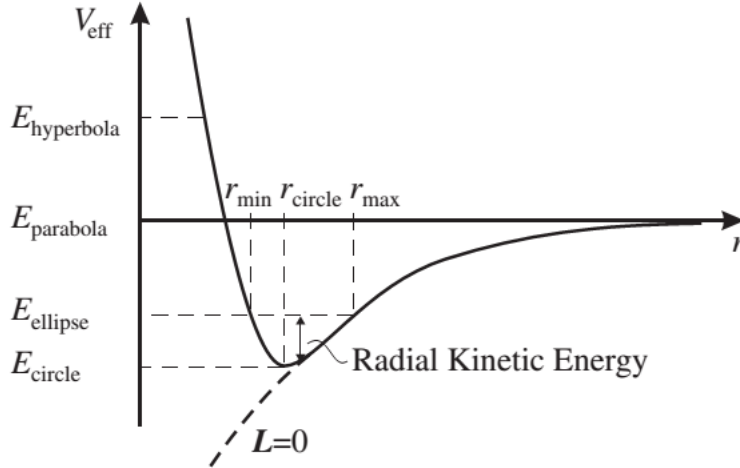


Figura 4.7: Andamento qualitativo del potenziale effettivo $U_l^{\text{eff}}(r)$

Così facendo, otteniamo l'equazione:

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dr^2} + U_l^{\text{eff}}(r) \right] \chi_{\ell l}(r) = \mathcal{E}_{\ell l} \chi_{\ell l}$$

che è l'equazione di Schrödinger stazionaria per una particella confinata in \mathbb{R}_+ in un potenziale unidimensionale³⁵

Di tale sistema, almeno qualitativamente, sappiamo già dire tutto. Facendo riferimento alla sezione 3.9.7, poiché $U_l^{\text{eff}}(0) = +\infty$ (e infatti si ha $\chi(0) = 0$) il moto

avviene in una regione semilimitata, ossia solo per $r > 0$. Ne segue che, fissati l e m , lo spettro $\sigma(H)$ è sempre non degenere. Ma questo è proprio quello che volevamo dimostrare: possiamo perciò affermare che $\{H, \vec{L}^2, L_3\}$ costituisca un ICOC per il sistema considerato.

ICOC
 $\{H, \vec{L}^2, L_3\}$

A seconda del valore di \mathcal{E} si presentano allora due casi:

- $V_{\min} < \mathcal{E} < 0$: le soluzioni $\chi_{el}(r) \in L^2(\mathbb{R})$, e perciò devono soddisfare la condizione $\chi(0) = 0$. Ciò corrisponde allo *spettro discreto* degli *stati legati*.
- $0 < \mathcal{E} < +\infty$: le soluzioni non sono più in L^2 , ossia non sono più “fisiche” - e ciò corrisponde allo *spettro continuo* del sistema. In tal caso avremo solo che $\psi = h_{el}(r)Y_l^m(\theta, \varphi) \in \mathcal{S}'(\mathbb{R}^3)$.

4.11.1 Particella libera in coordinate sferiche

Per proseguire oltre è necessario specificare l'espressione del potenziale centrale $U(r)$. Partiamo dal caso più semplice, in cui $U(r) \equiv 0$, che non è altro che la riscrittura delle funzioni d'onda di una particella libera in coordinate sferiche. Dato che deve essere $\mathcal{E} > V_{\min} = 0$, ci aspettiamo uno *spettro continuo* (dato che, del resto, la particella non è confinata), e *nondegenere*, come visto nella precedente discussione.

Infatti (4.90) diviene:

$$\left[\frac{d^2}{dr^2} - \frac{l(l+1)}{r^2} + \frac{2m}{\hbar^2} \mathcal{E} \right] \chi_{el}(r) = 0 \quad (4.91)$$

Per risolverla, supponiamo che $\chi_{el}(r)$ sia sufficientemente regolare da consentire un'espansione in serie di potenze attorno all'origine, dove il termine dominante sarà r^α per un certo $\alpha > 0$.

$$\chi_{el}(r) \underset{r \rightarrow 0}{\sim} r^\alpha \quad \alpha > 0 \quad (4.92)$$

Sostituiamo in (4.91) per esaminare l'equazione nel *punto singolare* per $r \rightarrow 0$, dove l'*energia centrifuga* domina e quindi trascuriamo il termine con \mathcal{E} :

$$\begin{aligned} \left[\frac{d^2}{dr^2} - \frac{l(l+1)}{r^2} \right] \chi_{el}(r) &= 0 \\ \alpha(\alpha-1)r^{(\alpha-2)} - l(l+1)r^{\alpha-2} &= 0 \\ \Rightarrow \alpha^2 - \alpha - l^2 - l &= 0 \Rightarrow \alpha_{1,2} = l+1, -l \end{aligned} \quad (4.93)$$

In principio, r^{l+1} e r^{-l} sono due soluzioni possibili. Di queste, tuttavia, r^{-l} è *spuria*, ed è introdotta dal fatto che, nel passaggio a coordinate sferiche, si ha una singolarità all'origine (dove la matrice jacobiana è singolare). Possiamo scartarla osservando che:

- Per lo *spettro discreto* deve essere $\chi_{el}(0) = 0$, e $r^{-l} \neq 0$ (o addirittura diverge) per $l \geq 0$.

³⁵ ∧ Ponendo $U_l^{\text{eff}}(r) = +\infty$ per $r < 0$ possiamo ricondurre il problema a tutto \mathbb{R} .

- Per lo *spettro continuo*, deve essere $\psi = hY \in \mathcal{S}'$, e perciò non possiamo immediatamente scartare r^{-l} . Notiamo però che per $l > 0$ la ψ non è normalizzabile in coordinate sferiche, mentre per $l = 0$ si ha che $\chi_{\epsilon 0} \sim 1$, da cui $h = \chi/r \sim r^{-1}$, e dato che $Y_0 \sim 1$, la $\psi = hY \sim r^{-1}$. Ma tale ψ non può essere (nemmeno in \mathcal{S}') la funzione d'onda di uno stato stazionario, poiché non verifica l'equazione di Schrödinger. Infatti, il suo laplaciano (che compare nell'equazione stazionaria) è una δ :

$$\Delta(hY) \sim \Delta \frac{1}{r} = \delta^{(3)}(\vec{x})$$

che non viene *cancellata* da alcun altro termine dell'equazione, dato che il potenziale $U(r)$ è ipotizzato regolare a meno di salti.

Escludendo il caso r^{-l} , giungiamo a:

$$\chi_{\epsilon l}(r) \underset{r \rightarrow 0}{\sim} r^{l+1}$$

Notiamo che questo risultato si estende anche agli $U(r)$ per cui $r^2 U(r) \underset{r \rightarrow 0}{\sim} 0$ (come nel caso del potenziale coulombiano) cosicché $U(r)$ è trascurabile rispetto a d^2/dr^2 e $l(l+1)/r^2$ per $r \rightarrow 0$, e perciò l'espressione (4.93) da cui siamo partiti vale senza modifiche.

*Estensione a
potenziali
"coulombiani"*

Per trovare esplicitamente le autofunzioni radiali partiamo esaminando il caso di $l = 0$, e cerchiamo di costruire tutto il resto partendo da lì. In tal caso, la (4.91) diviene l'equazione differenziale di un oscillatore armonico:

$$\left[\frac{d^2}{dr^2} + k^2 \right] \chi_{\epsilon 0}(r) = 0 \quad k = \sqrt{\frac{2m\mathcal{E}}{\hbar^2}} \quad (4.94)$$

Che ha due soluzioni indipendenti:

$$\chi_{\epsilon 0}(r) = \begin{cases} \sin(kr) \Rightarrow h_{\epsilon 0} = \frac{\sin(kr)}{r} & \chi_{\epsilon 0}(0) = 0 \\ \cos(kr) \Rightarrow h_{\epsilon 0} = \frac{\cos(kr)}{r} & \chi_{\epsilon 0}(0) \neq 0 \end{cases}$$

Per cui la condizione in $r = 0$ è soddisfatta solo dalla prima soluzione, che è l'unica accettabile: del resto sappiamo che gli autostati radiali sono *nondegeneri*.

Per $l > 0$ sappiamo che $\chi_{\epsilon l} \sim r^{l+1}$ attorno all'origine, e perciò $h = \chi/r$ sarà, in generale, il prodotto tra r^l e una opportuna funzione $h_{kl}(r)$ che *non ne modifichi* l'andamento nei pressi di 0:

$$h_{\epsilon l}(r) = r^l h_{kl}(r) \quad h_{kl}(r) \underset{r \rightarrow 0}{\longrightarrow} \text{cost} \neq 0 \quad (4.95)$$

Sostituendo (4.95) nell'equazione differenziale troviamo una condizione che deve rispettare la $h_{kl}(r)$, e che ci permetterà di determinarla:

$$\begin{aligned}
0 &= \left[\frac{d^2}{dr^2} + k^2 - \frac{l(l+1)}{r^2} \right] \underbrace{rh_{\epsilon l}}_{\chi_{\epsilon l}} \\
0 &= (rh_{\epsilon l})'' + k^2(rh_{\epsilon l}) - \frac{l(l+1)}{r^2}(rh_{\epsilon l}) \\
&\stackrel{(4.95)}{=} \cancel{l(l+1)r^{l-1}h_{kl}} + 2(l+1)r^l h'_{kl} + r^{l+1}h''_{kl} + k^2r^{l+1}h_{kl} - \cancel{l(l+1)r^{l-1}h_{kl}}
\end{aligned}$$

Dividendo per r^{l+1} otteniamo:

$$h''_{kl} + \frac{2(l+1)h'_{kl}}{r} + k^2h_{kl} = 0 \quad (4.96)$$

Prendendo la (4.96) e derivando rispetto a r :

$$(h'_{kl})'' + \frac{2(l+1)}{r}(h'_{kl})' + \left[-\frac{2(l+1)}{r^2} + k^2 \right] h'_{kl} = 0 \quad (4.97)$$

Notiamo ora che $rh_{k(l+1)}$ soddisfa la stessa equazione differenziale (4.97) di h'_{kl} . Infatti, inserendo nell'equazione di partenza $\chi_{\epsilon(l+1)} = rh_{\epsilon, l+1} = r(r^{l+1})h_{k, l+1}$ otteniamo:

$$\begin{aligned}
0 &= \left[\frac{d^2}{dr^2} + k^2 - \frac{(l+1)(l+2)}{r^2} \right] rh_{\epsilon, l+1} \\
0 &= (r^{l+1}rh_{k, l+1})'' + k^2r^{l+1}rh_{k, l+1} - \frac{(l+1)(l+2)}{r^2}r^{l+1}rh_{k, l+1} = \\
&= \cancel{l(l+1)r^{l-1}(rh_{k, l+1})} + 2(l+1)r^l(rh_{k, l+1})' + r^{l+1}(rh_{k, l+1})'' + k^2r^{l+1}(rh_{k, l+1}) \\
&\quad - \cancel{(l+2)(l+1)r^{l-1}rh_{k, l+1}} = \\
&\stackrel{(a)}{\Rightarrow} (rh_{k, l+1})'' + \frac{2(l+1)}{r}(rh_{k, l+1})' + \left[-\frac{2(l+1)}{r^2} + k^2 \right] (rh_{k, l+1})
\end{aligned}$$

dove in (a) si è diviso per r^{l+1} .

Abbiamo allora che $rh_{k, l+1}$ e h'_{kl} sono due soluzioni alla stessa equazione differenziale. Tuttavia, dalla nondegenerazione di $\sigma(H)$ sappiamo che non vi possono essere due autostati differenti per la stessa scelta degli autovalori ϵ, l , e allora le due funzioni devono essere la stessa, a meno di una costante che fissiamo a 1:

$$h'_{kl} = rh_{k(l+1)} \Rightarrow \frac{1}{r} \frac{d}{dr} h_{kl} = h_{k(l+1)}$$

Abbiamo allora trovato un modo per “incrementare” l’autovalore l di un autostato. Dato che conosciamo l’autostato per $l = 0$, da $h_{\epsilon 0} = r^0 h_{k0} = h_{k0}$, possiamo ricavare tutti gli altri reiterando:

$$\begin{aligned}
h_{\epsilon l}(r) &= r^l h_{kl}(r) = r^l \left(\frac{1}{r} \frac{d}{dr} \right)^l h_{k0}(r) = \\
&= r^l \left(\frac{1}{r} \frac{d}{dr} \right)^l h_{\epsilon 0} = r^l \left(\frac{1}{r} \frac{d}{dr} \right)^l \frac{\sin(kr)}{r}
\end{aligned}$$

La costante di normalizzazione A è fissata dalla condizione (propria di uno spettro continuo):

$$\int_0^\infty dr r^2 h_{\mathcal{E}l}(r) h_{\mathcal{E}'l}(r) = \delta(\mathcal{E} - \mathcal{E}')$$

e risulta:

$$A = \frac{(-1)^l}{k^l} \sqrt{\frac{2m}{k\pi}}$$

Moltiplicando e dividendo per k in opportuni punti, possiamo semplificare la forma finale per la soluzione radiale per la particella libera:

$$\begin{aligned} h_{\epsilon l} &= \frac{(-1)^l}{k^l} \sqrt{\frac{2m}{k\pi}} k^l r^l \left(\frac{1}{r} \frac{d}{dkr} \right)^l \frac{\sin(kr)}{kr} = \\ &= \sqrt{\frac{2m}{k\pi}} k (kr)^l (-1)^l \left(\frac{1}{kr} \frac{d}{dkr} \right)^l \frac{\sin(kr)}{kr} \equiv \sqrt{\frac{2m}{\pi}} j_l(kr) \end{aligned}$$

dove le $j_l(x)$ sono le **funzioni di Bessel** sferiche di ordine l .

$$j_l(x) \equiv (-1)^l x^l \left(\frac{1}{x} \frac{d}{dx} \right)^l \frac{\sin(x)}{x}$$

Per $r \rightarrow +\infty$ possiamo ottenere l'andamento dominante osservando che il termine che decresce meno rapidamente in $\left(-\frac{1}{r} \frac{d}{dr}\right)^l \frac{\sin(kr)}{r}$ è quello in cui tutte le derivate agiscono sul seno. Infatti, quando una derivata agisce su $1/r$ produce termini che decrescono più velocemente:

*Andamento
asintotico a
 $r \rightarrow +\infty$*

$$-\frac{d}{dr} \frac{\sin(kr)}{r} \sim \overbrace{-k \frac{\cos(kr)}{2}}^{-k \sin(kr - \pi/2)/2} + \cancel{\frac{\sin(kr)}{r^2}}$$

e tale ragionamento può essere esteso ad ogni derivata. Perciò:

$$h_{\epsilon l}(r) \underset{r \rightarrow \infty}{\sim} \frac{\sin(kr - l\frac{\pi}{2})}{r}$$

Tale funzione non³⁶ è in $L^2(\mathbb{R}_+, r^2 dr)$, ma è ancora una distribuzione in $\mathcal{S}'(\mathbb{R})$. Concludiamo perciò che per $U(r) \equiv 0$, lo spettro è continuo: $\sigma(H) = \sigma_C(H) = \mathbb{R}_+$, come del resto ci si aspetta, dato che la particella non è vincolata.

Oss: Notiamo che la scelta di $\sin(kr)/r$ rispetto a $\cos(kr)/r$ per $h_{\epsilon 0}$ era dettata dal comportamento in $r = 0$. Per l'andamento asintotico tale vincolo viene meno. Quindi per un potenziale $U(r)$ che non pone vincoli a $r = 0$ entrambe le soluzioni $\sin(kr)/r$ e $\cos(kr)/r$ per $h_{\epsilon 0}$ risultano accettabili.

In sintesi, è importante ricordare:

³⁶Infatti è della forma $\sin(x)/x$ che è localmente integrabile, ma non Lebesgue-integrabile - come si ricorda dai corsi di Analisi.

- Come ridurre un problema con potenziale centrale in 3D a uno in 1D tramite potenziale effettivo
- L'andamento delle autofunzioni a grandi distanze e piccole distanze dal centro.

(Lezione 31 del
04/12/2018)

4.12 Stati legati per l'atomo di idrogeno

Occupiamoci ora di studiare il caso di un potenziale $U(r)$ coulombiano, analizzando il caso specifico dell'*atomo di idrogeno*, e seguendo da vicino la trattazione di tale sistema data da Dirac.

L'atomo di idrogeno consiste in un protone di carica $+e$ e un elettrone di carica $-e$ che si *attraggono* tramite l'interazione elettromagnetica, il cui potenziale è:

$$U(r) = -\frac{e^2}{r}$$

Possiamo perciò subito scrivere l'equazione radiale:

$$\frac{d^2 \chi_{el}(r)}{dr^2}(r) + \left[\frac{2m}{\hbar^2} \left(\mathcal{E} + \frac{e^2}{r} \right) - \frac{l(l+1)}{r^2} \right] \chi_{el}(r) = 0 \quad (4.98)$$

Siamo interessati agli **stati legati**, ossia a quelli per cui $\mathcal{E} < 0$. Sappiamo che:

$$r h_{el}(r) = \chi_{el}(r)$$

e $h_{el}(r) \in L^2(\mathbb{R}_+, r^2 dr)$ oppure $\chi_{el}(r) \in L^2(\mathbb{R}_+, dr)$.

Dall'analisi della precedente lezione, sappiamo anche le condizioni al contorno e l'andamento per $r \rightarrow 0$ di χ_{el} :

$$\chi_{el}(0) = 0 \quad \chi_{el}(r) \underset{r \rightarrow 0}{\sim} r^{l+1}$$

Vogliamo allora trovare esplicitamente le soluzioni radiali all'equazione (4.98). Prima di farlo, riscriviamola in maniera più semplice. In maniera analoga a quanto fatto nel caso di $U(r) \equiv 0$, definiamo delle opportune costanti k e a :

$$\frac{d^2 \chi_{el}(r)}{dr^2} + \left[-k^2 + \frac{2}{ra} - \frac{l(l+1)}{r^2} \right] \chi_{el}(r) = 0 \quad k \equiv \sqrt{-\frac{2m}{\hbar^2} \mathcal{E}}; \quad a \equiv \frac{\hbar^2}{me^2}$$

dove $a \approx 0.52 \text{\AA}$ è detto *raggio di Bohr*.

Un'ulteriore semplificazione è data da un cambio di variabili che permette di "estrarre" una costante k^2 e rimuoverla dividendo per $4k^2$:

$$\begin{aligned} x \equiv 2kr &\Rightarrow r = \frac{x}{2k} \\ 4k^2 \frac{d^2 \chi_{el}(x)}{dx^2} + \left[-k^2 + \frac{4k}{xa} - \frac{l(l+1)}{x^2} 4k^2 \right] \chi_{el}(x) &= 0 \\ \Rightarrow_{/(4k^2)} \frac{d^2 \chi_{el}(x)}{dx^2} + \left[-\frac{1}{4} + \frac{\nu}{x} - \frac{l(l+1)}{x^2} \right] \chi_{el}(x) &= 0 \quad \nu \equiv \frac{1}{ka} \end{aligned} \quad (4.99)$$

Tale cambio di variabili è motivato dal fatto che la x così definita è adimensionale. Infatti:

$$[k] = \sqrt{\frac{\text{kg}}{\text{J}^2 \text{s}^2}} \text{J} = \sqrt{\frac{\text{kg}}{\text{Nm s}^2}} = \sqrt{\frac{1}{\text{m}^2}} = \frac{1}{\text{m}}; \quad [r] = \text{m} \Rightarrow [kr] = \mathcal{N}$$

In generale³⁷, quando si ha a che fare con lo studio di comportamenti asintotici (come vedremo nei prossimi paragrafi) è comodo lavorare in una scala “fondamentale” con quantità adimensionali.

Per risolvere la (4.99), partiamo analizzando i punti di possibile singolarità dell’equazione differenziale, che si possono avere per $x \rightarrow \infty$, oppure, data la presenza di denominatori, per $x \rightarrow 0$. La strategia è quella di risolvere “approssimazioni” dell’equazione in questi limiti in cui consideriamo solo i termini dominanti, e poi “cercare di raccordare” il tutto per ottenere una soluzione generale³⁸.

Per $x \rightarrow \infty$ l’unico termine dominante è $-1/4$, e l’equazione approssimata è quella di un *repulsore armonico*:

$$\frac{d^2}{dx^2} \chi_{el}(x) - \underbrace{\frac{1}{4}}_{\omega^2} \chi_{el}(x) = 0 \Rightarrow \chi_{el}(x) \underset{x \rightarrow \infty}{\sim} e^{\pm \omega x} = e^{\pm \frac{x}{2}} = e^{\pm kr}$$

Di queste due, e^{+kr} di sicuro non sta in L^2 (diverge esponenzialmente), e perciò solo e^{-kr} è accettabile:

$$\chi_{el}(x) \underset{x \rightarrow 0}{\sim} e^{-\frac{x}{2}}$$

Per $x \rightarrow 0$, invece, il problema si riconduce a quello già visto nella lezione precedente, dato che $U(r)$ ed \mathcal{E} sono trascurabili rispetto all’*energia centrifuga*:

$$\frac{d^2 \chi_{el}}{dx^2} - \frac{l(l+1)}{x^2} \chi_{el}(x) = 0 \Rightarrow \chi_{el}(x) \underset{x \rightarrow 0}{\sim} x^{l+1}$$

Dalle informazioni appena ricavate, ipotizziamo che una soluzione generale sia data dal prodotto tra le soluzioni “locali” appena trovate e una certa funzione $v_l(x)$ che non alteri il comportamento di $\chi_{el}(x)$ ai limiti $x \rightarrow 0$ e $x \rightarrow \infty$:

$$\chi_{el}(x) = \underbrace{x^{l+1}}_f \underbrace{e^{-\frac{x}{2}}}_g \underbrace{v_l(x)}_h$$

³⁷^CFR Lecture 24 di bit.ly/2GLsM66, mentre il processo è detto “Nondimensionalization”

³⁸^Il termine tecnico è **ansatz**, ossia un *educated guess* sulla soluzione di un’equazione differenziale difficile.

Inserendo nell'equazione (4.98), e calcolando la derivata seconda iterando più volte Leibniz:

$$\begin{aligned}
(fgh)'' &= (f'gh + fg'h + fgh')' \\
&= (f''gh + f'g'h + f'gh') + \\
&\quad (f'g'h + fg''h + fg'h') + \\
&\quad (f'gh' + fg'h' + fgh'') = \\
&= (f''g + 2f'g' + fg'')h + 2(f'g + 2fg')h' + fgh'' \\
(4.98) &= \left[\cancel{l(l+1)x^{l-1}} + 2(l+1)x^l \left(-\frac{1}{2}\right) + x^{l+1} \cancel{\left(-\frac{1}{2}\right)^2} \right] e^{-\frac{x}{2}} v_l(x) + \\
&\quad + 2 \left((l+1)x^l + x^{l+1} \left(-\frac{1}{2}\right) \right) e^{-\frac{x}{2}} v_l'(x) + x^{l+1} e^{-\frac{x}{2}} v_l''(x) + \\
&\quad + \left[-\cancel{\frac{1}{4}} + \frac{\nu}{x} - \cancel{\frac{l(l+1)}{x^2}} \right] e^{-\frac{x}{2}} x^{l+1} v_l(x) = 0
\end{aligned}$$

Dividendo per x^l e $e^{-x/2}$ giungiamo a:

$$xv_l'' + (2l+2-x)v_l' - ((l+1)-\nu)v_l = 0 \quad (4.100)$$

che dobbiamo risolvere per v_l . Possiamo farlo sviluppando $v_l(x)$ in serie:

$$v_l(x) = \sum_{p=0}^{\infty} a_p x^p \quad (4.101)$$

con $a_0 = 1$, dato che non vogliamo turbare l'andamento a $x \rightarrow 0$, e perciò deve essere $v_l(0) \neq 0$, dato che altrimenti non avremmo $\chi_{el}(x) \underset{x \rightarrow 0}{\sim} x^{l+1}$.

Inserendo (4.101) in (4.100) otteniamo:

$$\sum_{p=2}^{\infty} p(p-1)a_p x^{p-1} + \sum_{p=1}^{\infty} p[2(l+1)x^{p-1} - x^p]a_p - \sum_{p=0}^{\infty} ((l+1)-\nu)a_p x^p = 0$$

Conviene riportare tutto sotto la stessa sommatoria, sostituendo $p-1 \equiv p' \Rightarrow p = p' + 1$:

$$\sum_{p'=0}^{\infty} (p'+1)p'a_{p'+1}x^{p'} + \sum_{p'=0}^{\infty} a_{p'+1}(p'+1)(2(l+1))x^{p'} - \sum_{p=0}^{\infty} a_p p x^p - \sum_{p=0}^{\infty} a_p (l+1-\nu)x^p = 0$$

Nota: nella prima sommatoria evidenziata p' parte da 0, e non da 1 come si avrebbe dalla situazione. Possiamo far ciò notando che essa conteneva già un termine p' , e quindi partire da $p' = 0$ non fa altro che aggiungere in testa alla successione un termine nullo, che non modifica la serie. Un discorso analogo (stavolta per p e non per p') vale per la seconda sommatoria evidenziata.

Dato che p' e p sono indici muti, possiamo anche usare lo stesso indice per tutti e due: così facendo compariranno termini a_{p+1} oltre a quelli a_p . Potremo allora trovare una *relazione di ricorrenza* scrivendo a_{p+1} in funzione del termine precedente a_p , e usare tale relazione per *definire* la serie. Riarrangiando, otteniamo allora:

$$\Rightarrow \sum_{p=0}^{\infty} (p+1)(p+2(l+1))a_{p+1}x^p = \sum_{p=0}^{\infty} [p+l+1-\nu]a_p x^p$$

Tale uguaglianza deve essere vera $\forall x$, e perciò deve valere *potenza per potenza* (ossia togliendo le sommatorie):

$$(p+1)(p+2l+2)a_{p+1} = (p+l+1-\nu)a_p \quad (4.102)$$

“Diminuendo” p di 1 otteniamo una relazione leggermente più semplice:

$$p(p+2l+1)a_p = (p+l-\nu)a_{p-1} \quad (4.103)$$

Verifichiamone la convergenza, per esempio tramite il criterio del rapporto:

$$\frac{a_p}{a_{p-1}} = \frac{p+l-\nu}{p(p+2l+1)} \underset{p \rightarrow \infty}{\sim} \frac{1}{p} \Rightarrow a_p \underset{p \rightarrow \infty}{\sim} \frac{1}{p!}$$

La serie definitivamente converge, ma il suo andamento asintotico è esponenziale:

$$v_l(x) = \sum_{p=0}^{\infty} a_p x^p \underset{x \rightarrow \infty}{\sim} \sum_{p=0}^{\infty} \frac{x^p}{p!} = e^x$$

E sostituendo nell'espressione per la $\chi_{el}(x)$:

$$\chi_{el}(x) \sim x^{l+1} e^{-\frac{x}{2}} e^x \notin L^2$$

e questo è un gran problema, perché sappiamo (per il fatto che la materia esiste) che gli autostati legati dell'atomo di idrogeno sono “fisici”, e quindi devono stare in L^2 .

Perché ciò non succeda, la serie di $v_l(x)$ deve troncarsi, ossia se per un certo $p = n' \in \mathbb{N}$ opportuno si ha $a_{n'} \neq 0$, ma $a_{n'+1} = 0$, da cui, per ricorrenza, tutti i termini successivi sono nulli. Imponendo $a_{n'+1} = 0$ si ha che il coefficiente di $a_{n'} \neq 0$ in (4.102) deve annullarsi, ma ciò è possibile solo se ν assume valori interi:

$$n' + l + 1 - \nu = 0 \Rightarrow \nu = \{n' + l + 1, n' \in \mathbb{N}\} = \{n \in \mathbb{N}, n \geq l + 1\}$$

Ricordando l'espressione per ν :

$$\nu = \frac{1}{ka} = \frac{me^2}{\hbar^2} \sqrt{\frac{\hbar^2}{-2m\mathcal{E}}}$$

si ha che la condizione su ν si traduce in una condizione sull'energia. Indichiamo con \mathcal{E}_n il valore dell'energia a cui corrisponde $\nu = n$:

$$\mathcal{E}_n = -\frac{\hbar^2}{2m} \frac{1}{(\nu a)^2} = -\frac{\hbar^2}{2m} \left(\frac{me^2}{\hbar^2} \right)^2 \frac{1}{n^2} = -\frac{me^4}{2\hbar^2} \frac{1}{n^2} \approx -\frac{13.6 \text{ eV}}{n^2} \quad (4.104)$$

Ritroviamo allora le **energie di Bohr**, che erano state ricavate inizialmente imponendo artificialmente che il momento angolare fosse multiplo intero di una certa costante. Le \mathcal{E}_n spiegano anche le frequenze di Rydberg per l'atomo di idrogeno: i fotoni emessi hanno infatti un'energia pari a $\mathcal{E}_m - \mathcal{E}_n$, con $m < n$, dato che sono generati quando un elettrone passa da uno stato *eccitato* a uno a più bassa energia. Ecco allora perché, come avevamo osservato a inizio del corso, si scopre che tali frequenze sono sempre ottenibili come differenza di due numeri.

Come sono fatte le autofunzioni associate alle \mathcal{E}_n ?

Da (4.103) ricaviamo:

$$\frac{a_p}{a_{p-1}} = \frac{p + l - \nu}{p(p + 2l + 1)}$$

Fissando $a_0 = 1$, avremo allora che il coefficiente p -esimo è dato da:

$$a_p = \frac{a_p}{a_{p-1}} \frac{a_{p-1}}{a_{p-2}} \dots \frac{a_1}{a_0} 1$$

Sappiamo che $a_p \neq 0$ per p che vanno da 0 a $n' = \nu - l - 1 = n - l - 1$. L'espressione per $v_l(x)$ è quindi data espandendo la serie:

$$v_l(x) = \sum_{p=0}^{n'} a_p x^p$$

Sostituendola nell'espressione per la soluzione radiale $\chi_{nl}(x)$ otteniamo:

$$\chi_{nl}(x) = x^{l+1} e^{-x/2} \left[\sum_{p=0}^{n'-l-1} \frac{p + l - n}{p} \frac{(p + l - n - 1)}{(p - 1)(p - 1 + 2l + 1)} \dots \frac{(1 + l - n)}{1(1 + 2l + 1)} x^p \right]$$

Possiamo ricondurci ad una forma più familiare scrivendo:

$$\begin{aligned} (-1)^p \frac{(n - l - 1)!}{(n - l - 1 - p)!} &= (p + l - n)(p + l - n - 1) \dots (1 + l - n) \\ \frac{1}{p!} &= \frac{1}{p} \frac{1}{p - 1} \dots \frac{1}{1} \\ \frac{(2l + 1)!}{(p + 2l + 1)!} &= \frac{1}{(p + 2l + 1)} \frac{1}{(p - 1 + 2l + 1)} \dots \frac{1}{(1 + 2l + 1)} \end{aligned}$$

Dove il fattore alternante $(-1)^p$ presente nel termine evidenziato in giallo è una scelta convenzionale di fase, che permette di riscrivere le autofunzioni $\chi_{nl}(x)$:

$$\chi_{nl} = \sum_{p=0}^{n-l-1} (-1)^p \frac{(n - l - 1)!}{(n - l - 1 - p)!} \frac{(2l + 1)!}{(p + 2l + 1)!} \frac{1}{p!} x^p \quad (4.105)$$

utilizzando i cosiddetti **polinomi di Laguerre generalizzati**, che hanno la forma³⁹:

$$L_j^k(x) = \sum_{p=0}^j (-1)^p \frac{(j+k)!}{(j-p)!(k+p)!p!} x^p \quad (4.106)$$

Notiamo che il numeratore di (4.105) non dipende da p , ed è fissato una volta scelti n e l . Portandolo fuori dalla sommatoria possiamo considerarlo come un fattore moltiplicativo costante, che sarà fissato dalla normalizzazione. Così facendo, confrontando (4.105) con (4.106), ponendo $j = n - l - 1$ e $k = 2l + 1$ nel polinomio di Laguerre otteniamo (a meno del fattore moltiplicativo A al numeratore) χ_{nl} :

$$\chi_{nl}(x) = Ax^{l+1} e^{-\frac{x}{2}} L_{n-l-1}^{2l+1}(x)$$

Sappiamo dalla precedente sezione che H, \vec{L}, L_3 formano un ICOC, e perciò esistono delle autofunzioni comuni per tali operatori, che hanno la forma:

$$\psi_{nlm}(r, \theta, \varphi) = h_{nlm}(r) Y_l^m(\theta, \varphi)$$

dove $h = \chi/r$.

Ripercorrendo a ritroso i vari cambi di variabile effettuati si ha che:

$$x = 2kr \quad k = \sqrt{-\frac{2m}{\hbar^2} \mathcal{E}_n} \quad \mathcal{E}_n = -\frac{me^4}{2\hbar^2} \frac{1}{n^2} \quad a = \frac{\hbar^2}{me^2}$$

da cui otteniamo:

$$x = \frac{2r}{na}$$

Possiamo allora costruire l'espressione per $\psi(r, \theta, \varphi)$:

$$\begin{aligned} \psi(r, \theta, \varphi) &= B \frac{1}{r} \left(\frac{2r}{na} \right)^{l+1} \exp\left(-\frac{r}{na}\right) L_{n-l-1}^{2l+1}\left(\frac{2r}{na}\right) Y_l^m(\theta, \varphi) = \\ &\stackrel{(a)}{=} C \left(\frac{2r}{na} \right)^l \exp\left(-\frac{r}{na}\right) L_{n-l-1}^{2l+1}\left(\frac{2r}{na}\right) Y_l^m(\theta, \varphi) \end{aligned}$$

dove in (a) abbiamo semplificato l' r al denominatore inglobando il fattore rimanente nella costante C , il cui valore è fissato dalla normalizzazione⁴⁰:

$$\psi_{nlm}(r, \theta, \varphi) = \sqrt{\left(\frac{2}{na}\right)^3 \frac{(n-l-1)!}{2n(n+l)!}} \left(\frac{2r}{na}\right)^l \exp\left(-\frac{r}{na}\right) L_{n-l-1}^{2l+1}\left(\frac{2r}{na}\right) Y_l^m(\theta, \varphi) \quad (4.107)$$

In particolare per $n = l + 1$ (il massimo momento angolare l per n fissato), poiché il relativo polinomio di Laguerre è costante, $L_0^{2l+1}(x) = 1$, si ha che la componente radiale di ψ_{nlm} è proporzionale a:

$$h_{nl}(r) \propto \left(\frac{2r}{na}\right)^l \exp\left(-\frac{r}{na}\right) = \left(\frac{2r}{na}\right)^{n-1} \exp\left(-\frac{r}{na}\right)$$

³⁹^mathworld.wolfram.com/AssociatedLaguerrePolynomial.html

⁴⁰^CFR pag. 154 [3].

Quindi la probabilità di trovare la particella tra r e $r + dr$ è proporzionale al modulo quadro (con un r^2 dovuto alla misura in coordinate sferiche):

$$r^2 |h_{nl}(r)|^2 \propto r^2 \left(\frac{2r}{na}\right)^{2(n-1)} \exp\left(-\frac{2r}{na}\right) = \left(\frac{na}{2}\right)^2 \left(\frac{2r}{na}\right)^{2n} \exp\left(-\frac{2r}{na}\right)$$

Cerchiamone il massimo. Con la sostituzione $y = 2r/(na)$, imponendo l'annullamento della derivata prima giungiamo a:

$$\frac{d}{dy} (y^{2n} e^{-y}) = (2ny^{2n-1} - y^{2n}) e^{-y} \stackrel{!}{=} 0 \Rightarrow y = 2n \Rightarrow r_n = \frac{na}{2} y = n^2 a$$

Abbiamo allora ritrovato in r_n il **raggio di Bohr** dell'orbita n -esima.

Riepilogando: risolvendo l'equazione di Schrödinger stazionaria per gli stati legati dell'atomo di idrogeno ritroviamo le energie di Bohr, e nei casi in cui l è massimo, anche che la probabilità della posizione dell'elettrone si addensa attorno ai corrispondenti raggi di Bohr.

La dipendenza angolare dovuta alle $Y_l^m(\theta, \varphi)$ porta infine ad ottenere i “disegnini stravaganti” dei vari orbitali.

(Lezione 32 del
06/12/2018)

4.13 Esercizio 6

Si consideri una particella quantistica priva di spin di massa m e carica e ancorata a un piano in presenza di un campo magnetico uniforme e costante \vec{B} perpendicolare al piano. Si denoti con \vec{A} il potenziale vettore osservato a \vec{B} e si ponga $c = 1$.

1. Si scriva l'Hamiltoniana quantistica del sistema
2. Si determini in funzione di $\vec{X} = \{X, Y\}$ e $\vec{P} = \{P_x, P_y\}$ la velocità $\vec{V} = \{V_x, V_y\}$ e si calcoli il commutatore di V_x con V_y in rappresentazione $\{\vec{x}\}$.
3. Si determini lo spettro $\sigma(H)$ di H .
4. Definite le coordinate $X_0 = \frac{m}{eB} V_y + X$ e $Y_0 = -\frac{m}{eB} V_x + Y$ (analogo quantistico del centro dell'orbita di ciclotrone classica), si determini l'algebra dei commutatori di \vec{X}_0 e \vec{V}
5. Si determini un ICOC per il sistema e, utilizzando le \vec{X}_0 , la degenerazione di $\sigma(H)$.
6. Supponendo ora che la particella sia confinata (pareti impenetrabili) in un quadrato di lato L e scegliendo $A_x = -By$ e $A_y = 0$ per il potenziale vettore, si stimi la degenerazione di $\sigma(H)$ in questo caso.

4.13.1 Soluzione

1. Consideriamo la posizione della particella sul piano 2D come $\vec{X} = X, Y$.
L'Hamiltoniana *classica* di una particella di carica e che si muove all'interno di un campo elettromagnetico di potenziale vettore $\vec{A}(\vec{x})$ e scalare $\varphi(\vec{x})$ è data (usando le unità di Gauss per i campi):

$$H = \frac{1}{2m} \left(\vec{p} - \frac{q}{c} \vec{A}(\vec{x}, t) \right)^2 + q\varphi(\vec{x}, t)$$

In questo caso il campo è puramente magnetico e indipendente dal tempo, da cui $\varphi \equiv 0$. Inserendo allora le espressioni per gli operatori quantistici, e ponendo $c = 1$, otteniamo l'Hamiltoniana H in rappresentazione $\{\vec{x}\}$:

$$H = \frac{(-i\hbar\vec{\nabla} - e\vec{A}(\vec{x}))^2}{2m} = \frac{(\vec{P} - e\vec{A})^2}{2m} \quad (4.108)$$

2. La velocità \vec{V} è definita come la derivata temporale del vettore posizione. Usando allora la relazione ricavata esaminando la visuale di Heisenberg:

$$\vec{V} = \frac{d\vec{X}}{dt} = \frac{[\vec{X}, H]}{i\hbar}$$

Calcolando componente per componente:

$$\begin{aligned} \frac{dX_i}{dt} &= \frac{[X_i, H]}{i\hbar} = \frac{1}{i\hbar} \left[X_i, \frac{1}{2m} (\vec{P}^2 - e(\vec{A} \cdot \vec{P} + \vec{P} \cdot \vec{A}) + e^2 \vec{A}^2) \right] = \\ &\stackrel{(a)}{=} \frac{1}{2mi\hbar} \left[X_i, \vec{P}^2 - e(\vec{A} \cdot \vec{P} + \vec{P} \cdot \vec{A}) \right] = \\ &= \frac{1}{2mi\hbar} \left([X_i, \vec{P}^2] - e(\vec{A} \cdot [X_i, \vec{P}] + [X_i, \vec{P}] \cdot \vec{A}) \right) = \\ &\stackrel{(b)}{=} \frac{1}{2mi\hbar} \left(2i\hbar P_i - 2eA_i [X_i, P_i] \right) = \\ &= \frac{P_i}{m} - \frac{e}{mi\hbar} A_i i\hbar = \frac{P_i - eA_i}{m} \end{aligned} \quad (4.109)$$

In (a) si è usato il fatto che $[X_i, \vec{A}] = 0 = [X_i, \vec{A}^2]$, dato che il potenziale vettore è una normale funzione di \vec{x} . In (b), invece, si ha che $[X_i, P_j] = i\hbar\delta_{ij}$, e quindi l'unico termine non nullo è quello per $[X_i, P_i]$: ciò riduce la somma del prodotto scalare da tre a un unico termine.

Possiamo ora determinare il commutatore delle velocità:

$$\begin{aligned} [V_x, V_y] &= \frac{1}{m^2} [P_x - eA_x, P_y - eA_y] = \\ &= -\frac{e}{m^2} ([A_x, P_y] + [P_x, A_y]) \end{aligned} \quad (4.110)$$

dato che componenti di A e P lungo diversi assi commutano.

Per continuare i conti, come richiesto dal testo, passiamo in rappresentazione $\{\vec{x}\}$, dove $\vec{P} = -i\hbar\vec{\nabla}$. Dato che stiamo lavorando con operatori, conviene

applicare il tutto ad una funzione generica $\psi(\vec{x})$. Il primo commutatore diviene allora:

$$\begin{aligned} [A_x, P_y]\psi(\vec{x}) &= A_x(\vec{x}) \left(-i\hbar \frac{\partial}{\partial y} \psi(\vec{x}) \right) - \left(-i\hbar \frac{\partial}{\partial y} (A_x(\vec{x})\psi(\vec{x})) \right) = \\ &= \cancel{-i\hbar A_x(\vec{x}) \frac{\partial}{\partial y} \psi(\vec{x})} + \cancel{i\hbar A_x(\vec{x}) \frac{\partial}{\partial y} \psi(\vec{x})} + i\hbar \psi(\vec{x}) \frac{\partial}{\partial y} A_x(\vec{x}) = \\ &= i\hbar \psi(\vec{x}) \frac{\partial}{\partial y} A_x(\vec{x}) \Rightarrow [A_x, P_y] = i\hbar \frac{\partial A_x}{\partial y} \end{aligned}$$

E analogamente per il secondo si ottiene:

$$[P_x, A_y] = -[A_y, P_x] = -i\hbar \frac{\partial A_y}{\partial x}$$

Sostituendo tali risultati in (4.110) giungiamo a:

$$[V_x, V_y] = -\frac{e}{m^2} \left(i\hbar \frac{\partial}{\partial y} A_x - i\hbar \frac{\partial}{\partial x} A_y \right) = \frac{i\hbar e}{m^2} \left(\frac{\partial}{\partial x} A_y - \frac{\partial}{\partial y} A_x \right)$$

Espandendo $\vec{B} = (0, 0, B)^t = \nabla \times \vec{A}$, notiamo che il termine tra parentesi è proprio B :

$$[V_x, V_y] = \frac{i\hbar e}{m^2} B \quad (4.111)$$

Abbiamo perciò scoperto che la presenza del campo magnetico fa sì che le velocità in direzioni diverse non commutino tra loro.

3. Vogliamo ora trovare $\sigma(H)$. Per farlo, conviene ragionare per analogia. Osservando l'Hamiltoniana in (4.108) notiamo la presenza di un quadrato: proviamo allora ad espanderlo, magari usando le grandezze definite nei punti precedenti:

$$\begin{aligned} H &= \frac{1}{2m} [(P_x - eA_x)^2 + (P_y - eA_y)^2] = \frac{1}{2m} [m^2 V_x^2 + m^2 V_y^2] = \\ &= \frac{m}{2} [V_x^2 + V_y^2] \end{aligned}$$

L'Hamiltoniana così riscritta assomiglia a quella di un oscillatore armonico: la differenza principale è che qui abbiamo V_x e V_y , mentre nel caso dell'oscillatore compaiono X e P . Dalla teoria sappiamo però che lo spettro, nel caso dell'oscillatore, non dipende dalla natura dei singoli operatori, ma unicamente dalle relazioni di commutazione che intercorrono tra essi. Vorremmo allora che $[V_x, V_y] = [X, P] = i\hbar$. Da (4.111) sappiamo già che non è così, ma che la differenza è data solo un fattore **costante** eB/m^2 (dato che B è uniforme per ipotesi), e possiamo perciò sistamarla *riscaldando opportunamente* V_x e V_y . Per farlo possiamo allora distribuire il fattore reciproco m^2/eB in modo simmetrico, definiamo le velocità riscalate \tilde{V}_x, \tilde{V}_y come:

$$\tilde{V}_x \equiv \frac{m}{\sqrt{eB}} V_x \quad \tilde{V}_y \equiv \frac{m}{\sqrt{eB}} V_y$$

E in tal modo si ha:

$$[\tilde{V}_x, \tilde{V}_y] = i\hbar = [X, P]$$

Sostituendo nell'Hamiltoniana otteniamo:

$$H = \frac{m}{2} (V_x^2 + V_y^2) = \frac{eB}{2m} (\tilde{V}_x^2 + \tilde{V}_y^2)$$

Poiché allora il commutatore $[\tilde{V}_x, \tilde{V}_y]$ è lo stesso di $[X, P_x] = i\hbar$, si ha che l'algebra generata dalle due coppie di grandezze è la stessa. In altre parole, se sostituiamo a $\tilde{V}_x \rightarrow X$ e $\tilde{V}_y \rightarrow P_x$, otteniamo una nuova hamiltoniana che corrisponde ad un *sistema diverso*, ma ha lo stesso spettro della precedente⁴¹:

$$H' = \frac{eB}{2m} (X^2 + P_x^2); \quad \sigma(H') = \sigma(H)$$

Il vantaggio di questa H' è che è l'hamiltoniana di un oscillatore armonico, di cui conosciamo lo spettro. Confrontiamola con la forma “canonica” per trovare i valori delle costanti:

$$H' = \frac{eB}{2m} (X^2 + P_x^2) \stackrel{!}{=} \frac{1}{2M} (P^2 + M^2 \omega^2 X^2)$$

Da cui:

$$\frac{eB}{2m} = \frac{1}{2M}; \quad M^2 \omega^2 = 1 \Rightarrow \omega = \frac{1}{M} = \frac{eB}{m}$$

Dalla formula per lo spettro dell'oscillatore armonico giungiamo allo spettro cercato:

$$\sigma(H) = \sigma(H') = \hbar\omega \left(\mathbb{N} + \frac{1}{2} \right) = \frac{\hbar eB}{m} \left(\mathbb{N} + \frac{1}{2} \right)$$

Abbiamo scoperto che $\sigma(H)$ è unicamente discreto. Gli elementi di questo spettro sono detti, nello specifico, **livelli di Landau**.

Il collegamento tra particella in \vec{B} uniforme, e l'oscillatore armonico non è troppo strano: del resto sappiamo (classicamente) che l'effetto di \vec{B} è quello di far girare la particella nell'orbita (circolare) di ciclotrone, che vista proiettata sui due assi cartesiani corrisponde a due moti armonici. Quello che può risultare strano è che qui abbiamo *un solo moto* di oscillatore, e in un certo senso il comportamento di un sistema 2D è “codificato” in quello di un sistema 1D. I punti successivi gettano luce su questo aspetto, in quanto vedremo che, nel caso quantistico, le coordinate del centro dell'orbita X_0 e Y_0 **non** sono simultaneamente conoscibili (al contrario del caso classico). Si potrebbe allora dire, *euristicamente*, che il passaggio al mondo quantistico “dimezza” le osservabili simultaneamente conoscibili.

⁴¹ Dato che lo spettro dipende, in tali casi, unicamente dall'algebra dei commutatori, come ricavato per l'oscillatore armonico tramite l'uso degli operatori di creazione/distruzione

4. Definite le coordinate di centro dell'orbita:

$$X_0 \equiv \frac{m}{eB} V_y + X \quad Y_0 \equiv -\frac{m}{eB} V_x + Y$$

vogliamo trovare l'algebra dei commutatori di \vec{X}_0 e \vec{V} (ossia i commutatori di tutte le coppie possibili delle osservabili). Procediamo con i conti:

$$\begin{aligned} [X_0, Y_0] &= \left[\frac{m}{eB} V_y + X, -\frac{m}{eB} V_x + Y \right] = -\left(\frac{m}{eB} \right)^2 [V_y, V_x] + \frac{m}{eB} ([V_y, Y] - [X, V_x]) = \\ &\stackrel{(a)}{=} -\left(\frac{m}{eB} \right)^2 \left(-i\hbar \frac{eB}{m^2} \right) + \frac{m}{eB} \left(\left[\frac{P_y}{m}, Y \right] - \left[X, \frac{P_x}{m} \right] \right) = \\ &= i\hbar \frac{1}{eB} - \frac{i\hbar}{m} \frac{m}{eB} - \frac{i\hbar}{m} \frac{m}{eB} = -\frac{i\hbar}{eB} \\ [X_0, V_x] &= \left[\frac{m}{eB} V_y + X, V_x \right] = \frac{m}{eB} \left(-i\hbar \frac{eB}{m^2} \right) + \frac{i\hbar}{m} = 0 \\ [Y_0, V_y] &= \left[-\frac{m}{eB} V_x + Y, V_y \right] = -\frac{m}{eB} \left(i\hbar \frac{eB}{m^2} \right) + \frac{i\hbar}{m} = 0 \\ [X_0, V_y] &= \left[\frac{m}{eB} V_y + X, V_y \right] = 0 \\ [Y_0, V_x] &= 0 \end{aligned}$$

dove in (a) abbiamo usato la relazione *classica* $\vec{P} = m\vec{V}$.

5. Sappiamo che in un sistema 2D, sia $\{X, Y\}$ che $\{P_x, P_y\}$ sono ICOC, ma nessuno dei due contiene le variabili \vec{X}_0 che vorremmo. Notiamo che P_x è il momento coniugato di X , e infatti $[X, P_x] = i\hbar$, e P_y è coniugato a Y , infatti $[Y, P_y] = i\hbar$. Per *coniugazione* possiamo allora passare da un ICOC all'altro. Dai conti precedenti sappiamo che X_0 commuta con V_x . Scegliamo quindi come ICOC $\{X_0, V_x\}$. Il momento coniugato di X_0 è $\propto Y_0$, e quello coniugato a V_x è $\propto V_y$. Analogamente, perciò, potremmo prendere $\{Y_0, V_y\}$ come ICOC.

Vogliamo ora determinare la degenerazione di $\sigma(H)$, ossia il numero di autostati corrispondenti allo stesso valore di energia H . Un buon punto di partenza per far ciò è osservare la forma di H , e chiedersi quali operazioni la lascino invariata - in altre parole, *spesso* la degenerazione si nasconde dietro a una simmetria del sistema.

Partendo allora da:

$$H = \frac{m}{2} (V_x^2 + V_y^2)$$

possiamo notare che H resta invariata da traslazioni. Ci aspettiamo perciò, intuitivamente, che $\sigma(H)$ abbia degenerazione infinita: fissata un'energia \mathcal{E} , è possibile traslare il sistema come si vuole - in infiniti modi - e quindi avremo infinite funzioni d'onda differenti con la stessa \mathcal{E} .

Verifichiamo rigorosamente tutto ciò. L'idea è di trovare una base di autostati comuni di H e una coordinata che definisca la *posizione* del sistema, e

notare che gli infiniti autoket di quest'ultima corrispondono allo stesso autovalore di H . Potremmo per esempio usare X_0 . Dato che $[X_0, V_x] = 0$ e $[X_0, V_y] = 0$, si ha che:

$$[H, X_0] = 0$$

perciò H è compatibile con X_0 , e H *non dipende* da X_0 . Quindi H e X_0 hanno autovettori comuni (per la compatibilità), ma fissato l'autovalore di H , come $\mathcal{E}_n = \frac{\hbar e B}{m} \left(n + \frac{1}{2}\right)$, X_0 può assumere un qualsiasi autovalore. Perciò la degenerazione di $\sigma(H)$ è infinita.

Tutto ciò deriva, intuitivamente, dal fatto che, in termini classici, le particelle compiono orbite di ciclotrone a causa del campo magnetico. Ragionevolmente, il campo magnetico fissa il raggio dell'orbita, ma il centro si può scegliere *qualunque*, basta traslare il sistema. Da qui la degenerazione infinita (ammesso che si possa sempre traslare il sistema senza alterarlo, ossia che B agisca su un piano infinito).

6. Supponiamo invece che la particella sia vincolata ad appartenere a un quadrato di lato L . Scegliamo per il potenziale vettore (*fissando il gauge*):

$$A_x = -By \quad A_y = 0 \quad (4.112)$$

Vogliamo stimare la degenerazione di $\sigma(H)$ in questo caso. Ci aspettiamo che vi sia ancora degenerazione, ma forse non infinita, dato che ora il piano ha una dimensione finita. Esplicitiamo l'Hamiltoniana:

$$H = \frac{(P_x + eBY)^2}{2m} + \frac{(P_y)^2}{2m}$$

e si ha che:

$$[P_x, H] = 0$$

Ciò significa che P_x è *costante*, e che esiste una base comune di autoket di P_x e H .

P_x è il momento in $[0, L]$, e quindi ha **spettro discreto**. Infatti, imponendo la condizione di periodicità ai suoi autostati (che hanno la forma di *onde piane*) otteniamo gli autovalori:

$$\exp\left(i\frac{p}{\hbar}L\right) \stackrel{!}{=} \exp\left(i\frac{p}{\hbar}0\right) = 1 \Rightarrow \sigma(P_x) = \frac{2\pi\hbar}{L}n; \quad n \in \mathbb{N}$$

Il fatto che il momento possa assumere valori a salti discreti fa pensare che non tutte le orbite siano possibili. In particolare, ci aspettiamo che la posizione del centro non possa essere qualunque. Calcolando per esempio Y_0 :

$$Y_0 = -\frac{m}{eB}V_x + Y \stackrel{(4.109)}{=} -\frac{m}{eB}\left(\frac{P_x - eBY}{m}\right) + Y \stackrel{(4.112)}{=} -\frac{P_x}{eB} - Y + Y = -\frac{P_x}{eB}$$

Pertanto lo spettro di Y_0 è pari a:

$$\sigma(Y_0) = -\frac{\sigma(P_x)}{eB}$$

Sappiamo che $[H, Y_0] = 0$, e perciò, seguendo lo stesso ragionamento del punto precedente, si ha che fissato un autovalore $\mathcal{E}_n = \frac{\hbar e B}{m} \left(n + \frac{1}{2}\right)$ possiamo scegliere come vogliamo Y_0 . Perciò:

$$\sigma(Y_0) = -\frac{2\pi\hbar}{eBL}\mathbb{Z} = -\frac{2\pi\hbar}{eBL}k; \quad k \in \mathbb{Z}$$

Ma sappiamo che Y_0 rappresenta un centro di un'orbita, e dal vincolo sul piano si deve avere:

$$0 \leq \sigma(Y_0) \leq L \Rightarrow 0 \leq \frac{2\pi\hbar k}{eBL} \leq L \Rightarrow 0 \leq k \leq \frac{eBL^2}{2\pi\hbar} \quad (4.113)$$

e perciò la degenerazione di un \mathcal{E}_n è $\approx \frac{eBL^2}{2\pi\hbar}$. Il valore è solo approssimato, dato che vi sarebbero diverse sottigliezze da discutere, di cui ora non ci occupiamo. Ragionando in maniera classica, poiché r è inversamente proporzionale a B , si ha che la degenerazione di $\sigma(H)$ aumenta all'aumentare dell'intensità del campo magnetico B .

(Lezione 33 e 34
del (2019))

4.14 Scattering

Discutiamo ora il caso di diffusione elastica da un bersaglio con la finalità di ottenere informazioni sulle forze a priori ignote esercitate sulle particelle incidenti, stabilendo un legame tra potenziale incognito V e i dati sperimentali che si esprimono in termini di sezione d'urto.

Vediamo dunque l'esperimento di diffusione (scattering). Si consideri un fascio di particelle identiche con momenti abbastanza ben focalizzati attorno ad un valor medio \vec{p}_0 contro un bersaglio fisso con un certo numero di centri di diffusione; il bersaglio è posto ad una distanza D_0 rispetto al raggio d'azione delle forze, mentre i rivelatori sono posizionati a grande distanza D dal bersaglio (Figura 4.8).

Sia a l'ampiezza del centro diffusore.

Si assumono le seguenti ipotesi:

1. si considerano solo urti elastici;
2. si suppone che la distanza tra i centri diffusori sia maggiore di a e che il bersaglio sia abbastanza sottile in modo che la particella interagisca con *un solo* centro diffusore.

Ad una grande distanza $D \gg a$ e a diversi angoli sono disposti i rivelatori che conteggiano il numero di particelle diffuse dal bersaglio nell'unità di tempo sotto l'angolo solido $d\Omega(\theta, \varphi)$ nella direzione (θ, φ) in un sistema di riferimento che ha origine nel centro delle forze con coordinate polari (Figura 4.9).

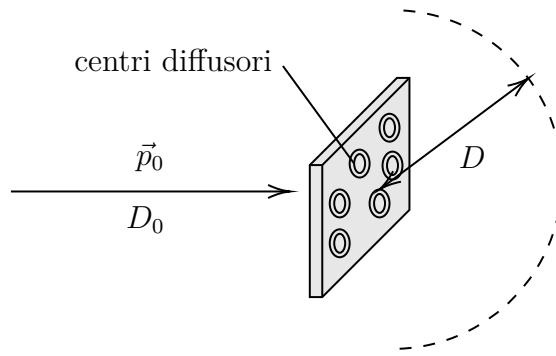


Figura 4.8: Schema esemplificativo dello scattering.

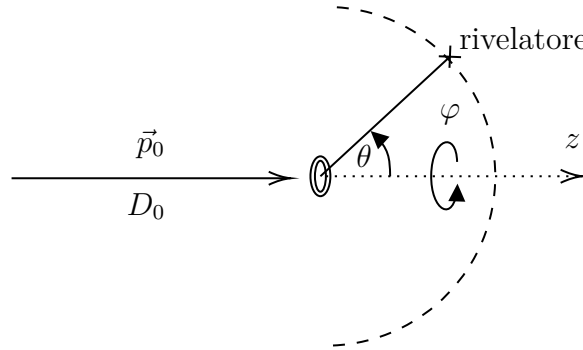


Figura 4.9: Schema delle coordinate angolari per lo scattering.

Sia Φ_i il *flusso di particelle incidenti* (pedice i) nell'unità di tempo per unità d'area. Denotiamo con \vec{j}_i la corrente corrispondente e con $\Phi_i = |\vec{j}_i|$. Si ha infatti

$$\Phi_i = \frac{dN_i}{dS dt} = \underbrace{\frac{dN_i}{dS dz}}_{dV} \frac{dz}{dt} = n_i v_i = j_i = |\vec{j}_i|$$

dove dN_i è il numero di particelle incidenti, dS l'unità d'area trasversa a \vec{p}_0 e n_i è la densità di particelle incidenti.

Sia $\Phi_d(\theta, \varphi)$ il *flusso di particelle diffuse* nell'unità di tempo e nell'unità di angolo solido nella direzione (θ, φ) . Denotiamo con \vec{j}_d la corrente associata al flusso tale che

$$\Phi_d(\theta, \varphi) = r^2 |\vec{j}_d \cdot \hat{r}| \quad \text{con} \quad \hat{r} = \frac{\vec{x}}{|\vec{x}|}$$

infatti si ha

$$\Phi_d(\theta, \varphi) = \frac{dN(\theta, \varphi)}{dt d\Omega} = \frac{dN(\theta, \varphi)}{d\Omega r^2 dr} \cdot r^2 \frac{dr}{dt} = r^2 |\vec{j}_d \cdot \hat{r}|$$

La sezione d'urto è il rapporto tra i flussi diffuso e incidente. La quantità

$$\sigma(\theta, \varphi) \equiv \frac{\Phi_d(\theta, \varphi)}{\Phi_i} = \frac{r^2 |\vec{j}_d \cdot \hat{r}|}{|\vec{j}_i|} \quad (4.114)$$

è detta *sezione d'urto differenziale*, ed ha le dimensioni di un'area. Si noti che il rapporto (4.114) è indipendente dal numero N di particelle incidenti che ho

mandato sul bersaglio, e le stesse definizioni e formule per i flussi e sezione d'urto valgono anche nel caso delle correnti di probabilità per particella singola. Si ha infatti

$$\begin{aligned}
 \sigma(\theta, \varphi) &= \frac{\text{N. di partic. diffuse per unità di tempo e angolo solido attorno a } (\theta, \varphi)}{\text{N. di partic. incidenti per unità di tempo e di superficie}} \\
 &= \frac{\text{Probabilità di diffusione per angolo solido attorno a } (\theta, \varphi)}{\text{Probabilità di incidenza per unità di superficie}} \\
 &= \frac{|\vec{j}_d \cdot \hat{r}| r^2}{|\vec{j}_i|} \quad (4.115)
 \end{aligned}$$

dove in quest'ultima formula \vec{j} assume appunto il significato di corrente di probabilità.

4.14.1 Soluzioni stazionarie

Supponiamo ora (e lo vedremo nel dettaglio in seguito) che una soluzione dell'equazione di Schrödinger stazionaria, fissato il momento incidente \vec{p} , a grande distanza dal bersaglio abbia la forma asintotica

$$\psi_{\text{as}}(\vec{x}, t) = e^{-\frac{i}{\hbar} \mathcal{E} t} (\psi_i + \psi_d) = e^{-\frac{i}{\hbar} \mathcal{E} t} \left[\underbrace{e^{\frac{i}{\hbar} \vec{p} \cdot \vec{x}}}_{=\psi_i} + \underbrace{f_{\vec{p}}(\theta, \varphi) \frac{e^{\frac{i}{\hbar} p r}}{r}}_{=\psi_d} \right] \quad (4.116)$$

dove $p = |\vec{p}|$. La ψ_i rappresenta la funzione d'onda della particella prima dell'urto, mentre la ψ_d è la funzione d'onda dopo l'urto, modulata da una funzione f che dipende dall'angolo di diffusione. Siccome a grande distanza dal bersaglio la particella è libera sappiamo perfettamente quanto vale l'energia $\mathcal{E} = p^2/2m$. Lo spin verrà trascurato nella discussione.

Consideriamo ora il caso di particella singola, dunque consideriamo le correnti di probabilità. Conoscendo la forma dell'onda di probabilità conosciamo anche la forma delle correnti di probabilità direttamente dalla loro definizione:

$$\begin{aligned}
 \vec{j}_i &\equiv \frac{\hbar}{2m} (\psi_i^* \vec{\nabla} \psi_i - \vec{\nabla} \psi_i^* \psi_i) \\
 &= \frac{\hbar}{2m} 2\text{Im}\{\psi_i^* \vec{\nabla} \psi_i\} \\
 &= \frac{\hbar}{m} \text{Im}\{e^{-i \frac{\vec{p} \cdot \vec{x}}{\hbar}} \vec{\nabla} e^{i \frac{\vec{p} \cdot \vec{x}}{\hbar}}\} \\
 &= \frac{\hbar}{m} \text{Im}\{e^{-i \frac{\vec{p} \cdot \vec{x}}{\hbar}} \frac{\vec{p}}{\hbar} e^{i \frac{\vec{p} \cdot \vec{x}}{\hbar}}\} \\
 &= \frac{\hbar}{m} \cdot \frac{\vec{p}}{\hbar} = \frac{\vec{p}}{m} = \vec{v}
 \end{aligned}$$

cioè si ottiene la velocità della particella. Per \vec{j}_d si ha analogamente

$$\begin{aligned}
|\vec{j}_d \cdot \hat{r}| &= \frac{\hbar}{m} \text{Im}[\psi_d^* \hat{r} \cdot \overbrace{\vec{\nabla}}^{\partial/\partial r} \psi_d] \\
&= \frac{\hbar}{m} \text{Im} \left[\psi_d^* \frac{\partial}{\partial r} \psi_d \right] \\
&= \frac{\hbar}{m} \text{Im} \left[\left(f_{\vec{p}}(\theta, \varphi) \frac{e^{i\frac{pr}{\hbar}}}{r} \right)^* \frac{\partial}{\partial r} \left(f_{\vec{p}}(\theta, \varphi) \frac{e^{i\frac{pr}{\hbar}}}{r} \right) \right] \\
&= \frac{\hbar}{m} \text{Im} \left[f_{\vec{p}}^*(\theta, \varphi) f_{\vec{p}}(\theta, \varphi) \frac{e^{-i\frac{pr}{\hbar}}}{r} \frac{\partial}{\partial r} \left(\frac{e^{i\frac{pr}{\hbar}}}{r} \right) \right] \\
&= \frac{\hbar}{m} \text{Im} \left[|f_{\vec{p}}(\theta, \varphi)|^2 \left(i \frac{p}{\hbar} \frac{1}{r^2} - \frac{1}{r^3} \right) \right] \\
&= \frac{\hbar}{m} |f_{\vec{p}}(\theta, \varphi)|^2 \frac{p}{\hbar} \frac{1}{r^2} = v \frac{|f_{\vec{p}}(\theta, \varphi)|^2}{r^2}
\end{aligned}$$

da cui si ottiene la sezione d'urto differenziale dalla definizione (4.115):

$$\sigma(\theta, \varphi) = \frac{|\vec{j}_d \cdot \hat{r}| r^2}{|\vec{j}_i|} = \frac{v \frac{|f_{\vec{p}}(\theta, \varphi)|^2}{r^2} r^2}{v} = |f_{\vec{p}}(\theta, \varphi)|^2 \quad (4.117)$$

4.14.2 Pacchetto d'onda incidente

La trattazione precedente tuttavia non riproduce un processo fisico, dato che tale processo non è stazionario. La situazione sperimentale è costituita da particelle incidenti che sono libere molto prima che avvenga l'urto ($t \rightarrow -\infty$), in cui siamo a grande distanza dal bersaglio, e nuovamente libere molto dopo che avvenga l'urto ($t \rightarrow +\infty$), diffuse a grande distanza. Tale funzione non è stazionaria. Notiamo inoltre che la funzione d'onda (4.116) che abbiamo costruito non è nemmeno in L^2 (non rappresenta dunque uno stato fisico).

Per riprodurre la situazione fisica dobbiamo quindi considerare dei *pacchetti d'onda* con le caratteristiche richieste: saranno perciò formati con soluzioni stazionarie e riprodurranno comunque la (4.117), anche se in più apparirà una relazione tra f e il potenziale V di interazione. La trattazione precedente, dunque, è quella “ingenua”, anche se porta allo stesso risultato. Vediamo in seguito la trattazione rigorosa.

Costruiamo il pacchetto con una funzione reale $g(\vec{x}) \in L^2(\mathbb{R}^3)$, piccata su $\vec{0}$, centrata in una regione di spazio di dimensioni longitudinali (rispetto al bersaglio) $\Delta x \sim \Delta y \sim l$ e trasversali $\Delta z \sim d$. Vogliamo poi che la sua trasformata di Fourier $\tilde{g}(\vec{p})$ sia reale e anch'essa piccata in $\vec{0}$. Una gaussiana 3D, ad esempio, soddisfa tutte queste proprietà, ma teniamoci più generali considerando una qualsiasi g . Se poi indichiamo con $\vec{p}_0 \parallel \vec{z}$ il momento medio del pacchetto incidente e con Δp_i la dispersione del momento incidente, richiediamo che $\Delta p_i \ll p_0$ in modo tale che tutto il pacchetto viaggi verso un'unica direzione. Si noti che per Heisenberg si ha che per le tante lunghezze d'onda che costituiscono il pacchetto vale

$$\Delta x_i \gg \Delta x_i \frac{\Delta p_i}{p_0} \gtrsim \frac{\hbar}{p_0} \sim \frac{1}{2\pi}$$

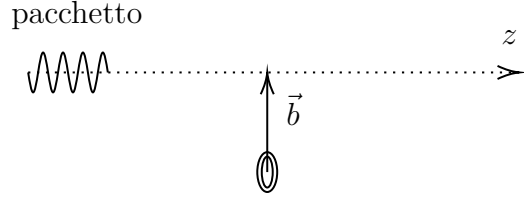


Figura 4.10: Schema sul significato del parametro d'impatto.

Introduciamo poi il parametro d'impatto $\vec{b} \perp \hat{z}$ (analogo a quello classico) diretto dal centro di scattering alla traiettoria incidente (Figura 4.10).

Ma com'è fatto e come si può scrivere il pacchetto d'onda incidente $\psi_i(\vec{x}, t)$? Per $t \rightarrow -\infty$, oltre alle richieste fatte finora, vorremmo ovviamente dei pacchetti liberi, pertanto la ψ_i dovrà essere composta da funzioni d'onda $e^{\frac{i}{\hbar}\vec{p} \cdot (\vec{x} - \vec{b})}$ corrispondenti alla particella libera, quindi

$$\psi_i(\vec{x}, t) = \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}(\vec{p} - \vec{p}_0) e^{\frac{i}{\hbar}\vec{p} \cdot (\vec{x} - \vec{b})} e^{-\frac{i}{\hbar} \frac{\vec{p}^2}{2m} t} \quad (4.118)$$

dove il termine $\frac{\vec{p}^2}{2m} = \mathcal{E}$ è l'energia all'infinito, che naturalmente è conservata nel processo di scattering. Si noti che la definizione utilizzata della trasformata di Fourier è quella con la convenzione “fisica” dei coefficienti. Della quantità (4.118) espando al prim'ordine attorno a \vec{p}_0 , cambiando poi variabile mediante la sostituzione $\vec{p} - \vec{p}_0 \rightarrow \vec{p}$ (oppure, equivalentemente $\vec{p} \rightarrow \vec{p} + \vec{p}_0$). Possiamo quindi riscrivere la (4.118) in modo approssimato come

$$\begin{aligned} \psi_i(\vec{x}, t) &= \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}(\vec{p} - \vec{p}_0) e^{\frac{i}{\hbar}\vec{p} \cdot (\vec{x} - \vec{b})} e^{-\frac{i}{\hbar} \frac{\vec{p}^2}{2m} t} \\ (\vec{p} \rightarrow \vec{p} + \vec{p}_0) &= \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}(\vec{p}) e^{\frac{i}{\hbar}(\vec{p} + \vec{p}_0) \cdot (\vec{x} - \vec{b})} e^{-\frac{i}{\hbar} \frac{(\vec{p} + \vec{p}_0)^2}{2m} t} \\ &\stackrel{(a)}{=} \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}(\vec{p}) e^{\frac{i}{\hbar}\vec{p} \cdot \vec{x}} e^{\frac{i}{\hbar}\vec{p}_0 \cdot \vec{x}} e^{-\frac{i}{\hbar}\vec{p} \cdot \vec{b}} e^{-\frac{i}{\hbar}\vec{p}_0 \cdot \vec{b}} e^{-\frac{i}{\hbar} \frac{\vec{p}^2}{2m} t} e^{-\frac{i}{\hbar} \frac{\vec{p}_0^2}{2m} t} e^{-\frac{i}{\hbar} \frac{\vec{p} \cdot \vec{p}_0}{m} t} \\ &= \underbrace{\int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}(\vec{p}) e^{\frac{i}{\hbar}\vec{p} \cdot (\vec{x} - \vec{b} - \frac{\vec{p}_0}{m} t)} e^{\frac{i}{\hbar}\vec{p}_0 \cdot \vec{x}} e^{-\frac{i}{\hbar} \frac{\vec{p}_0^2}{2m} t}}_{g\left(\vec{x} - \vec{b} - \frac{\vec{p}_0}{m} t\right)} e^{\frac{i}{\hbar}\vec{p}_0 \cdot \vec{x}} e^{-\frac{i}{\hbar} \frac{\vec{p}_0^2}{2m} t} \end{aligned}$$

Il termine $-\frac{i}{\hbar} \frac{\vec{p} \cdot \vec{p}_0}{m} t$ nel passaggio (a) è trascurato perché del second'ordine in \vec{p} . Pertanto il pacchetto d'onda incidente risulta della forma

$$\psi_i(\vec{x}, t) = g\left(\vec{x} - \vec{b} - \frac{\vec{p}_0}{m} t\right) e^{\frac{i}{\hbar}\vec{p}_0 \cdot \vec{x}} e^{-\frac{i}{\hbar} \frac{\vec{p}_0^2}{2m} t} \quad (4.119)$$

Notiamo che, dalla definizione di g , il pacchetto è diverso da zero se $z - \frac{p_0}{m} t \lesssim l$ e $|\vec{x}_\perp - \vec{b}| \lesssim d$ (Figura 4.11).

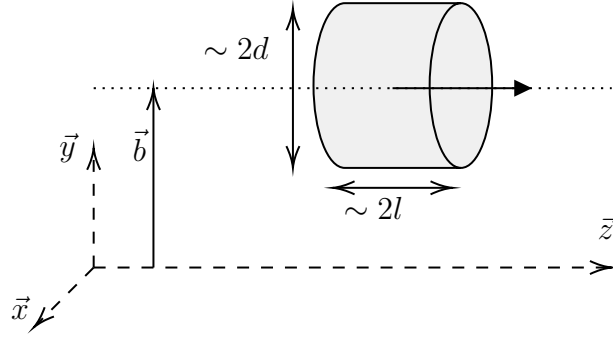


Figura 4.11: Schema del pacchetto d'onda incidente.

4.14.3 Pacchetto d'onda diffuso

Il risultato appena trovato è l'analogo della funzione $e^{\frac{i}{\hbar}\vec{p}\cdot\vec{x}}$ di prima, che descriveva il pacchetto molto prima dell'urto, quando la particella è libera.

Ora però vorremmo anche costruire l'analogo in versione di pacchetto d'onda della funzione asintotica (4.116), che per distinguerla dalla ψ_{as} la denoteremo con ψ_{in} (da interazione). Mentre le soluzioni stazionarie per una particella libera sono note, nel caso di una particella deviata da un potenziale V non lo sono. Per trovarla imponiamo che tale funzione obbedisca all'equazione della dinamica $H\psi_{\text{in}}(\vec{x}) = \mathcal{E}\psi_{\text{in}}(\vec{x})$ con l'hamiltoniana H “vera”, cioè quella completa di potenziale di interazione $V(\vec{x})$:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{x})\right)\psi_{\text{in}}(\vec{x}) = \frac{p^2}{2m}\psi_{\text{in}}(\vec{x}) \quad (4.120)$$

Dobbiamo dunque risolvere questa equazione per trovare la $\psi_{\text{in}}(\vec{x})$, da cui poi sarà possibile costruire il pacchetto. Tutto lo scopo dello scattering è ricavare da questa analisi la forma di $V(\vec{x})$ dalla sezione d'urto. Riscriviamo la (4.120) in una forma diversa:

$$\left(\frac{\hbar^2}{2m}\nabla^2 + \frac{p^2}{2m}\right)\psi_{\text{in}}(\vec{x}) = V(\vec{x})\psi_{\text{in}} \quad (4.121)$$

conosciuta come *Equazione di Lippmann-Schwinger*. In questa forma l'equazione è più semplice da risolvere per il fatto che la si può vedere in forma operatoriale come

$$A_p\psi_{\text{in}}(\vec{x}) = V(\vec{x})\psi_{\text{in}}(\vec{x}) \quad \text{con} \quad A_p = \frac{\hbar^2}{2m}\nabla^2 + \frac{p^2}{2m}$$

Osserviamo per prima cosa che l'onda piana $e^{\frac{i}{\hbar}\vec{p}\cdot\vec{x}}$ soddisfa l'omogenea associata $A_p e^{\frac{i}{\hbar}\vec{p}\cdot\vec{x}} = 0$, infatti applicando $\frac{\hbar^2}{2m}\nabla^2$ a $e^{\frac{i}{\hbar}\vec{p}\cdot\vec{x}}$ si ottiene $-\frac{p^2}{2m}$. Pertanto $e^{\frac{i}{\hbar}\vec{p}\cdot\vec{x}}$ è una soluzione dell'omogenea associata all'equazione di Lippmann-Schwinger, mentre una soluzione particolare si può trovare invertendo l'operatore A_p :

$$A_p\psi_{\text{in}} = V\psi_{\text{in}} \implies \psi_{\text{in}} = A_p^{-1}V\psi_{\text{in}}$$

La soluzione formale generale è la somma della soluzione dell'omogenea più quella particolare:

$$\psi_{\text{in}}(\vec{x}) = e^{\frac{i}{\hbar}\vec{p}\cdot\vec{x}} + (A_p^{-1}V\psi_{\text{in}})(\vec{x}) \quad (4.122)$$

Tale equazione è implicita (non determina esplicitamente ψ_{in}), e questo causa alcune complicazioni che vedremo più avanti. Come vedremo in seguito, il primo termine corrisponde alla funzione d'onda incidente, mentre il secondo a quella diffusa. La $\psi_{\text{in}}(\vec{x})$ è la funzione d'onda stazionaria con p fissato, mentre la soluzione più generale è il pacchetto di tutte le $\psi_{\text{in}}(\vec{x})$ integrate su p :

$$\psi(\vec{x}, t) = \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}(\vec{p} - \vec{p}_0) \psi_{\text{in}}(\vec{x}) e^{-\frac{i}{\hbar} \frac{\vec{p}^2}{2m} t}$$

Ora però vogliamo scegliere una forma particolare di A_p^{-1} tale che per $t \rightarrow -\infty$ il termine $A_p^{-1} V \psi_{\text{in}}$ della (4.122) non dia il contributo al pacchetto. In questo modo impongo che per $t \rightarrow -\infty$ mi rimanga solo il termine $e^{\frac{i}{\hbar} \vec{p} \cdot \vec{x}}$. Per far questo è però necessario invertire l'operatore differenziale A_p . Rappresentiamo innanzitutto A_p^{-1} in forma integrale, ovvero scrivo l'azione dell'operatore come un integrale:

$$A_p^{-1}(V \psi_{\text{in}}(\vec{x})) = \int d^3y A_p^{-1}(\vec{x} - \vec{y}) V(\vec{y}) \psi_{\text{in}}(\vec{y}) \quad (4.123)$$

detto anche *nucleo integrale*. La funzione $A_p^{-1}(\vec{x} - \vec{y})$ deve ovviamente soddisfare l'identità

$$A_p A_p^{-1}(\vec{x} - \vec{y}) = \delta^3(\vec{x} - \vec{y})$$

In modo che A_p^{-1} sia effettivamente l'inverso di A_p anche in forma integrale. Per trovare un'espressione integrale di $A_p^{-1}(\vec{x} - \vec{y})$ si utilizza la trasformata di Fourier:

$$A_p = \frac{\hbar^2}{2m} \vec{\nabla}^2 + \frac{p^2}{2m} \xrightarrow{\text{Fourier}} \tilde{A}_p(\vec{q}) = -\frac{\vec{q}^2}{2m} + \frac{\vec{p}^2}{2m}$$

Questa, in generale, è la “regola d'oro” per invertire operatori differenziali: siccome la trasformata di Fourier di operatori differenziali è sempre in forma polinomiale, ciò che si fa è applicare la trasformata ottenendo un polinomio, invertire tale polinomio e una volta fatto ciò si applica l'antitrasformata per ricavare il nucleo integrale. Ma nel nostro caso $\tilde{A}_p(\vec{q})$ è uno scalare, pertanto l'inverso di $\tilde{A}_p(\vec{q})$ sarà il suo reciproco:

$$\widetilde{A_p^{-1}}(\vec{q}) = \frac{1}{\tilde{A}_p(\vec{q})} = \frac{1}{\frac{\vec{p}^2}{2m} - \frac{\vec{q}^2}{2m}}$$

A questo punto rimarrebbe da invertire questa distribuzione. Tuttavia si ha una singolarità per $|\vec{p}| = |\vec{q}|$, quindi tale oggetto non è una distribuzione ben definita, e occorre una regolarizzazione. Ad esempio si può utilizzare la *parte principale*. Oppure, un altro metodo efficace di regolarizzazione è la *prescrizione* $\pm i\delta$:

$$\widetilde{A_p^{-1}}(\vec{q}) = \frac{1}{\frac{\vec{p}^2}{2m} - \frac{\vec{q}^2}{2m} + i\delta} \quad (\delta \rightarrow 0)$$

Quindi, applicando la trasformata inversa a quest'ultima espressione si ottiene

$$A_p^{-1}(\vec{x} - \vec{y}) = \int \frac{d^3q}{(2\pi\hbar)^3} \frac{e^{-i\frac{\vec{q}}{\hbar} \cdot (\vec{x} - \vec{y})}}{\frac{\vec{p}^2}{2m} - \frac{\vec{q}^2}{2m} + i\delta} \quad (4.124)$$

Dimostriamo dunque che il contributo $A_p^{-1}V\psi_{\text{in}}$ della soluzione stazionaria (e quindi anche di quello di un qualunque pacchetto d'onda) dell'equazione di Lippmann-Schwinger tende a zero per $t \rightarrow -\infty$. Inseriamo nella (4.123) la forma a pacchetto di ψ_{in} e la forma (4.124) appena ricavata di A_p^{-1} :

$$\begin{aligned} (A_p^{-1}V\psi_{\text{in}})(\vec{x}) &= \int d^3y A_p^{-1}(\vec{x} - \vec{y})V(\vec{y})\psi_{\text{in}}(\vec{y}) \\ &= \int d^3y \int \frac{d^3q}{(2\pi\hbar)^3} \frac{e^{-i\frac{\vec{q}}{\hbar}(\vec{x}-\vec{y})}}{\frac{p^2}{2m} - \frac{q^2}{2m} + i\delta} V(\vec{y}) \int \frac{d^3p}{(2\pi\hbar)^3} \underbrace{\tilde{g}(\vec{p} - \vec{p}_0)}_{\tilde{g}_{\vec{p}_0}(\vec{p})} e^{\frac{i}{\hbar}\vec{p}\cdot(\vec{y}-\vec{b})} e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} \end{aligned}$$

dove la ψ_{in} è stata scritta come un pacchetto di onde piane, ovvero la soluzione dell'omogenea associata dell'equazione di Lippman-Schwinger (shiftata in \vec{x} del parametro d'impatto \vec{b}). L'ultima espressione la possiamo riscrivere riarrangiando i termini all'interno degli integrali, come

$$(A_p^{-1}V\psi_{\text{in}})(\vec{x}) = \int d^3y V(\vec{y}) \int \frac{d^3q}{(2\pi\hbar)^3} e^{-i\frac{\vec{q}}{\hbar}(\vec{x}-\vec{y})} \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}_{\vec{p}_0}(\vec{p}) \frac{e^{\frac{i}{\hbar}\vec{p}\cdot(\vec{y}-\vec{b})} e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}}{\frac{p^2}{2m} - \frac{q^2}{2m} + i\delta}$$

A questo punto dimostriamo che il terzo integrale

$$\int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}_{\vec{p}_0}(\vec{p}) \frac{e^{\frac{i}{\hbar}\vec{p}\cdot(\vec{y}-\vec{b})} e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}}{\frac{p^2}{2m} - \frac{q^2}{2m} + i\delta} \quad (4.125)$$

è nullo per $t \rightarrow -\infty$. Per farlo dobbiamo calcolare l'integrale scrivendo la $\tilde{g}_{\vec{p}_0}(\vec{p})$ in coordinate polari ($|\vec{p}|, \theta_p, \varphi_p$) lungo l'asse $\vec{x} - \vec{b}$ ed eseguendo un cambio di variabile $p \rightarrow \mathcal{E} = \frac{p^2}{2m}$. Indicando con $d\Omega_p$ l'angolo solido in p si ottiene

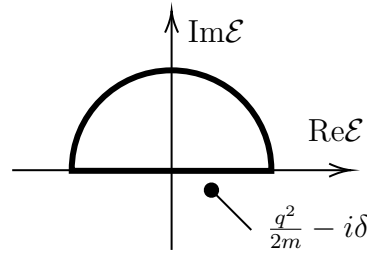
$$\begin{aligned} d^3p &\rightarrow p^2 dp \overbrace{d\theta_p d\varphi_p}^{d\Omega_p} \\ &= 2mEd \left(\sqrt{2mE} \right) d\Omega_p \\ &= 2mE\sqrt{2m} \frac{1}{2\sqrt{E}} dE d\Omega_p \\ &= \frac{(2m)^{\frac{3}{2}}}{2} \sqrt{\mathcal{E}} d\mathcal{E} d\Omega_p \end{aligned}$$

$$\tilde{g}(\vec{p}) = \tilde{g}(|\vec{p}|, \theta_p, \varphi_p) = \tilde{g}(\sqrt{2m\mathcal{E}}, \theta_p, \varphi_p)$$

L'integrale (4.125) diventa quindi

$$\frac{(2m)^{\frac{3}{2}}}{2} \int \frac{d\Omega_p}{(2\pi\hbar)^3} \int_{-\infty}^{+\infty} d\mathcal{E} \sqrt{\mathcal{E}} H(\mathcal{E}) \tilde{g}(\sqrt{2m\mathcal{E}}, \theta_p, \varphi_p) \frac{e^{\frac{i}{\hbar}\sqrt{2m\mathcal{E}}|\vec{x}-\vec{b}|\cos\theta_p} e^{-\frac{i}{\hbar}Et}}{\mathcal{E} - \left(\frac{q^2}{2m} - i\delta\right)}$$

dove H è la funzione di Heaviside (l'integrale in E senza la H andrebbe da 0 a $+\infty$). Usando il teorema dei residui, poiché il polo si trova in $\mathcal{E} = \frac{q^2}{2m} - i\delta$, per $t < 0$ in modulo sufficientemente grande possiamo chiudere il contorno nel semipiano complesso $\text{Im}\mathcal{E} > 0$:



Per $t \rightarrow -\infty$ domina il termine $e^{-\frac{i}{\hbar}Et}$. L'area scelta non contiene poli pertanto l'integrale è pari a zero, e così a $t \rightarrow -\infty$ il termine $A_p^{-1}V\psi_{\text{in}}$ è nullo e non contribuisce al pacchetto.

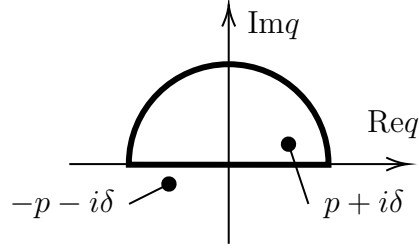
Rimane ora da mostrare che per $t \rightarrow +\infty$ il termine $A_p^{-1}V\psi_{\text{in}}$ *riproduce un pacchetto diffuso* e che *determina f in funzione di V* . Poniamoci innanzitutto nell'approssimazione $D \gg a, \Delta x$. Riscriviamo la (4.124) come

$$\begin{aligned} A_p^{-1}(\vec{x} - \vec{y}) &= \int \frac{d^3q}{(2\pi\hbar)^3} \frac{e^{\frac{i}{\hbar}\vec{q}\cdot(\vec{x}-\vec{y})}}{\frac{p^2}{2m} - \frac{q^2}{2m} + i\delta} \\ &= 2m \int \frac{d^3q}{(2\pi\hbar)^3} \frac{e^{\frac{i}{\hbar}\vec{q}\cdot(\vec{x}-\vec{y})}}{p^2 - q^2 + im\delta} \\ &= \frac{2m}{(2\pi\hbar)^3} \int d^3q \frac{e^{\frac{i}{\hbar}\vec{q}\cdot(\vec{x}-\vec{y})}}{p^2 - q^2 + i\delta} \end{aligned}$$

dove nell'ultimo passaggio si è operata la sostituzione $m\delta \rightarrow \delta$ dato che rappresenta comunque un termine che deve tendere a 0. Passando poi in coordinate polari $(|\vec{q}|, \theta_q, \varphi_q)$ si ottiene (ricordando lo jacobiano q^2):

$$\begin{aligned} A_p^{-1}(\vec{x} - \vec{y}) &= \frac{2m}{(2\pi\hbar)^3} \underbrace{\int_0^{2\pi} d\varphi_q}_{2\pi} \int_{-1}^1 d\cos\theta_q \int_0^{+\infty} dq \frac{e^{\frac{i}{\hbar}q|\vec{x}-\vec{y}|\cos\theta_q}}{p^2 - q^2 + i\delta} q^2 \\ &= \frac{2m}{\hbar^3} \frac{1}{(2\pi)^2} \int_{-1}^1 dk \int_0^{+\infty} dq \frac{e^{\frac{i}{\hbar}q|\vec{x}-\vec{y}|k}}{p^2 - q^2 + i\delta} q^2 \quad (k = \cos\theta_q) \\ &= \frac{2m}{\hbar^3} \frac{1}{(2\pi)^2} \int_0^{+\infty} dq \frac{\hbar}{i} \frac{q^2}{q|\vec{x} - \vec{y}|} \frac{e^{\frac{i}{\hbar}q|\vec{x}-\vec{y}|} - e^{-\frac{i}{\hbar}q|\vec{x}-\vec{y}|}}{p^2 - q^2 + i\delta} \\ &= i \frac{2m}{(2\pi\hbar)^2} \frac{1}{|\vec{x} - \vec{y}|} \int_0^{+\infty} dq \frac{qe^{\frac{i}{\hbar}q|\vec{x}-\vec{y}|} - qe^{-\frac{i}{\hbar}q|\vec{x}-\vec{y}|}}{p^2 - q^2 + i\delta} \\ &= i \frac{2m}{(2\pi\hbar)^2} \frac{1}{|\vec{x} - \vec{y}|} \left(\int_0^{+\infty} dq \frac{qe^{\frac{i}{\hbar}q|\vec{x}-\vec{y}|}}{p^2 - q^2 + i\delta} - \int_0^{+\infty} dq \frac{qe^{-\frac{i}{\hbar}q|\vec{x}-\vec{y}|}}{p^2 - q^2 + i\delta} \right) \\ &= i \frac{2m}{(2\pi\hbar)^2} \frac{1}{|\vec{x} - \vec{y}|} \left(\int_0^{+\infty} dq \frac{qe^{\frac{i}{\hbar}q|\vec{x}-\vec{y}|}}{p^2 - q^2 + i\delta} + \underbrace{\int_{-\infty}^0 dq \frac{qe^{+\frac{i}{\hbar}q|\vec{x}-\vec{y}|}}{p^2 - q^2 + i\delta}}_{q \rightarrow -q} \right) \\ &= i \frac{2m}{(2\pi\hbar)^2} \frac{1}{|\vec{x} - \vec{y}|} \int_{-\infty}^{+\infty} dq \frac{q}{q^2 - p^2 - i\delta} e^{\frac{i}{\hbar}q|\vec{x}-\vec{y}|} \end{aligned}$$

Si hanno dunque due poli in $\pm\sqrt{p^2 + i\delta}$, che posso espandere in $\pm(p + i\delta)$ per $\delta \rightarrow 0$. Chiudiamo il contorno per $\text{Im}q > 0$:



Utilizzando il teorema dei residui si ottiene infine

$$\begin{aligned}
 A_p^{-1}(\vec{x} - \vec{y}) &= i \frac{2m}{(2\pi\hbar)^2 |\vec{x} - \vec{y}|} \int_{-\infty}^{+\infty} dq \frac{q}{q^2 - p^2 - i\delta} e^{\frac{i}{\hbar} q |\vec{x} - \vec{y}|} \\
 &= i \frac{2m}{(2\pi\hbar)^2 |\vec{x} - \vec{y}|} 2\pi i \frac{p}{2p} e^{\frac{i}{\hbar} p |\vec{x} - \vec{y}|} \\
 &= -\frac{2m}{\hbar^2} \frac{1}{4\pi} \frac{e^{\frac{i}{\hbar} p |\vec{x} - \vec{y}|}}{|\vec{x} - \vec{y}|}
 \end{aligned}$$

Il potenziale V è non nullo solo vicino al centro diffusore. La \vec{y} è integrata sul supporto del potenziale $V(\vec{y})$, ovvero nei punti vicino all'origine, cioè per $|\vec{y}| \lesssim a$. Nell'equazione di Lippmann-Schwinger abbiamo dunque che $A_p^{-1}(\vec{x} - \vec{y})V(\vec{y})\psi_{\text{in}}(\vec{y}) \neq 0$ solo per $|\vec{y}| \lesssim a \ll D = |\vec{x}|$.

Si ha poi che $|\vec{x}| \sim D$, e dato che ci poniamo nell'approssimazione $a \ll D$, quindi $|\vec{y}| \ll |\vec{x}|$ si può espandere $|\vec{x} - \vec{y}|$ ottenendo

$$|\vec{x} - \vec{y}| = \sqrt{x^2 + y^2 - 2\vec{x} \cdot \vec{y}} \simeq |\vec{x}| \left(1 - \frac{1}{2} \cdot 2 \frac{\vec{x} \cdot \vec{y}}{x^2} \right) = |\vec{x}| - \hat{r} \cdot \vec{y} = r - \hat{r} \cdot \vec{y}$$

in cui abbiamo posto $|\vec{x}| = r$, e da questo si ottiene

$$\begin{aligned}
 A_p^{-1}(\vec{x} - \vec{y}) &= -\frac{2m}{\hbar^2} \frac{1}{4\pi} \frac{e^{\frac{i}{\hbar} p |\vec{x} - \vec{y}|}}{|\vec{x} - \vec{y}|} \\
 &\simeq -\frac{2m}{\hbar^2} \frac{1}{4\pi} \frac{e^{\frac{i}{\hbar} p (r - \hat{r} \cdot \vec{y})}}{r - \underbrace{\hat{r} \cdot \vec{y}}_{\ll r}} \\
 &\simeq -\frac{2m}{\hbar^2} \frac{1}{4\pi} \frac{e^{\frac{i}{\hbar} pr}}{r} e^{-\frac{i}{\hbar} p \hat{r} \cdot \vec{y}}
 \end{aligned} \tag{4.126}$$

A questo punto consideriamo il pacchetto $\psi_d(\vec{x}, t)$ costituito dalle soluzioni stazionarie $A_p^{-1}V\psi_{\text{in}}$ relative alla componente diffusa:

$$\psi_d(\vec{x}, t) = \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}(\vec{p} - \vec{p}_0) e^{-\frac{i}{\hbar} \vec{p} \cdot \vec{b}} e^{-\frac{i}{\hbar} \frac{p^2}{2m} t} (A_p^{-1}V\psi_{\text{in}})(\vec{x})$$

dove si è inserito anche lo shift di \vec{b} e il termine di evoluzione temporale.

A questo punto si inserisce nell'equazione precedente la forma integrale (4.123) di $A_p^{-1}V\psi_{\text{in}}$:

$$\psi_d(\vec{x}, t) = \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}_{\vec{p}_0}(\vec{p}) e^{-\frac{i}{\hbar}\vec{p}\cdot\vec{b}} e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} \int d^3y A_p^{-1}(\vec{x} - \vec{y}) V(\vec{y}) \psi_{\text{in}}(\vec{y})$$

e si inserisce poi l'approssimazione (4.126) di $A_p^{-1}(\vec{x} - \vec{y})$ all'interno del pacchetto ottenendo

$$\psi_d(\vec{x}, t) = \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}_{\vec{p}_0}(\vec{p}) e^{-\frac{i}{\hbar}\vec{p}\cdot\vec{b}} e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} \int d^3y \left(-\frac{2m}{\hbar^2} \frac{1}{4\pi} \frac{e^{\frac{i}{\hbar}pr}}{r} e^{-\frac{i}{\hbar}p\hat{r}\cdot\vec{y}} \right) V(\vec{y}) \psi_{\text{in}}(\vec{y})$$

Si possono isolare nell'integrale finale i termini che dipendono da (θ, φ) :

$$\psi_d(\vec{x}, t) = \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}_{\vec{p}_0}(\vec{p}) e^{-\frac{i}{\hbar}\vec{p}\cdot\vec{b}} e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} \frac{e^{\frac{i}{\hbar}pr}}{r} \underbrace{\int d^3y \left(-\frac{2m}{\hbar^2} \frac{1}{4\pi} \right) e^{-\frac{i}{\hbar}p\hat{r}(\theta, \varphi)\cdot\vec{y}} V(\vec{y}) \psi_{\text{in}}(\vec{y})}_{\equiv f_{\vec{p}}(\theta, \varphi)} \quad (4.127)$$

Il termine finale corrisponde alla funzione angolare $f_{\vec{p}}(\theta, \varphi)$ che avevamo scritto all'inizio della trattazione sullo scattering. Si ottiene quindi

$$\psi_d(\vec{x}, t) = \int \frac{d^3p}{(2\pi\hbar)^3} \tilde{g}(\vec{p} - \vec{p}_0) e^{-\frac{i}{\hbar}\vec{p}\cdot\vec{b}} e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} \frac{e^{\frac{i}{\hbar}pr}}{r} f_{\vec{p}}(\theta, \varphi) \quad (4.128)$$

A questo punto, esattamente come si è fatto per il pacchetto incidente, si espande in \vec{p}_0 (considerando quindi $p \ll p_0$) e si attua il cambio di variabile $\vec{p} \rightarrow \vec{p} + \vec{p}_0$. Facciamolo per ogni termine:

$$\begin{aligned} \tilde{g}(\vec{p} - \vec{p}_0) &\xrightarrow{\vec{p} \rightarrow \vec{p} + \vec{p}_0} \tilde{g}(\vec{p}) \\ -\frac{i}{\hbar}\vec{p} \cdot \vec{b} &\xrightarrow{\vec{p} \rightarrow \vec{p} + \vec{p}_0} -\frac{i}{\hbar}(\vec{p} + \vec{p}_0) \cdot \vec{b} \\ &= -\frac{i}{\hbar}(\vec{p} \cdot \vec{b} + \vec{p}_0 \cdot \vec{b}) \\ &= -\frac{i}{\hbar}\vec{p} \cdot \vec{b} \quad (\vec{p}_0 \cdot \vec{b} = 0) \\ -\frac{i}{\hbar}\frac{p^2}{2m}t &\xrightarrow{\vec{p} \rightarrow \vec{p} + \vec{p}_0} -\frac{i}{\hbar}\frac{(\vec{p} + \vec{p}_0) \cdot (\vec{p} + \vec{p}_0)}{2m}t \\ &= -\frac{i}{\hbar}\frac{p^2 + p_0^2 + 2\vec{p} \cdot \vec{p}_0}{2m}t \\ (p^2 \sim 0) &= -\frac{i}{\hbar}\frac{p_0^2}{2m}t - \frac{i}{\hbar}\frac{\vec{p} \cdot \vec{p}_0}{m}t \\ &\quad -\frac{i}{\hbar}\frac{p_0^2}{2m}t - \frac{i}{\hbar}(\vec{p} \cdot \hat{p}_0)\frac{p_0 t}{m} \\ \frac{i}{\hbar}pr &= \frac{i}{\hbar}\sqrt{\vec{p} \cdot \vec{p}}r \\ &\xrightarrow{\vec{p} \rightarrow \vec{p} + \vec{p}_0} \frac{i}{\hbar}\sqrt{(\vec{p} + \vec{p}_0) \cdot (\vec{p} + \vec{p}_0)}r \\ &= \frac{i}{\hbar}\sqrt{p^2 + p_0^2 + 2\vec{p}_0 \cdot \vec{p}}r \\ (p^2 \sim 0) &= \frac{i}{\hbar}\sqrt{p_0^2 + 2\vec{p}_0 \cdot \vec{p}}r \\ &= \frac{i}{\hbar}p_0\sqrt{1 + 2\hat{p}_0 \cdot \frac{\vec{p}}{p_0}}r \\ &\simeq \frac{i}{\hbar}p_0\left(1 + \hat{p}_0 \cdot \frac{\vec{p}}{p_0}\right)r \\ &= \frac{i}{\hbar}p_0r + \frac{i}{\hbar}\hat{p}_0 \cdot \vec{p}r \end{aligned}$$

$$\begin{aligned}
f_{\vec{p}}(\theta, \varphi) &\xrightarrow{\vec{p} \rightarrow \vec{p} + \vec{p}_0} f_{\vec{p} + \vec{p}_0}(\theta, \varphi) \\
&= |f_{\vec{p} + \vec{p}_0}(\theta, \varphi)| e^{i \arg f_{\vec{p} + \vec{p}_0}} \\
&\simeq |f_{\vec{p}_0}(\theta, \varphi)| e^{i \arg f_{p_0} + i (\nabla_p \arg f_p) |_{p_0} \cdot \vec{p}} \\
&= |f_{\vec{p}_0}(\theta, \varphi)| e^{i \arg f_{p_0} + \overbrace{i \hbar (\nabla_p \arg f_p) |_{p_0} \cdot \vec{p}}^{\equiv \vec{s}}} \\
&= |f_{\vec{p}_0}(\theta, \varphi)| e^{i \arg f_{p_0} + \frac{i}{\hbar} \vec{s} \cdot \vec{p}} \\
&= |f_{\vec{p}_0}(\theta, \varphi)| e^{i \arg f_{p_0}} e^{\frac{i}{\hbar} \vec{s} \cdot \vec{p}} \\
&= f_{\vec{p}_0}(\theta, \varphi) e^{\frac{i}{\hbar} \vec{s} \cdot \vec{p}}
\end{aligned}$$

dove si è definito

$$\vec{s} \equiv \hbar (\nabla_p \arg f_p) |_{p_0} \cdot \vec{p}_0$$

e vale $|\vec{s}| \simeq a \ll l, d$. Pertanto operando le sostituzioni appena viste nella (4.128) si ottiene

$$\begin{aligned}
\psi_d(\vec{x}, t) &= \int \frac{d^3 p}{(2\pi\hbar)^3} \tilde{g}(\vec{p}) e^{-\frac{i}{\hbar} \vec{p} \cdot \vec{b}} e^{-\frac{i}{\hbar} \frac{p_0^2}{2m} t - \frac{i}{\hbar} (\vec{p} \cdot \hat{p}_0) \frac{p_0 t}{m}} \frac{e^{\frac{i}{\hbar} p_0 r + \frac{i}{\hbar} \hat{p}_0 \cdot \vec{p} r}}{r} f_{\vec{p}_0}(\theta, \varphi) e^{\frac{i}{\hbar} \vec{s} \cdot \vec{p}} \\
&= \frac{e^{\frac{i}{\hbar} p_0 r}}{r} \int \frac{d^3 p}{(2\pi\hbar)^3} \tilde{g}(\vec{p}) e^{-\frac{i}{\hbar} \vec{p} \cdot \vec{b}} e^{-\frac{i}{\hbar} (\vec{p} \cdot \hat{p}_0) \frac{p_0 t}{m}} e^{\frac{i}{\hbar} \hat{p}_0 \cdot \vec{p} r} e^{\frac{i}{\hbar} \vec{s} \cdot \vec{p}} f_{\vec{p}_0}(\theta, \varphi) e^{-\frac{i}{\hbar} \frac{p_0^2}{2m} t} \\
&= \frac{e^{\frac{i}{\hbar} p_0 r}}{r} \int \frac{d^3 p}{(2\pi\hbar)^3} \tilde{g}(\vec{p}) e^{-\frac{i}{\hbar} \vec{p} \cdot \vec{b} - \frac{i}{\hbar} (\vec{p} \cdot \hat{p}_0) \frac{p_0 t}{m} + \frac{i}{\hbar} \hat{p}_0 \cdot \vec{p} r + \frac{i}{\hbar} \vec{s} \cdot \vec{p}} f_{\vec{p}_0}(\theta, \varphi) e^{-\frac{i}{\hbar} \frac{p_0^2}{2m} t} \\
&= \frac{e^{\frac{i}{\hbar} p_0 r}}{r} \int \frac{d^3 p}{(2\pi\hbar)^3} \tilde{g}(\vec{p}) e^{\frac{i}{\hbar} \vec{p} \cdot (\hat{p}_0 (r - \frac{p_0 t}{m}) - \vec{b} + \vec{s})} f_{\vec{p}_0}(\theta, \varphi) e^{-\frac{i}{\hbar} \frac{p_0^2}{2m} t} \\
&= \frac{e^{\frac{i}{\hbar} p_0 r}}{r} g\left(\hat{p}_0 \left(r - \frac{p_0 t}{m}\right) - \vec{b} + \vec{s}\right) f_{\vec{p}_0}(\theta, \varphi) e^{-\frac{i}{\hbar} \frac{p_0^2}{2m} t}
\end{aligned}$$

Pertanto la forma finale del pacchetto d'onda diffuso risulta

$$\psi_d(\vec{x}, t) = \frac{e^{\frac{i}{\hbar} p_0 r}}{r} g\left(\hat{p}_0 \left(r - \frac{p_0 t}{m}\right) - \vec{b} + \vec{s}\right) f_{\vec{p}_0}(\theta, \varphi) e^{-\frac{i}{\hbar} \frac{p_0^2}{2m} t} \quad (4.129)$$

Dalla definizione si ha che il valore del pacchetto è non trascurabile solo se $|\vec{b}| \lesssim l$ e $\left|r - \frac{p_0 t}{m}\right| \lesssim d$.

4.14.4 Approssimazione di Born per la sezione d'urto

A questo punto si calcola il modulo quadro della funzione d'onda (4.129) del pacchetto diffuso:

$$\begin{aligned}
|\psi_d(\vec{x}, t)|^2 &= \frac{|e^{\frac{i}{\hbar} p_0 r}|^2}{r^2} \left| g\left(\hat{p}_0 \left(r - \frac{p_0 t}{m}\right) - \vec{b} + \vec{s}\right) \right|^2 |f_{\vec{p}_0}(\theta, \varphi)|^2 \left| e^{-\frac{i}{\hbar} \frac{p_0^2}{2m} t} \right|^2 \\
&= \frac{1}{r^2} \left| g\left(\hat{p}_0 \left(r - \frac{p_0 t}{m}\right) - \vec{b} + \vec{s}\right) \right|^2 |f_{\vec{p}_0}(\theta, \varphi)|^2
\end{aligned}$$

Si può quindi calcolare la probabilità che la particella venga diffusa nell'unità di angolo solido lungo (θ, φ) come

$$\begin{aligned}\text{Pr}(\theta, \varphi) &= \int d^2b \int_0^{+\infty} dr r^2 |\psi_d(\vec{x}, t)|^2 \\ &= \int d^2b \int_0^{+\infty} dr r^2 |\psi_d(r, \theta, \varphi, t)|^2 \\ &= \int d^2b \int_0^{+\infty} \left| g\left(\hat{p}_0\left(r - \frac{p_0}{m}t\right) - \vec{b} + \vec{s}\right) \right|^2 |f_{\vec{p}_0}(\theta, \varphi)|^2 dr\end{aligned}$$

Tale probabilità ci interessa per il limite di $t \rightarrow +\infty$ e per r tra 0 e $+\infty$. Si esegue dunque un cambio di variabile $u = r - \frac{p_0}{m}t$, con $u \in] -\infty, +\infty[$, ottenendo

$$\begin{aligned}\text{Pr}(\theta, \varphi) &= \int d^2b \int_{-\infty}^{+\infty} du |g(\hat{p}_0 u - \vec{b} + \vec{s})|^2 |f_{\vec{p}_0}(\theta, \varphi)|^2 \\ (\vec{p}_0 \perp \vec{b}, \vec{s}) &= \int d^3\vec{x} |g(\vec{x})|^2 |f_{\vec{p}_0}(\theta, \varphi)|^2 \\ &= \|g\|^2 |f_{\vec{p}_0}(\theta, \varphi)|^2\end{aligned}$$

Il secondo passaggio è dato dal fatto che integrare su $\hat{p}_0 u - \vec{b} + \vec{s}$ con u e \vec{b} che variano sui loro domini equivale ad integrare su \mathbb{R}^3 . Infatti $\vec{b} + \vec{s}$ varia sul piano perpendicolare a \vec{p}_0 (cioè su \mathbb{R}^2) al variare di \vec{b} , mentre $\hat{p}_0 u$ varia sulla retta generata da \hat{p}_0 (cioè su \mathbb{R}) perpendicolare a \vec{b} .

La probabilità che la particella incida risulta invece (attuando lo stesso cambio di variabile, ma con z al posto di r):

$$\begin{aligned}\text{Pr}_{\text{inc}} &= \int d^2b \int_{-\infty}^0 dz |\psi_i(\vec{x}, t)|^2 \\ &= \int d^2b \int_{-\infty}^0 dz |\psi_i(0, 0, z, t)|^2 \\ &= \int d^2b \int_{-\infty}^0 dz \left| g\left(\hat{p}_0\left(z - \frac{p_0}{m}t\right) - \vec{b}\right) \right|^2 \\ \left(u = z - \frac{p_0}{m}t\right) &= \int d^2b \int_{-\infty}^{+\infty} du |g(\hat{p}_0 u - \vec{b})|^2 \\ (\vec{b} \perp \vec{p}_0) &= \int d^3\vec{x} |g(\vec{x})|^2 \\ &= \|g\|^2\end{aligned}$$

Si ottiene finalmente la sezione d'urto differenziale

$$\sigma(\theta, \varphi) = \frac{\text{Pr}(\theta, \varphi)}{\text{Pr}_{\text{inc}}} = \frac{|f_{\vec{p}_0}(\theta, \varphi)|^2 \|g\|^2}{\|g\|^2} = |f_{\vec{p}_0}(\theta, \varphi)|^2 \quad (4.130)$$

che risulta esattamente la formula iniziale.

Ora però, a differenza di prima, sappiamo che forma ha la $f_{\vec{p}}(\theta, \varphi)$ in funzione del potenziale, che si ricava dalla (4.127):

$$f_{\vec{p}}(\theta, \varphi) = -\frac{2m}{\hbar^2} \frac{1}{4\pi} \int d^3y V(\vec{y}) e^{-\frac{p}{\hbar} \hat{r} \cdot \vec{y}} \psi_{\text{in}}(\vec{y})$$

Il problema ora è che la forma analitica di ψ_{in} non è conosciuta, perché l'equazione che abbiamo risolto è implicita per ψ_{in} . Per questo utilizziamo l'*approssimazione di Born*: supponiamo cioè che V sia abbastanza debole in modo tale che in f sia possibile operare la seguente sostituzione:

$$\psi_{\text{in}} = e^{\frac{i}{\hbar}\vec{p}\cdot\vec{x}} + (A_p^{-1}V\psi^{\text{in}})(\vec{x}) \xrightarrow{V \text{ debole}} e^{\frac{i}{\hbar}\vec{p}\cdot\vec{x}}$$

Con l'approssimazione di Born si ritrova dunque

$$\begin{aligned} f_{\vec{p}_0}(\theta, \varphi)_{\text{Born}} &= -\frac{2m}{\hbar^2} \frac{1}{4\pi} \int d^3y V(\vec{y}) e^{i\left(\frac{\vec{p}_0\cdot\vec{y}}{\hbar} - p_0 \frac{\hat{r}\cdot\vec{y}}{\hbar}\right)} \\ &= -\frac{2m}{\hbar^2} \frac{1}{4\pi} \int d^3y V(\vec{y}) e^{i\left(\frac{p_0\hat{z} - p_0\hat{r}}{\hbar}\right)\cdot\vec{y}} \\ &= -\frac{2m}{\hbar^2} \frac{1}{4\pi} \tilde{V}\left(\frac{p_0}{\hbar}(\hat{z} - \hat{r})\right) \end{aligned}$$

da cui poi è facilmente calcolabile la sezione d'urto differenziale (4.130).

4.14.5 Applicazione a potenziali noti

Facciamo un esempio di applicazione. Consideriamo il *potenziale di Debye* della forma

$$V(\vec{x}) = \frac{Ae^{-B|\vec{x}|}}{|\vec{x}|} \quad (4.131)$$

che consiste di una generalizzazione del potenziale coulombiano. Infatti $V(\vec{x})$ diventa il potenziale di Coulomb se $B = 0$. Per $B \simeq m_\pi$ si ottiene invece il *potenziale di Yukawa*. Dalla (4.131) è possibile calcolare la trasformata del potenziale:

$$\tilde{V}(\vec{k}) = \frac{4\pi A}{\vec{k}^2 + B^2}$$

Calcolando il modulo quadro di $\hat{z} - \hat{r}$ si ottiene

$$(\hat{z} - \hat{r})^2 = \hat{z}^2 + \hat{r}^2 - 2\hat{z} \cdot \hat{r} = 2 - 2\hat{z} \cdot \hat{r} = 2(1 - \cos\theta) = \left(2 \sin \frac{\theta}{2}\right)^2$$

da cui si ricava

$$\left(\frac{p_0}{\hbar}(\hat{z} - \hat{r})\right)^2 = \frac{p_0^2}{\hbar^2}(\hat{z} - \hat{r})^2 = \frac{p_0^2}{\hbar^2} \left(2 \sin \frac{\theta}{2}\right)^2$$

Calcoliamo dunque la sezione d'urto:

$$\begin{aligned} f(\theta, \varphi)_{\text{Born}} &= -\frac{2m}{\hbar^2} \frac{1}{4\pi} \tilde{V}\left(\frac{p_0}{\hbar}(\hat{z} - \hat{r})\right) \\ &= -\frac{2m}{\hbar^2} \frac{A}{\left(\frac{p_0}{\hbar}(\hat{z} - \hat{r})\right)^2 + B^2} \\ &= -\frac{2m}{\hbar^2} \frac{A}{\frac{p_0^2}{\hbar^2} \left(2 \sin \frac{\theta}{2}\right)^2 + B^2} \\ &= -\frac{A}{\frac{p_0^2}{2m} \left(2 \sin \frac{\theta}{2}\right)^2 + \frac{\hbar^2}{2m} B^2} \end{aligned}$$

da cui si ottiene

$$\sigma(\theta, \varphi) = |f(\theta, \varphi)_{\text{Born}}|^2 = \frac{A^2}{\left[\frac{p_0^2}{2m} \left(2 \sin \frac{\theta}{2} \right)^2 + \frac{\hbar^2}{2m} B^2 \right]^2}$$

E per $B = 0$ ottengo dunque la *sezione d'urto di Rutherford*:

$$\sigma_{\text{Ruth}}(\theta, \varphi) = \sigma(\theta, \varphi, B = 0) = \frac{A^2}{\left[\frac{p_0^2}{2m} \left(2 \sin \frac{\theta}{2} \right)^2 \right]^2} = \left(\frac{A}{2E_0} \right)^2 \frac{1}{4 \sin^4 \frac{\theta}{2}}$$

Imponendo $B = 0$ in realtà il raggio di interazione diventa infinito e dunque l'approssimazione di Born non varrebbe più. Tuttavia la sezione d'urto di Rutherford è quella giusta pertanto anche con $B = 0$ si vede che l'approssimazione di Born rimane vera. Questo fatto è dovuto a proprietà particolari del potenziale coulombiano, e dunque non vale per potenziali generici.

(Lezione 35 del
12/12/2018)

4.15 Particelle identiche

In MC possiamo sempre **distinguere** due particelle **identiche**, dove per *particelle identiche* intendiamo particelle che hanno le *stesse proprietà indipendenti dallo stato* (stessa massa, carica, etc.).

Denotando con $q_t^{(j)}$ e $p_t^{(j)}$ posizione e momento al tempo t della particella j -esima, consideriamo ad un certo t_0 il punto nello spazio delle fasi che descrive lo *stato puro* del sistema. Se vale $(q_{t_0}^{(1)}, p_{t_0}^{(1)}) \neq (q_{t_0}^{(2)}, p_{t_0}^{(2)})$, ossia se le due particelle sono “distinte” in almeno una tra posizione e momento, per l'unicità della soluzione delle equazioni del moto con condizioni iniziali date, le loro traiettorie nello spazio delle fasi non si intersecheranno mai nella loro evoluzione temporale.

*In pratica, due particelle che sono identiche, es. due elettroni, possono essere “contrassegnate” dalla loro “storia”, ossia dal percorso che hanno compiuto nello spazio delle fasi. Ciò le rende, almeno in teoria, **distinguibili**, cioè in principio è possibile riferirsi nello specifico ad una di esse, individuandola senza ambiguità. Notiamo però che qui stiamo usando un'ipotesi forte: che la “storia” di ogni particella sia univocamente determinata.*

Distinguibilità in
MC

In MQ, tuttavia, anche supposto che all'istante iniziale le due particelle identiche abbiano due funzioni d'onda $\psi^{(1)}, \psi^{(2)}$ con supporti spaziali disgiunti, che descrivono gli stati delle particelle (1) e (2), la loro evoluzione tramite Schrödinger in generale può portare ad una *sovrapposizione* dei loro supporti, dopo la quale non c'è più modo di sapere, anche ammesso che i supporti ritornino disgiunti, a quale di essi corrispondano le particelle (1) o (2).

Indistinguibilità
in MQ

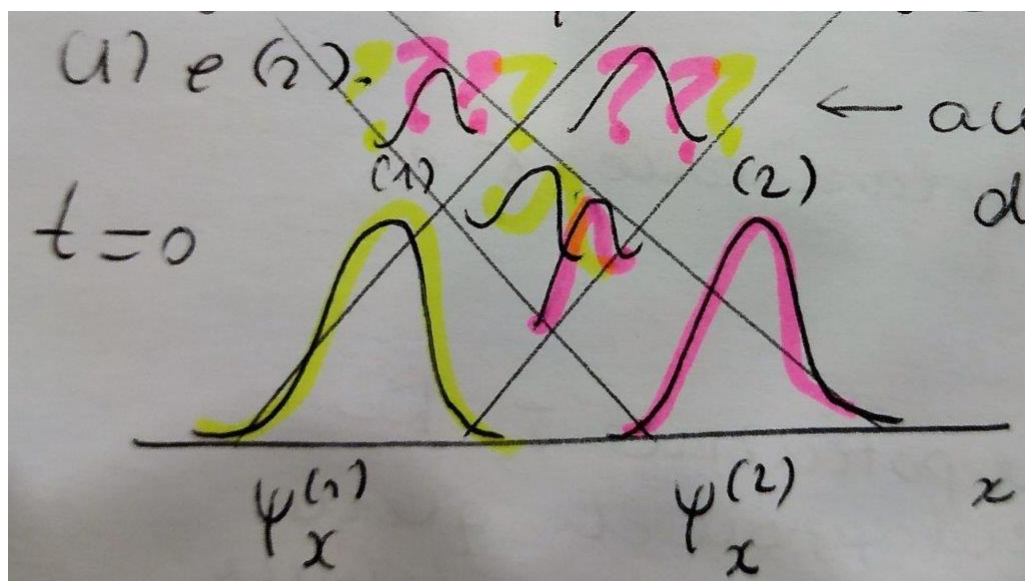


Figura 4.12: Dopo la sovrapposizione delle due funzioni d'onda (qui rappresentate come gaussiane) non è più possibile identificare esattamente a quale delle due particelle iniziali si riferiscano.

Graficamente, consideriamo due funzioni d'onda con supporto separato, che visualizziamo come due gaussiane che *non si sovrappongono*. Vogliamo, ad ogni istante, dare un *nome* alle due funzioni: se siamo in grado di farlo potremo considerare le due particelle *distinguibili*. Ma se nella loro evoluzione le due funzioni si *sovrappongono* e si separano di nuovo, non possiamo essere sicuri che la prima funzione abbia “rimbalzato” sulla seconda, oppure “le sia passata attraverso”. Abbiamo quindi una ambiguità sulla *label* da assegnare alle funzioni d'onda.

Sperimentalmente, in una camera a nebbia, date le “traiettorie” di due particelle identiche che si intersecano (che sono in realtà costituite da un insieme finito di goccioline), non possiamo individuare *quale “traiettoria” sia quella effettiva* dopo l'interazione.

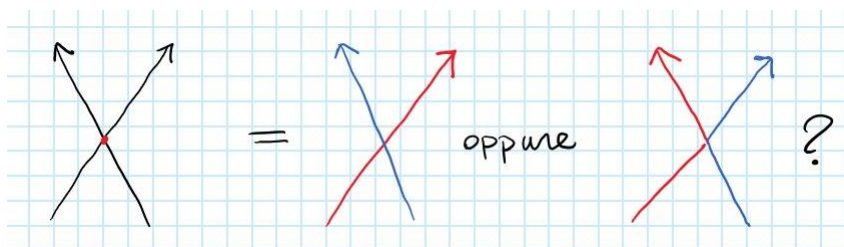


Figura 4.13: Ambiguità nel discernere la “traiettoria effettiva” partendo dai dati di “traiettoria” ottenuti da una camera a nebbia.

Perciò non esiste in linea di principio la possibilità di determinare una *distinzione* tra particelle identiche a tutti gli istanti, ovvero non è possibile ancorare la loro individualità a caratteristiche estrinseche, dipendenti dallo stato (come la loro

“storia” in senso classico).

Ciò ha una grossa conseguenza a livello fisico: ha infatti senso porsi **solo** domande che non dipendono dall’individualità delle particelle identiche, ad esempio la probabilità di trovare **un** elettrone in un fissato volume, ma non **un fissato** elettrone in quel volume.

Formalmente ne segue che i **valori medi** di osservabili (sperimentalmente misurabili) in stati che differiscono per **scambio di particelle identiche** devono avere lo stesso valore.

Conseguenza dell’indistinguibilità

Sappiamo che in MQ lo spazio di Hilbert di n particelle *distinte* è il prodotto tensore degli spazi di ogni particella \mathcal{H} . Quindi, se le n particelle sono identiche, gli stati dovranno trovarsi tra i raggi vettori di $\mathcal{H}^{\otimes N} \equiv \mathcal{H} \otimes \mathcal{H} \otimes \cdots \otimes \mathcal{H}$.

Scelta una base $\{e_j\}_{j \in J}$ per \mathcal{H} , sappiamo che $\{e_{j_1} \otimes e_{j_2} \otimes \cdots \otimes e_{j_N}, j_i \in J\}$ è una base in $\mathcal{H}^{\otimes N}$.

Consideriamo le permutazioni di N particelle identiche. Queste permutazioni costituiscono un gruppo⁴², denotato con S_N . Ogni elemento di S_N rappresenta perciò un possibile “riordinamento” delle N particelle. Intuitivamente, possiamo pensare ad ogni possibile permutazione come una serie di *scambi* effettuati uno dopo l’altro. In effetti, si può dimostrare che semplicemente scambiando solo “elementi con i loro vicini” si può *costruire* una permutazione generica.

Il gruppo S_N delle permutazioni

Andando nello specifico, denotiamo con $1, 2, \dots, N$ gli elementi su cui effettuiamo gli scambi e con σ_i (che non ha nulla a che fare con le matrici di Pauli) l’operazione di scambio dell’elemento i -esimo con l’elemento $(i+1)$ -esimo, dove $i = 0, \dots, N-1$. Si ha allora che una generica permutazione (elemento del gruppo S_N) si ottiene *moltiplicando* delle σ_i , con i seguenti vincoli:

1. $\sigma_i^2 = 1$. In altre parole, scambiare due volte di fila le stesse particelle (i -esima e $(i+1)$ -esima) non cambia nulla.

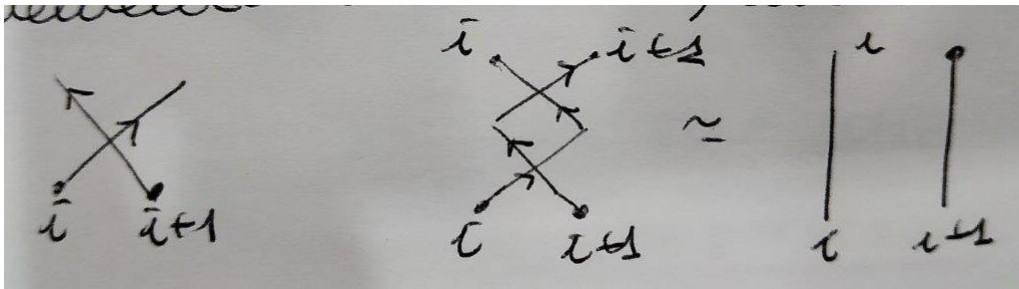


Figura 4.14: Illustrazione dell’effetto della permutazione $\sigma_i^2 = \text{id}$. L’idea del grafico sta nell’applicare le permutazioni *una alla volta*, collegando ad ogni passo i punti *di arrivo* di ciascuna particella con quelli *di partenza*. Per esempio, dopo l’applicazione di una σ_i , la particella i finisce nel posto $i+1$, e quella $i+1$ nel posto i . Una volta applicata la serie di permutazioni, si collega il punto di partenza di ogni particella con quello di arrivo finale, visualizzando così la *permutazione risultante*.

⁴² Dato che la composizione di permutazioni è ancora una permutazione, per ogni permutazione esiste la *permutazione inversa* che ne annulla l’azione, esiste la “permutazione nulla” che non modifica nessun ordine e vale la proprietà associativa per le permutazioni.

2. $\sigma_i \sigma_j = \sigma_j \sigma_i$ se $|i - j| \geq 2$: due permutazioni di particelle “non contigue” sono indipendenti l’una dall’altra, e perciò non conta l’ordine in cui le si effettua. L’idea intuitiva è che ciascuna permutazione agisce *localmente*, solo sull’ i -esimo e $(i + 1)$ -esimo elemento, e perciò applicarle ad elementi che distano “più di un posto” uno dall’altro fa sì che non vi sia interazione tra le due.

3. $\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}$

Detto A l’elemento di posto i -esimo, e B e C quelli di posti rispettivamente $i + 1$ e $i + 2$:

$$\begin{aligned} \sigma_i \sigma_{i+1} \sigma_i : A B C &\rightarrow B A C \rightarrow B C A \rightarrow C B A \\ \sigma_{i+1} \sigma_i \sigma_{i+1} : A B C &\rightarrow A C B \rightarrow C A B \rightarrow C B A \end{aligned}$$

Cioè scambiare prime due, poi seconde due e poi di nuovo prime due, oppure scambiare seconde due, poi prime due e poi di nuovo seconde due, produce la stessa cosa (ossia scambiare la prima con la terza).

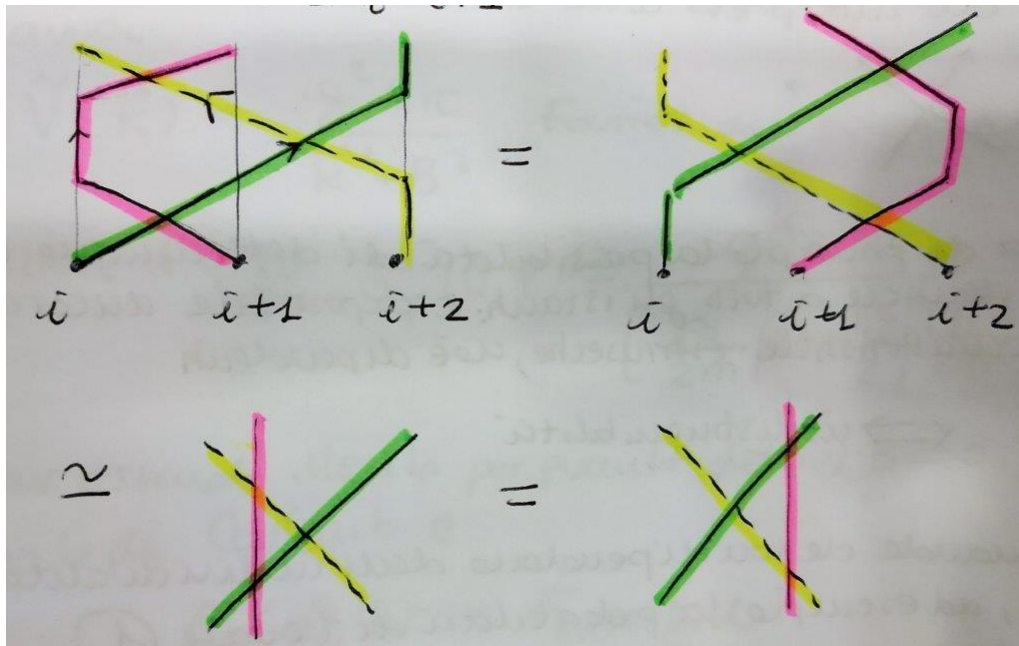


Figura 4.15: Graficamente, applichiamo in successione σ_i , σ_{i+1} e σ_i (figura a sinistra), oppure σ_{i+1} , σ_i e σ_{i+1} (figura a destra). Ad ogni passo colleghiamo con una linea (del colore della rispettiva particella) il *punto di arrivo* con quello di *partenza*: per esempio, nella figura in alto a destra il primo passo è σ_{i+1} , per cui la particella $i + 1$ finisce al posto $i + 2$, e quella $i + 2$ al posto $i + 1$, mentre quella i rimane in i . Una volta eseguiti tutti i passi, colleghiamo punti di arrivo all’ultimo passo con quelli di partenza delle rispettive particelle al primo: notiamo che l’ordine (dei colori) è lo stesso sia nella figura a sinistra che in quella a destra - e ciò dimostra l’equivalenza che cercavamo.

Le σ_i così viste sono tuttavia “operazioni astratte”, e se vogliamo applicarle ai problemi che ci interessano dovremo sceglierne una *rappresentazione* (analogamente

a quanto già fatto per traslazioni e rotazioni).

Per ogni $\sigma \in S_N$ definiamo allora l'operatore unitario:

$$U(\sigma) : e_{j_1} \otimes e_{j_2} \otimes \cdots \otimes e_{j_N} \mapsto e_{\sigma(j_1)} \otimes e_{\sigma(j_2)} \otimes \cdots \otimes e_{\sigma(j_N)}$$

esteso per linearità e continuità a $\mathcal{H}^{\otimes N}$.

Nella notazione $\sigma(x)$ rappresenta il vettore che prende il posto di x *dopo* aver effettuato la permutazione. In altre parole $U(\sigma)$ *scambia* tra loro le basi dei singoli \mathcal{H} che costituiscono $\mathcal{H}^{\otimes N}$, e costituisce quindi una rappresentazione unitaria di S_N in $\mathcal{H}^{\otimes N}$ (ossia una “concretizzazione” della nozione astratta di permutazione).

Poiché vogliamo che i valori medi *non* dipendano dalla permutazione, deve valere, $\forall |\psi\rangle \in \mathcal{H}^{\otimes N}$ e \forall osservabile A in $\mathcal{H}^{\otimes N}$ la condizione di **invarianza dei valori medi**:

$$\langle \psi | U^\dagger(\sigma) A U(\sigma) | \psi \rangle = \langle \psi | A | \psi \rangle \quad \forall \sigma \in S_N \Rightarrow U^\dagger(\sigma) A U(\sigma) = A \quad (4.132)$$

Perciò non tutti gli operatori autoaggiunti in $\mathcal{H}^{\otimes N}$ sono osservabili, ma solo quelli per cui vale (4.132). Poiché ogni operatore in $\mathcal{H}^{\otimes N}$ può scriversi come combinazione lineare (eventualmente infinita) di operatori $A_1 \otimes A_2 \otimes \cdots \otimes A_N$, ove gli A_i , detti operatori di particella singola, agiscono sul termine i -esimo \mathcal{H} in $\mathcal{H}^{\otimes N}$, la (4.132) ci dice che **sono osservabili solo operatori autoaggiunti in $\mathcal{H}^{\otimes N}$ simmetrici** per *ogni* scambio degli operatori di particella singola. In altre parole, per un sistema di N particelle identiche, calcolare il valore di A per la particella i -esima deve dare lo *stesso* risultato per ogni i , dato che se le particelle sono indistinguibili non vi è un modo univoco di identificare la particella i in questione.

Analogamente avremo una condizione di *invarianza degli stati*, per cui “scambiare gli stati di due particelle identiche non cambia lo stato del sistema”, ossia $U(\sigma) |\psi\rangle$, $\forall \sigma \in S_N$, $|\psi\rangle \in \mathcal{H}^{\otimes N}$ deve descrivere lo stesso stato di $|\psi\rangle$. Poiché le σ si possono scrivere come prodotti di σ_i , ci basta esaminare quest'ultimo caso.

Matematicamente, tale equivalenza si traduce nel fatto che $U(\sigma_i) |\psi\rangle$ è pari a $|\psi\rangle$ a meno di un fattore moltiplicativo $c(\sigma_i)$:

$$U(\sigma_i) |\psi\rangle = c(\sigma_i) |\psi\rangle$$

Per trovare $c(\sigma_i)$ ricordiamo la proprietà di “compatibilità per l'operazione di gruppo” per le rappresentazioni, ossia:

$$\sigma = \sigma_i \sigma_j \Rightarrow U(\sigma) = U(\sigma_i) U(\sigma_j)$$

Applicando allora due volte di fila la stessa permutazione σ_i^2 , sappiamo che il sistema rimane invariato ($\sigma_i^2 = e$, dove e è l'elemento neutro di S_N). Perciò, usando la rappresentazione unitaria di S_N :

$$U(\sigma_i^2) = U(\sigma_i) U(\sigma_i) = U(\sigma_i)^2$$

Inoltre, sempre dalla definizione di rappresentazione, U manda l'elemento neutro e nell'elemento neutro delle matrici, ossia \mathbb{I} :

$$U(\sigma_i^2) = U(e) = \mathbb{I}$$

Applicando ad una generica $|\psi\rangle$ (normalizzata) otteniamo:

$$U(\sigma_i)^2 |\psi\rangle = c(\sigma_i)^2 |\psi\rangle = U(e) |\psi\rangle = |\psi\rangle \Rightarrow c(\sigma_i)^2 = 1$$

Prendendo la radice quadrata troviamo due possibilità: $c(\sigma_i) = \pm 1$.

Perciò, $\forall \sigma_i$ deve succedere:

- $c(\sigma_i) = 1$, ossia la funzione d'onda/il vettore di stato sono *simmetrici per scambi* (non cambiano di segno dopo una permutazione). Le particelle che soddisfano questa condizione sono detti **bosoni**.
- $c(\sigma_i) = -1$, dove invece funzione d'onda/vettore di stato sono *antisimmetrici per scambi*. Le particelle che soddisfano questa proprietà sono i **fermioni**.

Tutto ciò porta a un fenomeno nuovo: i **settori di superselezione**.

Settori di
superselezione

Poiché le osservabili sono *simmetriche per scambi*, i sottospazi di $\mathcal{H}^{\otimes N}$ sede delle rappresentazioni simmetriche (bosoni) o antisimmetriche (fermioni) sono lasciati invariati dalle operazioni di evoluzione temporale e di misura. Infatti, il primo caso si realizza applicando l'esponenziale di un'osservabile – H – che è simmetrica, mentre il secondo è dato dall'uso di proiettori, che essendo osservabili sono simmetrici.

Quindi gli stati simmetrici e antisimmetrici non si mescolano mai, qualsiasi operazione si faccia. Descrivono quindi “mondi distinti”, che tecnicamente si chiamano **settori di superselezione**.

Un altro esempio, meno ovvio, di settore di superselezione è dato dalla carica elettrica. Non è possibile sovrapporre due stati di carica diversa, e qualsiasi operazione si faccia non è possibile cambiarla. Operatori che non rispettano tali condizioni non costituiscono osservabili.

Denotati con $|\psi_i\rangle$, $i = 1, \dots, N$ i ket delle singole particelle identiche che compongono il sistema, sappiamo perciò che lo stato “totale” $|\Psi\rangle$ del sistema deve essere simmetrico o antisimmetrico per scambi di particelle, a seconda che esse siano rispettivamente bosoni o fermioni. In generale, il prodotto tensore delle $|\psi_i\rangle$ non ha questa proprietà, e perciò dovremo considerare un'operazione più specifica: il prodotto tensore (anti)simmetrico.

L'idea è di combinare le $|\psi_i\rangle$ in modo che il risultante tensore sia (anti)simmetrico⁴³. Per farlo basta sommare su *tutte le permutazioni possibili* dei prodotti tensori delle $|\psi_i\rangle$ (eventualmente introducendo un fattore alternante): in questo modo l'effetto di una permutazione è quello di *scambiare* due termini di una somma, lasciando invariato il risultato (a meno del segno, nel caso sia presente il fattore alternante). Esplicitamente, avremo le seguenti espressioni per $|\Psi\rangle$:

$$|\Psi\rangle = \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} \prod_i |\psi_{\sigma(i)}\rangle \quad \text{ket per } N \text{ bosoni}$$

$$|\Psi\rangle = \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} (-1)^{|\sigma|} \prod_i |\psi_{\sigma(i)}\rangle \quad \text{ket per } N \text{ fermioni}$$

⁴³Un'operazione analoga è definita formalmente nella lezione 10_1 nelle dispense del corso di “Geometria Differenziale” tenuto dal professor Bottacin.

Dove con $|\sigma|$ indichiamo il numero di scambi effettuati per formare la permutazione σ (che a priori non è ben definito, ma è ben definita la sua *parità*, che è l'unica cosa che ci interessa nella formula di sopra).

Notiamo che tali stati per il sistema composto **non** sono il semplice prodotto tensore di ket di particelle singole $|\psi_i\rangle$, ma combinazioni lineari di tali prodotti, cioè sono “entangled”. Questa è la conseguenza *fisica* della *perdita dell'individualità*: il fatto che non si possa discutere di particelle singole (poiché non è possibile identificarle), fa sì che l'unico modo di descrivere il sistema sia in una “visione globale” che non può essere spezzata nelle singole parti.

4.15.1 Il principio di esclusione di Pauli

Una conseguenza immediata dell'antisimmetria per i fermioni è il seguente principio:

Teorema 4.15.1. *Non è possibile per un sistema di N fermioni identici avere due particelle con lo stesso insieme di autovalori di un ICOC.*

Principio di esclusione di Pauli

Ad esempio, consideriamo come \mathcal{H} lo spazio di Hilbert di particella singola per un atomo di idrogeno. Un ICOC è $\{H, L^2, L_z, S_z\}$. Indichiamo i rispettivi autovalori di tali osservabili con $\{n, l, m, s_z\}$.

Per il principio di esclusione si ha che è impossibile mettere due elettroni con gli stessi n, l, m, s_z nell'atomo di idrogeno - cioè al massimo vi sono due elettroni per *orbitale* (dove un orbitale è definito da i numeri quantici n, l, m).

Lo stesso ICOC vale anche per un qualsiasi sistema coulombiano - e in particolare per atomi più complessi. In particolare, in un atomo con molti elettroni siamo costretti a mettere elettroni in orbitali a energia crescente per rispettare il principio di Pauli, e da ciò deriva tutta la *chimica*.

In modo analogo, applicando il principio di esclusione di Pauli all'ICOC $\{\vec{X}, S_z\}$, risulta che (in modo “immaginifico”) due elettroni non possono avere neppure “potenzialmente” stessa posizione e spin.

Verifichiamolo. Ipotizziamo per assurdo che $\vec{x}_i, s_{z_i} = \vec{x}_j, s_{z_j}$, ossia che gli elettroni i -esimo e j -esimo abbiano stessa posizione e spin. Se partiamo da una funzione d'onda: $\psi(\vec{x}_1, s_{z_1}, \dots, \vec{x}_N, s_{z_N})$ e scambiamo i due elettroni i e j , sappiamo che il segno deve cambiare per l'antisimmetria (essendo gli elettroni fermioni):

$$\begin{aligned} \psi(\vec{x}_1, s_{z_1}, \dots, \vec{x}_i, s_{z_i}, \dots, \vec{x}_j, s_{z_j}, \dots, \vec{x}_N, s_{z_N}) = \\ = -\psi(\vec{x}_1, s_{z_1}, \dots, \vec{x}_j, s_{z_j}, \dots, \vec{x}_i, s_{z_i}, \dots, \vec{x}_N, s_{z_N}) \end{aligned}$$

Ma se i due termini scambiati sono uguali (per ipotesi), allora le due ψ devono essere uguali (scambiare gli stessi termini *non* modifica la ψ). Ma allora l'unica possibilità per cui contemporaneamente $\psi = -\psi$ e $\psi = \psi$ è che sia $\psi = 0$.

Poiché la funzione d'onda descrive le probabilità di misura, anche prima che essa venga effettuata, si ha che *anche prima di una misura* gli elettroni non possono occupare gli stessi stati di posizione-spin.

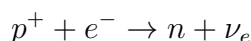
Qualitativamente tale fenomeno spiega l'*incompenetrabilità* della materia.

*L'incompenetrabilità
della materia*

Esaminando la struttura atomica, infatti, non risulta per nulla evidente ad un primo sguardo perché due atomi non possano “attraversarsi”. Il nucleo è infatti incredibilmente piccolo rispetto alle distanze a cui “orbitano” gli elettroni – per cui l’atomo è considerabile come quasi interamente “vuoto”, e anche la probabilità di un’interazione elettromagnetica significativa tra due e^- è relativamente bassa. Tuttavia, si ha che la sola *potenzialità* di occupazione degli stessi stati è impedita dal principio di esclusione – e perciò due atomi sono *costretti* a rimanere ben separati, dato che la probabilità di “attraversarsi” è **nulla**.

Il principio di esclusione è utile per comprendere anche la struttura delle **stelle di neutroni**. Se la forza di gravitazionale è sufficientemente intensa, gli orbitali sono “schiacciati” uno contro l’altro, ma non possono compenetrarsi per principio di esclusione. Tali condizioni favoriscono allora un processo che altrimenti è molto raro: un protone e un elettrone si *fondono* a formare un neutrone (liberando un neutrino elettronico).

*Le stelle di
neutroni*



I neutroni così prodotti possono ora “compattarsi” molto di più, dato che la distanza tra ciascuno di essi è dell’ordine di $10^{-15}m$. Ciò permetterebbe all’intera massa del Sole, che normalmente ha un raggio $\sim 700 \cdot 10^3 km$, di essere concentrata in un raggio di 10km. La densità della materia *degenere* così prodotta (detta “neutronio”) è tale che un solo cm^3 abbia una massa dell’ordine $\sim 2 \cdot 10^6 kg$.

4.15.2 Effetti in dimensioni inferiori

C’è un modo lievemente diverso di vedere la statistica (comportamento per scambi) che è rilevante in $d < 3$ (trascuriamo, per ora, lo spin).

Consideriamo lo scambio di posizione di due particelle x_i e x_j . Fisicamente possiamo pensarlo come un *cammino continuo* che porta $x_i \rightarrow x_j$ e $x_j \rightarrow x_i$. Ci chiediamo: conta l’ordine dello scambio? Se permutiamo x_i e x_j “in senso orario” otteniamo lo stesso risultato che nel caso “antiorario”?

In $d = 3$, i due *modi* di effettuare lo scambio sono equivalenti, dato che i due spostamenti possono avvenire contemporaneamente lungo curve che non si intersecano, come visibile in figura 4.16.

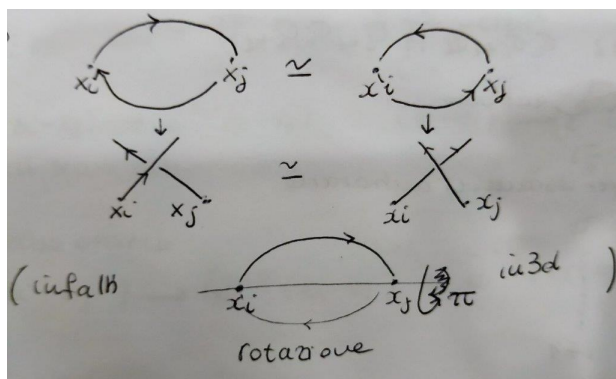


Figura 4.16: Scambio di x_i e x_j in $d = 3$: possiamo immaginare di far passare x_i sopra al piano e x_j sotto, e in tal modo i due percorsi di scambio non si intersecano, e perciò l'ordine di scambio non conta: possiamo vedere una delle due possibilità (es. x_i che passa sopra e x_j sotto) come la versione *ruotata* di π dell'altra possibilità (x_i sotto e x_j sopra).

Per far ciò, tuttavia, abbiamo dovuto usare *la terza dimensione spaziale*. Lo stesso trucco non funziona più in $d = 2$, dove i cammini devono per forza intersecarsi, come visibile in figura 4.17.

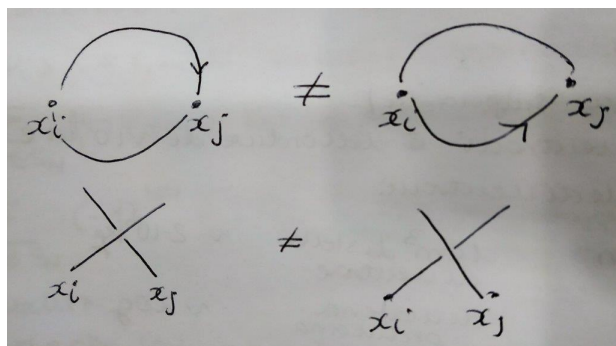


Figura 4.17: In $d = 2$ i due percorsi si intersecano, e non c'è modo di convertire una scelta di ordine nell'altra.

Infatti, poiché stiamo considerando fermioni, le due particelle non possono occupare la stessa posizione allo stesso tempo. Avremo allora due possibilità *non equivalenti*: uno scambio *orario* e uno *antiorario*. Non c'è modo di mappare uno nell'altro senza intersecare i cammini – cosa che è proibita dalla natura fermionica delle particelle.

Una conseguenza di ciò è che avremo due tipi di *permutazione* che scambiano gli elementi i -esimo e $i + 1$ -esimo: quella oraria σ_i e quella antioraria σ_i^{-1} , con:

$$\sigma_i \neq \sigma_i^{-1} \quad d = 2$$

Una conseguenza immediata è che $\sigma_i^2 \neq 1$: due permutazioni orarie consecutive *non* riportano allo stato iniziale.

Possiamo visualizzare ciò in modo pittoresco, schematizzando il cammino di ogni

particella durante la permutazione come un *filo colorato*, come rappresentato in figura 4.18.

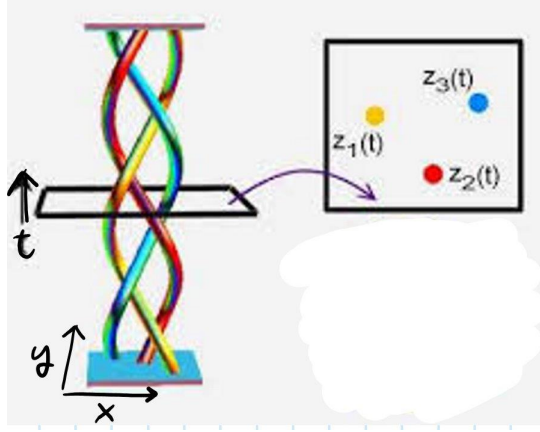


Figura 4.18: Schema dei cammini di 3 particelle durante una permutazione

In tale rappresentazione, scambiare due volte *nello stesso senso* (es. orario) due particelle, fa sì che i loro cammini *si intreccino*. Poiché non si può far passare un filo attraverso un altro (essendo le particelle fermioni), non è possibile *separare* i due cammini dopo i due scambi orari, che quindi sono *topologicamente* distinti dai due cammini “paralleli” (e non intrecciati) che avremmo avuto nel caso senza alcuna permutazione.

D’altro canto, applicare prima uno scambio orario e poi uno antiorario *riporta* effettivamente alla situazione iniziale, dato che l’intreccio dato dalla prima operazione è “sciolto” dalla seconda.

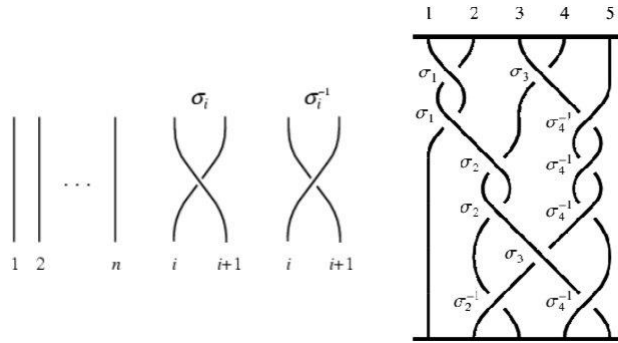


Figura 4.19: Esempio di intrecci generati dall’applicazione consecutiva di più permutazioni σ_j

Tutto ciò fa sì che non si possa più usare il gruppo S_N per descrivere l’indistinguibilità di particelle in $d = 2$. Al suo posto introduciamo il gruppo generato da prodotti di scambi *con orientazione*, detto **gruppo delle trecce** B_N per N oggetti⁴⁴, i cui elementi godono delle seguenti proprietà:

⁴⁴La B sta per *braid*

- $\sigma_i \sigma_j = \sigma_j \sigma_i$, $|i - j| > 2$
- $\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}$
- $\sigma_i^2 \neq 1$, $\sigma_i \sigma_i^{-1} = 1$, dove con σ_i si denota la permutazione *oraria*, e con σ_i^{-1} quella *antioraria*.

In particolare, per la rappresentazione unitaria di B_N valgono:

$$\begin{aligned} U(\sigma_i) |\psi\rangle &= e^{i\theta} |\psi\rangle \\ U(\sigma_i^{-1}) |\psi\rangle &= e^{-i\theta} |\psi\rangle \end{aligned} \quad (4.133)$$

per un certo $\theta \in [0, 2\pi)$ che dipende dalla natura della particella esaminata.

4.15.3 L'effetto di Hall quantistico

Quanto visto sembra una pura peculiarità matematica – eppure, come già visto per lo spin, la natura sembra sfruttare tutte le libertà possibili, e le differenze tra S_N e B_N si traducono in effetti sperimentali.

Storicamente, le basi teoriche appena discusse furono concepite negli anni '70 da Leinaas-Myrheim, e solo nel 1986 fu osservata la loro conseguenza *fisica* nell'**effetto Hall quantistico frazionario**. Si trova infatti che, per certe particelle, il θ delle relazioni (4.133) non è né 0 (come per i bosoni) né π (come per i fermioni), ma assume un valore *generico* – da cui il nome di **anyon**.

Richiami dell'effetto di Hall classico.

Consideriamo una lamina conduttrice di dimensioni L , w , t , rappresentata in 4.20. Un potenziale V_x (a cui corrisponde il campo elettrico ϵ_x) genera su di essa una corrente I_x , che scorre in direzione \hat{x} (con gli elettroni che procedono nel verso $-\hat{x}$ opposto). Si pone un campo magnetico B_z lungo \hat{z} , per cui gli elettroni subiscono l'effetto della forza di Lorentz (in sistema cgs con unità Gaussiane):

$$\vec{F}_l = q \frac{\vec{v}}{c} \times \vec{B} = -e \left(-\frac{v}{c} \hat{x} \right) \times B_z \hat{z} = \frac{evB_z}{c} (-\hat{y})$$

dove \vec{v} è la velocità di drift degli elettroni, nel verso contrario a quello (convenzionale) della corrente.

Gli elettroni tendono allora ad accumularsi lungo una delle facce laterali del conduttore, generando un campo elettrico di modulo ϵ_y diretto come $-\hat{y}$. Tale campo produce una forza elettrostatica $\vec{F}_e = (-e)\vec{\epsilon}_y$ sui singoli elettroni, che si oppone alla forza di Lorentz \vec{F}_l . All'equilibrio le due forze si compensano perfettamente, e si ha (ponendo $c = 1$):

$$\epsilon_y = vB_z \Rightarrow V_H = vB_z w \quad (4.134)$$

dove V_H è il potenziale di Hall, che (almeno classicamente) varia in modo continuo al variare di B_z .

Consideriamo ora l'equazione per I_x :

$$I_x = nAvq$$

dove A è la sezione trasversa del conduttore (attraversata dalla corrente), n è il numero di cariche elettriche disponibili alla conduzione per unità di volume, v è la loro velocità di drift (con il verso convenzionale) e q la loro carica. Sostituendo i valori specifici giungiamo a:

$$I_x = n(tw)(-v)(-e) \Rightarrow w = \frac{I_x}{nte}$$

Sostituendo in (4.134) giungiamo a:

$$V_H = \frac{I_x B_z}{nte}$$

Definiamo ora la **resistenza di Hall** R_H come il rapporto (in analogia alla legge di Ohm):

$$R_H = \frac{V_H}{I_x} = \frac{B_z}{nte}$$

Perciò, almeno *classicamente*, anche R_H dipende in modo continuo da B_z . Analogamente si definisce la relativa resistività, spesso indicata con ρ_{xy} , utilizzando nella formula di sopra il modulo del campo elettrico ϵ_y al posto del potenziale V_H . Definiamo, infine la *resistenza longitudinale* R dalla legge di Ohm:

$$R = \frac{V_x}{I_x}$$

che, nel caso classico, possiamo immaginare *costante e molto bassa* (nel caso di una lamina conduttrice). La relativa resistività è data da $\rho_{xx} = \epsilon_x/I_x$.

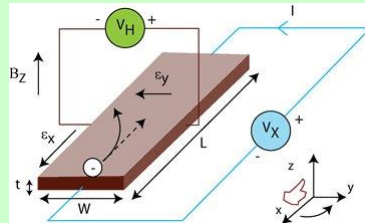


Figura 4.20: Schema dell'effetto di Hall classico

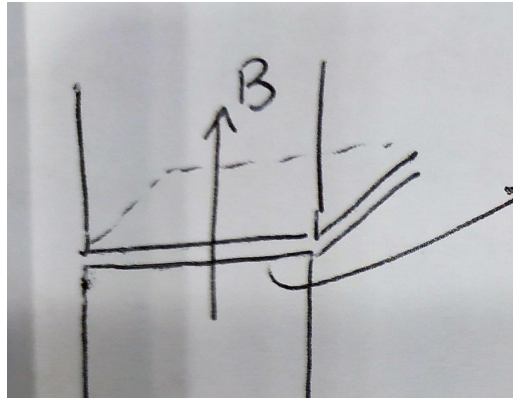


Figura 4.21: Schema dell'apparato sperimentale per l'effetto di Hall quantistico frazionario

Sperimentalmente, consideriamo un interfaccia tra semiconduttori – per esempio quella tra silicio e strato di ossido in un MOSFET. Introduciamo un campo magnetico \vec{B} perpendicolare all'interfaccia. Gli elettroni sulla superficie del semiconduttore seguono le orbite di ciclotrone, per cui abbiamo i livelli di Landau per le energie, dati da:

$$\mathcal{E}_n = \hbar\omega \left(n + \frac{1}{2} \right) \quad \omega = \frac{eB}{m_e} \quad (4.135)$$

dove ω è la pulsazione di ciclotrone.

Come osservato in un precedente esercizio i livelli di Landau sono fortemente degeneri, ossia è possibile avere un gran numero di elettroni alla stessa energia, che si dimostra essere proporzionale all'intensità di B . Perciò, all'aumentare del campo magnetico, aumenta sia la *distanza* tra un livello e il successivo (come visibile dalla formula (4.135)) che il numero di elettroni che possono occupare lo stesso livello. Poiché ad ogni \mathcal{E}_n è associato un gran numero di possibili stati, un grafico della *densità* degli stati in funzione dell'energia è dato da una successione di δ di Dirac equispaziate, come rappresentato in figura 4.22.

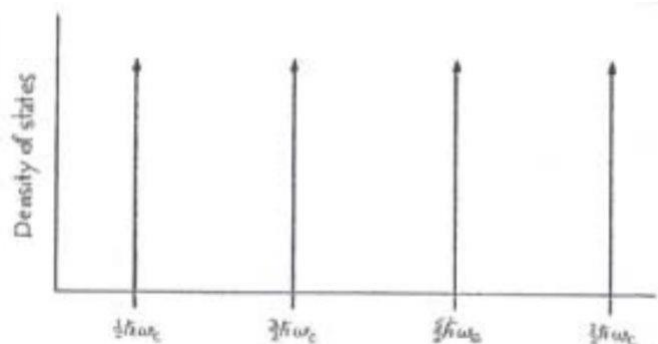


Figura 4.22: Densità degli stati in funzione dell'energia in un sistema ideale in $d = 2$

Come diretta conseguenza del principio di esclusione di Pauli, N elettroni si disporranno a riempire progressivamente i livelli di Landau partendo da quelli con energia più bassa – e per temperature T molto basse e B molto alti possiamo considerare trascurabili le eccitazioni termiche tra un livello e un altro, dato che le particelle non hanno abbastanza energia per superare gap sempre più grandi. Risulta perciò ben definita una *energia massima*, detta **energia di Fermi**, che corrisponde a quella dell'ultimo livello occupato.

In realtà, tuttavia, i livelli di Landau non sono ben definiti a causa delle *impurezze* dell'interfaccia. Tali impurezze causano delle “buche di potenziale localizzate”, dove gli elettroni possono rimanere intrappolati, come rappresentato in figura 4.23.

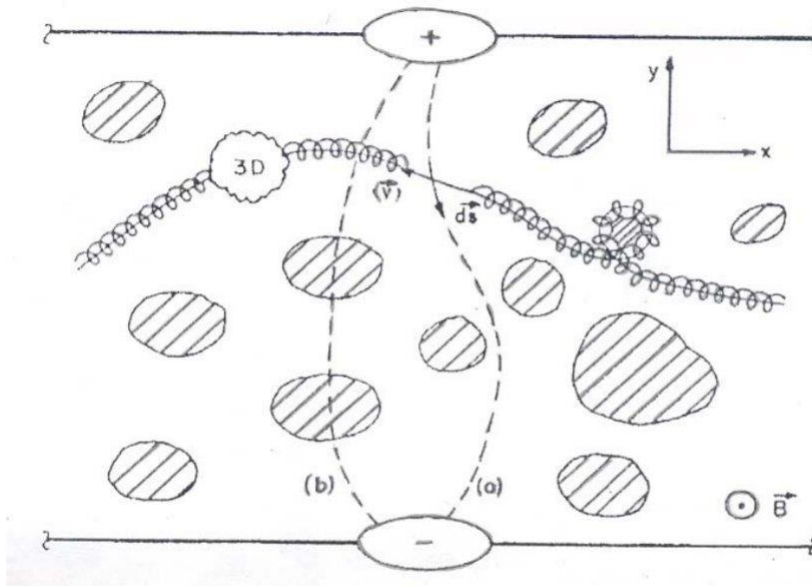


Figura 4.23: Le impurezze nel cristallo sono rappresentate come regioni tratteggiate. Gli elettroni, nel loro moto sulla superficie, possono ritrovarsi in *stati localizzati* attorno alle impurezze, con un'energia intermedia tra due livelli di Landau

Avremo perciò una densità di stati non trascurabile anche ad energie tra \mathcal{E}_n consecutivi. Un grafico più realistico è riportato in figura 4.24.

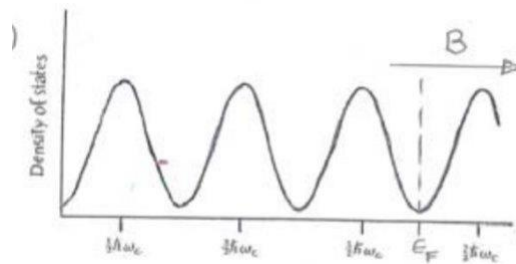


Figura 4.24: Densità degli stati in funzione dell'energia in un cristallo con impurezze

Possiamo ora spiegare l'effetto di Hall intero, scoperto nel 1983. Per $T \sim 10$ K e $B \sim 10$ T, si realizza un sistema in cui solo *pochi* livelli di Landau sono riempiti dagli elettroni (ossia si trovano al di sotto dell'energia di Fermi).

Detta n la densità di elettroni e d la degenerazione di ogni livello, il numero ν di livelli riempiti è dato da:

$$n = \nu d$$

Per d usiamo la stima già trovata in (4.113), per cui arriviamo a:

$$\frac{1}{\nu} = \frac{nh}{eB}$$

Si scopre sperimentalmente che ν assume *valori esattamente interi* non solo per i B_ν che verificano tale condizione, ma anche per B “leggermente diversi”. Questo è l'effetto delle impurità, per cui alcuni elettroni *rimangono intrappolati* in *stati localizzati* ad energie intermedie tra i singoli livelli, e in tal modo “stabilizzano” il sistema, facendo sì che vi siano sempre un *numero intero* di livelli riempiti.

La conseguenza di ciò è che la *resistività di Hall* ρ_{xy} è quantizzata:

$$\rho_{xy} = \frac{B}{ne} = \nu \frac{h}{e^2}$$

In figura 4.25 è possibile notare come, aumentando progressivamente il campo magnetico, la resistività ρ_{xy} , espressa in multipli di h/e^2 “salga a salti”, rimanendo costante per un certo range di valori di B .

Allo stesso tempo, la resistività longitudinale ρ_{xx} tende a 0 durante i singoli *plateau* di ρ_{xy} . Ciò è conseguenza del fatto che ρ_{xx} misura la perdita di energia per scattering inelastico degli elettroni, ossia quanta energia perdono “per attrito” attraversando la lamina. Ma, se esattamente ν livelli sono pieni, gli elettroni non hanno posto dove andare, e perciò non perdono energia spostandosi: in effetti, si dimostra che in questi casi anche la conduttanza σ_{xx} è nulla.

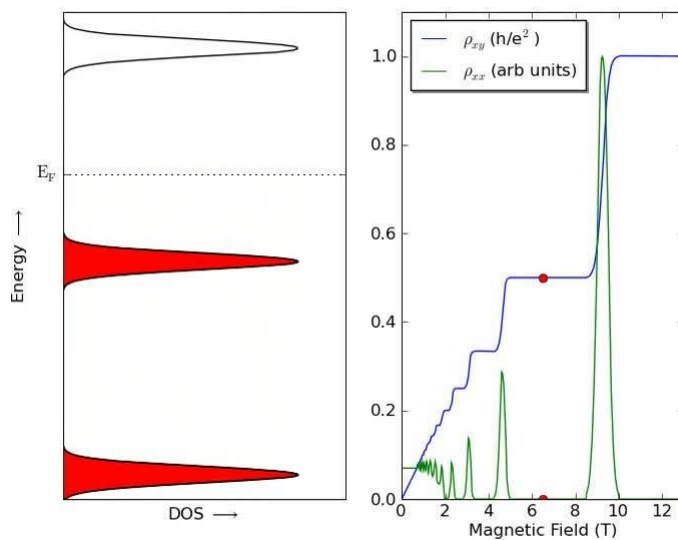


Figura 4.25: Grafico delle resistività ρ_{xy} e ρ_{xx} in funzione di B

Il punto importante è che l'esistenza in primo luogo dei *plateau* delle resistività è dovuta a imperfezioni del sistema, che hanno effetti di natura *topologica*, nel senso che le funzioni d'onda degli elettroni "si arrotolano" attorno alle impurità, rimanendo ivi *localizzate*. Questa situazione rende il sistema molto *stabile*, ed offre possibili applicazioni nella realizzazione di porte logiche per computer quantistici.

Poiché tuttavia siamo in $d = 2$, si apre la possibilità che rotazioni orarie e anti-orarie attorno a impurità *non siano equivalenti*, e che quindi entri in gioco una statistica *di trecce* e non di normali permutazioni. In effetti, questo è proprio quello che succede diminuendo ulteriormente la temperatura a $T \sim 1\text{K}$, come fu scoperto nel 1986.

In questo nuovo sistema si trovò che gli elettroni avevano uno stato di energia minimo uniforme, e i livelli eccitati erano costituiti da eccitazioni che obbedivano alla statistica di trecce, con $\theta = 2\pi/3$.

Se x, y sono le coordinate del piano, definiamo $z = x + iy$. Allora le funzioni d'onda di N di queste eccitazioni sono della forma:

$$\psi(z_1, \dots, z_N) = \prod_{i < j} (z_i - z_j)^{\frac{1}{3}} \prod_{i=1}^N e^{-c|z_i|^2}$$

Scambiando in senso antiorario due particelle, normalmente avremmo $(z_i - z_j) \rightarrow e^{i\pi}(z_i - z_j)$, ma ora abbiamo:

$$(z_i - z_j)^{\frac{1}{3}} \rightarrow (e^{i\pi}(z_i - z_j))^{\frac{1}{3}} = e^{i\frac{\pi}{3}}(z_i - z_j)^{\frac{1}{3}}$$

Analogamente, uno scambio orario sarebbe $(z_i - z_j) \rightarrow e^{-i\pi}(z_i - z_j)$, ma qui diviene:

$$(z_i - z_j)^{\frac{1}{3}} \rightarrow (e^{-i\pi}(z_i - z_j))^{\frac{1}{3}} = e^{-i\frac{\pi}{3}}(z_i - z_j)^{\frac{1}{3}}$$

La differenza appena notata è all'origine di *plateau* della resistività ρ_{xy} anche a valori frazionari di h/e^2 , cosa che non succedeva nel caso dell'effetto di Hall intero.

(Lezione 36 del
13/12/2018)

4.16 Potenziali periodici

In generale, risolvere l'equazione di Schrödinger per un potenziale generico $V(x)$ è estremamente difficile. A volte è però possibile sfruttare eventuali regolarità di $V(x)$ per semplificare il problema. Uno di questi casi, molto importante per lo studio di *cristalli*, è dato da un potenziale $V(x)$ periodico, per cui possiamo usare il **teorema di Bloch**.

Teorema di Bloch

Teorema 4.16.1. Sia $H = \frac{P^2}{2m} + V(\vec{X})$ una Hamiltoniana in \mathbb{R}^n con $V(\vec{x})$ periodico di periodo \vec{a} , cioè $V(\vec{x} + \vec{a}) = V(\vec{x})$. Allora le soluzioni dell'equazione di Schrödinger stazionaria per H , in rappresentazione $\{\vec{x}\}$, hanno la forma:

$$\psi_{\vec{k}}(\vec{x}) = e^{i\vec{k} \cdot \vec{x}} u_{\vec{k}}(\vec{x})$$

dove $u_{\vec{k}}(\vec{x})$ è una funzione periodica di periodo \vec{a} e \vec{k} , con componenti $k_i \in \left[-\frac{\pi}{a_i}, \frac{\pi}{a_i}\right]$, è detto **vettore d'onda di Bloch**.

Dimostrazione

Ragioniamo, per semplicità di notazione, nel caso in $d = 1$.

Introduciamo gli operatori di traslazione di multipli di a :

$$(T_n \psi)(x) = (U(na)\psi)(x) = \psi(x + na) \quad n \in \mathbb{Z}$$

Perciò applicare T_n *trasla spazialmente* la funzione d'onda di un multiplo intero della periodicità del potenziale $V(x)$. Possiamo comporre più traslazioni di questo tipo sfruttando la *regola di gruppo* delle traslazioni:

$$T_n T_m = T_m T_n = T_{n+m}$$

per cui i T_n commutano tra loro. Notiamo inoltre che commutano con l'Hamiltoniana H :

$$H = \frac{P^2}{2m} + V(X)$$

Sappiamo infatti che traslazioni spaziali lasciano invariato il *momento*:

$$T_n^\dagger P T_n = P = T_n^\dagger T_n P \Rightarrow [T_n, P] = 0$$

e ciò deriva dal fatto che i T_n sono generati dal momento P , e quindi sono funzioni di quest'ultimo:

$$T_n = \exp\left(i \frac{na}{\hbar} P\right)$$

Analogamente, verifichiamo che i T_n non variano il *potenziale*, dato che traslano in maniera “sincrona” con la sua periodicità:

$$T_n^\dagger V(X) T_n = V(x + na) = V(X) = T_n^\dagger T_n V(X) \Rightarrow [T_n, V(X)] = 0$$

Abbiamo perciò che:

$$[T_n, H] = 0 \quad \forall n \in \mathbb{Z}$$

Per teorema della compatibilità si ha allora che H e $\{T_n\}_{n \in \mathbb{Z}}$ sono compatibili e quindi hanno una base di autovettori $\psi_{\mathcal{E}}$ (eventualmente generalizzati) comuni. In rappresentazione $\{x\}$ valgono le seguenti equazioni agli autovalori:

$$\begin{aligned} H \psi_{\mathcal{E}}(x) &= \mathcal{E} \psi_{\mathcal{E}}(x) \\ T_n \psi_{\mathcal{E}}(x) &= \gamma_n \psi_{\mathcal{E}}(x) \end{aligned}$$

I T_n sono operatori unitari, e quindi non modificano il modulo di $\psi_{\mathcal{E}}(x)$. Vale cioè:

$$|T_n \psi_{\mathcal{E}}(x)| = |\gamma_n| |\psi_{\mathcal{E}}(x)| = |\psi_{\mathcal{E}}(x)| \Rightarrow |\gamma_n| = 1$$

Da cui scopriamo che γ_n è al più una fase, di forma:

$$\gamma_n = e^{i\beta_n} \quad \beta_n \in \mathbb{R}$$

β_n deve essere tale da soddisfare la *regola di gruppo*, per cui la composizione di due traslazioni è ancora una traslazione:

$$\begin{aligned} T_n T_m \psi_{\mathcal{E}}(x) &= T_{n+m} \psi_{\mathcal{E}}(x) \\ e^{i\beta_n} e^{i\beta_m} &= e^{i\beta_{n+m}} \Rightarrow \beta_n + \beta_m = \beta_{n+m} \end{aligned}$$

Perciò i β_n sono, a meno di un fattore costante, elementi di un gruppo additivo di numeri interi. In altre parole, la regola appena trovata è verificata solamente se ogni singolo β_n è il multiplo intero di un *angolo* θ , che fissiamo nell'intervallo $[-\pi, \pi]$ grazie alla periodicità 2π dell'esponenziale e^{ix} . In particolare scriviamo $\theta = ka$, dove a è la periodicità del potenziale $V(x)$, e k appartiene all'intervallo $[-\pi/a, +\pi/a]$. Riepilogando, giungiamo a:

$$\beta_n = n\theta = nka; \quad k \in \left[-\frac{\pi}{a}, \frac{\pi}{a}\right]$$

Sostituendo nell'equazione agli autovalori otteniamo:

$$T_n \psi_{\mathcal{E}}(x) = \gamma_n \psi_{\mathcal{E}}(x) = e^{inka} \psi_{\mathcal{E}}(x)$$

Perciò, di per sé, $\psi_{\mathcal{E}}(x)$ è “periodica a meno di una fase”. Possiamo evidenziare la sua “parte periodica”, estraendo allora un fattore e^{ikx} da $\psi_{\mathcal{E}}(x)$, e definendo quindi una nuova funzione $u_k(x)$:

$$\psi_{\mathcal{E}}(x) = e^{ikx} u_k(x) \Rightarrow u_k(x) \equiv e^{-ikx} \psi_{\mathcal{E}}(x)$$

Applicando ora T_n a $u_k(x)$:

$$T_n u_k(x) = e^{-ik(x+na)} e^{ikna} \psi_{\mathcal{E}}(x) = e^{-ikx} \psi_{\mathcal{E}}(x) = u_k(x)$$

Troviamo che $u_k(x)$ è periodica di periodo a . Perciò $\psi_{\mathcal{E}}(x) = e^{ikx} u_k(x)$ con $k \in [-\pi/a, +\pi/a]$, come si voleva dimostrare.

□

4.16.1 Esercizio 7

Sia $H = \frac{P^2}{2m} + V_0 \sum_{n \in \mathbb{Z}} \delta(x + na)$ l'Hamiltoniana di un sistema unidimensionale di una particella con massa m definita in un opportuno dominio $D(H)$ di autoaggiuntezza (che non ci interessa specificare, dato che non è semplice da trovare).

Nota: la δ non è un operatore, dato che applicandolo ad una $|\psi\rangle \in \mathcal{H}$ produce un risultato che è ancora una δ , e quindi non sta più in \mathcal{H} . Tuttavia si può dimostrare che la somma di $P^2/(2m)$ e le varie δ produce un operatore ben definito - dato che le singolarità, in qualche modo, sono “cancellate”.

Si considerino le soluzioni $\psi_{\mathcal{E}}(x)$ dell'equazione di Schrödinger stazionaria, con $\mathcal{E} \in \sigma(H)$.

1. Si calcoli la condizione di raccordo delle derivate $\psi'_\varepsilon(x)$ nelle posizioni $x_n = na$ integrando l'equazione di Schrödinger stazionaria in un intorno di ampiezza $\xi > 0$ di tali punti di ampiezza e considerando poi il limite $\xi \rightarrow 0$. Si assuma (cosa non ovvia da dimostrare) che le $\psi_\varepsilon(x)$ siano continue in tutto \mathbb{R} .

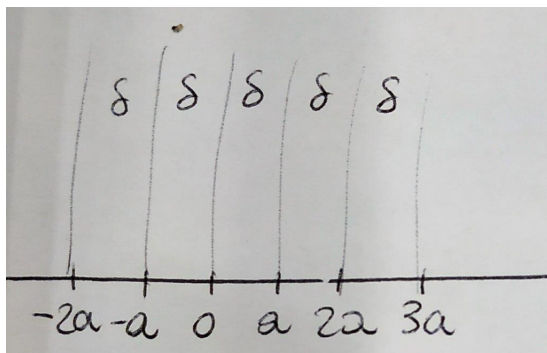


Figura 4.26: Grafico qualitativo del potenziale $V(x)$

2. Utilizzando il *teorema di Bloch* e le condizioni di raccordo in $x_n = na$ per la funzione d'onda ψ_ε e la sua derivata prima ψ'_ε , si determini l'equazione che collega il vettore d'onda k di Bloch allo spettro $\sigma(H)$ dell'energia.
3. Si discuta qualitativamente la struttura di $\sigma(H)$ determinandone la degenerazione.
4. Si supponga ora che l'Hamiltoniana H sia definita in $[0, Na] \subset \mathbb{R}$ (invece che su tutto \mathbb{R}), con $N \in \mathbb{N}$ e condizioni periodiche al bordo, restringendo la sommatoria \sum_n nella definizione di H a $n \in [0, N] \cap \mathbb{Z}$. Si discuta qualitativamente la struttura di $\sigma(H)$ in questo caso.

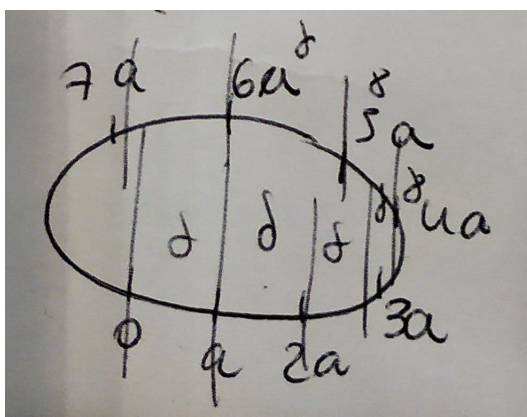


Figura 4.27: Schema qualitativo del potenziale *con condizioni periodiche al bordo*

4.16.2 Soluzione

1. Partiamo scrivendo l'equazione di Schrödinger stazionaria:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi_{\mathcal{E}}(x) + V_0 \sum_{n \in \mathbb{Z}} \delta(x + na) \psi_{\mathcal{E}}(x) = \mathcal{E} \psi_{\mathcal{E}}(x)$$

Seguendo il suggerimento, integriamola in un intervallo $[na - \frac{\xi}{2}, na + \frac{\xi}{2}]$:

$$\begin{aligned} & \int_{na - \frac{\xi}{2}}^{na + \frac{\xi}{2}} \left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi_{\mathcal{E}}(x) dx + \int_{na - \frac{\xi}{2}}^{na + \frac{\xi}{2}} V_0 \sum_{m \in \mathbb{Z}} \delta(x + ma) \psi_{\mathcal{E}}(x) dx \right] = \int_{na - \frac{\xi}{2}}^{na + \frac{\xi}{2}} \mathcal{E} \psi_{\mathcal{E}}(x) dx \\ \Rightarrow & -\frac{\hbar^2}{2m} \left(\psi'_{\mathcal{E}} \left(na + \frac{\xi}{2} \right) - \psi'_{\mathcal{E}} \left(na - \frac{\xi}{2} \right) \right) + V_0 \psi_{\mathcal{E}}(na) = \mathcal{E} \int_{na - \frac{\xi}{2}}^{na + \frac{\xi}{2}} \psi_{\mathcal{E}}(x) dx \end{aligned}$$

Prendendo il lim per $\xi \rightarrow 0$:

$$-\frac{\hbar^2}{2m} (\psi'_{\mathcal{E}}(na^+) - \psi'_{\mathcal{E}}(na^-)) + V_0 \psi_{\mathcal{E}}(na) = 0 \quad (4.136)$$

Dove abbiamo usato la continuità di $\psi_{\mathcal{E}}(x)$ (data per ipotesi), per affermare che un suo integrale su un intervallo di misura nulla è nullo. La (4.136) così ottenuta è la relazione di raccordo, in forma implicita, per le derivate $\psi'_{\mathcal{E}}(x)$ in $x_n = na$.

2. Dal teorema di Bloch sappiamo che $\psi_{\mathcal{E}}(x)$ è “periodica a meno di una fase”. Infatti, vale $\psi_{\mathcal{E}}(x) = e^{ikx} u_{k,\mathcal{E}}(x)$, dove tale $u_{k,\mathcal{E}}(x)$ è una funzione periodica di periodo a . Perciò:

$$\begin{aligned} \psi_{\mathcal{E}}(x + na) &= e^{ik(x+na)} u_{k,\mathcal{E}}(x + na) \stackrel{(a)}{=} e^{ik(x+na)} u_{k,\mathcal{E}}(x) = \\ &\stackrel{(b)}{=} e^{ik(x+na)} e^{-ikx} \psi_{\mathcal{E}}(x) = e^{ikna} \psi_{\mathcal{E}}(x) \end{aligned} \quad (4.137)$$

dove in (a) abbiamo usato la periodicità di $u_{k,\mathcal{E}}(x)$, e in (b) la definizione $u_{k,\mathcal{E}}(x) = e^{-ikx} \psi_{\mathcal{E}}(x)$.

Possiamo allora concentrarci sul trovare $\psi_{\mathcal{E}}(x)$ solo per $x \in [0, a]$, e usare la relazione (4.137) e le condizioni di raccordo per trovare la $\psi_{\mathcal{E}}(x)$ per $x \in \mathbb{R}$. Date le $\psi_{\mathcal{E}}(x)$ conosciamo lo spettro $\sigma(H)$ (da cui sono indicizzate), che possiamo poi mettere in relazione a k .

Notiamo ora che per $x \in]0, a[$ il potenziale $V(x) = 0$ è quello di una particella libera, di cui possiamo ricavare immediatamente le autofunzioni:

Autofunzioni
 $\psi_{\mathcal{E}}(x)$, $x \in]0, a[$

$$\begin{aligned} -\frac{\hbar^2}{2m} \psi''_{\mathcal{E}}(x) &= \mathcal{E} \psi_{\mathcal{E}}(x) \Rightarrow \psi''_{\mathcal{E}}(x) + \overbrace{\frac{2m\mathcal{E}}{\hbar^2}}^{q^2} \psi_{\mathcal{E}}(x) = 0 \\ \Rightarrow \psi_{\mathcal{E}}(x) &= A e^{iqx} + B e^{-iqx} \quad q = \sqrt{\frac{2m\mathcal{E}}{\hbar^2}} \end{aligned} \quad (4.138)$$

dove $A, B \in \mathbb{C}$ dipendono dalle condizioni al contorno su $\psi_{\mathcal{E}}(x)$ e $\psi'_{\mathcal{E}}(x)$.

Partiamo esaminando quelle su $\psi_{\mathcal{E}}(x)$. Per trovarle, facciamo uso del teorema di Bloch. Sappiamo che $u_{\mathcal{E},k}(x) = e^{-ikx}\psi_{\mathcal{E}}(x)$ per Bloch è periodica di periodo a , ed è anche continua, dato che lo è $\psi_{\mathcal{E}}(x)$ per ipotesi. Perciò deve essere:

$$u_{\mathcal{E},k}(0^-) = u_{\mathcal{E},k}(0^+) \quad (4.139)$$

La $\psi_{\mathcal{E}}(x)$ ricavata, tuttavia, vale solo in $]0, a[$, e mentre per $a^- \in]0, a[$ non ci sono problemi, si ha $0^- \notin]0, a[$, e quindi la (4.139) non è immediatamente utilizzabile. Possiamo però aggirare il problema usando la periodicità:

$$u_{\mathcal{E},k}(0^-) = u_{\mathcal{E},k}(a^-) \Rightarrow e^{-ik0}\psi_{\mathcal{E}}(0) = e^{-ika}\psi_{\mathcal{E}}(a) \quad (4.140)$$

Solo ora possiamo sostituire l'espressione trovata per $\psi_{\mathcal{E}}(x)$ in (4.138) e trovare quindi la prima condizione di raccordo:

Raccordo per
 $\psi_{\mathcal{E}}(x)$

$$A + B = Ae^{i(q-k)a} + Be^{-i(q+k)a} \quad (4.141)$$

D'altro canto, l'equazione di raccordo per la derivata $\psi'_{\mathcal{E}}(x)$ è già stata ricavata in (4.136), e basta esplicitarla per $n = 0$:

$$\psi'_{\mathcal{E}}(0^+) - \psi'_{\mathcal{E}}(0^-) = \frac{2mV_0}{\hbar^2}\psi_{\mathcal{E}}(0)$$

Di nuovo notiamo che $0^- \notin [0, a]$, per cui dobbiamo usare la periodicità. Usando l'analoga di (4.140) per la derivata:

$$\psi'_{\mathcal{E}}(0^-) = e^{-ika}\psi'_{\mathcal{E}}(a^-)$$

Otteniamo ora un'espressione in cui possiamo sostituire le derivate della $\psi_{\mathcal{E}}(x)$ trovata in (4.138):

Raccordo per
 $\psi'_{\mathcal{E}}(x)$

$$\begin{aligned} \psi'_{\mathcal{E}}(0^+) - e^{-ika}\psi'_{\mathcal{E}}(a^-) &= \frac{2mV_0}{\hbar^2}\psi_{\mathcal{E}}(0) \\ iq(A - B) - e^{-ika}iq(Ae^{iqa} - Be^{-iqa}) &= \frac{2mV_0}{\hbar^2}(A + B) \end{aligned} \quad (4.142)$$

Infine, mettiamo a sistema le condizioni di raccordo per $\psi_{\mathcal{E}}(x)$ (4.141) e la derivata $\psi'_{\mathcal{E}}(x)$ (4.142). Da questa troveremo la relazione cercata tra vettore di Bloch e spettro dell'energia.

$$\begin{cases} A(1 - e^{i(q-k)a}) + B(1 - e^{-i(q+k)a}) = 0 \\ A(iq)(1 - e^{i(q-k)a}) - B(iq)(1 - e^{-i(q+k)a}) = \frac{2mV_0}{\hbar^2}(A + B) \end{cases}$$

Ponendo:

$$\hat{A} = (1 - e^{i(q-k)a}); \quad \hat{B} = (1 - e^{-i(q+k)a})$$

Otteniamo:

$$\begin{cases} A\hat{A} + B\hat{B} \\ iq(A\hat{A} - B\hat{B} = \frac{2mV_0}{\hbar^2}(A + B) \end{cases} \Rightarrow \begin{cases} B = -A\hat{A}/\hat{B} \\ 2A\hat{A}(iq) = \frac{2mV_0}{\hbar^2}A \left(1 - \frac{\hat{A}}{\hat{B}}\right) \end{cases}$$

Svolgendo i conti nella seconda equazione:

$$\begin{aligned} \hat{A}\hat{B}(iq) &= \frac{mV_0}{\hbar^2}(\hat{B} - \hat{A}) \\ \Rightarrow (iq)[1 - e^{i(q-k)a} - e^{-i(q+k)a} + e^{-2ika}] &= \frac{mV_0}{\hbar^2}(e^{i(q-k)a} - e^{-i(q+k)a}) \end{aligned}$$

Moltiplicando entrambi i membri per e^{ika} e dividendo per $2i$:

$$\begin{aligned} iq(e^{ika} + e^{-ika} - e^{iqa} - e^{-iqa}) &= \frac{mV_0}{\hbar^2}(e^{iqa} - e^{-iqa}) \\ \Rightarrow q(\cos(ka) - \cos(qa)) &= \frac{mV_0}{\hbar^2} \sin(qa) \end{aligned}$$

Poiché l'autovalore dell'energia è contenuto in q , per trovare la relazione $k(\mathcal{E})$ cercata basta ora isolare il termine con k :

$$\cos(ka) = \cos(qa) + \frac{mV_0}{\hbar^2} \frac{\sin(qa)}{q} \quad q = \sqrt{\frac{2m\mathcal{E}}{\hbar^2}} \quad (4.143)$$

3. Com'è fatto, qualitativamente, lo spettro $\sigma(H)$, e qual è la sua degenerazione?

Partiamo da (4.143), dove imponiamo che (dal teorema di Bloch) $k \in \left[-\frac{\pi}{a}, \frac{\pi}{a}\right]$, da cui $-1 \leq \cos(ka) \leq +1$, e perciò:

$$-1 \leq \underbrace{\cos(qa) + \frac{mV_0}{\hbar^2} \frac{\sin(qa)}{q}}_{y(q)} \leq 1$$

Si ha allora che non tutti i valori di $q(\mathcal{E})$ sono, a priori, permessi, ma solo quelli per cui $-1 < y(q) < +1$, come rappresentato in figura 4.28.

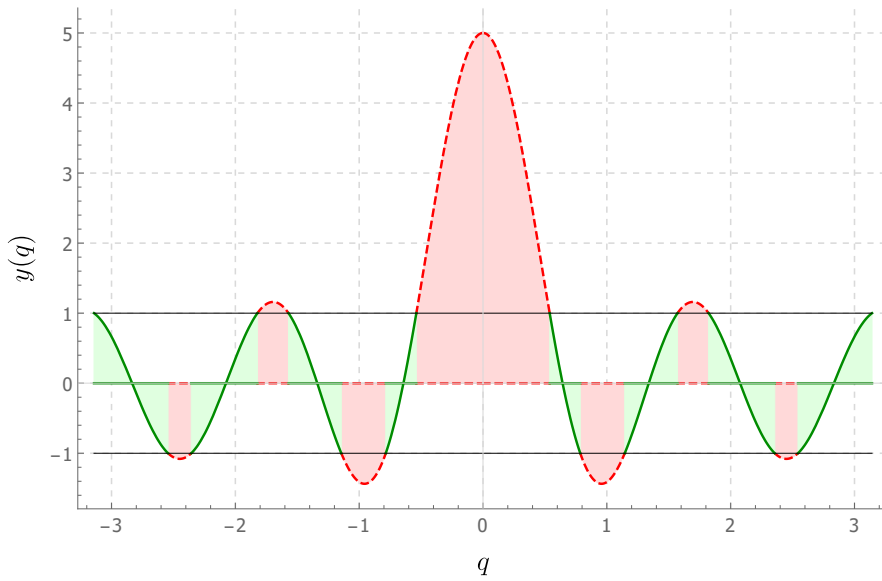


Figura 4.28: Spettro a bande: solo i valori di $q(\mathcal{E})$ evidenziati in verde sono fisicamente permessi (per $mV_0/\hbar^2 = 1$, $a = 4$)

Escludiamo le regioni in cui la funzione esce dal range $-1 \leq y \leq +1$.

Per le q ammesse, la $\psi_{\mathcal{E}}(x)$ non sta in L^2 (oscilla all'infinito), e perciò $\sigma(H)$ è continuo. Il supporto di $\sigma(H)$ è l'unione di intervalli disgiunti che chiameremo *bande* (i tratti evidenziati in verde sull'asse x del grafico 4.28). Quanti modi sono possibili a energia fissata? Notiamo che $\sigma(H)$ dipende univocamente da q . Fissare l'energia significa impostare:

*Spettro continuo
a bande*

*Degenerazione di
 $\sigma(H)$*

$$\mathcal{E} = \frac{\hbar^2 q^2}{2m}$$

e per una certa \mathcal{E} sono possibili 2 valori di q : q e $-q$. Perciò la degenerazione di $\sigma(H)$ è 2.

4. Consideriamo ora lo stesso sistema confinato in una regione finita, con $x \in [0, Na]$. Dalle condizioni al contorno periodiche per la funzione d'onda $\psi_{\mathcal{E}}(x)$:

$$\psi_{\mathcal{E}}(0) = \psi_{\mathcal{E}}(Na) \underset{(a)}{=} e^{ikNa} \psi_{\mathcal{E}}(0) \Rightarrow e^{ikNa} = 1$$

dove in (a) si è applicato il teorema di Bloch. Si ha perciò una condizione ulteriore sull'esponente kNa :

$$kNa \in 2\pi\mathbb{Z} \Rightarrow k \in \frac{2\pi\mathbb{Z}}{Na} \cap \left[-\frac{\pi}{a}, \frac{\pi}{a}\right]$$

Ne segue che k può assumere solo valori in un set discreto e finito. Poiché k è legato a $\sigma(H)$ come visto al punto precedente, ricaviamo che in questo caso l'energia ha uno *spettro discreto*. Infatti, la $\psi_{\mathcal{E}}(x)$ è ora ristretta a un

intervallo compatto, e perciò non “oscilla più all’infinito”: è quindi modulo-quadro sommabile e si trova in L^2 :

$$\psi_{\mathcal{E}}(x) \in L^2([0, Na], dx) \Rightarrow \sigma(H) \text{ è discreto}$$

E in effetti:

$$\left\{ k | k \in \frac{2\pi\mathbb{Z}}{Na} \cap \left[-\frac{\pi}{a}, \frac{\pi}{a} \right] \right\} \text{ è discreto}$$

Perciò nelle *bande* di figura 4.28 solo alcuni valori *discreti* sono ammessi: invece di avere tratti evidenziati “continui” per i q ammessi si ha una sequenza di *singoli punti* (seppur molto fini, dato che ogni banda può contenere un numero di stati dell’ordine di 10^{23}). Chiamiamo questo tipo di spettro uno *spettro discreto a bande*: si hanno quindi intervalli “larghi” che contengono un gran numero di autovalori molto vicini tra loro, separati da *gap* in cui non ve ne è nessuno.

Spettro discreto a bande

L’esercizio appena fatto ha un’utile interpretazione fisica, in quanto costituisce un *toy model* (modello semplificato) per un cristallo 1D.

Cristalli in $d = 1$

Consideriamo una *fila* di atomi *equispaziati*. Ciascuno di essi ha un potenziale coulombiano, e avremo degli stati discreti corrispondenti agli stati legati degli elettroni (i “consueti” orbitali). Un elettrone può, per *effetto tunnel*, attraversare lo spazio tra un atomo e l’altro, schematizzabile come una barriera (le δ nell’esempio appena fatto). Se il cristallo è infinito, la funzione d’onda che incorpora tutte le *onde evanescenti* di trasferimento da un atomo all’altro *non* appartiene a L^2 . *Fisicamente*, ciò significa che gli elettroni possono *spostarsi arbitrariamente lontano*. Tali stati costituiscono “un’onda”.

In un reticolo periodico, l’elettrone si comporta perciò “come un’onda”, da cui lo *spettro continuo a bande*.

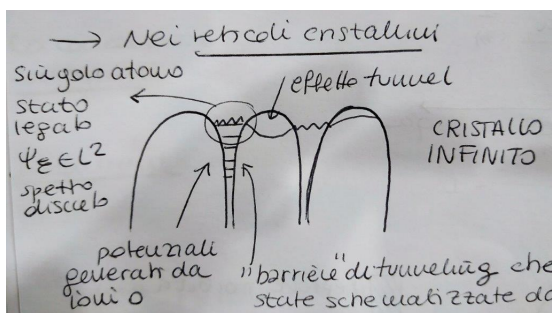


Figura 4.29: La possibilità degli elettroni di passare da un atomo all’altro del reticolo cristallino per *effetto tunnel* fa sì che $\sigma(H)$ sia continuo

Se il cristallo è finito, avremo invece uno *spettro discreto a bande*.

In un *reticolo ionico* con N_e elettroni, l’Hamiltoniana delle particelle singole è data da:

$$H_i = \frac{P_i^2}{2m} + V(x_i)$$

con P_i, X_i dell'elettrone i -esimo.

L'Hamiltoniana del sistema composto è data, ignorando le interazioni tra i vari elettroni:

$$H \approx \sum_{i=1}^{N_e} H_i$$

Nota: H è simmetrica per scambio di variabili, come vogliamo dato che gli elettroni sono indistinguibili.

Per il principio di esclusione di Pauli, due elettroni non possono avere gli stessi valori per un ICOC di particella singola, che in questo caso è dato da $\{q\}$:

$$q = \pm \sqrt{\frac{2mH}{\hbar^2}}$$

Gli elettroni prendono i vari q possibili, *dal basso verso l'alto*. Dato che N_e è finito, si ha, come diretta conseguenza del principio di Pauli, che esiste un $q_{max} \equiv q_F$, detto **momento di Fermi**.

Avremo allora due possibilità:

- q_F sta all'interno della banda *permessa*. Se allora applichiamo una differenza di potenziale, l'energia (cinetica media) degli elettroni *aumenta*. Dato che esistono q di energia permessa sopra a q_F , gli elettroni possono occuparli, e quindi spostarsi. Tale processo descrive la *conduzione* nei metalli.
- Se invece q_F sta alla *fine* della banda permessa (ossia tutti i valori permessi sono *pieni*), anche applicando una differenza di potenziale gli elettroni non possono muoversi, dato che non hanno livelli energetici *permessi* subito sopra che possano occupare. Questo è la descrizione dei materiali *isolanti*.

L'idea alla base dell'elettronica è quella di trovare un *gap* tra due bande che sia *sufficientemente piccolo* in modo da poter essere superato applicando un opportuno potenziale, ma *sufficientemente alto* affinché l'agitazione termica non basti a far ciò.

Perciò il funzionamento dell'elettronica è un altro esempio di applicazione del principio di esclusione di Pauli.

Momento di
Fermi

(Lezione 37 del
14/12/2018)

4.17 Teoria delle perturbazioni

Spesso, in MQ, ci si trova di fronte a problemi che non sono risolvibili analiticamente, ma “che non si discostano eccessivamente” da altri problemi che invece lo sono. Uno strumento utile per gestire tali situazioni è dato dalla *teoria perturbativa*, di cui ci occuperemo in questa sezione, limitatamente al caso di perturbazioni **indipendenti dal tempo**.

Perturbazioni
indipendenti dal
tempo

L'idea di fondo sta nell'esaminare come cambiano gli autostati di una hamiltoniana data dalla somma di due termini: uno che si sa risolvere in modo esatto, e un

altro, più piccolo del primo, che “lo perturba”. Un esempio è l’atomo di idrogeno inserito in un campo magnetico debole.

Matematicamente, consideriamo un’Hamiltoniana H scomponibile in due termini:

$$H = H_0 + \lambda V \quad (4.144)$$

dove di H_0 sappiamo autovalori e autostati, e λ è un parametro “piccolo” rispetto agli altri.

Denotiamo gli autovalori di H_0 con \mathcal{E}_n^0 , che assumeremo inizialmente **ortonormali** e **non degeneri**, e i rispettivi autostati con $|\mathcal{E}_n^0\rangle$. Per H_0 valgono allora le equazioni agli autovalori:

$$H_0 |\mathcal{E}_n^0\rangle = \mathcal{E}_n^0 |\mathcal{E}_n^0\rangle$$

*Caso 1:
Autovalori \mathcal{E}_n^0
non degeneri*

Denotiamo invece con $|\mathcal{E}_n\rangle$ gli autostati dell’hamiltoniana H “completa”, per cui:

$$H |\mathcal{E}_n\rangle = \mathcal{E}_n |\mathcal{E}_n\rangle$$

Al limite di una perturbazione “nulla”, questi ultimi autovalori \mathcal{E}_n tendono a quelli \mathcal{E}_n^0 conosciuti:

$$\lim_{\lambda \rightarrow 0} \mathcal{E}_n = \mathcal{E}_n^0$$

Supponiamo (**ipotesi perturbativa**) di poter espandere sia $|\mathcal{E}_n\rangle$ che \mathcal{E}_n in una serie di Mac-Laurin in λ :

*Ipotesi
perturbativa*

$$|\mathcal{E}_n\rangle = |\mathcal{E}_n^0\rangle + \lambda |\mathcal{E}_n^1\rangle + \lambda^2 |\mathcal{E}_n^2\rangle + \dots \quad (4.145)$$

$$\mathcal{E}_n = \mathcal{E}_n^0 + \lambda \mathcal{E}_n^1 + \lambda^2 \mathcal{E}_n^2 + \dots \quad (4.146)$$

$$\mathcal{E}_n^k = \frac{1}{k!} \frac{d^k \mathcal{E}_n}{d\lambda^k} \Big|_{\lambda=0} \quad |\mathcal{E}_n^k\rangle = \frac{1}{k!} \frac{d^k |\mathcal{E}_n\rangle}{d\lambda^k} \Big|_{\lambda=0}$$

Assumiamo come normalizzazione per gli autovettori $|\mathcal{E}_n\rangle$ la seguente espressione:

$$\langle \mathcal{E}_n^0 | \mathcal{E}_n \rangle \stackrel{!}{=} 1 \quad (4.147)$$

Espandendo $|\mathcal{E}_n\rangle$ nella normalizzazione (4.147) tramite la serie di potenze in (4.145) otteniamo:

$$1 \stackrel{(4.147)}{=} \langle \mathcal{E}_n^0 | \mathcal{E}_n \rangle = \underbrace{\langle \mathcal{E}_n^0 | \mathcal{E}_n^0 \rangle}_{=1} + \lambda \langle \mathcal{E}_n^0 | \mathcal{E}_n^1 \rangle + \lambda^2 \langle \mathcal{E}_n^0 | \mathcal{E}_n^2 \rangle + \dots \quad \forall \lambda$$

Poiché tale espressione deve valere per ogni scelta di λ , si che i coefficienti dei termini dello stesso ordine devono essere uguali. Ma poiché il membro a sinistra non contiene potenze di λ (se non quella di ordine 0, ricaviamo che:

$$\langle \mathcal{E}_n^0 | \mathcal{E}_n^k \rangle = 0 \quad k = 1, 2, 3, \dots$$

Nota: la normalizzazione appena adottata è *non canonica*, nel senso che così facendo in genere $\langle \mathcal{E}_n | \mathcal{E}_n \rangle \neq 1$. Come vedremo nei conti seguenti, tuttavia, si dimostra una scelta utile al nostro scopo.

Per trovare gli autovalori \mathcal{E}_n^k dell'espansione perturbativa partiamo dall'equazione agli autovalori generale:

$$(H_0 + \lambda V) |\mathcal{E}_n\rangle = \mathcal{E}_n |\mathcal{E}_n\rangle$$

Inserendo gli sviluppi in potenze (4.145) e (4.146):

$$(H_0 + \lambda V) \sum_{k=0}^{\infty} \lambda^k |\mathcal{E}_n^k\rangle = \left(\sum_{l=0}^{\infty} \lambda^l \mathcal{E}_n^l \right) \left(\sum_{m=0}^{\infty} \lambda^m |\mathcal{E}_n^m\rangle \right) \quad (4.148)$$

$$\begin{aligned} (H_0 + \lambda V) (|\mathcal{E}_n^0\rangle + \lambda |\mathcal{E}_n^1\rangle + \lambda^2 |\mathcal{E}_n^2\rangle + \dots) &= (\mathcal{E}_n^0 + \lambda \mathcal{E}_n^1 + \lambda^2 \mathcal{E}_n^2 + \dots) \cdot \\ &\cdot (|\mathcal{E}_n^0\rangle + \lambda |\mathcal{E}_n^1\rangle + \lambda^2 |\mathcal{E}_n^2\rangle + \dots) \end{aligned} \quad (4.149)$$

Nuovamente, l'espressione vale $\forall \lambda$, e quindi i coefficienti di termini dello stesso ordine devono essere uguali.

- All'ordine $k = 0$, riotteniamo l'equazione agli autovalori per H_0 :

$$H_0 |\mathcal{E}_n^0\rangle = \mathcal{E}_n^0 |\mathcal{E}_n^0\rangle$$

- All'ordine $k = 1$, si ha invece:

$$H_0 |\mathcal{E}_n^1\rangle + V |\mathcal{E}_n^0\rangle = \mathcal{E}_n^0 |\mathcal{E}_n^1\rangle + \mathcal{E}_n^1 |\mathcal{E}_n^0\rangle$$

Prendendo il prodotto scalare con $|\mathcal{E}_n^0\rangle$ otteniamo:

$$\underbrace{\langle \mathcal{E}_n^0 | H_0 | \mathcal{E}_n^1 \rangle}_{\mathcal{E}_n^0 \langle \mathcal{E}_n^0 | \mathcal{E}_n^1 \rangle = 0} + \langle \mathcal{E}_n^0 | V | \mathcal{E}_n^0 \rangle = \mathcal{E}_n^0 \underbrace{\langle \mathcal{E}_n^0 | \mathcal{E}_n^1 \rangle}_{=0} + \mathcal{E}_n^1$$

Notando che $\langle \mathcal{E}_n^0 | H = \mathcal{E}_n^0 \langle \mathcal{E}_n^0 |$ (per la versione duale dell'equazione di Schrödinger) e che $\langle \mathcal{E}_n^0 | \mathcal{E}_n^1 \rangle = 0$ per la normalizzazione adottata, giungiamo a:

$$\mathcal{E}_n^1 = \langle \mathcal{E}_n^0 | V | \mathcal{E}_n^0 \rangle$$

E perciò \mathcal{E}_n , al primo ordine in λ è dato da:

$$\mathcal{E}_n \approx \mathcal{E}_n^0 + \lambda \mathcal{E}_n^1 = \langle \mathcal{E}_n^0 | H_0 | \mathcal{E}_n^0 \rangle + \lambda \langle \mathcal{E}_n^0 | V | \mathcal{E}_n^0 \rangle \stackrel{(4.144)}{=} \langle \mathcal{E}_n^0 | H | \mathcal{E}_n^0 \rangle$$

Perciò, al I ordine in λ l'autovalore \mathcal{E}_n è il valor medio di H nello stato imperturbato $|\mathcal{E}_n^0\rangle$.

Più in generale, all'ordine k :

$$H_0 |\mathcal{E}_n^k\rangle + V |\mathcal{E}_n^{k-1}\rangle = \sum_{l=0}^k \mathcal{E}_n^l |\mathcal{E}_n^{k-l}\rangle \quad (4.150)$$

Prendendo nuovamente il prodotto scalare con $|\mathcal{E}_n^0\rangle$:

Espressione per \mathcal{E}_n^k

$$\underbrace{\langle \mathcal{E}_n^0 | H_0 | \mathcal{E}_n^k \rangle + \langle \mathcal{E}_n^0 | V | \mathcal{E}_n^{k-1} \rangle}_{\mathcal{E}_n^0 \langle \mathcal{E}_n^0 | \mathcal{E}_n^k \rangle = 0} = \mathcal{E}_n^k \Rightarrow \mathcal{E}_n^k = \langle \mathcal{E}_n^0 | V | \mathcal{E}_n^{k-1} \rangle \quad k > 0 \quad (4.151)$$

dato che al secondo membro *sopravvive* solo il termine con $k-l=0 \Rightarrow l=k$, dato che i prodotti $\langle \mathcal{E}_n^0 | \mathcal{E}_n^{k-l} \rangle = \delta_{0,k-l}$ per la normalizzazione adottata.

Notiamo che otteniamo una *formula chiusa*. Riscrivendo il membro a sinistra di (4.148) e prendendo il prodotto scalare con $|\mathcal{E}_n^0\rangle$:

$$\begin{aligned} \langle \mathcal{E}_n^0 | (H_0 + \lambda V) \left(\sum_{l=0}^{\infty} \lambda^l |\mathcal{E}_n^l\rangle \right) &= \langle \mathcal{E}_n^0 | H_0 | \mathcal{E}_n^0 \rangle + \lambda \langle \mathcal{E}_n^0 | V | \mathcal{E}_n^0 \rangle + \sum_{l=1}^{\infty} \lambda^{l+1} \langle \mathcal{E}_n^0 | V | \mathcal{E}_n^l \rangle = \\ &\stackrel{(a)}{=} \mathcal{E}_n^0 + \lambda \mathcal{E}_n^1 + \sum_{l=1}^{\infty} \lambda^{l+1} \mathcal{E}_n^{l+1} = \mathcal{E}_n \end{aligned}$$

dove in (a) abbiamo usato le relazioni appena trovate, notando che $\langle \mathcal{E}_n^0 | H | \mathcal{E}_n^0 \rangle = \langle \mathcal{E}_n^0 | H_0 | \mathcal{E}_n^0 \rangle = \mathcal{E}_n^0$ e, per $l \geq 1$ $\langle \mathcal{E}_n^0 | V | \mathcal{E}_n^l \rangle = \mathcal{E}_n^l$. Ritroviamo perciò, nel membro a destra, l'espansione in serie di potenze di \mathcal{E}_n .

Ricordando poi che:

$$\sum_{l=0}^{\infty} \lambda^l |\mathcal{E}_n^l\rangle = |\mathcal{E}_n\rangle$$

Da cui:

$$\langle \mathcal{E}_n^0 | H | \mathcal{E}_n \rangle = \mathcal{E}_n^0 + \lambda \mathcal{E}_n^1 + \sum_{l=1}^{\infty} \lambda^{l+1} \mathcal{E}_n^{l+1} = \mathcal{E}_n$$

Cerchiamo ora un'espressione per i vettori $|\mathcal{E}_n^k\rangle$. Un'idea è determinarli nella base degli autovettori $|\mathcal{E}_m^0\rangle$ che conosciamo. Matematicamente, partiamo usando la completezza di Dirac per la base $\{|\mathcal{E}_m^0\rangle\}_{m \in \mathbb{N}}$, nell'ipotesi di **autovalori non degeneri**, da cui otteniamo l'espansione di un generico $|\mathcal{E}_n^k\rangle$:

Ricerca degli autovettori $|\mathcal{E}_n^k\rangle$

$$|\mathcal{E}_n^k\rangle = \mathbb{I} |\mathcal{E}_n^k\rangle = \sum_m |\mathcal{E}_m^0\rangle \langle \mathcal{E}_m^0 | \mathcal{E}_n^k \rangle \quad k > 0$$

Notiamo che $\langle \mathcal{E}_m^0 | \mathcal{E}_n^k \rangle$ è nullo se $m \neq n$ (per la normalizzazione introdotta). Riscriviamo quindi la somma come:

$$|\mathcal{E}_n^k\rangle = \sum_{m \neq n} |\mathcal{E}_m^0\rangle \langle \mathcal{E}_m^0 | \mathcal{E}_n^k \rangle$$

Prendendo ora il prodotto scalare con $|\mathcal{E}_m^0\rangle$ dell'equazione agli autovalori di ordine k (4.150):

$$\begin{aligned} \langle \mathcal{E}_m^0 | H_0 | \mathcal{E}_n^k \rangle + \langle \mathcal{E}_m^0 | V | \mathcal{E}_n^{k-1} \rangle &= \sum_{l=0}^k \mathcal{E}_n^l \langle \mathcal{E}_m^0 | \mathcal{E}_n^{k-l} \rangle \\ \mathcal{E}_m^0 \langle \mathcal{E}_m^0 | \mathcal{E}_n^k \rangle + \langle \mathcal{E}_m^0 | V | \mathcal{E}_n^{k-1} \rangle &= \mathcal{E}_n^0 \langle \mathcal{E}_m^0 | \mathcal{E}_n^k \rangle + \sum_{l=1}^k \mathcal{E}_n^l \langle \mathcal{E}_m^0 | \mathcal{E}_n^{k-l} \rangle \end{aligned}$$

Isolando $\langle \mathcal{E}_m^0 | \mathcal{E}_n^k \rangle$:

$$\langle \mathcal{E}_m^0 | \mathcal{E}_n^k \rangle = \frac{1}{\mathcal{E}_n^0 - \mathcal{E}_m^0} \left(\langle \mathcal{E}_m^0 | V | \mathcal{E}_n^{k-1} \rangle - \sum_{l=1}^k \mathcal{E}_n^l \langle \mathcal{E}_m^0 | \mathcal{E}_n^{k-l} \rangle \right)$$

E, sostituendo nella completezza:

$$|\mathcal{E}_n^k\rangle = \sum_{m \neq n} \frac{|\mathcal{E}_m^0\rangle}{\mathcal{E}_n^0 - \mathcal{E}_m^0} \left(\langle \mathcal{E}_m^0 | V | \mathcal{E}_n^{k-1} \rangle - \sum_{l=1}^k \mathcal{E}_n^l \langle \mathcal{E}_m^0 | \mathcal{E}_n^{k-l} \rangle \right) \quad (4.152)$$

L'espressione è ben definita dato che, per ipotesi di nondegenerazione, $\mathcal{E}_n^0 \neq \mathcal{E}_m^0$ per $m \neq n$ (e per $m = n$ $\langle \mathcal{E}_m^0 | \mathcal{E}_m^k \rangle = \delta_{0,k}$ dalla normalizzazione).

Notiamo ora che il termine di destra contiene solo termini “precedenti” nell'espansione, del tipo $|\mathcal{E}_n^s\rangle$ con $s < k$. Anche gli autovalori \mathcal{E}_n^l che compaiono richiedono, per essere calcolati, termini *precedenti*, data la formula trovata in (4.151):

$$\mathcal{E}_n^k = \langle \mathcal{E}_n^0 | V | \mathcal{E}_n^{k-1} \rangle$$

Perciò possiamo procedere *iterativamente*, calcolando $|\mathcal{E}_n^{k+1}\rangle$ a partire dai $|\mathcal{E}_n^k\rangle$ di ordine inferiore già determinati, e così ricavare tutta l'espansione.

Sostituendo le espressioni per $|\mathcal{E}_n^k\rangle$ nell'espansione per $|\mathcal{E}_n\rangle$ in (4.145) giungiamo a:

Formula per gli autovettori $|\mathcal{E}_n\rangle$

$$|\mathcal{E}_n\rangle = |\mathcal{E}_n^0\rangle + \lambda \sum_{m \neq n} |\mathcal{E}_m^0\rangle \frac{\overbrace{\langle \mathcal{E}_m^0 | V | \mathcal{E}_n^0 \rangle}^{\langle \mathcal{E}_m^0 | \mathcal{E}_n^1 \rangle}}{\mathcal{E}_n^0 - \mathcal{E}_m^0} + \dots$$

1. Il primo termine:

$$\begin{aligned} \mathcal{E}_n^1 &= \langle \mathcal{E}_n^0 | V | \mathcal{E}_n^0 \rangle \\ |\mathcal{E}_n^1\rangle &\stackrel{(4.152)}{=} \sum_{m \neq n} |\mathcal{E}_m^0\rangle \frac{\langle \mathcal{E}_m^0 | V | \mathcal{E}_n^0 \rangle - \cancel{\mathcal{E}_n^1 \langle \mathcal{E}_m^0 | \mathcal{E}_n^0 \rangle}}{\mathcal{E}_m^0 - \mathcal{E}_n^0} = \sum_{m \neq n} |\mathcal{E}_m^0\rangle \frac{\langle \mathcal{E}_m^0 | V | \mathcal{E}_n^0 \rangle}{\mathcal{E}_m^0 - \mathcal{E}_n^0} \end{aligned} \quad (4.153)$$

2. Per il secondo termine, invece, otteniamo:

$$\mathcal{E}_n^2 \stackrel{(4.151)}{=} \langle \mathcal{E}_n^0 | V | \mathcal{E}_n^1 \rangle = \sum_{m \neq n} \frac{\langle \mathcal{E}_n^0 | V | \mathcal{E}_m^0 \rangle \langle \mathcal{E}_m^0 | V | \mathcal{E}_n^0 \rangle}{\mathcal{E}_n^0 - \mathcal{E}_m^0} = \sum_{m \neq n} \frac{|\langle \mathcal{E}_n^0 | V | \mathcal{E}_m^0 \rangle|^2}{\mathcal{E}_n^0 - \mathcal{E}_m^0}$$

Finora abbiamo utilizzato una normalizzazione *non canonica*, per cui $\langle \mathcal{E}_n | \mathcal{E}_n^0 \rangle = 1$. Se vogliamo passare alla normalizzazione “usuale” è perciò necessario un passaggio di *rinormalizzazione*. Denotando il risultato con $|\mathcal{E}_n\rangle^N$, tale che ${}^N\langle \mathcal{E}_n | \mathcal{E}_n \rangle^N = 1$, otteniamo:

Normalizzazione canonica

$$|\mathcal{E}_n\rangle^N = z^{\frac{1}{2}} |\mathcal{E}_n\rangle \quad \langle \mathcal{E}_n | \mathcal{E}_n^0 \rangle = 1$$

z è detta *costante di rinormalizzazione* della funzione d'onda.

Si trova $z = 1$ per il primo ordine, mentre per $k = 2$:

$$\lambda^2 \Rightarrow z_2 = 1 - \lambda^2 \sum_{m \neq n} \frac{|\langle \mathcal{E}_m^0 | V | \mathcal{E}_n^0 \rangle|^2}{(\mathcal{E}_n^0 - \mathcal{E}_m^0)^2}$$

Possiamo ora calcolare agilmente le probabilità di misura. Se il sistema è in un autostato $|\mathcal{E}_n\rangle$ di H , la probabilità che una misura la trovi nello stato $|\mathcal{E}_n^0\rangle$ è:

$$p = \left| {}^N\langle \mathcal{E}_n | \mathcal{E}_n^0 \rangle \right|^2$$

Dove la N ad apice a sinistra indica che stiamo usando il *bra* normalizzato canonicamente, duale di $|\mathcal{E}_n\rangle^N$.

Cosa succede se rimuoviamo l'ipotesi che H_0 abbia solo autovalori nondegeneri? Supponiamo che \mathcal{E}_n^0 sia *degenere* e che $\{|\mathcal{E}_n^0, i\rangle$ autoket con $i = 1, \dots, d(n) = \deg(\mathcal{E}_n^0)\}$ sia una base dell'autospazio \mathcal{H}_n corrispondente all'autovalore \mathcal{E}_n^0 di H_0 . Nei calcoli precedenti era cruciale che $\mathcal{E}_n^0 \neq \mathcal{E}_m^0$, per evitare l'annullarsi del denominatore. In presenza di degenerazione, tuttavia, ciò accade. Infatti, se sostituiamo alla completezza di Dirac la sua versione *con degenerazione*:

Caso 2:
Autovalori \mathcal{E}_n^0
degeneri

$$\mathbb{I} = \sum_m |\mathcal{E}_m^0\rangle \langle \mathcal{E}_m^0| \rightarrow \mathbb{I} = \sum_m \sum_{i=1}^{d(m)} |\mathcal{E}_m^0, i\rangle \langle \mathcal{E}_m^0, i|$$

ripetendo i passaggi di prima, giungiamo all'espressione:

$$\mathcal{E}_m^0 \langle \mathcal{E}_m^0, i | \mathcal{E}_n^k \rangle + \langle \mathcal{E}_m^0, i | V | \mathcal{E}_m^{k-1} \rangle = \mathcal{E}_n^0 \langle \mathcal{E}_m^0, i | \mathcal{E}_n^k \rangle + \sum_{l=1}^k \mathcal{E}_n^l \langle \mathcal{E}_m^0 | \mathcal{E}_n^{k-l} \rangle$$

da cui non possiamo più ricavare $\langle \mathcal{E}_m^0, i | \mathcal{E}_n^k \rangle$, dato che dovremmo dividere per $\mathcal{E}_n^0 - \mathcal{E}_m^0$ che non sempre è $\neq 0$.

Per eliminare questi problemi osserviamo che $1/(\mathcal{E}_n^0 - \mathcal{E}_m^0)$ compare, per esempio in (4.153), moltiplicato per un fattore che, nel caso con degenerazione, è del tipo $\langle \mathcal{E}_n^0, \alpha | V | \mathcal{E}_m^0, \beta \rangle$.

Poiché abbiamo libertà nella scelta degli autovettori generalizzati (basta che siano una qualsiasi base dell'autospazio dell'autovalore esaminato), possiamo sceglierli in modo che gli elementi di matrice:

$$\langle \mathcal{E}_n^0, \alpha | V | \mathcal{E}_m^0, \beta \rangle = 0 \Leftrightarrow \alpha \neq \beta$$

Perciò, scegliendo così i ket $\{|\mathcal{E}_n^0, \alpha\rangle\}$ come una base dell'autospazio \mathcal{H}_n di autovalore \mathcal{E}_n^0 in cui V è diagonale, il problema precedente si risolve perché ogni termine $1/(\mathcal{E}_n^0 - \mathcal{E}_m^0)$ che divergerebbe ha come numeratore $\langle \mathcal{E}_n^0, \alpha | V | \mathcal{E}_m^0, \beta \rangle = 0$ per $\beta \neq \alpha$. Perciò gli stessi argomenti del caso non degenerare si applicano al caso degenerare con questa scelta di base, che è possibile se V è **simmetrico**. Perciò, riadattando le formule precedentemente ricavate:

$$\mathcal{E}_n^\alpha = \mathcal{E}_n^0 + \lambda \langle \mathcal{E}_n^0, \alpha | V | \mathcal{E}_n^0, \alpha \rangle + \lambda^2 \sum_{m \neq n} \sum_{\beta} \frac{|\langle \mathcal{E}_m^0, \beta | V | \mathcal{E}_n^0, \alpha \rangle|^2}{\mathcal{E}_n^0 - \mathcal{E}_m^0} + \dots \quad (4.154)$$

$$|\mathcal{E}_n, \alpha\rangle = |\mathcal{E}_n^0, \alpha\rangle + \lambda \sum_{m \neq n} \sum_{\beta} |\mathcal{E}_m^0, \alpha\rangle \frac{\langle \mathcal{E}_m^0, \alpha | V | \mathcal{E}_n^0, \beta \rangle}{\mathcal{E}_n^0 - \mathcal{E}_m^0} + \dots \quad (4.155)$$

Con una scelta delle basi $|\mathcal{E}_n, \alpha\rangle$ tale che:

$$\begin{aligned} \langle \mathcal{E}_m^0, \beta | V | \mathcal{E}_n^0, \alpha \rangle &= 0 \quad m \neq n \\ \langle \mathcal{E}_n^0, \beta | V | \mathcal{E}_n^0, \alpha \rangle &= 0 \quad \alpha \neq \beta \end{aligned}$$

4.17.1 Conseguenze delle perturbazioni

In generale, se a \mathcal{E}_n^0 degenerare sono associati $d(n)$ autovettori (di H_0) generalizzati $|\mathcal{E}_n^0, \alpha\rangle$, con $\alpha = 1, \dots, d(n)$, i relativi autovettori della H “completa” $|\mathcal{E}_n, \alpha\rangle$ non saranno più degeneri, ma corrisponderanno a diversi autovalori $\mathcal{E}_n^\alpha, \mathcal{E}_n^\beta, \dots$, come si può notare dalla formula (4.154).

Un'**applicazione** particolarmente interessante si ottiene partendo dall'Hamiltoniana H_0 dell'atomo di idrogeno, e perturbandola introducendo un campo elettrico uniforme lungo \hat{z} :

$$H = H_0 + \lambda V = H_0 + \lambda e E z$$

dove ci interessa il caso per $\lambda = 1$.

Il sistema così descritto consente di analizzare il cosiddetto **effetto Stark**⁴⁵.

Effetto Stark

Detti $|n, l, m\rangle$ gli autostati conosciuti di H_0 (ignoriamo lo spin degli elettroni), concentriamoci sul caso (semplice) con $n = 2$, per cui avremo 4 orbitali degeneri, dato che (ignorando la struttura fine) l'energia dipende solo da n :

$2s^0$	$2p^{-1}$	$2p^0$	$2p^{+1}$
$ 2, 0, 0\rangle$	$ 2, 1, -1\rangle$	$ 2, 1, 0\rangle$	$ 2, 1, 1\rangle$

Al primo ordine in λ , gli autovalori dell'Hamiltoniana H si ottengono dalla (4.154):

$$\mathcal{E}_n^\alpha = \mathcal{E}_n^0 + \lambda \langle \mathcal{E}_n^0, \alpha | V | \mathcal{E}_n^0, \alpha \rangle \quad (4.156)$$

⁴⁵Fonti utili: <http://bit.ly/2SsPHYD>, <http://bit.ly/2D5teHj>

dove per $|\mathcal{E}_n^0, \alpha\rangle$ dobbiamo scegliere la base in cui V è diagonale, ossia tale che:

$$\langle \mathcal{E}_n^0, \alpha | V | \mathcal{E}_n^0, \beta \rangle = 0 \quad \forall \alpha \neq \beta$$

Nel caso degli autostati dell'idrogeno, tale relazione **non** è soddisfatta dalla base di autoket $|n, l, m\rangle$ da cui siamo partiti, cioè:

$$eE \langle n, l, m | z | n, l', m' \rangle \neq 0 \quad \forall l \neq l', m \neq m'$$

Infatti, calcolando gli elementi di matrice del termine evidenziato non si ottiene una matrice diagonale. Dovremo quindi diagonalizzarla e considerare la base dei suoi autovettori, e solo alla fine applicare la formula delle perturbazioni.

Svolgiamo allora i conti. Fissando l , osserviamo che:

$$\langle n, l, m | z | n, l, m' \rangle = 0 \quad \forall m \neq m'$$

Infatti gli m sono gli autovalori di L_3 , con cui z commuta, dato che L_3 genera le rotazioni attorno a \hat{z} : $[z, L_3] = 0$.

Ciò si nota esplicitamente passando in coordinate polari:

$$z = r \cos \theta; \quad L_3 = -i\hbar \frac{\partial}{\partial \varphi} \Rightarrow [z, L_3] = 0$$

Perciò, per $m \neq m'$ a l fissato:

$$\begin{aligned} 0 &= \langle n, l, m | \overbrace{[z, L_3]}^{=0} | n, l, m' \rangle = \\ &= \langle n, l, m | z L_3 | n, l, m' \rangle - \langle n, l, m | L_3 z | n, l, m' \rangle = \\ &= m' \langle n, l, m | z | n, l, m' \rangle - m \langle n, l, m | z | n, l, m' \rangle = \\ &= \underbrace{(m - m')}_{\neq 0} \langle n, l, m | z | n, l, m' \rangle \Rightarrow \langle n, l, m | z | n, l, m' \rangle = 0 \end{aligned}$$

Notiamo poi che anche gli elementi diagonali (ossia con $m = m'$ e $l = l'$) della matrice sono nulli, poiché le autofunzioni $|n, l, m\rangle$ sono **simmetriche**: sono infatti date da polinomi di Legendre, che sono o *pari* o *dispari*⁴⁶. Se calcoliamo allora $\langle n, l, m | z | n, l, m \rangle$ in rappresentazione $\{\vec{x}\}$:

$$\langle n, l, m | z | n, l, m \rangle = \int_{\mathbb{R}} z \psi^*(x, y, z) \psi(x, y, z)$$

dove $\psi(x, y, z)$ è o pari o dispari. Il prodotto di due funzioni pari è pari, e anche quello di due funzioni dispari, e perciò $\psi^* \psi$ è sempre pari. Moltiplicando allora per z , che è dispari, si ha un'integrazione di una funzione dispari ($z \psi^* \psi$) su un dominio simmetrico rispetto all'origine, che è perciò nulla.

La stessa considerazione, usando solo la notazione di Dirac, si ottiene trasformando

⁴⁶ Ciò deriva da un fatto generale, per cui in caso di un potenziale simmetrico - come quello Coulombiano per l'idrogeno - è sempre possibile scegliere una base di autofunzioni costituita da autofunzioni simmetriche, ossia sempre o pari o dispari.

per parità l'operatore z , con $\mathcal{P}z\mathcal{P} = -z$. Se $|n, l, m\rangle$ è pari avremo poi $\langle n, l, m| \mathcal{P} = +1 \langle n, l, m|$, mentre se è dispari $\langle n, l, m| \mathcal{P} = -1 \langle n, l, m|$, con relazioni analoghe per i ket. Perciò:

$$\begin{aligned} \langle n, l, m| \mathcal{P}z\mathcal{P} |n, l, m\rangle &= (\pm 1)^2 \langle n, l, m| z |n, l, m\rangle \\ - \langle n, l, m| z |n, l, m\rangle &= \langle n, l, m| z |n, l, m\rangle \Rightarrow \langle n, l, m| z |n, l, m\rangle = 0 \end{aligned}$$

Gli unici elementi di matrice non nulli, per cui $n = 2$, $m = m'$ e $l \neq l'$ sono quindi:

$$\langle 2, 1, 0| z |2, 0, 0\rangle, \langle 2, 0, 0| z |2, 1, 0\rangle$$

In forma matriciale, prendendo come base dell'autospazio $\mathcal{H}_{n=2}$ (in ordine) $\{|2, 0, 0\rangle, |2, 1, -1\rangle, |2, 1, 0\rangle, |2, 1, 1\rangle\}$, $\langle n, l, m| z |n, l', m'\rangle$ è dato da:

$$\begin{pmatrix} \langle 2, 0, 0| z |2, 0, 0\rangle & \langle 2, 1, -1| z |2, 0, 0\rangle & \langle 2, 1, 0| z |2, 0, 0\rangle & \langle 2, 1, 1| z |2, 0, 0\rangle \\ \langle 2, 0, 0| z |2, 1, -1\rangle & \langle 2, 1, -1| z |2, 1, -1\rangle & \langle 2, 1, 0| z |2, 1, -1\rangle & \langle 2, 1, 1| z |2, 1, -1\rangle \\ \langle 2, 0, 0| z |2, 1, 0\rangle & \langle 2, 1, -1| z |2, 1, 0\rangle & \langle 2, 1, 0| z |2, 1, 0\rangle & \langle 2, 1, 1| z |2, 1, 0\rangle \\ \langle 2, 0, 0| z |2, 1, 1\rangle & \langle 2, 1, -1| z |2, 1, 1\rangle & \langle 2, 1, 0| z |2, 1, 1\rangle & \langle 2, 1, 1| z |2, 1, 1\rangle \end{pmatrix}$$

Solo i termini “doppiamente evidenziati” (in verde) sono non nulli:

$$\langle n, l, m| z |n, l', m'\rangle = \begin{pmatrix} 0 & 0 & \langle 2, 1, 0| z |2, 0, 0\rangle & 0 \\ 0 & 0 & 0 & 0 \\ \langle 2, 0, 0| z |2, 1, 0\rangle & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Procediamo con i conti, usando la forma esplicita per le autofunzioni $|n, l, m\rangle$ vista in (4.107). Passando allora in rappresentazione $\{r, \theta, \varphi\}$ in $\mathcal{H} = L^2(\mathbb{R}_+, r^2 dr) \otimes L^2(S^2, d\Omega)$ calcoliamo il prodotto scalare $\langle 2, 0, 0| z |2, 1, 0\rangle$.

Gli autostati che ci interessano sono, esplicitamente:

$$\begin{aligned} \psi_{2,0,0}(r, \theta, \varphi) &= \frac{1}{\sqrt{8a^3}} \exp\left(-\frac{r}{2a}\right) L_1^1\left(\frac{r}{a}\right) Y_0^0(\theta, \varphi) \\ \psi_{2,1,0}(r, \theta, \varphi) &= \frac{1}{\sqrt{8a^3}} \frac{1}{\sqrt{3}} \left(\frac{r}{a}\right) \exp\left(-\frac{r}{2a}\right) L_0^3\left(\frac{r}{a}\right) Y_1^0(\theta, \varphi) \end{aligned}$$

con le espressioni per le armoniche sferiche e i polinomi generalizzati di Laguerre dati da:

$$\begin{aligned} Y_0^0(\theta, \varphi) &= \frac{1}{\sqrt{4\pi}} & Y_1^0(\theta, \varphi) &= \sqrt{\frac{3}{4\pi}} \cos \theta \\ L_0^a(x) &= 1 & L_1^a(x) &= 1 + a - x \end{aligned}$$

Con $z = r \cos \theta$ l'integrale diviene:

$$\begin{aligned} \langle 2, 0, 0 | z | 2, 1, 0 \rangle &= \int_0^\infty r^2 dr \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\varphi \psi_{2,0,0}(r, \theta, \varphi) z \psi_{2,1,0}(r, \theta, \varphi) = \\ &= \frac{1}{32\pi a^3} \int_0^\infty dr \frac{r^4}{a} \left(2 - \frac{r}{a}\right) \exp\left(-\frac{r}{a}\right) \underbrace{\int_0^\pi \cos^2(\theta) \sin(\theta) d\theta}_{2/3} \underbrace{\int_0^{2\pi} d\varphi}_{2\pi} = \\ &\stackrel{(a)}{=} \frac{a}{24} \int_0^\infty dy y^4 (2 - y) e^{-y} = -3a \end{aligned}$$

dove in (a) si è effettuata la sostituzione $y = r/a$, da cui $dr = a dy$.

Si ricava poi $\langle 2, 1, 0 | z | 2, 0, 0 \rangle = (\langle 2, 0, 0 | z | 2, 1, 0 \rangle)^*$: in questo caso i due coincidono in quanto puramente reali.

Reintroducendo il fattore eE del potenziale, la matrice completa $\langle n, l, m | V | n, l, m \rangle$, nella base di $\mathcal{H}_{n=2}$ (in ordine) $\{|2, 0, 0\rangle, |2, 1, -1\rangle, |2, 1, 0\rangle, |2, 1, 1\rangle\}$, diviene:

$$\langle n, l, m | V | n, l, m \rangle = \begin{pmatrix} 0 & 0 & -3aeE & 0 \\ 0 & 0 & 0 & 0 \\ -3aeE & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Per poter applicare la teoria perturbativa, è necessario passare nella base in cui $\langle n, l, m | V | n, l, m \rangle$ è diagonale - che certamente esiste perché V è simmetrico, e quindi i suoi elementi di matrice formano una matrice simmetrica e diagonalizzabile.

In questo caso troviamo come autovalori 0 con molteplicità 2, e $\lambda_\pm = \pm 3aeE$. I rispettivi autovettori, scritti in termini della base, sono dati da:

$$\begin{aligned} |u_0, 1\rangle &= |2, 1, -1\rangle; \quad |u_0, 2\rangle = |2, 1, 1\rangle; & \lambda &= 0 \\ |u_0, 3\rangle &= |u_+\rangle = \frac{1}{\sqrt{2}}(|2, 0, 0\rangle - |2, 1, 0\rangle); & \lambda_+ &= +3aeE \\ |u_0, 4\rangle &= |u_-\rangle = \frac{1}{\sqrt{2}}(|2, 0, 0\rangle + |2, 1, 0\rangle); & \lambda_- &= -3aeE \end{aligned}$$

Possiamo ora applicare la teoria perturbativa. Al primo ordine, da (4.156), ponendo $\lambda = 1$:

$$\mathcal{E}_n^\alpha = \mathcal{E}_n^0 + \langle \mathcal{E}_n^0, \alpha | V | \mathcal{E}_n^0, \alpha \rangle$$

dove i $|\mathcal{E}_n^0, \alpha\rangle = |u_0, \alpha\rangle$ sono gli autovettori appena calcolati, che finalmente rispettano $\langle \mathcal{E}_n^0, \alpha | V | \mathcal{E}_n^0, \beta \rangle = 0$ per $\alpha \neq \beta$. α è poi l'indice che scorre sulla degenerazione (in H_0) dei $|\mathcal{E}_n^0, \alpha\rangle$ di autovalore (in H_0) pari a \mathcal{E}_n^0 , e che in questo caso va da 1 a 4. Svolgendo i conti:

$$\begin{aligned} \mathcal{E}_n^{\alpha=1} &= \mathcal{E}_n^{\alpha=2} = \mathcal{E}_n^0 \\ \mathcal{E}_n^{\alpha=3} &= \mathcal{E}_n^0 + 3aeE \\ \mathcal{E}_n^{\alpha=4} &= \mathcal{E}_n^0 - 3aeE \end{aligned}$$

Perciò per $n = 2$ si ha che gli stati che prima erano degeneri, tutti e quattro con energia \mathcal{E}_n^0 , “si splittano” e assumono tre energie diverse, come visibile in figura 4.30.

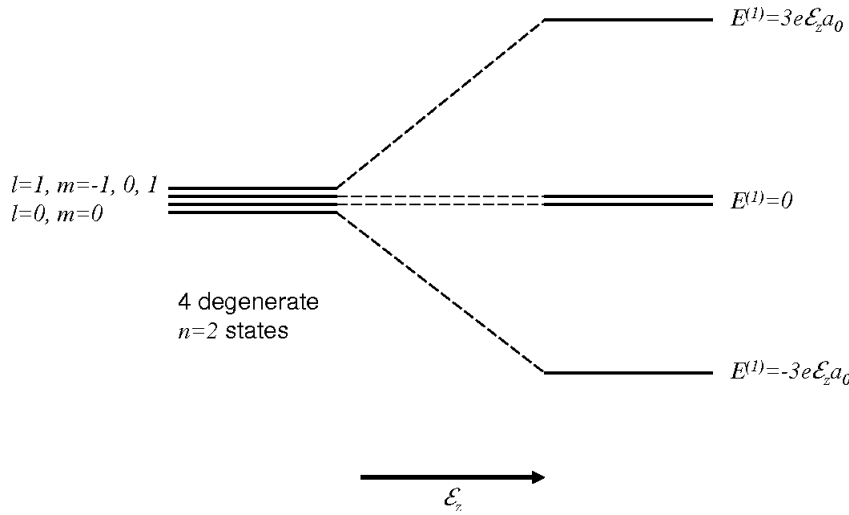


Figura 4.30: Effetto Stark per il livello $n = 2$ dell'idrogeno: in presenza di un campo elettrico esterno, stati che prima erano degeneri non lo sono più.

Le conclusioni ottenute tramite la teoria perturbativa al primo ordine sono vicine a quanto si ottiene sperimentalmente, ma sono parziali.

Oppenheimer, infatti, osservò che, sovrapponendo il potenziale elettrico con quello *effettivo* coulombiano, per esempio esaminando il diagramma dell'energia in una sezione parallela all'asse \hat{z} come in figura 4.31, si presenta la possibilità che un elettrone “di core”, ossia *legato* “in profondità” all'atomo di idrogeno, per effetto tunnel si muova in una regione *non legata*, e cioè a spettro continuo. L'esistenza di tale effetto, perciò, invaliderebbe l'esistenza stessa di stati legati per tale sistema. Tale osservazione, teoricamente valida, non è predetta né dalla teoria perturbativa (anche ad ordini maggiori), né è osservata sperimentalmente. Cerchiamo di capirne il perché.

L'osservazione di Oppenheimer

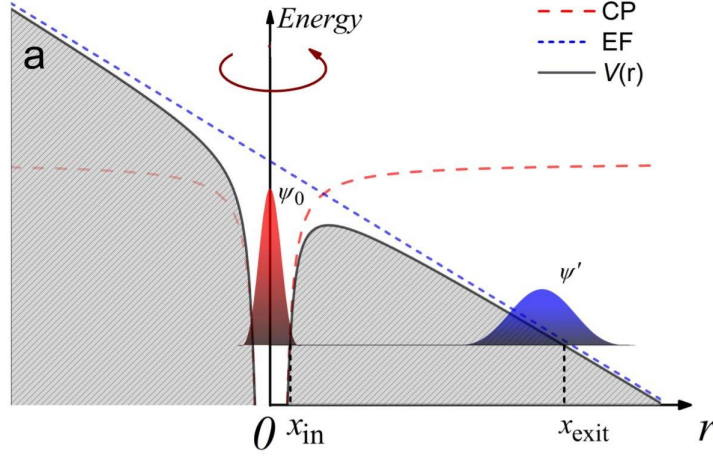


Figura 4.31: Il potenziale centrale (CP) coulombiano, sovrapposto a quello (EF) di un campo elettrico uniforme produce un $V(r)$, che, sufficientemente lontano dal nucleo, è più basso dell'energia degli elettroni negli stati *legati* interni, che quindi hanno la possibilità di uscire per effetto tunnel.

Per una particella di massa m ed energia \mathcal{E} che urta contro una barriera di potenziale \bar{V} larga L la probabilità di oltrepassarla per effetto tunnel è dell'ordine ricavato in (3.134):

$$\Pi \sim \exp\left(\frac{-2m\sqrt{\bar{V} - \mathcal{E}}}{\hbar} 2L\right)$$

Stimiamo l'energia di un elettrone in uno stato legato nell'atomo di idrogeno (immerso nel campo elettrico) usando l'espressione per le *energie di Bohr* (4.104):

$$\mathcal{E}_n^0 \sim -\frac{me^4}{2\hbar^2} \frac{1}{n^2}$$

Poiché la buca di potenziale del nucleo di idrogeno è molto stretta, per un campo elettrico sufficientemente basso è considerabile *simmetrica*, per cui la sua altezza $\bar{V} \sim 0$. Otteniamo quindi la stima per Π , con $n = 2$ (stato “profondamente legato”):

$$\Pi \sim \exp\left(-\frac{2L}{\hbar} \sqrt{\frac{2m^2e^4}{4\hbar^2}}\right) = \exp\left(-L \frac{me^2}{\hbar^2}\right) = \exp\left(-\frac{L}{a}\right) \quad a = \frac{\hbar^2}{me^2}$$

dove a è il raggio di Bohr.

Sempre poiché la *buca* del potenziale coulombiano è molto stretta, stimiamo L come l'ascissa dell'intersezione tra il potenziale $y = V(x)$ e $y = -|\mathcal{E}|$ (che corrisponde a x_{exit} in figura 4.31, dove consideriamo $x_{in} \approx 0$). Otteniamo allora:

$$-eEz - \frac{e^2}{z} = -|\mathcal{E}| \Rightarrow L = \frac{|\mathcal{E}| + \sqrt{|\mathcal{E}|^2 - 4e^3E}}{2eE} \underset{E \ll 1}{\sim} \frac{|\mathcal{E}|}{eE}$$

Da cui otteniamo:

$$\Pi \sim \exp\left(-\frac{1}{a} \frac{|\mathcal{E}|}{eE}\right) \quad (4.157)$$

che ha una singolarità essenziale per $E = 0$, e quindi non è qui espandibile in una serie di Mac-Laurin come si era supposto per applicare la teoria perturbativa. Ecco allora perché l'effetto di *ionizzazione tunnel* suggerito da Oppenheimer risulta “invisibile” agli sviluppi perturbativi.

Del resto, sperimentalmente, il fenomeno non è osservato in quanto estremamente raro per campi elettrici “ragionevoli”.

Proviamo infatti ad inserire qualche numero in (4.157). Prendendo l'energia dello stato fondamentale $\mathcal{E}/e = -13.6\text{V}$, $a = 0.5 \times 10^{11}\text{m}$ e per $E = 10^5\text{V/m}$ (generato per esempio in un condensatore piano con lastre distanti 1m e con una $\Delta V = 100\text{kV}$ tra di esse), si ha:

$$\Pi \sim 10^{-6 \cdot 10^6}$$

che è infinitamente piccolo (per confronto, le particelle nell'intero universo osservabile sono stimate attorno a 10^{75}). Perciò, seppur matematicamente lo spettro di un atomo di idrogeno in campo elettrico sia continuo, nella pratica ne osserviamo solo uno spettro discreto - e l'effetto di *ionizzazione tunnel* è rilevabile solo usando appositi apparati, con campi elettrici estremamente intensi, per cui però la teoria perturbativa è inapplicabile in partenza.

(Lezione 38 del
17/12/2018)

4.18 Esercizio 8

Si consideri un *oscillatore armonico* unidimensionale di massa m e frequenza angolare ω . All'istante $t = 0$ l'oscillatore si trova nello stato descritto dalla funzione d'onda ψ_α , con $\alpha \in \mathbb{C} \setminus \{0\}$, determinata univocamente (a meno della normalizzazione) dall'essere autostato dell'operatore a :

$$a = \frac{1}{\sqrt{2}}(X' + iP') \quad X' = \sqrt{\frac{m\omega}{\hbar}}X, \quad P' = \sqrt{\frac{1}{m\omega\hbar}}P$$

di autovalore α :

$$a|\psi_\alpha\rangle = \alpha|\psi_\alpha\rangle$$

Nota: i $|\psi_\alpha\rangle$ sono detti *stati coerenti*, e hanno applicazioni nei laser.

1. Si determini se ψ_α è autostato dell'Hamiltoniana H .
2. Si determinino i valori medi di X , P e H in $|\psi_\alpha\rangle$.
3. Si determini $(\Delta X)_{\psi_\alpha}(\Delta P)_{\psi_\alpha}$

4. Si dimostri che l'evoluto temporale di $|\psi_\alpha(t)\rangle$ è ancora autostato di a , cioè:

$$a |\psi_\alpha(t)\rangle = \alpha(t) |\psi_\alpha(t)\rangle$$

e si determini $\alpha(t)$.

5. Si esprima $|\psi_\alpha\rangle$ in funzione degli autostati di H *non normalizzati* $|n\rangle \equiv (a^\dagger)^n |0\rangle$, dove $a |0\rangle = 0$.

4.18.1 Soluzione

1. Gli autostati di H per un oscillatore armonico sono le $|\psi_n\rangle$ (che in rappresentazione $\{x\}$ corrispondono alle funzioni di Hermite). Per rispondere alla domanda basta allora verificare quale $|\psi_n\rangle$ è anche autostato di a con operatore α , e cioè corrisponde a $|\psi_\alpha\rangle$.
Dalla teoria sappiamo che a è l'operatore di distruzione (o annichilazione) che “abbassa” gli autovettori:

$$a |\psi_n\rangle = \sqrt{n} |\psi_{n-1}\rangle$$

Perciò l'unica $|\psi_n\rangle$ che è anche autostato di a è “quella che non può essere ulteriormente abbassata”, e cioè $|0\rangle$, con autovalore $\alpha = 0$. Tuttavia $|0\rangle$ non può essere $|\psi_\alpha\rangle$, poiché si è espressamente richiesto $\alpha \in \mathbb{C} \setminus \{0\}$. Perciò $|\psi_\alpha\rangle$ non è autostato di H .

Nota: chiedere $\alpha \in \mathbb{C} \setminus \{0\}$, e non in $\mathbb{R} \setminus \{0\}$ è sensato, poiché a non è un operatore hermitiano, e perciò i suoi autovalori non è detto che siano reali.

2. Partiamo da X e scriviamolo in termini di operatori di creazione e distruzione a e a^\dagger :

$$X = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger) \quad (4.158)$$

In tal modo possiamo usare le relazioni (duali) dell'equazione agli autovalori:

$$a |\psi_\alpha\rangle = \alpha |\psi_\alpha\rangle \leftrightarrow \langle\psi_\alpha| a^\dagger = \alpha^* \langle\psi_\alpha| \quad (4.159)$$

per calcolare il valor medio in $|\psi_\alpha\rangle$:

$$\begin{aligned} \langle\psi_\alpha| X |\psi_\alpha\rangle &= \sqrt{\frac{\hbar}{2m\omega}} \langle\psi_\alpha| (a + a^\dagger) |\psi_\alpha\rangle = \\ &= \sqrt{\frac{\hbar}{2m\omega}} (\langle\psi_\alpha| a |\psi_\alpha\rangle + \langle\psi_\alpha| a^\dagger |\psi_\alpha\rangle) = \\ &= \sqrt{\frac{\hbar}{2m\omega}} (\alpha \langle\psi_\alpha| \psi_\alpha\rangle + \alpha^* \langle\psi_\alpha| \psi_\alpha\rangle) = \sqrt{\frac{\hbar}{2m\omega}} (\alpha + \alpha^*) \end{aligned}$$

Facciamo la stessa cosa per P e H , dove:

$$P = \sqrt{\frac{m\hbar\omega}{2}} \frac{1}{i} (a - a^\dagger) \quad H = a^\dagger a + \frac{1}{2} \quad (4.160)$$

e i valori medi sono dati da:

$$\begin{aligned}\langle \psi_\alpha | P | \psi_\alpha \rangle &= \sqrt{\frac{m\omega\hbar}{2}} \langle \psi_\alpha | \left(\frac{a - a^\dagger}{i} \right) | \psi_\alpha \rangle = \sqrt{\frac{m\omega\hbar}{2}} \frac{\alpha - \alpha^*}{i} \\ \langle \psi_\alpha | H | \psi_\alpha \rangle &= \hbar\omega \langle \psi_\alpha | a^\dagger a + \frac{1}{2} | \psi_\alpha \rangle = \hbar\omega \left(|\alpha|^2 + \frac{1}{2} \right)\end{aligned}$$

Nota: per $a^\dagger a$ si è applicato a^\dagger al bra a sinistra, e a al ket a destra:

$$\langle \psi_\alpha | a^\dagger a | \psi_\alpha \rangle = \alpha^* \langle \psi_\alpha | \alpha | \psi_\alpha \rangle = \alpha^* \alpha \langle \psi_\alpha | \psi_\alpha \rangle$$

a e a^\dagger non hanno autovettori comuni

Ciò si può fare solo in questo caso, in presenza di sia bra e ket di ψ_α . Se avessimo $a^\dagger a | \psi_\alpha \rangle$ il risultato sarebbe molto più complesso, dato che a e a^\dagger **non** hanno autovettori comuni non nulli: per di più a^\dagger proprio non ne ha (e non è tenuto ad averli, dato che - come a - non è hermitiano). Lo si dimostra tramite i commutatori. Sia ipoteticamente $|\psi\rangle$ autovettore comune di a e a^\dagger , per cui:

$$a |\psi\rangle = \alpha |\psi\rangle \quad a^\dagger |\psi\rangle = \beta |\psi\rangle$$

Sappiamo dalla teoria che $[a, a^\dagger] = 1$. Se lo applichiamo a $|\psi\rangle$:

$$\begin{aligned}[a, a^\dagger] |\psi\rangle &= (aa^\dagger - a^\dagger a) |\psi\rangle = (\beta\alpha - \alpha\beta) |\psi\rangle = 0 \\ [a, a^\dagger] |\psi\rangle &= \mathbb{I} |\psi\rangle = |\psi\rangle\end{aligned}$$

e perciò l'unico autovettore comune è di nuovo $|\psi\rangle = |0\rangle$.

3. Analogamente al punto precedente, partiamo calcolando X^2 in termini di a e a^\dagger ricordando (4.158):

$$\begin{aligned}X^2 &= \frac{\hbar}{2m\omega} (a + a^\dagger)(a + a^\dagger) = \frac{\hbar}{2m\omega} (a^2 + a^{\dagger 2} + a^\dagger a + aa^\dagger) = \\ &\stackrel{(a)}{=} \frac{\hbar}{2m\omega} (a^2 + a^{\dagger 2} + 2a^\dagger a + 1)\end{aligned}$$

dove in (a) abbiamo riscritto $aa^\dagger = [a, a^\dagger] + a^\dagger a = 1 + a^\dagger a$ in modo da poter usare le relazioni per i bra e i ket di $|\psi_\alpha\rangle$ viste in (4.159) nel calcolo del valor medio:

$$\langle \psi_\alpha | X^2 | \psi_\alpha \rangle = \frac{\hbar}{2m\omega} (\alpha^2 + \alpha^{*2} + 2|\alpha|^2 + 1)$$

Da cui si ricava la fluttuazione cercata usando il valor medio calcolato al punto precedente:

$$\begin{aligned}(\Delta X)_{\psi_\alpha}^2 &= \langle X^2 \rangle_{\psi_\alpha} - \langle X \rangle_{\psi_\alpha}^2 = \\ &= \frac{\hbar}{2m\omega} (\alpha^2 + \alpha^{*2} + 2|\alpha|^2 + 1 - (\alpha + \alpha^*)^2) = \frac{\hbar}{2m\omega}\end{aligned}$$

Seguiamo la stessa procedura per calcolare $(\Delta P)_{\psi_\alpha}$, ricordando (4.160):

$$\begin{aligned}
P^2 &= \frac{m\omega\hbar}{2} \left(\frac{a - a^\dagger}{2} \right)^2 = -\frac{m\omega\hbar}{2} (a^2 + a^{\dagger 2} - \textcolor{yellow}{aa^\dagger} - a^\dagger a) = \\
&\stackrel{(a)}{=} -\frac{m\omega\hbar}{2} (a^2 + a^{\dagger 2} - 2a^\dagger a - 1) \\
\langle \psi_\alpha | P^2 | \psi_\alpha \rangle &= -\frac{m\omega\hbar}{2} (\alpha^2 + \alpha^{*2} - 2|\alpha|^2 - 1) \\
(\Delta P)_{\psi_\alpha}^2 &= \langle P^2 \rangle_{\psi_\alpha} - \langle P \rangle_{\psi_\alpha}^2 = \\
&= \frac{m\omega\hbar}{2} \left(\alpha^2 + \alpha^{*2} - 2|\alpha|^2 - 1 - \left(\frac{\alpha - \alpha^*}{1} \right)^2 \right) = \frac{m\omega\hbar}{2}
\end{aligned}$$

Da cui calcoliamo infine il prodotto cercato:

$$(\Delta X)_{\psi_\alpha} (\Delta P)_{\psi_\alpha} = \sqrt{\frac{\hbar}{2m\omega}} \frac{m\omega\hbar}{2} = \frac{\hbar}{2}$$

Notiamo allora che sugli *stati coerenti* dell'oscillatore armonico il principio di Heisenberg è realizzato al suo *minimo*.

4. Lo stato $|\psi_\alpha(t)\rangle$ al tempo t si ottiene dalla formula dell'evoluzione temporale unitaria:

$$|\psi_\alpha(t)\rangle \equiv \exp\left(-\frac{itH}{\hbar}\right) |\psi_\alpha\rangle$$

Valutare l'esponenziale risulta in tal caso molto difficile, poiché non è banale scrivere $|\psi_\alpha\rangle$ come combinazione di autostati di H $|\psi_n\rangle$, e anche facendolo si ottiene una serie infinita difficile da gestire.

Un procedimento *più furbo* sta nell'utilizzare la formula di Hadamard, che è applicabile per espandere un'esponenziale nei casi più generali. Dati due operatori A e B si ha infatti:

$$\begin{aligned}
e^A B e^{-A} &= B + [A, B] + \frac{1}{2!} [A, [A, B]] + \dots = \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \underbrace{[A, \dots, [A, B]]}_{n \text{ volte}} \underbrace{\dots}_{n \text{ volte}}
\end{aligned} \tag{4.161}$$

dove non è neanche richiesto che A sia hermitiano: e infatti, in questo caso, $A = a$ non lo è.

Cerchiamo allora di ricondurci alla forma di (4.161). Supponiamo che $|\psi_\alpha(t)\rangle$ rimanga autovettore di a , con un autovalore $\alpha(t)$ che in principio dipenderà dal tempo:

$$a |\psi_\alpha(t)\rangle = \alpha(t) |\psi_\alpha(t)\rangle$$

Se troviamo $\alpha(t)$ avremo dimostrato che lo è. Esplicitando allora $|\psi_\alpha(t)\rangle$ giungiamo a:

$$a \exp\left(-\frac{itH}{\hbar}\right) |\psi_\alpha\rangle = \alpha(t) \exp\left(-\frac{itH}{\hbar}\right) |\psi_\alpha\rangle$$

Basta ora moltiplicare a sinistra entrambi i membri per $\exp(itH/\hbar)$, e sfruttare il fatto che $\alpha(t)$ è scalare per semplificare gli esponenziali a destra. Otteniamo così una forma a cui applicare la (4.161):

$$\exp\left(\frac{itH}{\hbar}\right) a \exp\left(-\frac{itH}{\hbar}\right) |\psi_\alpha\rangle = \alpha(t) |\psi_\alpha\rangle$$

Per calcolarla, usiamo la formula:

Scrivendo $B = H$ in termini di a e a^\dagger per facilitare il calcolo dei commutatori:

$$H = \hbar\omega \left(a^\dagger a + \frac{1}{2}\right)$$

Giungiamo a:

$$\begin{aligned} \exp\left(\frac{itH}{\hbar}\right) a \exp\left(-\frac{itH}{\hbar}\right) &= \exp\left(it\omega \left(a^\dagger a + \frac{1}{2}\right)\right) a \exp\left(-it\omega \left(a^\dagger a + \frac{1}{2}\right)\right) = \\ &= a + it\omega \left[a^\dagger a + \frac{1}{2}, a\right] + \dots = \\ &= a - it\omega a + \dots = a \sum_n \frac{(-i)^n (t\omega)^n}{n!} = e^{-it\omega} a \end{aligned}$$

E sostituendo nell'espressione di sopra, troviamo:

$$e^{-it\omega} a |\psi_\alpha\rangle = e^{-it\omega} \alpha |\psi_\alpha\rangle = \alpha(t) |\psi_\alpha\rangle$$

da cui $\alpha(t) = \alpha e^{-i\omega t}$, che verifica l'ipotesi che abbiamo fatto in partenza. Abbiamo allora dimostrato che $|\psi_\alpha(t)\rangle$ è ancora autovettore di a .

5. Gli autostati $|n\rangle = (a^\dagger)^n |0\rangle$ costituiscono, a meno della normalizzazione, gli autostati di H , “costruiti” a partire dallo stato fondamentale $|0\rangle$ (ripercorrendo quanto fatto nell'analisi teorica dell'oscillatore armonico), e quindi costituiscono una base ortogonale.

Vogliamo determinare i coefficienti c_n dell'espansione di $|\psi_\alpha\rangle$ in questa base:

$$|\psi_\alpha\rangle = \sum_{n=0}^{+\infty} c_n |n\rangle$$

Applicando a ad entrambi i membri:

$$a |\psi_\alpha\rangle = \sum_{n=0}^{+\infty} c_n a |n\rangle = \sum_{n=1}^{+\infty} c_n \sqrt{n} |n-1\rangle \quad (4.162)$$

si ha che a “abbassa” gli autovettori dell'espansione. D'altro canto, poiché $|\psi_\alpha\rangle$ è α -autovalore di a :

$$a |\psi_\alpha\rangle = \alpha |\psi_\alpha\rangle = \alpha \sum_n c_n |n\rangle \quad (4.163)$$

Uguagliando (4.162) e (4.163):

$$\alpha \sum_{n=0}^{+\infty} c_n |n\rangle = \sum_{n=1}^{+\infty} c_n \sqrt{n} |n-1\rangle$$

Shiftiamo l'indice al secondo membro per riportare le due sommatorie "in pari":

$$\alpha \sum_{n=0}^{+\infty} c_n |n\rangle = \sum_{n=0}^{+\infty} \sqrt{n+1} c_{n+1} |n\rangle$$

L'uguaglianza deve valere per i singoli $|n\rangle$, e quindi:

$$\alpha c_n = \sqrt{n+1} c_{n+1} \Rightarrow c_{n+1} = \frac{\alpha}{\sqrt{n+1}} c_n$$

che è la relazione di ricorsione che definisce la successione dei coefficienti c_n che stiamo cercando. Reiterandola n volte:

$$c_n = \frac{\alpha}{\sqrt{n}} c_{n-1} = \frac{\alpha^2}{\sqrt{n(n-1)}} c_{n-2} = \cdots = \frac{\alpha^n}{\sqrt{n!}} c_0$$

dove c_0 è individuato dalla normalizzazione. Fermandoci qui, abbiamo trovato l'espansione:

$$|\psi_\alpha\rangle = c_0 \sum_{n=0}^{+\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle$$

Imponendo la normalizzazione (cosa non richiesta nell'esercizio):

$$\sum_{n=0}^{+\infty} |c_n|^2 \stackrel{!}{=} 1 \Rightarrow \sum_{n=0}^{+\infty} \frac{(|\alpha|^2)^n}{n!} |c_0|^2 = 1$$

in cui riconosciamo lo sviluppo di un'esponenziale:

$$e^{|\alpha|^2} |c_0|^2 = 1 \Rightarrow |c_0|^2 = e^{-|\alpha|^2} \Rightarrow c_0 = (e^{-|\alpha|^2})^{1/2} = \exp\left(-\frac{|\alpha|^2}{2}\right)$$

Perciò, lo sviluppo *normalizzato* degli stati coerenti nella base (normalizzata) $|n\rangle$ degli autostati di H dell'oscillatore armonico è:

$$|\psi_\alpha\rangle = \sum_{n=0}^{+\infty} c_n |n\rangle = \exp\left(-\frac{|\alpha|^2}{2}\right) \sum_{n=0}^{+\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle$$

Alternativamente, un modo (più teorico e generale) per risolvere lo stesso punto è quello riportato qui di seguito.

Partiamo dal *suggerimento implicito*:

$$a |\psi_\alpha\rangle = \alpha |\psi_\alpha\rangle \Rightarrow (a - \alpha) |\psi_\alpha\rangle = 0$$

Allora:

$$\exists S \text{ t.c. } SaS^{-1} = a - \alpha \Rightarrow SaS^{-1}S|0\rangle = 0 \Rightarrow S|0\rangle = \psi_\alpha$$

Da cui:

$$e^{\alpha a^\dagger} a e^{-\alpha a^\dagger} = a + [\alpha a^\dagger, a] + \dots$$

e ricaviamo:

$$|\psi_\alpha\rangle = e^{\alpha a^\dagger} |0\rangle = \sum_{n=0}^{\infty} \frac{1}{n!} \alpha^n (a^\dagger)^n |0\rangle = \sum_{n=0}^{\infty} \frac{1}{n!} \alpha^n |n\rangle$$

(Lezione 39 del
19/12/2018)

4.19 Paradosso Einstein-Podolski-Rosen (1935)

Disclaimer. Ci occuperemo in questa parte degli aspetti *paradossali* e *filosofici* della MQ, che potrebbero far dubitare della sua validità o completezza. Si tenga conto del fatto che finora **nessun esperimento** contraddice la MQ, su cui anzi si basano gran parte delle tecnologie più recenti.

4.19.1 Realismo e controfattualità

Prima di dedicarci ad analizzare gli aspetti *fisici* del paradosso EPR, è necessario dare qualche definizione di carattere filosofico, esplicitando le “assunzioni inconse” che vengono fatte nell’analisi scientifica. Solo in questo modo sarà possibile apprezzare appieno la natura *paradossale*, nel senso etimologico di “contro l’esperienza comune”, dell’esperimento proposto da Einstein-Podolsky-Rosen nel 1935.

Con **realismo scientifico** intendiamo la corrente filosofica, generalmente condivisa nella nostra società, che la conoscenza scientifica sia **rigorosa**, **oggettiva** e trovi **corrispondenza** con la realtà tramite esperimenti, e che la scienza ci dia informazioni sulle cose che esistono, non solo sugli oggetti di cui abbiamo percezione diretta tramite i sensi, ma anche su oggetti che non percepiamo direttamente ma compaiono nelle teorie, come *atomi*, *elettroni*, etc. (che, in principio, non sono “osservabili con gli occhi”).

*Realismo
scientifico*

All’opposto, l’**antirealismo scientifico** (o *strumentalismo*) è la corrente filosofica che ritiene che le teorie scientifiche siano solo *strumenti per predire i fenomeni*, ma non per descrivere il mondo come realmente è. Anzi, il mondo esterno potrebbe “non esistere”, o comunque non avremmo alcun accesso certo ad esso⁴⁷.

Strumentalismo

Realismo e strumentalismo sono due visioni *metafisiche*, nel senso che riguardano principalmente l’*attitudine* di uno scienziato nel discutere una teoria. Per un realista gli *elettroni* “esistono”, cioè “sono veri”, e hanno le proprietà (carica, massa,

⁴⁷ ∧ CFR l’esperimento mentale del “brain in a vat”

spin, etc.) date dalla teoria. Ogni nuova teoria fisica, perciò, non è altro che un'immagine "più dettagliata" di come l'universo è realmente. Per uno strumentalista, gli "elettroni" sono solo un utile concetto - un nome da associare a formule matematiche che consentono di spiegare con precisione i fenomeni che percepiamo. Si può dire che un realista sia concentrato su come l'universo è, indipendentemente da ogni cosa, mentre uno strumentalista, notando che possiamo percepire ogni cosa solo attraverso i sensi, e quindi da una prospettiva decisamente *biased*, si limita a studiare le *interazioni* tra universo e mente umana. Per un realista un albero che cade in mezzo ad una foresta senza che nessuno lo senta *fa rumore*, mentre per uno strumentalista non ha senso porsi il problema dell'esito di una misura senza che essa sia stata svolta.

Se realismo e strumentalismo sono essenzialmente materia di preferenza personale, non è detto che i *framework* da esse generati siano equivalenti.

In particolare, un realista tende - quasi inconsciamente - ad assumere una proprietà *epistemologica* (cioè legata alla natura del conoscere) detta **definitezza controfattuale**⁴⁸ (CFD). L'idea della CFD è che la realtà "esista" mentre non la si osserva, e che porsi domande su esiti di misure *non svolte* di "proprietà reali" abbia senso. Se la Luna è *reale*, allora continua ad *esistere* anche se nessuno può vederla perché un cataclisma ha sterminato l'umanità intera.

Definitezza
controfattuale

Per applicare la definizione di CFD è necessario definire un *modo* per riconoscere gli elementi della **realtà oggettiva**, che sono in principio distinti dai **concetti fisici** utilizzati da una teoria scientifica per spiegarli. A tal proposito, ci accontenteremo della seguente **condizione sufficiente**: *se senza disturbare in nessun modo un sistema possiamo prevedere con certezza (cioè con probabilità unitaria) il valore di una quantità fisica A, allora esiste un elemento di realtà fisica corrispondente a questa quantità fisica*.

Elemento di
realtà oggettiva

Accettando la CFD, possiamo aggiungere che A, essendo reale, avrebbe avuto lo stesso valore anche se la misura non fosse stata eseguita.

In altre parole, per la CFD le *proprietà reali* di un sistema (date per esempio dalla condizione di sopra) hanno *un unico valore ben definito* in ogni istante: per esempio la massa di un elettrone è *definita* anche se non la stiamo misurando.

Il problema del paradosso EPR sta proprio nel fatto che la CFD, ingenuamente accettata in ottica realista e perfettamente compatibile con la MC, non è rispettata in MQ. Tale inaspettata conseguenza è stata effettivamente verificata sperimentalmente con ottima precisione⁴⁹.

4.19.2 I principi del paradosso EPR

Il paradosso EPR sorge combinando due principi *ragionevoli* e comunemente assunti:

⁴⁸^In inglese: counterfactual definiteness

⁴⁹^<https://www.nature.com/articles/nature15759>

- **Principio di realtà \mathcal{R} :** Se, senza intervenire su un dato sistema, è possibile prevedere con certezza (cioè con probabilità unitaria) il valore di una grandezza fisica, a questa corrisponde una **proprietà oggettiva** del sistema, cioè una proprietà indipendente da eventuali osservatori esterni che (CFD) avrebbe lo stesso valore anche se la misura non fosse stata effettuata.

Per esempio, se in MC misuriamo (q, p) , sappiamo che la particella avrebbe avuto posizione e momento (q, p) anche se non fossero stati misurati, poiché si può assumere *idealmente* che il processo di misura *non disturbi* il sistema. In MQ ciò è applicabile *solo agli autostati*. Ad esempio, supponiamo che lo stato di spin sia dato da $|+\rangle$ autostato di S_z di autovalore $+\hbar/2$, allora diremo che l'elettrone “ha” spin $\hbar/2$, indipendentemente dal fatto che lo misuriamo o meno, poiché questa è una *proprietà oggettiva* di quell'elettrone.

Nota: Già in questa prima accezione notiamo come la MQ differisca dalla MC, dato che nella prima possiamo applicare l'idea di *proprietà oggettiva* esclusivamente agli *autostati*, e non a tutti gli stati.

- **Principio di località** (alla Einstein) \mathcal{L} .

Località

Dati due sistemi fisici e supposto che durante un certo intervallo di tempo Δt essi rimangano *isolati* tra loro (cosa realizzata in relatività se la distanza tra di essi è più grande di $c\Delta t$), allora l'evoluzione delle proprietà fisiche (determinate dagli “elementi di realtà”) di uno di essi durante tale intervallo di tempo non può essere influenzata da operazioni eseguite sull'altro.

In parole povere, il principio di località afferma che ogni oggetto è influenzato direttamente solo dalle sue immediate vicinanze. L'idea è che ogni interazione abbia una velocità *finita* (che è $\leq c$ per la relatività), e perciò perturbare un oggetto A non può influenzare istantaneamente un altro oggetto B posto ad una certa distanza da esso. Per esempio, se il Sole esplodesse, la Terra non subirebbe alcun effetto per circa 8 minuti, poiché tale è il tempo impiegato dalla luce per percorrere la distanza Sole-Terra, e per \mathcal{L} nessuna interazione può propagarsi più velocemente di c . Analogamente, spingere su una ipotetica barra lunga 1 anno luce produce un'onda di compressione su di essa che si propaga alla velocità del suono, che è sicuramente $< c$, e perciò l'effetto della spinta giunge all'altro estremo non istantaneamente.

La località è molto importante sia per garantire che la causalità sia rispettata (se la velocità di ogni segnale è limitata è automaticamente impossibile una *macchina del tempo* che comunica con il passato), che per *semplificare* la trattazione dei sistemi fisici. Nel corso di un esperimento, infatti, vorremmo dover considerare solo un numero finito di corpi che possono interagire con l'apparato in questione, e non rassegnarci a dover tener conto di *tutto l'universo* - cosa che ha il potenziale di negare buona parte delle capacità predittive della fisica.

4.19.3 Versione classica di EPR

Verifichiamo la compatibilità di $\mathcal{L} + \mathcal{R}$ in MC con un sistema *analogo classico* dell'EPR.

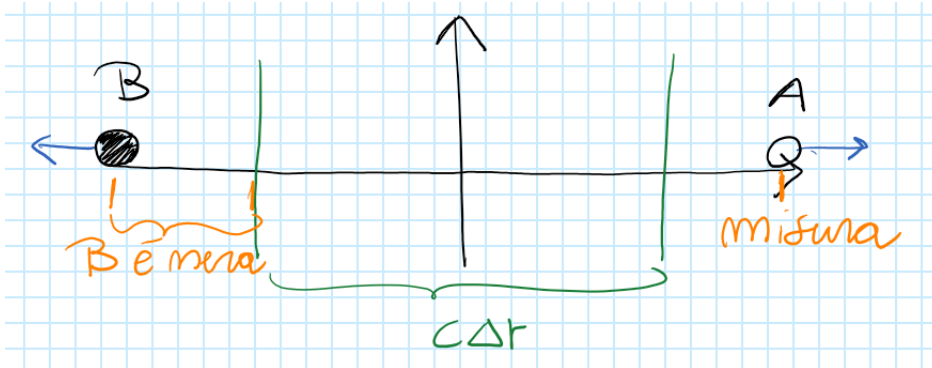


Figura 4.32: Schema dell'analogo classico del paradosso EPR

Consideriamo un insieme di coppie di palle classiche che si muovono una verso destra (palle A) e una verso sinistra (palle B), e di cui sappiamo che una è bianca e l'altra è nera, ma non sappiamo *quale*, e la distribuzione dei colori è casuale.

A e B partono dall'origine del sdr del laboratorio a $t = 0$. Per semplicità, consideriamo che A e B possano interagire fino ad un certo tempo T , dopodiché, per $t > T$, data la loro distanza, siano considerabili indipendenti (ossia si trovino a distanza di tipo *spazio* relativamente alla scala temporale Δt dell'esperimento). Per esempio potremmo aspettare che A e B si trovino a 1 anno luce di distanza, per cui ogni modifica fatta su A possa influenzare B solo dopo un anno, ossia molto dopo la conclusione dell'esperimento.

A $T + t_0$, dove t_0 è dell'ordine di Δt (e quindi $\ll 1$ anno), osserviamo A , che supponiamo essere bianca. Poiché sappiamo che, per ipotesi, osservando A e B simultaneamente (rispetto al sdr del laboratorio), una è nera e l'altra è bianca, deduciamo che B a $T + t_0$ deve essere **nera**.

Abbiamo allora ottenuto una misura del colore di B senza perturbarla in alcun modo. Per \mathcal{R} si ha che il *colore* è una proprietà oggettiva di B , intesa come *reale*. Segue allora, per la CFD, che B deve avere un colore *definito* ad ogni istante: dato che stiamo parlando di un qualcosa di "reale", non può assumere proprietà contraddittorie! Per $t < T + t_0$ abbiamo allora due possibilità mutualmente esclusive:

- B era bianca
- B era nera

Il principio di località \mathcal{L} ci porta a scartare la prima. Se B fosse stata bianca prima di $T_0 + t$, allora significherebbe che la misura di A ha modificato una proprietà fisica di B , nonostante A e B fossero *separate* e *non interagenti*, ma questo è escluso da \mathcal{L} . Deve allora essere che, per $t < T + t_0$, B era nera.

In realtà, prima di T ammettiamo la possibilità di interazioni tra A e B , che

possono modificarne i rispettivi stati. Perciò, da $\mathcal{R} + \mathcal{L}$ deduciamo che, data la misura di A bianca a $T + t_0$, B deve essere nera almeno per $T < t < T + t_0$, dove T è il tempo dopo il quale A e B sono considerabili *non interagenti* nella scala temporale dell'esperimento.

Tutto ciò non stupisce. L'indeterminatezza iniziale è di “tipo classico”, nel senso che deriva da un'ignoranza dello sperimentatore: lo stato iniziale del sistema non era infatti di conoscenza massimale. Dopo la misura di A , e la relativa deduzione per B tramite $\mathcal{R} + \mathcal{L}$, la nostra conoscenza del sistema si è accresciuta, compatibilmente con una **realtà preesistente**.

Possiamo perciò dire che la MC è compatibile con il **realismo locale** ($\mathcal{R} + \mathcal{L}$).

4.19.4 Versione quantistica di EPR

Passiamo ora alla versione quantistica dell'esperimento, in cui al posto di palle consideriamo elettroni, e invece del colore lo *spin*.

Sia dato un numero $N \gg 1$ di particelle di spin 0 che decadono a $t = 0$ in coppie di particelle di spin 1/2 in uno stato *entangled* di *singoletto*, ossia tale che:

$$S^{tot} = 0 \quad S_z^{tot} = 0$$

Che deriva dalla *conservazione dello spin* (analoga all'usuale conservazione del momento angolare).

Perciò le N coppie di elettroni assumono a $t = 0$ lo stato puro:

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

dove con $|+\rangle$ e $|-\rangle$ denotiamo gli autostati, rispettivamente della particella A e B , dell'operatore S_z .

Consideriamo una di queste coppie, dove indichiamo con A l'elettrone emesso verso destra, e con B quello verso sinistra.

Ad un certo istante $t_0 > 0$, i supporti delle funzioni d'onda spaziali dei due elettroni si trovano a una distanza tale da far sì che non possano interagire tra loro senza violare \mathcal{L} .

Dopodiché a $t_1 > t_0$ l'osservatore A misura lo spin S_z^A . Dato che gli autostati di A sono una combinazione *in parti uguali* di $|+\rangle_A$ e $|-\rangle_A$, A osserverà per circa metà delle coppie uno spin $\hbar/2$, e per un'altra metà uno spin $-\hbar/2$.

Nel primo caso, per il postulato di proiezione di Von Neumann, lo stato di spin delle coppie dopo la misura diventa:

$$|\psi'\rangle = (|+\rangle_A \langle +|_A \otimes \mathbb{I}_B) |\psi\rangle = \frac{1}{\sqrt{2}} |+\rangle_A |-\rangle_B$$

che, una volta normalizzato, corrisponde allo stato puro:

$$|\psi'\rangle = |+\rangle_A |-\rangle_B$$

Siamo allora *certi* che se facessimo una misura di S_z^B troveremmo $-\hbar/2$, visto che $|\psi'\rangle$ è *autostato* di S_z^B . Sappiamo che la misura su A non può aver disturbato B . Infatti, per ipotesi di *separabilità*⁵⁰ di A e B , nel senso che i due elettroni hanno “realità distinte” e non sono un unico oggetto, ha senso parlare di una misura solo su uno di essi (A) e per località ogni perturbazione non può propagarsi istantaneamente. Si verifica allora la *condizione sufficiente* di \mathcal{R} , da cui segue che lo spin è una *proprietà oggettiva* di B .

Per \mathcal{L} , la misura appena fatta non può aver prodotto una proprietà di B , che quindi doveva essere già così, almeno dal momento in cui A e B non potevano più interagire. Perciò, per $t_0 < t \leq t_1$ sappiamo che B aveva uno **spin definito**, e pari a $-\hbar/2$.

Si ha perciò che la coppia considerata si trova nello stato $|\psi\rangle$ per $t \in [0, t_0]$, e in $|\psi'\rangle$ per $t \in [t_0, t_1]$.

In generale, ripetendo l'esperimento più volte, si avrà che l'esito della misura a t_1 , da cui deduciamo per $\mathcal{R} + \mathcal{L}$ lo stato $|\psi'\rangle$ per $t \in [t_0, t_1]$, sarà equamente $+\hbar/2$ o $-\hbar/2$. Ciò significa che, per molti esperimenti, lo stato per $t \in [t_0, t_1]$ sarà una “miscela” delle due possibilità $|\psi'_+\rangle = |+\rangle_A |-\rangle_B$ e $|\psi'_-\rangle = |-\rangle_A |+\rangle_B$, ciascuna con probabilità $1/2$, ossia uno **stato misto** ρ :

$$\begin{aligned}\rho &= \frac{1}{2} |\psi'_+\rangle \langle \psi'_+| + \frac{1}{2} |\psi'_-\rangle \langle \psi'_-| = \\ &= \frac{1}{2} (|+\rangle_A \langle +| \otimes |-\rangle_B \langle -|) + \frac{1}{2} (|-\rangle_A \langle -| \otimes |+\rangle_B \langle +|)\end{aligned}$$

Il problema sorge osservando che, per $t \in [0, t_1]$, la misura non è stata ancora effettuata (lo sarà a t_1), e perciò lo stato del sistema *non è collassato* per proiezione di von Neumann, e deve essere $|\psi(t)\rangle$ dato dall'evoluzione unitaria:

$$|\psi(t)\rangle = \exp\left(-\frac{i}{\hbar}tH\right) |\psi\rangle$$

Si dimostra che l'evoluzione unitaria non può portare uno stato inizialmente puro $|\psi\rangle$ ad uno stato misto ρ .

Infatti, partendo da uno stato puro generico $|\psi\rangle \langle \psi|$ e calcolandone l'evoluto temporale ($U |\psi\rangle = |\psi(t)\rangle$) si ottiene sempre uno stato puro:

$$U |\psi\rangle \langle \psi| U^\dagger = |\psi(t)\rangle \langle \psi(t)|$$

Perciò affermare (grazie a $\mathcal{R} + \mathcal{L}$) che per $t \in [t_0, t_1]$ lo stato della coppia è ρ è incompatibile con l'evoluzione unitaria dello stato iniziale $|\psi\rangle$, che prevederebbe, sempre per $t \in [t_0, t_1]$ uno stato $|\psi(t)\rangle$ fondamentalmente diverso.

Si ha infatti che ρ e $|\psi(t)\rangle$ producono valori medi differenti per misure di correlazione di osservabili che non sono diagonali nella base $\{|+\rangle, |-\rangle\}$ degli autostati di S_z utilizzata per definire $|\psi(t)\rangle$ e ρ .

⁵⁰Questa ipotesi era contenuta implicitamente nel primo enunciato di \mathcal{L} , e la esplicitiamo qui per completezza. Si noti che, senza separabilità, non sarebbe possibile parlare di località.

Possiamo usare, per esempio, S_x , dato che $[S_x, S_z] \neq 0$ e perciò sicuramente S_x non è diagonale nella base degli autoket di S_z . In particolare, l'autoket di autovalore $\hbar/2$ di S_x è dato da:

$$|+\rangle_x = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) \quad (4.164)$$

Una misura di correlazione di S_x consiste nel misurare *contemporaneamente* S_x su entrambe le particelle A e B al tempo $t \in [t_0, t_1]$.

Con una scelta opportuna della fase iniziale, $|\psi(t)\rangle = |\psi\rangle$, che riscriviamo come:

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|+\rangle_A |-\rangle_B - |-\rangle_A |+\rangle_B \right) = \frac{1}{\sqrt{2}} (|\psi_1\rangle - |\psi_2\rangle) \quad (4.165)$$

Denotiamo con P_+ il proiettore sull'autospazio di $S_x^A \otimes S_x^B$ di autovalore $\hbar/2, \hbar/2$ dato da:

$$\begin{aligned} P_+ &= |+\rangle_x^A \langle +| \otimes |+\rangle_x^B \langle +| = \\ &\stackrel{(4.164)}{=} \left(\frac{|+\rangle_A + |-\rangle_A}{\sqrt{2}} \frac{|+\rangle_B + |-\rangle_B}{\sqrt{2}} \right) \otimes \left(\frac{|+\rangle_B + |-\rangle_B}{\sqrt{2}} \frac{|+\rangle_A + |-\rangle_A}{\sqrt{2}} \right) = \\ &= \frac{1}{2} [|+\rangle \langle +| + |-\rangle \langle -| + |+\rangle \langle -| + |-\rangle \langle +|] \otimes \frac{1}{2} [|+\rangle \langle +| + |-\rangle \langle -| + |+\rangle \langle -| + |-\rangle \langle +|] \end{aligned}$$

La probabilità cercata si ottiene allora da:

$$\begin{aligned} W_{|\psi\rangle}^{S_x^A S_x^B} &= \|P_+ |\psi\rangle\|^2 \stackrel{(4.165)}{=} \left\| P_+ \frac{1}{\sqrt{2}} (|\psi_1\rangle - |\psi_2\rangle) \right\|^2 = \\ &\stackrel{(a)}{=} \frac{1}{2} \|P_+ |\psi_1\rangle\|^2 + \frac{1}{2} \|P_+ |\psi_2\rangle\|^2 - \frac{1}{2} \langle \psi_2 | P_+ | \psi_1 \rangle - \frac{1}{2} \langle \psi_1 | P_+ | \psi_2 \rangle \quad (4.166) \end{aligned}$$

dove in (a) abbiamo sviluppato la norma, usando la linearità di P_+ .

Otteniamo:

$$\begin{aligned} \|P_+ |\psi_1\rangle\|^2 &= \left| \frac{1}{2} [|+\rangle_A + |-\rangle_A] \otimes \frac{1}{2} [|+\rangle_B + |-\rangle_B] \right|^2 = \frac{1}{4} (1+1) \frac{1}{4} (1+1) = \frac{1}{4} \\ \langle \psi_1 | P_+ | \psi_2 \rangle &= \langle \psi_1 | \left(\frac{1}{2} [|+\rangle_A + |-\rangle_A] \otimes \frac{1}{2} [|+\rangle_B + |-\rangle_B] \right) = \frac{1}{2} \frac{1}{2} = \frac{1}{4} \end{aligned}$$

Per simmetria (notando che P_+ non cambia per scambio di particelle) si ha $\|P_+ |\psi_2\rangle\|^2 = 1/4$ e $\langle \psi_2 | P_+ | \psi_1 \rangle = 1/4$. Inserendo i risultati in (4.166) giungiamo a:

$$W_{|\psi\rangle}^{S_x^A S_x^B} = \frac{1}{4} + \frac{1}{4} - \frac{1}{4} - \frac{1}{4} = 0$$

D'altro canto, se usassimo ρ come stato, si avrebbe:

$$W_{\rho}^{S_x^A S_x^B} = \text{Tr}(\rho P_+)$$

che corrisponde ad una *media* delle probabilità risultanti dalle misure sugli stati puri che compongono ρ , *pesata* dalle probabilità di tali stati. Perciò, poiché ρ corrisponde con $p = 1/2$ a $|\psi_1\rangle$, e con $p = 1/2$ a $|\psi_2\rangle$, avremo:

$$W_{\rho}^{S_x^A S_x^B} = \frac{1}{2} \|P_+ \psi_1\|^2 + \frac{1}{2} \|P_+ \psi_2\|^2 = \frac{1}{2} \frac{1}{4} + \frac{1}{2} \frac{1}{4} = \frac{1}{4}$$

Perciò ρ e $|\psi\rangle$ sono in contraddizione:

$$W_{|\psi\rangle}^{S_x^A S_x^B} \neq W_{\rho}^{S_x^A S_x^B}$$

La differenza tra le due probabilità sta nel fatto che in ρ non sono presenti i *prodotti misti*, cioè i termini di interferenza tra autoket diversi.

Per la visione realista, dopo t_0 A e B sono in uno stato definito, ossia in autoket (opposti) di S_z , che *non interagiscono* (mancano i termini di interferenza). Perciò misurare S_x per ciascuna produce due “lanci di moneta” completamente indipendenti, che produrranno lo stesso risultato (es. spin \uparrow) $1/4$ delle volte.

D’altro canto, accettare l’evoluzione unitaria comporta l’esistenza di un’interferenza *non locale* tra A e B . In altre parole, potremmo dire che “le due monete” che corrispondono alle misure di S_x *restano collegate*, e quindi la probabilità che una misura dia \uparrow dipende - per qualche effetto superluminale - dall’esito della misura dell’altra.

4.19.5 L’argomento di EPR

In realtà c’è una terza opzione, che permette di salvare la località.

Continuando con l’esempio precedente, sappiamo che la MQ prescrive, per determinare l’esito di una misura di S_x , una sorta di “lancio di moneta”, ossia una procedura indeterministica. Precisamente, possiamo dire che la funzione d’onda *non contiene* un risultato definito per una singola misura di S_x su un autostato di S_z . Se l’indeterminazione è fondamentale, allora sorge il problema di correlazioni non locali.

Tuttavia potremmo pensare che questa non sia l’immagine completa. Continuando l’analogia, la moneta potrebbe essere *truccata*: l’idea è che dopo t_0 , la misura di S_x per A si ottiene lanciando una moneta con due teste o due croci, e quella per B con la combinazione opposta. Poiché non sappiamo prima del lancio quale delle due monete truccate sia capitata ad A , a noi sembra di lanciare una moneta “fair”, e poiché B ha sempre il risultato opposto di A , ci chiediamo come le due monete, *distinte e lontane* possano “comunicare tra loro” per mantenere la correlazione. Se però, invero, il gioco era truccato fin dall’inizio⁵¹ la correlazione non implica alcuna interazione non locale.

Affermare questo, però, significa negare la *completezza* della MQ. Vediamo precisamente in che senso.

- **Completezza \mathcal{C} .** Se ogni elemento di **realtà oggettiva** (inteso come elemento che rispetta la condizione sufficiente riportata all’inizio) ha una controparte tra i **concetti** utilizzati da una teoria per spiegare la realtà, allora tale teoria è **completa**.

Nel paper del 1935, l’argomento di EPR si basa sulla premessa che una sola di queste due opzioni può essere vera ad un dato momento:

⁵¹ \wedge Truth is, the game was rigged from the start (cit).

1. La MQ è non completa
2. La MQ è completa, e perciò quantità fisiche che sono descritte da operatori non commutanti non possono essere *contemporaneamente reali*

L'implicazione in (2) si ricava immediatamente: con “elemento reale” intendiamo qualcosa di *definito*, con un valore preciso, e perciò in MQ gli elementi reali di un sistema nello stato $|\psi\rangle$ possono essere solo le osservabili di cui $|\psi\rangle$ è autostato. Ma due osservabili non compatibili non hanno una base di autoket comuni, e quindi non possono essere reali allo stesso tempo.

Consideriamo ora il sistema delle coppie di particelle con spin opposto visto nel precedente esempio. Si ha che una misura su A , secondo la MQ, cambia lo stato del sistema, con conseguenze anche su B . Se misuriamo S_z per A , allora S_z sarà “reale” anche per B (con valore opposto a quanto trovato in A), cioè B sarà in un autostato di S_z , mentre se misuriamo S_x su A , immediatamente sappiamo che B è in un autostato di S_x . Ma, per località, la misura su A non può aver cambiato “proprietà reali” di B , che perciò dovevano essere preesistenti. Segue che le due autofunzioni che può assumere B a seguito di misure di S_z o S_x in A devono avere la stessa realtà.

Ricaviamo allora che, assumendo (1) falsa, ossia che la MQ sia completa, risulta che anche (2) deve essere falsa, perchè per località è possibile che due osservabili che non commutano (come S_x e S_z) abbiano realtà simultanea. Ma almeno una tra (1) e (2) deve essere vera, e quindi l'assunzione di partenza è assurda. Perciò (1) è vera, e la MQ è incompleta.

In definitiva, EPR dimostra che, per la MQ, $\mathcal{C} + \mathcal{R} + \mathcal{L}$ sono **contraddittorie**.

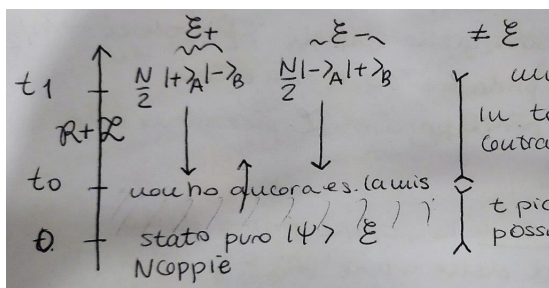


Figura 4.33: L'incompatibilità tra \mathcal{R} , \mathcal{L} e \mathcal{C}

4.19.6 Risolvere il paradosso

Abbiamo allora 3 possibilità “minimali” per risolvere la situazione, che corrispondono a rinunciare a uno dei tre principi: \mathcal{R} , \mathcal{L} o \mathcal{C} (trascuriamo, per ora, il caso *brutto* in cui dovremmo rinunciare a *più di un* principio).

- **Rinunciamo a \mathcal{R}** (ovvero a CFD) dichiarando che anche sapendo con certezza il valore che si otterrebbe misurando un'osservabile *senza* disturbare il sistema, a questo valore *non* si può attribuire il carattere di una proprietà fisica oggettiva indipendente dalla misura. Non possiamo in particolare estrapolare l'informazione prima della misura, anche conoscendo la dinamica del sistema. In tal modo non possiamo “portare indietro” il risultato ottenuto in t_1 , e quindi non si ha contraddizione in $[t_0, t_1]$.

L'abbandonare \mathcal{R} significa in sostanza sostenere che la funzione d'onda non è una proprietà del sistema fisico indipendente dall'osservazione, neppure per gli autostati, ma è un oggetto matematico che sintetizza un insieme massimale di informazioni sul sistema. La MQ è perciò un paradigma che consente di conoscere l'evoluzione temporale dell'informazione e di fare predizioni sui risultati (*strumentalismo* per gli stati ψ).

\mathcal{L} è allora preservato dal fatto che le osservazioni modificano le nostre *conoscenze sui sistemi*, ma non le loro proprietà *oggettive* (di cui ψ non fa più parte). Perciò a cambiare sono le predizioni, anche nel caso di conoscenza massimale, analogamente a quanto succede in MC quando si ha conoscenza non massimale. *Si potrebbe dire che la misura “crea” lo stato.*

- Rinunciamo a \mathcal{L} . Osserviamo che se ora applichiamo \mathcal{R} , lo stato diventa una *proprietà oggettiva*. Infatti $|\psi\rangle\langle\psi|$ è un'osservabile, e supponendo di produrre lo *stesso* sistema *più volte* nello stato $|\psi\rangle$ (incognito), misurando i suoi elementi di matrice possiamo determinarlo esattamente (nel caso in esame con 4×4 misure, effettuate con tecniche di *tomografia quantistica*) lasciando le altre coppie senza disturbo. Secondo \mathcal{R} senza \mathcal{L} una misura può modificare *istantaneamente* la funzione d'onda a distanze arbitrariamente grandi. Non si può perciò prolungare all'indietro l'informazione neppure se i sistemi sono a distanze di tipo spazio, dato che anche in tal caso potrebbero, in principio, influenzarsi a vicenda: in altre parole non esiste più un t_0 dopo il quale i due sistemi sono completamente non interagenti.

Fortunatamente si può dimostrare che, usando tali modifiche, non si possono trasmettere informazioni *più velocemente* di c . Infatti, poiché ogni misura trova in maniera *casuale* uno dei due risultati, B vede una sequenza casuale di $+$ e $-$, che senza l'informazione su A è *perfettamente compatibile*, ad esempio, con l'eventualità in cui su A non è stata fatta alcuna misura. Solo quando A e B *confrontano* le loro misure è possibile notare che ogni $+$ di A corrisponde a un $-$ di B , e quindi si nota la correlazione. Ma per avere l'informazione che rivela la correlazione a distanze di tipo spazio A deve inviare un'informazione a B , e tale comunicazione è per forza limitata da c .

Perciò, anche rinunciando a \mathcal{L} vale ancora la relatività, e la causalità è salva. Il realismo \mathcal{R} può essere mantenuto abbandonando il *riduzionismo*, che sostiene che un sistema è semplicemente la somma dei suoi costituenti. Senza \mathcal{L} , infatti, il risultato di un esperimento dipende non solo dai suoi costituenti, ma potenzialmente dallo stato di tutto l'universo, e la realtà diventa una *foresta di funzioni d'onda entangled*.

- Rinunciamo a \mathcal{C} (posizione sostenuta da Einstein). Il fatto che (per $\mathcal{R} + \mathcal{L}$),

dopo la misura di spin di A , gli spin di B “abbiano” individualmente valore $+\hbar/2$ o $-\hbar/2$ *senza* che siano stati disturbati può far supporre che tale proprietà l'avessero anche *prima* della misura (esattamente come il colore delle palle classiche), solo che la $|\psi\rangle$, che è la stessa per tutte le coppie prima della misura, semplicemente *non aveva* quest'informazione. *In altre parole, gli spin di A e B erano già definiti in precedenza, ma non erano conoscibili tramite $|\psi\rangle$, che quindi costituiva un'informazione incompleta del sistema.* Quindi la ψ dà solo informazioni probabilistiche come *media di variabili “nascoste”* che costituiscono le “vere” variabili reali.

Un'analogia è la temperatura in Meccanica Statistica, che è semplicemente una media su un numero molto grande di particelle, i cui stati non sono misurati individualmente.

(Lezione 40 del
20/12/2018)

4.19.7 Le disuguaglianze di Bell

Rinunciando a \mathcal{C} , supponiamo che esistano *variabili nascoste* che assumono valori λ , che assegnate al sistema (ma a noi sconosciute), determinino **univocamente** il valore delle osservabili. In altre parole, ogni particella *contiene dei parametri interni* che definiscono il suo comportamento a seguito di ogni misura. Il modello più semplice per una particella di spin $1/2$ si può ottenere immaginando che esista un vettore unitario $\vec{\lambda}$, associato alla particella, e tale che una misura di spin lungo una direzione \hat{u} sia generata da:

$$(\vec{S} \cdot \hat{u})(\vec{\lambda}) = \frac{\hbar}{2} \text{sgn}(\vec{\lambda} \cdot \hat{u})$$

per cui se la misura avviene su un asse \hat{u} che “è circa diretto come” $\vec{\lambda}$, allora si otterrà uno spin $\hbar/2$, mentre in caso fosse “circa diretto in maniera opposta” a $\vec{\lambda}$ si avrebbe $-\hbar/2$, con un'ambiguità nel caso $\vec{\lambda} \cdot \hat{u} = 0$, che però avviene per un set di punti con misura nulla.

Così facendo è possibile mantenere il **realismo locale**: le particelle hanno sempre stati definiti, e l'indeterminazione nasce dal fatto che i parametri da cui essi dipendono non sono conoscibili.

Notiamo che è possibile scegliere λ in infiniti modi, e inventarsi infinite *regole* che colleghino le variabili nascoste ai risultati sperimentali. Ci chiediamo: tra tutte queste possibilità ne esiste almeno una che è compatibile, in tutto e per tutto, con quanto determinato sperimentalmente?

Poiché, almeno finora, nessun esperimento ha contraddetto la MQ, ci aspettiamo che le teorie a variabili nascoste possano replicare tutti i risultati della MQ: ma come esserne sicuri?

Tale problema, apparentemente insormontabile, fu risolto da un'osservazione *furba* fatta da Bell nel 1965, che consentì di creare un *esperimento* per testare direttamente tale eventualità.

Partiamo considerando una *generica teoria a variabili nascoste*, dove i “parametri nascosti” sono denotati con λ , e possono essere di tutti i tipi - possiamo pensare a λ come a un vettore di variabili continue/discrete con qualsivoglia spettro e relazione interna. Supponiamo che λ codifichi i risultati di ogni possibile misura. Perciò, conoscere λ significa conoscere il risultato $A(\lambda)$ di una misura di A (come avviene in MC).

Possiamo pensare che, ottenuta una collezione di N particelle nello stato $|\psi\rangle$, i valori di λ differiscano per ogni particella, e che si distribuiscano⁵² secondo una certa $\rho_\psi(\lambda)$, che gode delle usuali proprietà di una distribuzione di probabilità:

$$\rho_\psi(\lambda) \geq 0 \quad \int_{\mathbb{R}} \rho_\psi(\lambda) d\lambda = 1 \quad (4.167)$$

In tal modo si spiega perché particelle nello stesso stato iniziale possano dare esiti diversi per date misure, per esempio nel caso $|\psi\rangle$ sia un autoket di S_z e si scelga di misurare S_x , dato che $[S_x, S_z] \neq 0$.

Consideriamo allora la misura che ha causato problemi in primo luogo, ossia la correlazione di spin misurati lungo assi diversi relativi ad una coppia di particelle a spin 1/2 nello stato di singoletto.

Esplicitamente, detti \hat{u} e \hat{u}' due versori, misuriamo il prodotto (correlazione) dello spin della particella A lungo \hat{u} , e della particella B lungo \hat{u}' . Se esiste una teoria alle variabili nascoste *valida*, allora esiste una scelta di $\rho(\lambda)$ che è compatibile con il risultato della MQ per ogni scelta dello stato iniziale $|\psi\rangle$:

$$\langle \psi | (\vec{S} \cdot \hat{u})_A \otimes (\vec{S} \cdot \hat{u}')_B | \psi \rangle \stackrel{?}{=} \int d\lambda \rho(\lambda) (\vec{S} \cdot \hat{u})_A(\lambda) (\vec{S} \cdot \hat{u}')_B(\lambda) \quad \forall |\psi\rangle$$

Il trucco sta nel modificare leggermente la misura, in modo che il risultato della teoria a variabili nascoste non cambi, mentre si ottengano previsioni diverse dalla MQ.

Consideriamo due sperimentatori, Alice (A) e Bob (B), che effettuano misure di spin in laboratori a distanza di tipo spazio per la durata dell'esperimento, per cui durante ciascuna misura nessun segnale può andare da A a B o viceversa. A seguito di N misure, Alice e Bob si ritrovano e confrontano i risultati ottenuti.

Alice possiede due *detector* di spin, orientati lungo \vec{u}_0 e \vec{u}_3 generici (e non uguali), mentre Bob può effettuare misure lungo \vec{u}_1 e \vec{u}_2 . Supponiamo per semplicità (ma senza perdita di generalità) che \vec{u}_i si trovino tutti sullo stesso piano, per esempio $\hat{x}\hat{z}$.

L'idea è quella di misurare la seguente quantità:

$$B(\{\vec{u}_i\}) = (\vec{S}_A \cdot \vec{u}_0) \otimes [\vec{S}_B \cdot \vec{u}_1 - \vec{S}_B \cdot \vec{u}_2] + (\vec{S}_A \cdot \vec{u}_3) \otimes [\vec{S}_B \cdot \vec{u}_1 + \vec{S}_B \cdot \vec{u}_2]$$

⁵²Questa è già un'ipotesi più debole rispetto alla CFD. Infatti, conoscere tutte le λ implica conoscere il risultato, con una $\rho(\lambda)$ che si riduce a una delta di Dirac. L'idea è che anche introducendo una sorgente di variabilità, una teoria alle variabili nascoste, come vedremo, non è compatibile con i risultati sperimentali.

che, se sapessimo le variabili nascoste λ , avrebbe un valore già determinabile:

$$B(\{\vec{u}_i\}, \lambda) = (\vec{S}_A \cdot \vec{u}_0)(\lambda)[(\vec{S}_B \cdot \vec{u}_1)(\lambda) - (\vec{S}_B \cdot \vec{u}_2)(\lambda)] + (\vec{S}_A \cdot \vec{u}_3)(\lambda)[(\vec{S}_B \cdot \vec{u}_1)(\lambda) + (\vec{S}_B \cdot \vec{u}_2)(\lambda)] \quad (4.168)$$

L'unica supposizione che facciamo su λ è quella della **località (di Bell)**, ossia che il valore di spin $(\vec{S}_A \cdot \vec{u}_i)(\lambda)$ misurato da Alice sia indipendente da quello $(\vec{S}_B \cdot \vec{u}_i)(\lambda)$ misurato da Bob, dato che le due misure avvengono a distanza di tipo spazio. Sappiamo, inoltre, che - per quante misure si facciano - i possibili risultati di una misura di spin lungo un asse qualsiasi sono sempre e solo due:

$$(\vec{S}_B \cdot \vec{u}_i)(\lambda) = \pm \frac{\hbar}{2} \quad (4.169)$$

Perciò, nell'ipotesi di località di Bell, i possibili risultati della misura di $B(\{\vec{u}_i\}, \lambda)$ nella teoria delle variabili nascoste, si ottengono *esaurendo* le possibilità. Ponendo $\hbar = 1$ da qui in poi, consideriamo:

- Se Bob ottiene due spin uguali, ossia:

$$(\vec{S}_B \cdot \vec{u}_1)(\lambda) = (\vec{S}_B \cdot \vec{u}_2)(\lambda)$$

Allora da (4.168) e (4.169) gli unici valori possibili di B sono:

$$B(\{\vec{u}_1\}, \lambda) = \pm \frac{1}{2}[0] \pm \frac{1}{2}[\pm 1] = \pm \frac{1}{2}$$

- Altrimenti, se Bob ha misurato spin diversi:

$$(\vec{S}_B \cdot \vec{u}_1)(\lambda) = -(\vec{S}_B \cdot \vec{u}_2)(\lambda)$$

Da (4.168) otteniamo lo stesso range:

$$B(\{\vec{u}_2\}, \lambda) = \pm \frac{1}{2}[\pm 1] \pm \frac{1}{2}[0] = \pm \frac{1}{2}$$

Unendo i due risultati, otteniamo la disuguaglianza:

$$-\frac{1}{2} \leq B(\{\vec{u}_i\}, \lambda) \leq \frac{1}{2} \quad (4.170)$$

Replicando $N \gg 1$ volte l'esperimento, otterremo un valor medio per $B(\{\vec{u}_i\}, \lambda)$ che dipenderà dalla forma specifica di $\rho(\lambda)$, e che quindi non possiamo conoscere a priori. L'unica cosa che possiamo fare è porre limiti su di esso: infatti da (4.170) e ricordando (4.167), deduciamo:

$$-\frac{1}{2} \leq \underbrace{\int d\rho_\psi(\lambda) B(\{\vec{u}_i\}, \lambda)}_{\langle B(\{\vec{u}_i\}) \rangle_\psi} \leq \frac{1}{2} \quad (4.171)$$

Il valore medio $\langle B(\{\vec{u}_i\}) \rangle_\psi$ è determinabile sperimentalmente, e perciò tale disuguaglianza è *verificabile*. Inoltre, poiché non abbiamo fatto alcuna ipotesi a riguardo, tali limiti valgono $\forall |\psi\rangle$.

Esaminiamo ora a cosa si giunge svolgendo invece i conti tramite la MQ, nel caso specifico di uno stato *massimamente correlato*, detto **stato di Bell**. Un esempio, nel caso degli spin che stiamo esaminando, è lo stato di singoletto $|\psi\rangle$:

$$|\psi\rangle = \frac{|+\rangle_A |-\rangle_B - |-\rangle_A |+\rangle_B}{\sqrt{2}}$$

dove $\{|+\rangle, |-\rangle\}$ sono gli autoket di S_z . Poiché vogliamo calcolare esiti di misure lungo direzioni generici, conviene passare alla base in direzione $\hat{u}(\theta, \varphi) = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$ generica. Partendo da (4.55):

$$\begin{aligned} |+\rangle_u &= \cos \frac{\theta}{2} e^{-i\varphi/2} |+\rangle + \sin \frac{\theta}{2} e^{+i\varphi/2} |-\rangle \\ |-\rangle_u &= -\sin \frac{\theta}{2} e^{-i\varphi/2} |+\rangle + \cos \frac{\theta}{2} e^{i\varphi/2} |-\rangle \end{aligned}$$

E invertendo⁵³:

$$\begin{aligned} |+\rangle &= \cos \frac{\theta}{2} e^{i\varphi/2} |+\rangle_u - \sin \frac{\theta}{2} e^{i\varphi/2} |-\rangle_u \\ |-\rangle &= \sin \frac{\theta}{2} e^{-\varphi/2} |+\rangle_u + \cos \frac{\theta}{2} e^{-i\varphi/2} |-\rangle_u \end{aligned}$$

Usando ora le equazioni agli autovalori:

$$(\vec{S} \cdot \hat{u}) |\pm\rangle_u = \pm \frac{1}{2} |\pm\rangle_u$$

Calcoliamo gli elementi di matrice nella base dove S_z è diagonale:

$$\begin{aligned} \langle + | (\vec{S} \cdot \hat{u}) | + \rangle &= \langle + | \left(+\frac{1}{2} \cos \frac{\theta}{2} e^{i\varphi/2} |+\rangle_u - \left(-\frac{1}{2}\right) \sin \frac{\theta}{2} e^{i\varphi/2} |-\rangle_u \right) = \\ &= \frac{1}{2} \cos^2 \frac{\theta}{2} - \frac{1}{2} \sin^2 \frac{\theta}{2} = \frac{1}{2} \cos \theta \\ \langle - | (\vec{S} \cdot \hat{u}) | - \rangle &= \langle - | \left(+\frac{1}{2} \sin \frac{\theta}{2} e^{-i\varphi/2} |+\rangle_u - \frac{1}{2} \cos \frac{\theta}{2} e^{-i\varphi/2} |-\rangle_u \right) = \\ &= \frac{1}{2} \sin^2 \frac{\theta}{2} - \frac{1}{2} \cos^2 \frac{\theta}{2} = -\frac{1}{2} \cos \theta \\ \langle + | (\vec{S} \cdot \hat{u}) | - \rangle &= \langle + | \left(\frac{1}{2} \sin \frac{\theta}{2} e^{-i\varphi/2} |+\rangle_u - \frac{1}{2} \cos \frac{\theta}{2} e^{-i\varphi/2} |-\rangle_u \right) = \\ &= \frac{1}{2} \cos \frac{\theta}{2} \sin \frac{\theta}{2} e^{-i\varphi} + \frac{1}{2} \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{-i\varphi} = \frac{1}{2} \sin \theta e^{-i\varphi} \\ \langle - | (\vec{S} \cdot \hat{u}) | + \rangle &= \frac{1}{2} \sin \theta e^{+i\varphi} \end{aligned}$$

⁵³ \wedge Basta sommare o sottrarre membro a membro dopo aver moltiplicato per $\sin \theta/2$ o $\cos \theta/2$

Sintetizzando:

$$\langle \pm | \vec{S} \cdot \vec{u} | \pm \rangle = \pm \frac{1}{2} \cos \theta \quad \langle \pm | \vec{S} \cdot \vec{u} | \mp \rangle = \frac{1}{2} \sin \theta e^{\mp i\varphi}$$

Usando i risultati appena ottenuti possiamo calcolare il valor medio di una singola misura di correlazione ad angoli generici, con $\hat{u}'(\theta', \varphi)$ (varia l'angolo θ , ma non φ):

$$\begin{aligned} \langle \psi | \overbrace{(\vec{S}_A \cdot \hat{u}) \otimes (\vec{S}_B \cdot \hat{u}')}^S | \psi \rangle &= \frac{1}{\sqrt{2}} ({}_A \langle + | {}_B \langle - | - {}_A \langle - | {}_B \langle + |) S \frac{1}{\sqrt{2}} (| + \rangle_A | - \rangle_B - | - \rangle_A | + \rangle_B) = \\ &= \frac{1}{2} [{}_A \langle + | {}_B \langle - | S | + \rangle_A | - \rangle_B - {}_A \langle + | {}_B \langle - | S | - \rangle_A | + \rangle_B + \\ &\quad - {}_A \langle - | {}_B \langle + | S | + \rangle_A | - \rangle_B + {}_A \langle - | {}_B \langle + | S | - \rangle_A | + \rangle_B] \end{aligned}$$

Per ciascun termine basta usare gli opportuni elementi di matrice già calcolati. Per esempio, per il primo caso:

$${}_A \langle + | {}_B \langle - | S | + \rangle_A | - \rangle_B = \left(\frac{1}{2} \cos \theta \right) \left(-\frac{1}{2} \cos \theta' \right)$$

abbiamo usato $\langle + | S_u | + \rangle$ per A , e $\langle - | S_{u'} | - \rangle$ per B . Svolgendo i conti:

$$\begin{aligned} \langle \psi | S | \psi \rangle &= \frac{1}{2} \left[\left(\frac{1}{2} \cos \theta \right) \left(-\frac{1}{2} \cos \theta' \right) - \left(\frac{1}{2} \sin \theta e^{-i\varphi} \right) \left(\frac{1}{2} \sin \theta' e^{i\varphi} \right) + \right. \\ &\quad \left. - \left(\frac{1}{2} \sin \theta e^{-i\varphi} \right) \left(\frac{1}{2} \sin \theta' e^{+i\varphi} \right) + \left(-\frac{1}{2} \cos \theta \right) \left(+\frac{1}{2} \cos \theta' \right) \right] = \\ &= -\frac{1}{4} (\cos \theta \cos \theta' - \sin \theta \sin \theta') = -\frac{1}{4} \cos(\theta - \theta') \end{aligned}$$

Sintetizzando:

$$\langle \psi | (\vec{S}_A \cdot \vec{u}) (\vec{S}_B \cdot \vec{u}') | \psi \rangle = -\frac{1}{4} \cos(\theta - \theta') \quad (4.172)$$

Possiamo finalmente calcolare il valor medio di $B(\{\vec{u}_i\})$ che ci interessa:

$$\begin{aligned} \langle \psi | B(\{\vec{u}_i\}) | \psi \rangle &= \langle \psi | \left[(\vec{S}_A \cdot \vec{u}_0) \otimes (\vec{S}_B \cdot \vec{u}_1) - (\vec{S}_A \cdot \vec{u}_0) \otimes (\vec{S}_B \cdot \vec{u}_2) + \right. \\ &\quad \left. + (\vec{S}_A \cdot \vec{u}_3) \otimes (\vec{S}_B \cdot \vec{u}_1) + (\vec{S}_A \cdot \vec{u}_3) \otimes (\vec{S}_B \cdot \vec{u}_2) \right] | \psi \rangle \end{aligned}$$

Detti α_i gli angoli (riferiti a $+\hat{x}$) che definiscono i \vec{u}_i unitari, possiamo applicare (4.172) per continuare i calcoli:

$$\begin{aligned} \langle \psi | B(\{\vec{u}_i\}) | \psi \rangle &= -\frac{1}{4} \left[\cos(\overbrace{\alpha_1 - \alpha_0}^{\theta_1}) - \cos(\overbrace{\alpha_2 - \alpha_0}^{\theta_2}) + \cos(\overbrace{\alpha_3 - \alpha_1}^{\theta_4}) + \cos(\overbrace{\alpha_2 - \alpha_3}^{\theta_3}) \right] = \\ &= -\frac{1}{4} [\cos \theta_1 - \cos \theta_2 + \cos \theta_4 + \cos \theta_3] \end{aligned}$$

Nota: le differenze degli α_i possono essere invertite senza cambiare il valore dei cos di cui sono argomento. Con questa scelta dei segni si ha che i θ_i definiti dalle differenze rispettano:

$$\theta_2 = \theta_1 + \theta_3 + \theta_4$$

in quanto:

$$(\alpha_2 - \alpha_0) = (\cancel{\alpha_1} - \alpha_0) + (\alpha_2 - \cancel{\alpha_3}) + (\cancel{\alpha_3} - \cancel{\alpha_1})$$

Per un'opportuna scelta degli angoli:

$$\begin{aligned}\theta_1 = \theta_3 = \theta_4 \equiv \theta \Rightarrow \theta_2 = 3\theta \\ \langle \psi | B(\{\vec{u}_i\}) | \psi \rangle = -\frac{1}{4}(3 \cos \theta - \cos 3\theta)\end{aligned}$$

Siamo interessati al *range* di valori che può assumere il valor medio appena calcolato. Cerchiamo quindi i punti critici derivando rispetto a θ :

Ottimizziamo rispetto a θ :

$$\frac{\partial}{\partial \theta} \langle \psi | B(\{\vec{u}_i\}) | \psi \rangle = 0 = \frac{1}{4}(3 \sin \theta - 3 \sin 3\theta) \Rightarrow \sin \theta = \sin(3\theta)$$

Tale equazione ha due classi di soluzioni, una per angoli uguali, con $\theta = 3\theta + 2k\pi \Rightarrow \theta = k\pi$, $k \in \mathbb{Z}$, e una per angoli supplementari $\theta = \pi - 3\theta + 2k\pi \Rightarrow \theta = (2k+1)\frac{\pi}{4}$. Data la periodicità basta studiare l'intervallo $\theta \in [-\pi, +\pi]$, e osservando che $B(\{\vec{u}_i\})$ è pari, ci limitiamo a $\theta \in [0, \pi]$. Si verifica che $B(\theta = 0, \pi) = \mp 1/2$, mentre per gli altri due angoli otteniamo un risultato interessante:

$$\begin{aligned}\langle \psi | B\left(\theta = \frac{\pi}{4}\right) | \psi \rangle &= -\frac{1}{4} \left(3 \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} \right) = -\frac{1}{\sqrt{2}} \\ \langle \psi | B\left(\theta = \frac{3\pi}{4}\right) | \psi \rangle &= -\frac{1}{4} \left(-3 \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} \right) = \frac{1}{\sqrt{2}}\end{aligned}$$

Perciò la media su N esperimenti avrà un valore compreso tra due limiti *più larghi* di quelli dell'equivalente con variabili nascoste (4.171):

$$-\frac{1}{\sqrt{2}} \leq \langle \psi | B(\{\vec{u}_i\}) | \psi \rangle \leq \frac{1}{\sqrt{2}} \quad (4.173)$$

Si ha quindi che MQ e teorie a variabili nascoste locali *non* sono compatibili. Sperimentalmente, l'esecuzione dell'esperimento (di Bell-Aspect) falsifica (4.171) a 40 deviazioni standard.

Perciò, non è possibile rinunciare a \mathcal{C} , dato che le teorie alternative non sono valide, e ne segue che la MQ è - allo stato attuale - una teoria completa.

Come soluzioni minimali all'EPR rimangono solo il rinunciare a \mathcal{R} o a \mathcal{L} : perciò o non vale la controfattualità o non vale il microgettivismo dei sottosistemi.

4.19.8 Il gatto di Schrödinger

Supponiamo, come ora si ritiene usualmente corretto, che tutta la realtà fisica sia descritta quantisticamente.

Possiamo allora considerare l'unione di **sistema misurato** e **apparato misuratore**, che costituisce a sua volta un **sistema isolato** (volendo, si può includere

anche l'ambiente in tale descrizione). Perciò, tale sistema dovrebbe evolvere deterministicamente, in modo unitario, come descritto dall'equazione di Schrödinger, con una Hamiltoniana come:

$$H = H_{\text{sistema}} + H_{\text{apparato}} + H_{\text{interazione s.a.}}$$

Consideriamo ad esempio l'esperimento di Stern-Gerlach per gli elettroni, in cui tramite dei magneti si misura lo spin di tali particelle.

Denotiamo con $|\psi_+\rangle$ e $|\psi_-\rangle$ funzioni d'onda del sistema misurato (ossia gli autostati di spin dell' e^- lungo \hat{z}), con $|0\rangle$ lo stato dell'apparato prima della misura, e con $|+\rangle$, $|-\rangle$ i possibili stati dell'apparato dopo la misura (della componente \hat{z} dello spin).

Se nel fascio iniziale abbiamo solo elettroni con spin $+1/2$, ossia nello stato $|\psi_+\rangle$, avremo solo deviazione verso l'alto, e quindi l'apparato - dopo la misura, sarà nello stato $|+\rangle$, che indica l'avvenuta rilevazione di spin verso l'alto:

$$|\psi_+\rangle \otimes |0\rangle \xrightarrow{\text{misura}} |\psi_+\rangle \otimes |+\rangle$$

D'altro canto, per elettroni tutti con spin $-1/2$, e quindi nello stato $|\psi_-\rangle$:

$$|\psi_-\rangle \otimes |0\rangle \xrightarrow{\text{misura}} |\psi_-\rangle \otimes |-\rangle$$

Per **linearità** dell'evoluzione (sistema isolato), possiamo unire i due risultati appena ottenuti per capire cosa succede partendo da elettroni che sono in una combinazione di $|\psi_+\rangle$ e $|\psi_-\rangle$. Da:

$$|\psi\rangle = \alpha_+ |\psi_+\rangle + \alpha_- |\psi_-\rangle$$

Giungiamo a:

$$|\psi\rangle \otimes |0\rangle \xrightarrow{\text{misura}} \alpha_+ |\psi_+\rangle |+\rangle + \alpha_- |\psi_-\rangle |-\rangle \quad (4.174)$$

Che questa situazione diventi paradossale lo intuirà Schrödinger aggiungendo all'apparato di rivelazione un *gatto*, chiuso in una scatola in cui si trova una boccetta di cianuro inizialmente sigillata. Se la posizione dell'elettrone sullo schermo è $|+\rangle$ scatta un martello che rompe la boccetta di cianuro, provocando la morte del gatto. Se invece la posizione è $|-\rangle$, il martello non scatta e il gatto rimane vivo.

Stiamo perciò correlando perfettamente $|+\rangle$ allo stato di $|\text{gatto morto}\rangle$, e $|-\rangle$ a quello di $|\text{gatto vivo}\rangle$. In tal caso, possiamo riscrivere la (4.174) come:

$$\alpha_+ |\psi_+\rangle |+\rangle |\text{gatto morto}\rangle + \alpha_- |\psi_-\rangle |-\rangle |\text{gatto vivo}\rangle \quad (4.175)$$

Aperto la scatola troveremo con probabilità $|\alpha_+|^2$ il gatto vivo, e con probabilità $|\alpha_-|^2$ morto.

Il problema è che prima dell'apertura della scatola, ma dopo l'esperimento, il gatto è potenzialmente vivo e morto come dice (4.175).

Cioè sotto l'assunzione di evoluzione unitaria, la sovrapposizione quantistica microscopica di $|\psi_+\rangle$ e $|\psi_-\rangle$ produce tramite *entanglement* una sovrapposizione macroscopica di gatto morto e gatto vivo!

Visto che, all'apertura della scatola, il risultato è per $|\alpha_+|^2$ gatto vivo e per $|\alpha_-|^2$ gatto morto e non abbiamo nessuna interferenza tra i due, potremmo essere tentati di affermare che lo stato dopo la misura, ma prima dell'apertura della scatola, sia uno stato misto, in cui gli stati di *gatto morto* e *gatto vivo* sono “pesati” dalle probabilità $|\alpha_+|^2$ e $|\alpha_-|^2$:

$$\rho = |\alpha_+|^2 |\psi_+\rangle \langle \psi_+| \otimes |\text{gatto morto}\rangle \langle \text{gatto morto}| + |\alpha_-|^2 |\psi_-\rangle \langle \psi_-| \otimes |\text{gatto vivo}\rangle \langle \text{gatto vivo}|$$

cioè il gatto è vivo o morto, ma lo sperimentatore non lo sa, e aprendo la scatola *acquisisce* un'informazione, come succede *in senso classico* quando la posizione di una particella è determinata precisamente all'interno di un range di possibilità. Dopo l'apertura, perciò, lo stato diventa “puro” e sappiamo se il gatto è vivo o è morto.

Ma ρ è uno stato misto, e si dimostra (teorema di Hepp) che nessuna evoluzione unitaria può portare da uno stato puro a uno stato misto. Perciò, tale soluzione “semplice” al paradosso non è percorribile.

In definitiva, non c'è una soluzione condivisa a tale paradosso, ma la MQ funziona benissimo ugualmente!

4.19.9 Riepilogo finale

(trascrizione delle diapositive disponibili qui: <http://www2.pd.infn.it/~march/IstFisTeoMQ.pdf>)

La fisica quantistica ci obbliga a rivedere alcune nozioni sulla “realtà fisica” che la nostra esperienza, formalizzata dalla fisica classica, ci aveva abituato a pensare come “naturali”.

Tali nuove proprietà sono:

1. **Indeterminismo:** Una conoscenza massimale del sistema ad un istante è data da un'onda di probabilità e non ci consente la previsione certa dei risultati di misure: la “realtà quantistica” in ambito predittivo è intrinsecamente indeterministica!

In MQ (interpretazione ortodossa) la struttura probabilistica è non epistemica (cioè non deriva da una nostra ignoranza). La natura ondulatoria genera, per esempio, il fenomeno dell'*effetto tunnel*.

Formalismo matematico: lo **stato** del sistema che caratterizza la massima informazione (*stato puro*) è descritto da un *raggio vettore* di uno spazio di Hilbert complesso separabile o equivalentemente da un proiettore in un suo sottospazio unidimensionale.

- Le quantità misurabili, le *osservabili*, sono descritte da *operatori autoaggiunti* nello spazio di Hilbert degli stati.
- I possibili risultati delle misure (*spettro*) sono descritti dagli autovalori (eventualmente generalizzati) di tali operatori.
- Due osservabili possono essere misurate simultaneamente senza disturbo reciproco se e solo se commutano. In tal caso si dicono *osservabili compatibili*.
- Un insieme completo (massimale) di osservabili compatibili (ICOC) definisce una *rappresentazione*, cioè un modo possibile di rappresentare lo stato tramite una funzione d'onda che dipende solo dagli autovalori dell'ICOC.

2. **Interferenza delle probabilità:** La probabilità di ottenere un risultato in una misura di un'osservabile è determinata dal modulo quadro della funzione d'onda del sistema (in una rappresentazione associata a un ICOC compatibile con tale osservabile). In caso di alternative non osservate, ma che sarebbero mutualmente esclusive se fossero osservate (come ad esempio, nel caso delle due fenditure, la possibilità dell'elettrone di passare da una o dall'altra fenditura) presenta il fenomeno di interferenza delle probabilità, incomprensibile anche con la probabilità classica (epistemica).

Formalismo matematico: La probabilità che misurando una osservabile A in uno stato puro ψ si ottenga un risultato in un insieme reale Δ è data dal valor medio in ψ del proiettore in Δ della famiglia spettrale di A : $(\psi, P^A(\Delta, \psi))$. Nel caso di più osservabili compatibili A_i si usa invece il valor medio del prodotto di $P^{A_i}(\Delta_j)$

3. **Osservazione = disturbo.** Lo stato di un sistema non osservato evolve deterministicamente secondo l'equazione di Schrödinger. Il sistema osservato si comporta in modo diverso da come si comporterebbe se non fosse osservato: per esempio, nel caso in cui è osservato l'interferenza delle probabilità scompare, la probabilità diventa epistemica, cioè l'informazione non è più massimale, ovvero il sistema dal punto di vista probabilistico dei risultati è descritto da uno stato misto.

Formalismo matematico: L'evoluzione deterministica dello stato ψ non osservato è descritta applicando a ψ la famiglia a un parametro di operatori unitari:

$$U(t) = \exp\left(-i\frac{Ht}{\hbar}\right)$$

- Uno stato *misto* quantistico può essere visto come una distribuzione di probabilità su un insieme numerabile di stati puri, ovvero una somma di tali stati puri rappresentati come proiettori unidimensionali pesati con le relative probabilità. Più in generale è descritto da una matrice densità ρ che è un operatore autoaggiunto positivo di traccia pari a 1.

- Dopo una misura ideale di un'osservabile A in uno stato puro ψ che ha dato un risultato in un insieme reale Δ lo stato è descritto da $P^A(\Delta)\psi$ (*proiezione di von Neumann*)
- Se il risultato della misura non è conosciuto, l'informazione sul sistema è descritta da uno stato misto ottenuto pesando con le relative probabilità gli stati puri corrispondenti ai possibili risultati.

4. **Indeterminazione** Non possiamo conoscere simultaneamente con precisione arbitraria tutte le grandezze fisiche; in particolare posizione e momento delle particelle. Non possiamo allora definire *traiettorie*. Matematicamente, vale il principio di indeterminazione di Heisenberg, che per le osservabili posizione X e momento P diviene:

$$(\Delta X)_\psi (\Delta P)_\psi \geq \frac{\hbar}{2}$$

Formalismo matematico: Se due operatori autoaggiunti A e B che descrivono osservabili non commutano, allora il prodotto delle fluttuazioni delle due osservabili in uno stato (eventualmente in un opportuno dominio) sono maggiori o uguali al modulo del valor medio di $[A, B]/2$ nello stato.

- Poiché i commutatori non nulli di un insieme irriducibile di osservabili, come posizione, momento e spin per una particella elementare, sono proporzionali a \hbar , la natura indeterministica della fisica quantistica dipende da \hbar .
- La presenza di commutatori non nulli comporta l'esistenza di spettri discreti proporzionali ad \hbar , ad esempio per il momento angolare e per l'energia degli stati legati, il che assicura la stabilità della materia, derivabile anche dalla natura di onde stazionarie degli autostati dell'energia.
- La natura di raggi vettori degli stati puri consente, tramite i *teoremi di Wigner e Bargmann*, l'esistenza di spin seminteri.

5. **Entanglement** Esistono stati di sistemi composti in cui le proprietà dei sottosistemi non hanno realtà indipendenti. Applicazioni: teletrasporto, crittografia e computer quantistici, la probabile tecnologia del futuro.

Formalismo matematico: Lo spazio di Hilbert di un sistema composto di sottosistemi non identici è il prodotto tensore degli spazi di Hilbert dei sottosistemi. Le funzioni d'onda del sistema composto possono essere rappresentate come funzioni degli autovalori di ICOC dei sottosistemi. Poiché lo spazio di Hilbert totale è uno spazio vettoriale, esistono degli stati in esso che non possono essere scritti come prodotto di stati dei sottosistemi, questi sono gli stati *entangled*.

6. **EPR+Bell:** Non è possibile avere sia la proprietà di controfattualità che di località (cioè realismo locale per gli stati)

7. **Perdita dell'individualità delle particelle identiche:** La sparizione delle traiettorie comporta che l'identità delle particelle porta alla loro indistinguibilità. La natura di raggi vettori degli stati puri consente poi l'esistenza di statistiche fermioniche che caratterizzano le particelle della materia. Una conseguenza di tale statistica è che esse non possono avere, neppure *potenzialmente*, la stessa posizione e verso dello spin, come affermato dal *principio di esclusione di Pauli*. Implicazioni: impenetrabilità dei corpi, chimica, proprietà dei solidi.

Bibliografia

- [1] Luca Martucci, Roberto Volpato. *Appunti di Istituzioni di Metodi Matematici per la Fisica*. Unipd, 10 maggio 2018.
- [2] Kenichi Konishi, Giampiero Paffuti *Meccanica quantistica: nuova introduzione*. Pisa university press, ed. 2012
- [3] David J. Griffiths *Introduction to Quantum Mechanics*
- [4] *Note del corso di "Fisica Moderna"* tenuto dal Professor Flavio Seno presso l'Università di Padova. Disponibili a bit.ly/2Cf1qS8
- [5] <https://www.reed.edu/physics/courses/P342.S10/Physics342/page1/page1.html>, <https://www3.nd.edu/~johnson/Class01F/chap2a.pdf>, http://web.mst.edu/~parris/QuantumTwo/Class_Notes/CentralPotential.pdf, <http://electron6.phys.utk.edu/qm2/modules/m1-3/central.htm>
- [6] <https://farside.ph.utexas.edu/teaching/qm/Quantum/node61.html>
- [7] <https://www.theoretical-physics.net/0.1/src/math/other.html#dirac-notation>