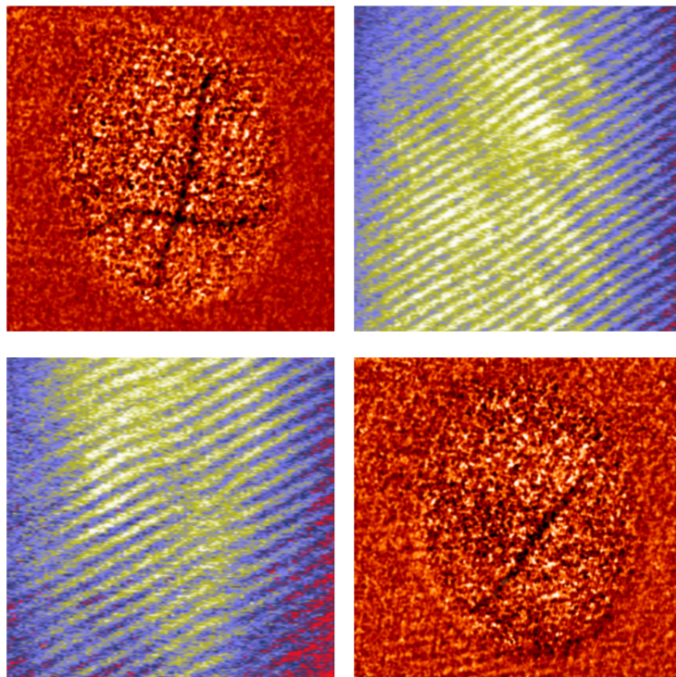


Franco Dalfovo

Dipartimento di Fisica, Università di Trento

Meccanica Quantistica

Appunti delle lezioni



Indice

Premessa	i
1 Le basi della teoria	1
1.1 Quello che già sappiamo	1
1.2 L'esperimento di Stern-Gerlach	3
1.3 L'analogia della luce polarizzata	7
1.4 Stati fisici, vettori di stato, osservabili	12
1.4.1 Vettori di stato, operatori e osservabili	12
1.4.2 L'esempio dello spin	13
1.4.3 Prodotto interno, norma e ortogonalità	14
1.4.4 Operatori lineari e hermitiani	15
1.4.5 Prodotto esterno	18
1.5 Stati di base e rappresentazione matriciale	19
1.5.1 Vettori e basi	19
1.5.2 Operatori e matrici	21
1.5.3 L'esempio dello spin	23
1.6 Definizione di misura	25
1.6.1 Misura	25
1.6.2 L'esempio dello spin (ancora una volta)	29
1.7 Osservabili compatibili e incompatibili	31
1.8 Relazione di indeterminazione	36
1.9 Cambiamenti di base	38
1.10 Spettri continui, posizione	40
1.11 Traslazioni spaziali e momento	43
1.12 Funzioni d'onda	48
2 Dinamica quantistica	55
2.1 Evoluzione temporale	55
2.2 Equazione di Schrödinger	57
2.3 Energia e costanti del moto	61
2.3.1 Stati stazionari e costanti del moto	61
2.3.2 Precessione dello spin	64
2.4 Rappresentazione di Heisenberg	67
2.5 Teorema di Ehrenfest e limite classico	71
2.6 L'oscillatore armonico	74
2.6.1 Lo spettro dell'oscillatore	74
2.6.2 Evoluzione temporale e stati coerenti	79
2.7 Equazione di continuità	81

3	Momento angolare	85
3.1	Definizione, commutatori e spettro	85
3.2	Momento angolare orbitale e rotazioni	89
3.3	Particella in un campo centrale	97
3.4	Atomo d'idrogeno	99
3.5	Spin	103
3.6	Addizione di momenti angolari	107
4	Metodi approssimati	113
4.1	Soluzioni esatte o approssimate	113
4.2	Perturbazioni indipendenti dal tempo	114
4.2.1	Spettro non degenerare	114
4.2.2	Esempio: perturbazione all'oscillatore armonico	118
4.2.3	Spettro degenerare	121
4.2.4	Esempio: perturbazione alla buca quadra	124
4.3	Perturbazioni dipendenti dal tempo	128
4.3.1	Probabilità di transizione	128
4.3.2	Perturbazione costante accesa a $t = 0$	132
4.3.3	Perturbazione armonica accesa a $t = 0$	133
4.3.4	Regola aurea di Fermi	134
4.4	Metodi variazionali	136
5	Più particelle e matrice densità	139
5.1	Due o più particelle	139
5.1.1	Due particelle isolate	141
5.1.2	Particelle identiche	142
5.1.3	Entanglement	151
5.2	Matrice densità	153
6	Conclusioni	161

Premessa

Questi sono gli appunti delle lezioni di Meccanica Quantistica che ho il piacere di tenere al terzo anno della laurea in fisica a Trento. Lo scopo del corso è fornire gli strumenti per comprendere le basi della meccanica quantistica non relativistica, cogliendone i principi fisici e acquisendo familiarità con il suo formalismo matematico. Per seguire questo corso con profitto è necessario conoscere bene la meccanica classica newtoniana e l'elettromagnetismo. Inoltre si suppone che chi legge abbia già visto quali sono i limiti della fisica classica e quali sono stati i tentativi di superare tali limiti tramite l'introduzione dei quanti di energia, il modello di Bohr e i successivi sviluppi fino all'equazione di Schrödinger. Il concetto di funzione d'onda dovrebbe essere noto, almeno a livello elementare, così come l'esistenza di spettri discreti e continui di energia nelle soluzioni dell'equazione di Schrödinger. Nel programma per la laurea triennale a Trento, questi argomenti vengono trattati al secondo anno nelle lezioni di Fisica Generale III. Dal punto di vista matematico, chi affronta queste lezioni dovrebbe conoscere bene l'analisi e l'algebra lineare, oltre alla meccanica analitica e i metodi matematici insegnati al secondo anno del corso di laurea. Parte degli strumenti matematici utili al corso vengono insegnati nel corso di Complementi matematici della meccanica quantistica, che viene tenuto in parallelo e nel quale vengono proposti esercizi con applicazioni semplici della teoria.

Essendo focalizzato sui concetti di base della teoria, questo testo riporta solo raramente i valori numerici di grandezze fisiche con le loro unità di misura. Quando accade, le unità di misura utilizzate sono quelle del sistema CGS di Gauss, anziché quelle del Sistema Internazionale basato sull'uso del kilogrammo, il metro e il secondo, ecc.; l'uso del sistema CGS nei libri di meccanica quantistica è un'abitudine che riflette la riluttanza dei fisici teorici a scrivere fattori del tipo $1/(4\pi\epsilon_0)$ nelle equazioni dell'elettromagnetismo e altri 4π che comparirebbero un po' ovunque nelle leggi fisiche. Dal punto di vista dello sviluppo e della presentazione della teoria non cambia nulla, ma è importante ricordarsi le tabelle di conversione tra le unità quando si confrontano testi diversi o quando si confrontano le predizioni della teoria con i dati sperimentali, che di regola sono espressi in unità SI. Peraltro, vale la pena sottolineare che nei testi in letteratura si usano anche sistemi basati sulle cosiddette unità naturali, come le unità atomiche, ad esempio; per non parlare di tutti quegli articoli in cui i teorici arrivano al punto di porre a 1 le costanti universali c , $\hbar$, o k_B per amore dell'eleganza o eccesso di pigrizia. Si tratta comunque di scelte convenzionali che non cambiano la fisica.

Le mie lezioni si basano soprattutto sul libro J.J. Sakurai, Jim Napolitano, *Meccanica Quantistica Moderna*, che nel seguito chiameremo semplicemente "il Sakurai". Lo suggerisco come libro di testo. Questi appunti, invece, sono una sorta di registrazione di quanto dico a lezione; servono a me per non perdere il filo del ragionamento e a chi segue le lezioni come punto di partenza per lo studio individuale. In molte parti gli appunti seguono da vicino il percorso tracciato dal Sakurai, ma di tanto in tanto se ne allontanano seguendo e miei gusti e tenendo conto delle ore disponibili e della necessità di armonizzare i contenuti del corso a

quelli degli altri insegnamenti del corso di laurea. Il Sakurai contiene molte più informazioni, con maggiore dettaglio, e per questo vi incoraggio a leggerlo. Lo schema espositivo è il cosiddetto “Stern-Gerlach first”, in cui è chiesto a chi lo legge di acquisire fin da subito il formalismo dei vettori di stato nella notazione di Dirac. È sempre utile consultare anche altri testi, specialmente quelli in cui la stessa teoria viene esposta in modo diverso. Tra i tanti, mi sento di suggerire questi: C. Cohen-Tannoudji, B. Diu e F. Laloe, *Quantum Mechanics*, Vol 1 e 2; A. Messiah, *Quantum Mechanics*; L.D. Landau, E.M. Lifshits, *Meccanica quantistica. Teoria non relativistica*; L.E. Picasso, *Lezioni di Meccanica Quantistica*; D.J. Griffiths, *Introduzione alla Meccanica Quantistica*. Infine, nel 2018 è stato pubblicato da Zanichelli il bel libro *Fisica Quantistica* di S.Forte e L.Rottoli che, mantenendo un’impostazione simile a quella del Sakurai, riesce a sintetizzare in modo efficace i concetti essenziali della teoria con un linguaggio più vicino a quello della ricerca attuale negli ambiti delle applicazioni della meccanica quantistica e dell’informazione quantistica.

Buona lettura.

Franco Dalfovo
Trento, 7 ottobre 2024

Capitolo 1

Le basi della teoria

1.1 Quello che già sappiamo

La fisica classica è una collezione di teorie che spiegano un vastissimo insieme di fenomeni naturali, dalla gravitazione all'elettromagnetismo, dalla meccanica dei fluidi all'ottica e l'acustica e molto altro. La fisica classica si basa su una visione corpuscolare della materia. Il mondo è fatto di particelle che interagiscono tra loro. L'interazione tra le particelle avviene tramite forze. La configurazione di un sistema di particelle in un dato istante è assegnata qualora si conoscano i valori delle coordinate spaziali e delle corrispondenti velocità di ciascuna di esse. Se le particelle sono N e sono libere di muoversi ovunque nello spazio tridimensionale, conoscere la loro configurazione corrisponde a conoscere $6N$ numeri, nelle opportune unità di misura; se le particelle sono soggette a vincoli, ad esempio perché possono muoversi solo lungo una data superficie curva, oppure perché sono legate tra loro per formare corpi rigidi, allora il numero di coordinate necessarie si riduce alle sole coordinate generalizzate e ai loro momenti coniugati, come ci insegna la meccanica analitica. A partire da una data configurazione il sistema evolve seguendo leggi deterministiche, tipicamente espresse nella forma di equazioni differenziali. Inoltre ad ogni configurazione di un sistema possiamo assegnare un valore della quantità di moto, dell'energia meccanica e del momento angolare e possiamo riformulare il problema dinamico in termini di leggi di conservazione o di principi variazionali. Le forze tra le particelle possono essere espresse in termini di campi, come il campo elettromagnetico, e possiamo attribuire quantità di moto, energia e momento angolare anche ai campi. Il fatto che un campo possa trasportare energia e quantità di moto nello spazio fa emergere il concetto di *radiazione*, che si aggiunge al concetto di *materia*, per completare il quadro classico del mondo fisico. Se le particelle di un sistema sono in numero troppo grande per poter determinare la loro configurazione istantanea in modo univoco, allora si ricorre ad assunzioni sulla natura statistica del problema e, tramite il calcolo di opportune funzioni di distribuzione e di valori medi di grandezze misurabili, si ottengono predizioni compatibili con i principi della termodinamica. In questo modo, l'applicazione di questi principi, del tutto generali, ci permette di ricavare informazioni sulla natura delle particelle del sistema e sulle loro interazioni.

Tutto ciò è bellissimo, ma non basta. Sappiamo che ci sono fenomeni che sfuggono a questa descrizione. Il primo ad accorgersene, suo malgrado, è Planck nel 1900 quando introduce un elemento nuovo, il quanto di energia $h\nu$, incompatibile con la fisica classica, per spiegare l'emissione di radiazione da parte di un corpo nero. L'interpretazione del quanto di energia come particella associata al campo elettromagnetico (quello che oggi chiamiamo fotone) è opera di Einstein che, in questo modo, nel 1905, spiega l'effetto fotoelettrico. È ancora lui a proporre di quantizzare anche le vibrazioni degli atomi nei solidi per predire

il comportamento del calore specifico in funzione della temperatura. Bohr invece introduce la quantizzazione dell'energia nella struttura degli atomi allo scopo di garantire la stabilità delle orbite degli elettroni e di dare una spiegazione agli spettri a righe di emissione e assorbimento. Bohr si accorge anche che la condizione per la stabilità delle orbite può essere espressa come quantizzazione del momento angolare. Gli sviluppi di queste idee portano alla cosiddetta “vecchia teoria dei quanti”.

Nei primi anni venti del '900 la situazione era insoddisfacente. Era chiaro che molti fenomeni non potevano più essere spiegati dalla fisica classica, ma le spiegazioni alternative, benché di successo in casi specifici, poggiavano tutte su ipotesi che venivano aggiunte *ad hoc* all'impianto concettuale esistente. Questa modalità di tappare le falle non poteva durare a lungo. Era decisamente preferibile tentare di riformulare i principi fisici di base al fine di ottenere una nuova teoria più generale, che fosse in grado di spiegare anche i nuovi fenomeni e che fosse riconducibile alla fisica classica in un qualche limite opportuno.

Un passo avanti nella direzione giusta lo compie Heisenberg a partire da una revisione critica del modello di Bohr per gli atomi. Un punto dolente della teoria era l'impossibilità di determinare la posizione e la velocità di un elettrone in un dato istante e, di conseguenza, l'impossibilità di osservarne la traiettoria. Heisenberg ritiene che debbano essere rimossi dalla formulazione del problema tutti i concetti che non siano collegati a quantità effettivamente misurabili. Per questa ragione, sceglie di lavorare unicamente con le frequenze della radiazione emessa e assorbita, cercando di definire gli stati fisici dell'atomo in termini di tali frequenze. Il suo sforzo porta alla cosiddetta *meccanica matriciale*, che già contiene in sé buona parte dei concetti che poi confluiranno in quella che oggi chiamiamo meccanica quantistica.

In parallelo, Schrödinger affronta lo stesso problema a partire dall'idea di de Broglie di attribuire un carattere ondulatorio alle particelle materiali, in analogia con il dualismo onda-particella già introdotto da Einstein per la radiazione elettromagnetica. Usando congetture ispirate all'ottica, Schrödinger arriva a scrivere un'equazione differenziale per una funzione complessa $\Psi(\mathbf{r}, t)$, detta funzione d'onda, associata in qualche modo alla delocalizzazione spaziale delle particelle. La soluzione dell'equazione nel caso dell'elettrone in un atomo d'idrogeno porta ai risultati corretti per le frequenze di assorbimento ed emissione, senza la necessità di invocare concetti classici ed ipotesi *ad hoc*. La meccanica basata sull'equazione di Schrödinger passa sotto il nome di *meccanica ondulatoria*. Le due meccaniche, quella ondulatoria e quella matriciale, si dimostrano essere del tutto equivalenti nei risultati che producono, potendo essere quindi considerate come due formulazioni alternative di un'unica teoria.

Tutti i ragionamenti fatti per arrivare alla nuova teoria in quegli anni riflettono le difficoltà che si incontrano quando si vuole andare oltre i confini noti, ma portandosi dietro il bagaglio di tutte le nozioni che sono date per conosciute e assodate a priori, senza sapere quali, tra le cose vecchie, si riveleranno ancora utili e quali no. Da qui nasce la difficoltà ad esempio di dare un senso alle domande del tipo: dove si trova in questo istante la tal particella? Cosa fa un elettrone mentre salta da uno stato stazionario all'altro? Se il senso che si cerca di dare è qualcosa di esprimibile in termini di nozioni classiche, la risposta a queste domande non c'è e le domande stesse sono mal poste.

Buona parte del dibattito sull'interpretazione della meccanica quantistica, che ha influenzato non poco la cultura scientifica e filosofica del XX secolo, si basa su domande mal poste, almeno se viste con gli occhi di un fisico. Tutto ciò che la fisica può e vuole fare è stabilire relazioni logiche tra fenomeni osservabili, che, tradotto in linguaggio concreto, significa definire ciò che è osservabile, codificarne la misura, scrivere relazioni matematiche tra gli esiti delle misure (numeri) e, infine, produrre schemi logico-deduttivi (teorie) che leghino assieme in modo coerente, a partire da alcuni principi base, tutte le conoscenze su una certa classe di fenomeni che siano verificate empiricamente entro i margini degli errori

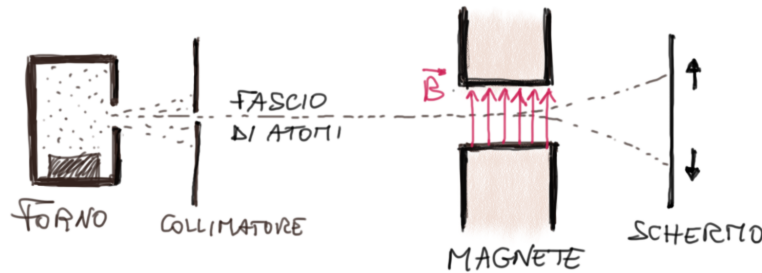
sperimentali. Ovviamente, quando si introduce una teoria nuova, si procede un po' a tentoni, appoggiandosi a qualche ragionevole sostegno. Ad esempio, Bohr seguiva un *principio di corrispondenza*, secondo cui la nuova teoria che stava elaborando doveva dare gli stessi risultati della fisica classica in certi limiti noti. A posteriori, quando tutta la nuova teoria è pronta e sufficientemente validata, il principio di corrispondenza rimane un buon approccio (ma non un principio) per individuare in quali limiti la fisica classica vale ancora, essendo un'approssimazione di quella quantistica. Trovandosi a discutere spesso con colleghi scettici riguardo allo strano comportamento delle particelle in regime quantistico, Bohr aveva anche escogitato un *principio di complementarità* secondo cui non è possibile rivelare, in una sola misura, sia il carattere corpuscolare sia il carattere ondulatorio di una particella, i due aspetti essendo, appunto, complementari. Bohr si era tanto convinto della bontà di questo principio, da trasformarlo in una sorta di talismano, al punto da metterlo alla base di una sua teoria, alquanto visionaria, sulla vita. L'autorevolezza di Bohr ha poi fatto sì che il principio di complementarità venisse letto, da chi fisico non è, come uno dei principi base della meccanica quantistica e tuttora venga proposto da eminenti scienziati come possibile soluzione di problemi complessi (come nel caso della relazione tra mente e coscienza). In realtà non è affatto un principio della teoria; si tratta piuttosto di un'idea, più o meno fertile a seconda dei casi, nata per rispondere a domande complicate riguardo al dualismo onda-particella, nel contesto del dibattito sull'interpretazione (in senso classico) della teoria. Dato che sono passati quasi cent'anni, possiamo permetterci il lusso di schivare le insidie delle interpretazioni e andare direttamente al sodo nella descrizione della teoria per come la teoria è e non per quello che appare alla luce di un linguaggio classico, inadeguato. In particolare, possiamo tralasciare per un po' il concetto di dualismo onda-particella, che è una sorta di relitto dei tempi in cui regnava confusione su cosa facessero i fotoni, gli elettroni e gli atomi, ma che non ha efficacia predittiva, né aiuta nella formulazione matematica della meccanica quantistica.

Gli studenti e le studentesse che hanno studiato per bene la fisica classica, con i suoi pregi e i suoi difetti, hanno uno zaino pieno di cose (concetti, strumenti, ...) che gli sono servite per arrivare fino a qui. Ora si tratta di incamminarsi oltre e, prima di partire, bisogna stabilire cosa conviene portarsi appresso e in quale punto guardare il fiume. La fisica si serve di idee e di esperimenti. Tra i tanti esperimenti ne scegliamo uno che ci sembra adatto, il cui esito sia semplice da descrivere a parole ma allo stesso tempo impossibile da interpretare con la fisica classica. Sulla base di questo esperimento cerchiamo di capire quali sono gli elementi essenziali che una teoria deve possedere per spiegare ciò che osserviamo. Cominciamo così. L'esperimento che scegliamo è quello di Stern-Gerlach.

1.2 L'esperimento di Stern-Gerlach

L'esperimento consiste nella misura del momento magnetico di un atomo sfruttando la sua interazione con un campo magnetico non uniforme. Se un atomo possiede un momento magnetico $\boldsymbol{\mu}$ e si trova in un campo magnetico $\mathbf{B}$, la sua energia conterrà un contributo pari a $E = -\boldsymbol{\mu} \cdot \mathbf{B}$. Che un atomo possa avere un momento magnetico non nullo ce lo aspettiamo dal fatto che è costituito da un insieme, globalmente neutro, di cariche elettriche. La distribuzione degli elettroni attorno al nucleo può produrre un momento angolare totale non nullo e, in analogia con le cariche che scorrono in una spira di corrente classica, possiamo ipotizzare che il momento magnetico sia proporzionale al momento angolare. Indipendentemente dalla natura di $\boldsymbol{\mu}$, quello che ci aspettiamo, sempre dalla fisica classica, è che un atomo con $\boldsymbol{\mu} \neq 0$ che si trovi in un campo magnetico non uniforme risenta di una forza $\mathbf{F} = -\nabla E = \nabla(\boldsymbol{\mu} \cdot \mathbf{B})$. Un magnete può essere configurato in modo tale da generare un campo magnetico $\mathbf{B}$ che abbia, ad esempio, $(dB_z/dz) \neq 0$. Un atomo che passi tra i poli

di tale magnete sentirà una forza nella direzione z , pari a $F = \mu_z(dB/dz)^1$. Supponiamo che un atomo si muova inizialmente nel vuoto lungo l'asse y ad una certa velocità e che entri in una regione di spazio dove agisce un campo magnetico con un valore assegnato di dB/dz ; la forza indotta dal campo lo devierà² verso il basso o verso l'alto in base al modulo e l'orientazione di μ . La misura della deflessione è di fatto una misura di μ_z .



Nel 1922 Otto Stern and Walter Gerlach realizzarono l'esperimento usando un fascio di atomi di argento. Gli atomi vengono estratti dal solido in un forno, vengono collimati in un fascio, fatti passare tra i poli di un magnete opportunamente sagomato e raccolti su uno schermo. Dal punto di vista classico, visto che gli atomi del fascio provengono da una sorgente che non privilegia alcuna direzione particolare del momento magnetico, essi avranno una distribuzione statisticamente isotropa di μ e, dunque, la macchia che andranno a formare sullo schermo dopo il passaggio attraverso il campo magnetico sarà una macchia allungata nella direzione z , gli atomi con μ nel verso positivo di z essendo quelli massimamente deflessi in alto, quelli in verso opposto analogamente in basso. In mezzo ci si aspetta una distribuzione continua, con un massimo a deflessione nulla.

L'esito dell'esperimento è diverso: si formano due sole macchie, una sopra e una sotto, simmetriche rispetto a $z = 0$; metà degli atomi vanno su, metà vanno giù, avendo subito la stessa accelerazione nel campo magnetico, a meno del segno. L'immagine originale in realtà era più simile a una striscia un po' sfumata, che si sdoppia lasciando in mezzo un vuoto di un paio di millimetri, ma questi sono dettagli legati alle difficoltà tecniche dell'esperimento di Stern e Gerlach. Esperimenti di questo tipo sono diventati routine e la precisione è ovviamente migliorata. Perciò, quando diciamo che si formano due macchie, lo intendiamo nel senso di esperimenti fatti al meglio delle tecniche disponibili, con il fascio ben collimato, gli atomi che viaggiano inizialmente con una distribuzione di velocità nota, il campo magnetico non uniforme nel modo voluto, eccetera. Poi va anche detto che Stern e Gerlach non erano affatto stupiti del risultato: anzi, era quello che si aspettavano e ne erano felici. Infatti, la teoria di Bohr e Sommerfeld prevedeva che il momento angolare $\mathbf{L}$ di un atomo potesse assumere solo orientazioni discrete nello spazio e, dato che μ è proporzionale a $\mathbf{L}$, anche μ_z doveva assumere valori discreti. Stern e Gerlach interpretarono il loro risultato come

¹In realtà, dovremmo tener conto che qualsiasi $\mathbf{B}$ deve obbedire alla legge $\nabla \cdot \mathbf{B} = 0$ e, dunque, ogni variazione di $\mathbf{B}(x, y, z)$ in una direzione spaziale comporta una variazione anche nelle altre direzioni; ma questo non è un problema, perché, con un'opportuna scelta di $\mathbf{B}$ e del suo gradiente si può rendere trascurabile l'effetto delle derivate lungo le direzioni diverse da z .

²Oltre a questa deviazione, l'atomo subirà una precessione dello momento magnetico attorno alla direzione di $\mathbf{B}$, detta precessione di Larmor, di cui parleremo in seguito, ma che non ha effetti sulla componente z del momento e, quindi, sulla deflessione verticale degli atomi; anzi, se il campo $\mathbf{B}$ è sufficientemente intenso, questa precessione favorisce la misura di μ_z , dato che riduce gli effetti delle deviazioni laterali dovute ai gradienti di $\mathbf{B}$ nel piano xy (vedi nota precedente), mediandoli a zero su tempi brevi rispetto ai tempi tipici del moto degli atomi.

la prima conferma diretta della teoria Bohr e Sommerfeld e come una misura del momento angolare orbitale degli elettroni nell'atomo di argento. A posteriori sappiamo che non è vero. La teoria di Bohr e Sommerfeld non è in grado di predire le due macchie in quell'esperimento e il momento angolare misurato non è quello orbitale, ma è lo *spin*.

Forse vi chiederete perché, se le cose stanno così, abbiamo scelto proprio questo esperimento per iniziare. Be', intanto della teoria di Bohr e Sommerfeld non ci importa un granché; sapevamo già che andava sostituita e ne vogliamo formulare una nuova; per questo scopo l'esperimento di Stern-Gerlach, come vedremo, va benissimo. Poi, il fatto che si parli di spin anziché di momento angolare non ha alcun rilievo per le considerazioni che vogliamo fare in questo capitolo. Quello che misuriamo nell'esperimento è μ_z ; questa quantità è una grandezza associata allo stato in cui si trova l'atomo quando passa nel campo magnetico e può essere sempre scritta come un qualche momento angolare associato all'atomo stesso moltiplicato per una costante che contiene parametri propri dell'atomo, come la massa e la carica degli elettroni ad esempio, e eventuali altre costanti universali. In questo capitolo non importa quale simbolo usiamo per questo momento angolare. Per ragioni storiche lo indichiamo con la lettera **S** e lo chiamiamo spin, in modo che μ sia direttamente proporzionale³ a **S**.

C'è un ultimo *caveat* prima di arrivare al cuore della faccenda. Nell'esperimento di Stern-Gerlach c'è una parte classica e una parte quantistica. La parte classica riguarda il modo con cui si prepara e si osserva il fascio di atomi e il modo con cui gli atomi interagiscono con il campo magnetico e con lo schermo posto all'uscita del magnete. In particolare, gli atomi vengono preparati in modo da entrare con una certa distribuzione di velocità nel magnete dove risentono della forza dovuta al gradiente del campo magnetico; tale forza induce una deviazione della traiettoria a seconda del valore di μ_z e assumiamo che la deviazione segua le leggi classiche. Possiamo usare la fisica classica, non relativistica, per varie ragioni: la velocità degli atomi è piccola rispetto a quella della luce; allo stesso tempo la loro massa e la loro velocità sono abbastanza grandi da far sì che la lunghezza d'onda di de Broglie degli atomi, $\lambda = h/p$ (dove h è la costante di Planck e $p = mv$), sia trascurabile rispetto a tutte le distanze in gioco nell'esperimento⁴. Inoltre la densità di atomi nel fascio è sufficientemente piccola da poter trascurare l'interazione tra di loro. La parte quantistica del problema sta nei gradi di libertà interni all'atomo che determinano il valore del momento magnetico di ciascuno di essi o, equivalentemente, il loro spin. Nel complesso si tratta di una misura di una grandezza soggetta a regole quantistiche tramite l'uso di un apparato che segue leggi classiche. Questo sarà un *leitmotiv* di tutte le misure, sulla cui importanza torneremo in seguito.

Bene, abbiamo spiegato come si esegue l'esperimento e abbiamo chiarito alcune questioni preliminari. Ora discutiamone le conseguenze e cerchiamo di formalizzare gli esiti introducendo il minimo numero di assunzioni.

Torniamo agli atomi che escono dal forno con un certo momento magnetico o, equivalentemente, un certo spin. Ci aspettiamo che l'orientazione dello spin sia casuale, con distribuzione di probabilità isotropa. Nel passaggio attraverso il magnete, con il campo magnetico orientato lungo z e un gradiente del campo nella stessa direzione, il fascio si divide in due formando sullo schermo due macchie uguali: quando ciascun atomo arriva allo schermo, o contribuisce alla macchia in alto o a quella in basso. Questo ci dice che i valori misurati della componente z dello spin, S_z , sono soltanto due. Il loro valore numerico non ci interessa

³La costante di proporzionalità si chiama rapporto giromagnetico; di cosa sia lo spin e di come si misura il rapporto giromagnetico ne parleremo più avanti.

⁴Un motivo valido per scegliere atomi pesanti come l'argento è proprio quello di garantire la bontà dell'approssimazione classica per la loro traiettoria; gli atomi hanno anche il vantaggio di essere neutri e non subire deviazioni dovute alla forza di Lorentz, che, per particelle cariche, sarebbero dominanti rispetto alla deviazione indotta dal gradiente del campo **B**.

al momento, ci basta indicarli con un simbolo che li distingua, ad esempio $+$ e $-$, oppure $\uparrow$ e $\downarrow$.

Dobbiamo essere molto cauti nel tentare di dedurre dall'esperimento informazioni sull'orientamento dello spin dell'atomo prima della misura, cioè prima che l'atomo stesso si schianti sullo schermo in una o nell'altra macchia. Se un atomo uscendo dal magnete va a finire in alto sullo schermo, fornendo così l'esito della misura del suo spin $S_z \uparrow$, potremmo essere tentati di dedurre che effettivamente l'atomo aveva lo spin $S_z \uparrow$ anche prima della misura, e altrettanto per gli atomi con $S_z \downarrow$ finiti nella macchia in basso; ma allora non si capisce bene per quale motivo non ci sia traccia di atomi con lo spin lungo x , ad esempio, che non verrebbero deflessi dal magnete.

Come verifica possiamo ruotare lo stesso magnete a 90° in modo da disporre il campo $\mathbf{B}$ e il suo gradiente nella direzione x . Ripetiamo lo stesso esperimento con lo stesso forno, lo stesso collimatore, lo stesso schermo. Troveremo due macchie come prima, ma stavolta nella direzione x , apparentemente senza traccia di atomi con spin lungo z . Lo stesso accade se giriamo il foro del forno e il collimatore in modo da propagare gli atomi lungo x , mettendo il campo magnetico lungo y . Troveremo due macchie nella direzione y . Di fatto recuperiamo l'isotropia della distribuzione iniziale, nel senso che l'esito dell'esperimento è lo stesso indipendentemente dalla direzione del fascio e dell'orientazione del magnete. Ma quello che non funziona è il tentativo di inferire dalla misura dello spin in una direzione il valore dello spin nelle altre due direzioni. Per capire meglio la situazione usiamo più apparati di Stern-Gerlach in sequenza.

Prendiamo lo schermo dell'apparato precedente, quello con il campo lungo z , e facciamo un buco in corrispondenza dell'area in cui si formava la macchia superiore. Ripetiamo l'esperimento. Stavolta solo gli atomi con $S_z \downarrow$ si fermeranno sullo schermo, mentre quelli con $S_z \uparrow$ proseguiranno oltre. Facciamoli passare attraverso un nuovo magnete, identico al precedente e con lo stesso campo magnetico con il gradiente lungo z . Mettiamo uno schermo dopo il secondo apparato. In questo esperimento osserviamo una sola macchia (non è una congettura, è proprio quello che si osserva se si fa l'esperimento): tutti gli atomi entranti nel secondo magnete escono verso l'alto. Nulla di strano: a differenza di prima, quando all'ingresso del magnete avevamo atomi con spin casuale, stavolta sappiamo che all'ingresso del secondo magnete ci sono solo atomi con spin $S_z \uparrow$, che inevitabilmente subiranno una deflessione verso l'alto. In un certo senso, il primo apparato è servito a preparare un insieme di atomi tutti nello stesso stato di spin, mentre il secondo ha semplicemente verificato che la preparazione era buona⁵.

Fin qui tutto bene. Ora però ruotiamo il secondo magnete di 90° in modo da allineare il campo e il suo gradiente nella direzione x . Abbiamo appena detto che tutti gli atomi uscenti dal primo magnete hanno spin $S_z \uparrow$. Dunque ci aspettiamo che non risentano del campo magnetico, dato che il gradiente del campo è ortogonale alla direzione del loro momento magnetico. Non dovrebbero essere deflessi e dovremmo vedere una sola macchia a $x = 0$. Invece ne troviamo due uguali alla solita distanza tra loro, la stessa delle macchie del primo apparato ma di intensità dimezzata⁶.

Un modo per venirne fuori potrebbe consistere nell'ipotizzare che gli atomi uscenti dal primo magnete avessero, oltre ad una componente positiva di spin lungo z , anche una componente lungo x , per metà positiva e per metà negativa. Questo accadrebbe se lo spin fosse un vettore orientato lungo la bisettrice del primo quadrante nel piano x - z per metà degli

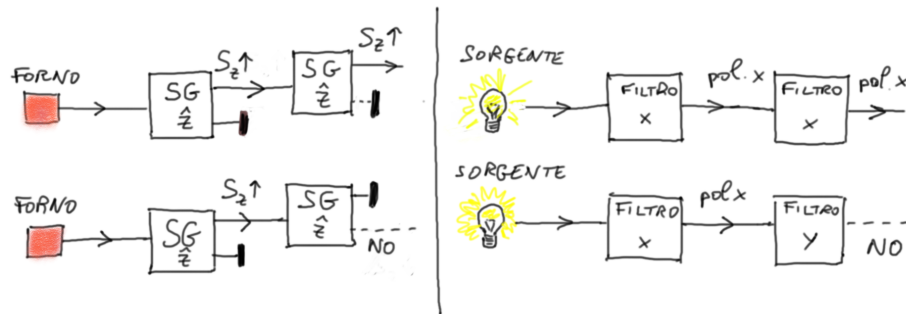
⁵Anche in fisica classica si usano opportune procedure per preparare un sistema in una configurazione iniziale desiderata, per poi studiarne il comportamento. Nella nuova teoria che stiamo introducendo, la preparazione degli stati fisici dei sistemi è un aspetto importante e non banale; l'apparato di Stern-Gerlach è un buon prototipo di tali procedure.

⁶Stiamo assumendo di usare magneti identici eccetto per l'orientazione spaziale e di tenere fisse anche le distanze tra magneti e schermo; quindi deviazioni del fascio di uguale entità implicano un uguale valore dello spin.

atomi e lungo la bisettrice del secondo quadrante per l'altra metà. Per quanto strana sia l'ipotesi, funzionerebbe nel dare una sola macchia lungo z e due macchie opposte lungo x , tutte alla stessa distanza dalla direzione originale del fascio. Per verificare se funziona davvero usiamo un terzo apparato in sequenza. Stavolta, in uscita dal secondo apparato blocchiamo gli atomi in corrispondenza della macchia con spin $S_x \downarrow$ e lasciamo andare avanti quelli con spin $S_x \uparrow$. Questi li facciamo passare attraverso il terzo magnete, con il campo lungo z . Se è vera l'ipotesi che abbiamo appena fatto, vedremo una sola macchia, perché tutti gli atomi che avevano passato il primo apparato avevano spin $S_z \uparrow$, indipendentemente dal valore della componente x . Invece no. Non succede. Quello che si osserva sono ancora due macchie uguali. Metà degli atomi in uscita dal terzo magnete hanno spin $S_z \downarrow$. Da dove vengono se li avevamo bloccati prima? Pare che la misura di S_x abbia distrutto le informazioni che avevamo prima su S_z .

1.3 L'analogia della luce polarizzata

Il comportamento del momento magnetico degli atomi in esperimento con sequenze di apparati di Stern-Gerlach ha qualche analogia con il comportamento della luce che passa attraverso filtri polarizzatori. Si tratta di analogie puramente formali, la fisica essendo completamente diversa. Nella polarizzazione della luce tutto è classico: non ci sono atomi, né spin; ci sono solo campi elettromagnetici che si propagano secondo le leggi di Maxwell. I filtri polarizzatori sono particolari dispositivi, anch'essi descrivibili con le equazioni di Maxwell, che filtrano la luce a seconda del tipo di polarizzazione. Nulla di misterioso. Cos'abbiamo in comune con Stern-Gerlach lo possiamo capire analizzando esperimenti con sequenze di filtri.



Supponiamo di avere un fascio di luce che si propaga in direzione z il cui campo elettrico è diretto lungo x e oscilla in questo modo:

$$\mathbf{E} = \hat{\mathbf{x}} E_0 \cos(kz - \omega t). \quad (1.1)$$

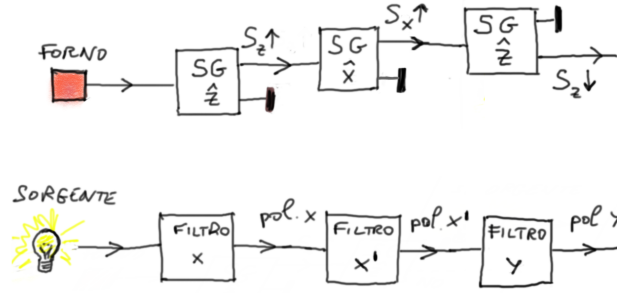
In questo caso diciamo che la luce è polarizzata linearmente lungo x . Per ottenere una luce polarizzata lungo x a partire da un fascio di luce non polarizzata, basta prendere un filtro polarizzatore, che è un semplice dispositivo ottico in grado di far passare solo e soltanto la luce polarizzata lungo una certa direzione. Supponiamo che sia un buon filtro, di quelli che lasciano passare il 100% della luce con la giusta polarizzazione e bloccano tutta l'altra. Se ruotiamo il filtro di 90° nel piano x - y , la luce che passa sarà quella polarizzata lungo y :

$$\mathbf{E} = \hat{\mathbf{y}} E_0 \cos(kz - \omega t). \quad (1.2)$$

Se usiamo due filtri in sequenza, uno lungo x e l'altro ancora lungo x , il secondo farà passare tutta la luce che arriva, dato che era stata preventivamente selezionata dal primo con la

giusta polarizzazione. Se invece il secondo filtro viene ruotato lungo y , allora non passa nulla. Guardando il secondo filtro da dietro, si vedrà nero. Questo comportamento è simile a quello di due apparati Stern-Gerlach in sequenza, il primo ad esempio con il campo lungo z e con un blocco all'uscita per gli atomi con $S_z \downarrow$, e l'altro con il campo nella stessa direzione, ma con il blocco per gli atomi con $S_z \downarrow$ oppure in $S_z \uparrow$, come in figura. L'equivalenza dei risultati, in termini di intensità dei fasci, si ottiene sostituendo gli stati di spin $\uparrow$ e $\downarrow$ lungo z con gli stati di polarizzazione x e y della luce.

Per ottenere l'equivalente dei tre apparati di Stern-Gerlach in sequenza ci serve un filtro polarizzatore che sia equivalente all'apparato di Stern-Gerlach lungo x , da mettere nel mezzo della sequenza di due apparati lungo z . La soluzione è semplice: basta ruotare un filtro polarizzatore di 45° rispetto al primo. Chiamiamo x' e y' le due coordinate ruotate di 45° rispetto a x e y . Allora la sequenza da utilizzare è quella nella figura qui sotto.



In entrambe le sequenze, se non ci fosse l'apparato o il filtro intermedio, il fascio osservato alla fine sarebbe di intensità nulla. Invece, con l'apparato (o il filtro) intermedio inserito, l'intensità del fascio osservata in uscita dalla sequenza è pari a un $1/4$ di quella in uscita dal primo apparato (o dal primo filtro). Questo è l'esito di entrambi gli esperimenti.

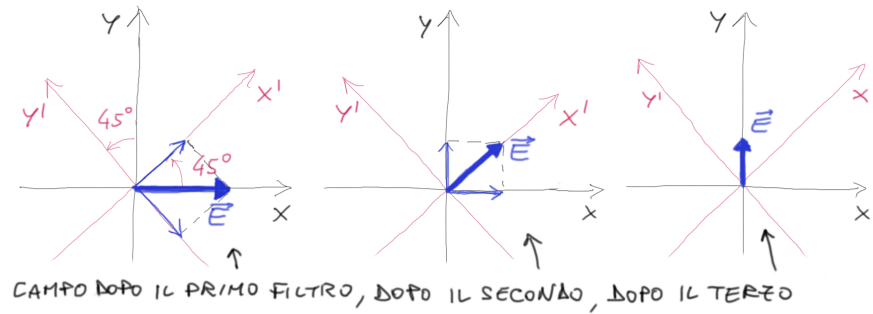
Nel caso della sequenza di filtri l'interpretazione è semplice. All'uscita del primo filtro c'è luce polarizzata lungo x , che corrisponde ad un campo elettrico oscillante nel tempo lungo l'asse x . Il secondo filtro non fa altro che decomporre il vettore $\mathbf{E}$ nelle sue componenti nelle direzioni ruotate x' e y' , facendo passare solo quella lungo x' . Si tratta di un problema di somma di vettori. Essendo l'angolo di 45° , la componente di $\mathbf{E}$ lungo x' è $1/\sqrt{2}$ più piccola del vettore originario, in modulo. All'ingresso del terzo filtro c'è un campo elettrico orientato lungo x' e il filtro decompone questo vettore lungo le direzioni x e y , lasciando passare solo la componente y ; in modulo questo significa un'altra riduzione di un fattore $1/\sqrt{2}$. Così la riduzione complessiva tra l'uscita del primo e l'uscita del terzo è di $1/2$ per il campo elettrico e, dato che l'intensità di un fascio di luce va come il quadrato del campo elettrico, la riduzione corrispondente di intensità del fascio è di un quarto, come quella osservata.

Tutto si basa sul fatto che il campo elettromagnetico soddisfa un *principio di sovrapposizione*: un campo elettrico in una data direzione può sempre essere scritto come composizione lineare di due campi elettrici in direzioni diverse. Ad esempio, il campo dopo il secondo filtro lo possiamo scrivere così

$$\hat{\mathbf{x}}' E_0 \cos(kz - \omega t) = \frac{1}{\sqrt{2}} \hat{\mathbf{x}} E_0 \cos(kz - \omega t) + \frac{1}{\sqrt{2}} \hat{\mathbf{y}} E_0 \cos(kz - \omega t) \quad (1.3)$$

e il terzo filtro ci dà in uscita la prima delle due componenti. Analogamente, un campo elettrico nella direzione y' , possiamo scomporlo in questo modo:

$$\hat{\mathbf{y}}' E_0 \cos(kz - \omega t) = -\frac{1}{\sqrt{2}} \hat{\mathbf{x}} E_0 \cos(kz - \omega t) + \frac{1}{\sqrt{2}} \hat{\mathbf{y}} E_0 \cos(kz - \omega t). \quad (1.4)$$



In effetti, possiamo anche vedere gli esperimenti con i filtri polarizzatori come un'ulteriore prova empirica del principio di sovrapposizione applicato all'elettromagnetismo classico.

Nel caso della sequenza di Stern-Gerlach il risultato è lo stesso per quanto riguarda l'intensità del fascio osservato (stavolta però fatto di atomi) ma l'interpretazione è meno banale. Verrebbe naturale pensare che il momento magnetico μ si comporti esso stesso semplicemente come $\mathbf{E}$, cioè come un vettore di cui gli apparati di Stern-Gerlach filtrano una o l'altra delle componenti. Ma non è così. Teniamo presente infatti che la rotazione degli apparati di Stern-Gerlach successivi nella sequenza è di 90° e non di 45° come nella sequenza di filtri polarizzatori. Se prendiamo un dato μ e ne cerchiamo le componenti a 90° , otteniamo sempre zero e l'apparato successivo non trova nulla da decomporre. Dunque, il principio di sovrapposizione che porta al risultato osservato non è quello per μ . Conviene piuttosto ragionare nel modo seguente. La misura eseguita con il filtro di polarizzazione ci dà informazioni su una particolare caratteristica dello stato fisico della luce: ci dice qual è il suo stato di polarizzazione in una certa direzione. Dato che lo stato di polarizzazione ha una corrispondenza uno-a-uno con la direzione del campo elettrico e dato che il campo elettrico obbedisce ad un principio di sovrapposizione, allora anche gli stati di polarizzazione della luce obbediscono allo stesso principio. Per analogia diciamo che l'apparato di Stern-Gerlach $SG\hat{z}$ dà informazioni sullo stato di spin di un atomo eseguendo una misura di S_z , mentre un apparato $SG\hat{x}$ dà informazioni sullo stato di spin dell'atomo eseguendo una misura di S_x . L'analogia con i filtri ci induce a congetturare che gli stati di spin obbediscano ad un principio di sovrapposizione e che i coefficienti della decomposizione degli stati coinvolti in sequenze di Stern-Gerlach siano gli stessi della decomposizione di $\mathbf{E}$ nelle sequenze equivalenti di filtri polarizzatori.

Per formalizzare l'idea ci serve un simbolo che rappresenti lo stato di spin di un atomo come un vettore in un nuovo spazio vettoriale, da definire. La scelta del simbolo ha una certa importanza (deve essere fatta in modo che il simbolo sia efficace e non ambiguo), ma è puramente convenzionale e non cambia la sostanza. Qui scegliamo di usare la notazione introdotta da Dirac. Se un atomo ha la componente dello spin S_z positiva, indichiamo lo stato di spin dell'atomo in questo modo:

$$|S_z \uparrow\rangle; \quad (1.5)$$

se ce l'ha negativa usiamo

$$|S_z \downarrow\rangle. \quad (1.6)$$

Analogamente indichiamo con $|S_x \uparrow\rangle$ e $|S_x \downarrow\rangle$ gli stati in uscita da uno Stern-Gerlach $SG\hat{x}$, e $|S_y \uparrow\rangle$ e $|S_y \downarrow\rangle$ quelli in uscita da uno $SG\hat{y}$.

Ora siamo pronti ad usare l'equivalenza dei risultati della sequenza di tre Stern-Gerlach e di tre filtri descritta in precedenza, dove il passaggio da $SG\hat{x}$ a $SG\hat{z}$ corrisponde dalla

rotazione del filtro dalle coordinate $x'-y'$ alle coordinate $x-y$. Guardando le due equazioni (1.3) e (1.4), ci si convince facilmente che tutto funziona come osservato nella sequenza di Stern-Gerlach se

$$\begin{cases} |S_x \uparrow\rangle = \frac{1}{\sqrt{2}}|S_z \uparrow\rangle + \frac{1}{\sqrt{2}}|S_z \downarrow\rangle \\ |S_x \downarrow\rangle = -\frac{1}{\sqrt{2}}|S_z \uparrow\rangle + \frac{1}{\sqrt{2}}|S_z \downarrow\rangle. \end{cases} \quad (1.7)$$

e, inoltre, se l'intensità del fascio di atomi è proporzionale al quadrato del coefficiente della combinazione lineare⁷. Le relazioni inverse sono

$$\begin{cases} |S_z \uparrow\rangle = \frac{1}{\sqrt{2}}|S_x \uparrow\rangle - \frac{1}{\sqrt{2}}|S_x \downarrow\rangle \\ |S_z \downarrow\rangle = \frac{1}{\sqrt{2}}|S_x \uparrow\rangle + \frac{1}{\sqrt{2}}|S_x \downarrow\rangle. \end{cases} \quad (1.8)$$

Nella sequenza considerata finora il primo apparato seleziona lo stato $|S_z \uparrow\rangle$ da un insieme di atomi con spin casuale (in senso statistico); il secondo decompone questo stato nelle componenti $|S_x \uparrow\rangle$ e $|S_x \downarrow\rangle$ lasciando passare solo la $|S_x \uparrow\rangle$; il terzo decompone questo stato nelle componenti $|S_z \uparrow\rangle$ e $|S_z \downarrow\rangle$ lasciando passare solo la $|S_z \downarrow\rangle$. Nel passaggio attraverso ciascuno degli ultimi due apparati il fascio si riduce in intensità della metà, come osservato. Per assicurarci che l'accordo non sia un accidente dovuto ad una scelta particolare della sequenza possiamo divertirci a cambiare l'ordine degli apparati $SG\hat{x}$ a $SG\hat{z}$ e la scelta di quale foro aprire in ciascuno schermo. L'esito di questi esperimenti è sempre in accordo con le decomposizioni degli stati di spin scritta qui sopra. È anche utile osservare che l'esperimento ci dice qual è il rapporto tra l'intensità del fascio in uscita e quello in ingresso. Per questo nelle (1.7), oltre ad aver scritto i coefficienti con modulo uguale, li abbiamo anche scelti in modo che la somma dei loro quadrati fosse 1, che è uguale al quadrato del coefficiente dello stato a sinistra dell'uguale. Dal punto di vista fisico questo corrisponde al fatto che il terzo apparato di Stern-Gerlach suddivide il fascio in due fasci uguali, ma la somma degli atomi che entrano e escono rimane la stessa. Moltiplicare entrambi i membri per una costante (incluso il valore -1 che equivale ad un cambio globale di segno), non cambia lo stato fisico di un atomo, l'esperimento rimane lo stesso, con gli stessi esiti.

Per completare il quadro ricordiamoci che abbiamo ancora un'altra direzione indipendente lungo la quale possiamo orientare il magnete dell'apparato di Stern-Gerlach, la direzione y . Ad esempio possiamo far uscire il fascio dal forno nella direzione z e farlo passare in una sequenza di $SG\hat{x}$ a $SG\hat{y}$; oppure lo possiamo far uscire lungo x e usare una sequenza di $SG\hat{y}$ e $SG\hat{z}$. Queste sequenze coinvolgeranno gli stati di spin $|S_y \uparrow\rangle$ e $|S_y \downarrow\rangle$ che, se è vera la nostra congettura, devono potersi scrivere come combinazioni lineari di $|S_z \uparrow\rangle$ e $|S_z \downarrow\rangle$ analoghe alle (1.7). I coefficienti non potranno essere gli stessi, perché un atomo che abbia lo spin lungo x si trova in uno stato fisicamente diverso da un atomo che ha lo spin lungo y . Tuttavia devono avere uguali i loro quadrati. Di questo ci possiamo convincere eseguendo un esperimento con il fascio di atomi che si propaga lungo x , trova un apparato $SG\hat{z}$ che lascia passare lo stato $|S_z \uparrow\rangle$, poi un secondo apparato a $SG\hat{y}$ che lascia passare lo stato $|S_y \uparrow\rangle$ e infine un terzo apparato $SG\hat{z}$. L'esperimento ci darà due macchie uguali in uscita in corrispondenza di $|S_z \uparrow\rangle$ e $|S_z \downarrow\rangle$ ⁸. Dunque i quadrati dei coefficienti devono essere uguali, ma gli stati diversi. Moltiplicare i coefficienti per una costante o cambiare tutti i segni non vale. Abbiamo già detto che questo non cambia lo stato fisico. Se rimaniamo nel campo dei numeri reali non ne veniamo fuori! Infatti, volendo scrivere i due stati $|S_y \uparrow\rangle$ e $|S_y \downarrow\rangle$ come combinazioni di $|S_z \uparrow\rangle$ e $|S_z \downarrow\rangle$ con i vincoli appena detti sui quadrati dei coefficienti, l'unica soluzione è scrivere uno dei due stati come somma e l'altro come differenza, con coefficienti

⁷Più avanti adotteremo una convenzione per cui $|S_x \downarrow\rangle$ avrà i segni opposti rispetto a quelli nell'espressione (1.7), ma questo non avrà effetti sui risultati fisici. Qui stiamo lavorando a livello di congetture; i dettagli li sistemeremo dopo.

⁸Notiamo che questa sequenza si ottiene da quella discussa finora solo per una rotazione di 90° del secondo apparato e il risultato per le macchie sullo schermo non cambia.

$1/\sqrt{2}$, ma così facendo si riottiene, a meno di una costante moltiplicativa globale, gli stessi stati $|S_x \uparrow\rangle$ e $|S_x \downarrow\rangle$.

Per venirne fuori conviene ricorrere ancora una volta all'analogia con la luce polarizzata. Anche in questo caso, infatti, possiamo individuare un terzo tipo di polarizzazione, indipendente dalla polarizzazione lineare in x - y e x' - y' , che dà risultati equivalenti a quelli di sequenze di Stern-Gerlach. Si tratta della polarizzazione circolare, in cui il vettore $\mathbf{E}$ ruota nel piano ortogonale alla direzione di propagazione, avendo le componenti x e y sfasate di $\pm\pi/2$:

$$\mathbf{E} = \frac{1}{\sqrt{2}}\hat{\mathbf{x}}E_0 \cos(kz - \omega t) + \frac{1}{\sqrt{2}}\hat{\mathbf{y}}E_0 \cos(kz - \omega t \pm \frac{\pi}{2}). \quad (1.9)$$

A seconda del segno dello sfasamento la luce si dice *levogira* o *destrogira*. La luce polarizzata circolarmente si ottiene facendo passare luce polarizzata linearmente attraverso una *lamina a quarto d'onda*. Se facciamo passare luce polarizzata circolarmente attraverso filtri polarizzatori lungo x o y otteniamo fasci in uscita di uguale intensità, come nel caso di un fascio polarizzato linearmente lungo x' o y' , ma la luce polarizzata circolarmente ha proprietà diverse da quella polarizzata linearmente. Per il vettore $\mathbf{E}$ la differenza è espressa da una fase nella funzione coseno. Se vogliamo usare la stessa combinazione anche per gli stati di spin, sfruttando una possibile equivalenza tra gli stati $|S_y \uparrow\rangle$ e $|S_y \downarrow\rangle$ e i due stati di polarizzazione circolare, dobbiamo sbarazzarci della funzione coseno, che per gli atomi nell'esperimento di Stern-Gerlach non ha alcun significato fisico, e passare al campo dei numeri complessi. Ad esempio, possiamo introdurre un campo complesso $\mathcal{E}$ nella forma

$$\mathcal{E} = \frac{1}{\sqrt{2}}\hat{\mathbf{x}}E_0 e^{i(kz - \omega t)} \pm \frac{i}{\sqrt{2}}\hat{\mathbf{y}}E_0 e^{i(kz - \omega t)}, \quad (1.10)$$

avendo sfruttato il fatto che $i = \exp(i\pi/2)$. La parte reale di $\mathcal{E}$ è il campo elettrico di prima, ma passando alla notazione complessa abbiamo ottenuto un suggerimento sui coefficienti da usare per gli stati di spin S_y :

$$\begin{cases} |S_y \uparrow\rangle = \frac{1}{\sqrt{2}}|S_z \uparrow\rangle + \frac{i}{\sqrt{2}}|S_z \downarrow\rangle \\ |S_y \downarrow\rangle = \frac{1}{\sqrt{2}}|S_z \uparrow\rangle - \frac{i}{\sqrt{2}}|S_z \downarrow\rangle \end{cases} \quad (1.11)$$

Il fatto di avere coefficienti complessi, ci impone di cambiare la ricetta per mettere in relazione i coefficienti con l'intensità dei fasci; la soluzione è semplice: invece dei quadrati, basta usare i moduli quadri.

Ora, eseguendo sequenze di Stern-Gerlach con campi magnetici in tutte e tre le direzioni indipendenti e schermi forati in corrispondenza di qualsiasi macchia, riusciamo a interpretare correttamente le intensità dei fasci, effettivamente osservate, in uscita da ogni apparato. L'intensità relativa tra uscita e ingresso è fissata dal modulo quadro del coefficiente della decomposizione degli stati di ingresso nella coppia di stati d'uscita. Per come li abbiamo definiti, gli stati di spin sono vettori in uno spazio vettoriale astratto di dimensione 2 i cui elementi possono essere rappresentati da coppie di numeri complessi.

In sintesi, il messaggio principale che esce da questi nostri ragionamenti è che gli stati fisici di un sistema, determinati dai valori misurati di certe grandezze fisiche, sembrano obbedire ad un principio di sovrapposizione analogo a quello dei vettori nello spazio euclideo delle coordinate, con la differenza che nel caso degli stati fisici si tratta di uno spazio complesso e astratto. Tutto ciò è solo una congettura ispirata da una singola tipologia di esperimenti, quelli con apparati di Stern-Gerlach sequenziali, e dall'analogia formale con la polarizzazione del campo elettromagnetico classico. Avendo imboccato una pista promettente, proviamo ad estendere questi concetti a situazioni più generali per trasformare la congettura in uno schema teorico soddisfacente.

1.4 Stati fisici, vettori di stato, osservabili

1.4.1 Vettori di stato, operatori e osservabili

Dato un generico sistema che abbiamo scelto di studiare (una particella, un insieme di particelle o una qualsiasi porzione di universo che ci interessa) ne possiamo indagare le proprietà misurando una o più grandezze fisiche, ciascuna definita tramite opportune procedure operative. Per estensione dei ragionamenti precedenti, così come nel caso del momento magnetico degli atomi di argento, assumiamo che per una qualsiasi grandezza fisica si possano individuare degli stati del sistema tali che, eseguendo la misura della grandezza in ognuno di questi, si troverà un valore ben definito, univocamente associato a quello stato. Tale valore è il numero reale, seguito dalle opportune unità di misura, che leggiamo sugli strumenti di misura avendo preparato il sistema preventivamente in quello stato. Indichiamo con α uno di questi valori e lo usiamo per etichettare lo stato corrispondente

$$|\alpha\rangle. \quad (1.12)$$

Assumiamo che lo stato $|\alpha\rangle$ sia un elemento di uno spazio vettoriale complesso⁹. Lo chiamiamo *vettore di stato* e, nel gergo di Dirac, si dice che è un *ket*. La dimensione dello spazio dipende dal numero degli esiti possibili della misura di quella grandezza. Nel caso del momento magnetico dell'atomo di argento, la dimensione è 2, perché lo spin misurato in qualsiasi direzione, può assumere solo due valori diversi. Altre grandezze possono ammettere spazi di vettori di stato di dimensione N , con N finito, o infinito e numerabile, oppure di dimensione infinita ma non numerabile, con α variabile continua e reale.

Come anticipato nella descrizione dell'esperimento di Stern e Gerlach, assumiamo che due vettori di stato che differiscono per una costante complessa non nulla rappresentino lo stesso stato fisico. È come dire che, ai fini della descrizione fisica di un sistema, nello spazio vettoriale che lo rappresenta conta solo la “direzione” dei vettori e non il loro modulo¹⁰. Il prodotto di un generico vettore $|\alpha\rangle$ per una costante numerica c è un vettore di stato che indichiamo in questo modo:

$$c|\alpha\rangle \quad (1.13)$$

e assumiamo che l'ordine dei simboli non sia rilevante, potendo scrivere lo stesso prodotto anche come $|\alpha\rangle c$. Nel caso particolare $c = 0$, il prodotto $c|\alpha\rangle$ coincide, per definizione, con il *vettore nullo*. Pur appartenendo allo stesso spazio vettoriale degli altri, il vettore nullo è l'unico che non rappresenta alcuno stato fisico.

Generalizzando il concetto di sovrapposizione lineare già discusso per lo spin, assumiamo che una combinazione lineare di vettori di stato sia ancora un vettore di stato. In altri termini, lo spazio vettoriale che contiene i possibili stati fisici del sistema non contiene solo gli stati per i quali è noto con certezza il valore che si otterrebbe misurando una certa grandezza, ma anche le loro combinazioni lineari. Gli spazi vettoriali che useremo, quindi, sono lineari nel senso che, se $|\alpha\rangle$ e $|\beta\rangle$ rappresentano stati fisici del sistema e c_1 e c_2 sono due costanti complesse qualsiasi, allora anche

$$|\gamma\rangle = c_1|\alpha\rangle + c_2|\beta\rangle, \quad (1.14)$$

⁹La scelta di usare l'esito della misura come etichetta all'interno del simbolo del *ket* $|\cdot\rangle$ è comoda, ma non necessaria; nel seguito useremo $|\alpha\rangle$, $|\beta\rangle$, ecc., per indicare vettori generici. In molti testi si usa il simbolo $|\Psi\rangle$, che richiama anche il concetto di funzione d'onda, ma qui vogliamo evitare di creare questa associazione mentale automatica tra *ket* e funzione d'onda, dato che quest'ultima è solo una delle possibili rappresentazioni dei vettori di stato, non l'unica e a volte nemmeno la migliore.

¹⁰Con un linguaggio matematico più preciso, si può dire che gli stati fisici sono in corrispondenza uno-a-uno con i “raggi”, non con i vettori.

rappresenta uno stato fisico ammissibile¹¹.

In uno spazio vettoriale possono agire degli *operatori* che, agendo su un vettore ne producono un altro. Indichiamo l'azione di un generico operatore A su uno vettore di stato $|\alpha\rangle$ con la notazione

$$A|\alpha\rangle, \quad (1.15)$$

che corrisponde ad un altro vettore. Quando il nuovo vettore $A|\alpha\rangle$ coincide con $|\alpha\rangle$ a meno di una costante moltiplicativa, allora $|\alpha\rangle$ è detto *autovettore* o *autostato* di A e la costante moltiplicativa è il corrispondente *autovalore*¹²:

$$A|\alpha\rangle = \alpha|\alpha\rangle. \quad (1.16)$$

Si badi bene che, in quest'ultima relazione, il primo α a destra dell'uguale è un numero, mentre il secondo α è solo una parte del simbolo $|\alpha\rangle$ che identifica un vettore.

Tutto questo ci porta ad identificare le grandezze fisiche, che d'ora in poi chiameremo *osservabili*, con operatori che agiscono sui vettori di stato di un sistema. Gli autostati di una osservabile sono quei particolari vettori di stato che non vengono modificati dal processo di misura; il sistema rimane nello stesso stato e la misura semplicemente registra l'autovalore corrispondente. In quanto tale, l'autovalore di un'osservabile deve essere reale; qualsiasi strumento di misura, infatti, produce numeri reali.

1.4.2 L'esempio dello spin

Nel caso del momento magnetico degli atomi l'osservabile misurata da un apparato di Stern-Gerlach è la componente di $\boldsymbol{\mu}$ nella direzione del gradiente del campo magnetico. Abbiamo visto che questa grandezza ammette solo due valori, che abbiamo indicato con $\uparrow$ e $\downarrow$. In realtà, conoscendo la massa degli atomi, la loro velocità, l'intensità del gradiente del campo magnetico e le distanze in gioco, possiamo ricavare con buona precisione il valore numerico del momento magnetico osservando la posizione delle macchie sullo schermo. Questo può essere tradotto in una grandezza dimensionalmente equivalente ad un momento angolare, che chiamiamo spin, in modo tale che i due valori ammessi per lo spin in una direzione qualsiasi sono $\pm\hbar/2$, dove $\hbar$ è la costante di Planck divisa per 2π . Questo lo diamo come un fatto empirico¹³ e qui non ci interessa nemmeno sapere il valore della costante che lega lo spin al momento magnetico. Vedremo a tempo debito qual è l'origine di questi valori dello spin e perchè salti fuori un $\hbar/2$ e non semplicemente $\hbar$ come verrebbe in mente pensando al modello

¹¹Sottolineo qui che questa affermazione è un'assunzione forte che avrà implicazioni sorprendenti. Il fatto che un sistema possa trovarsi in uno stato di sovrapposizione di altri stati, ad esempio stati in cui una certa grandezza assume valori diversi, sta al cuore di tutta la meccanica quantistica. Questo concetto lo stiamo formalizzando tramite la linearità degli spazi vettoriali che descrivono gli stati di un sistema e lo facciamo per estensione di quanto congetturato parlando degli esiti degli esperimenti del tipo Stern-Gerlach. In altri testi, come ad esempio il Forte-Rottoli, lo stesso concetto viene introdotto parlando della diffrazione di particelle da fenditure, che era anche l'esempio prediletto da Richard Feynman nelle sue famose lezioni; in tal caso la sovrapposizione di stati dà luogo a fenomeni di interferenza simili a quelli delle onde. Una ragione per ritenere preferibile Stern-Gerlach in queste mie lezioni è che così evito di coinvolgere da subito nozioni legate alla misura della posizione delle particelle e alla loro funzione d'onda, che preferisco discutere successivamente, come caso particolare di un approccio più generale.

¹²In altri termini, un autovettore di A è un elemento dello spazio vettoriale che viene trasformato da A parallelamente a se stesso.

¹³Il fatto che nell'esperimento di Stern e Gerlach le macchie siano due e i valori misurati dello spin solo $\pm\hbar/2$ è anche una conseguenza della scelta dell'argento come atomo su cui fare le misure. In un atomo di argento, infatti, lo spin complessivo dell'atomo è determinato interamente dallo spin di un singolo elettrone spaiato, dato che gli spin degli altri si sommano a zero. Questo aiuta l'interpretazione degli esiti e rende l'esperimento particolarmente efficace dal punto di vista didattico. Atomi diversi generalmente ammettono più macchie, a cui corrispondono più valori di spin e, dunque, spazi di Hilbert con dimensione maggiore di due.

di Bohr per l'atomo d'idrogeno. Ci basta sapere che, se usiamo un'apparato di Stern-Gerlach $SG\hat{z}$ per selezionare gli atomi nello stato $|S_z \uparrow\rangle$ e poi su questo stato eseguiamo la misura di S_z con un altro $SG\hat{z}$, troveremo il valore numerico $+\hbar/2$ per la componente z dello spin e gli atomi rimarranno tutti nello stato $|S_z \uparrow\rangle$, come prima della misura. Se invece prepariamo gli atomi nello stato $|S_z \downarrow\rangle$ e poi eseguiamo la misura di S_z , l'esito della misura sarà $-\hbar/2$ e gli atomi del fascio rimarranno tutti nello stato $|S_z \downarrow\rangle$. Dal punto di vista formale traduciamo tutto questo dicendo che esiste un operatore S_z , associato alla componente z dello spin, tale che $|S_z \uparrow\rangle$ e $|S_z \downarrow\rangle$ sono i suoi autostati e $\pm\hbar/2$ sono i corrispondenti autovalori:

$$S_z|S_z \uparrow\rangle = +\frac{\hbar}{2}|S_z \uparrow\rangle \quad (1.17)$$

$$S_z|S_z \downarrow\rangle = -\frac{\hbar}{2}|S_z \downarrow\rangle. \quad (1.18)$$

Relazioni simili le possiamo poi scrivere per le osservabili S_x e S_y e i loro autostati. Abbiamo quindi tre osservabili i cui autostati sono vettori dello stesso spazio vettoriale bidimensionale. Viene naturale chiedersi, a questo punto, cosa convenga utilizzare come base per la rappresentazione dei vettori di stato. Per rispondere occorre definire prima un prodotto interno e i concetti di ortogonalità e norma dei vettori.

1.4.3 Prodotto interno, norma e ortogonalità

Dato uno spazio di vettori di stato (*ket*) che descrivono un sistema in relazione ad una grandezza misurabile, conviene introdurre un nuovo spazio vettoriale, detto *spazio duale*, in modo che per ogni *ket* $|\alpha\rangle$ esista un vettore nello spazio duale, che chiamiamo *bra* e che indichiamo con il simbolo $\langle\alpha|$. La corrispondenza

$$|\alpha\rangle \longleftrightarrow \langle\alpha| \quad (1.19)$$

è detta *corrispondenza duale*. Per le combinazioni lineari vale

$$a|\alpha\rangle + b|\beta\rangle \longleftrightarrow a^*\langle\alpha| + b^*\langle\beta| \quad (1.20)$$

dove a^* e b^* sono i complessi coniugati delle costanti a e b .

A questo punto definiamo un *prodotto interno* (o prodotto scalare) di due vettori di stato $|\alpha\rangle$ e $|\beta\rangle$ come il numero complesso che si ottiene dal prodotto del vettore $|\alpha\rangle$ per il duale di $|\beta\rangle$

$$\langle\beta|\alpha\rangle = (\langle\beta|)(|\alpha\rangle) \quad (1.21)$$

tale che

$$\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^*. \quad (1.22)$$

Quando moltiplichiamo un vettore per se stesso, la definizione di prodotto interno implica

$$\langle\alpha|\alpha\rangle = \langle\alpha|\alpha\rangle^* \quad (1.23)$$

e quindi $\langle\alpha|\alpha\rangle$ è un numero reale. Assumiamo che sia anche non negativo

$$\langle\alpha|\alpha\rangle \geq 0, \quad (1.24)$$

dove l'uguaglianza vale solo se $|\alpha\rangle$ è un vettore nullo. Questa assunzione è chiamata *postulato della metrica definita positiva* e sarà essenziale più avanti per dare significato alla misura di osservabili in termini di probabilità. La quantità $\sqrt{\langle\alpha|\alpha\rangle}$ è detta *norma* del vettore $|\alpha\rangle$. Un vettore è detto *normalizzato* se

$$\langle\alpha|\alpha\rangle = 1; \quad (1.25)$$

se non lo è, allora è sempre possibile normalizzarlo eseguendo la seguente operazione

$$|\beta\rangle = \frac{1}{\sqrt{\langle\alpha|\alpha\rangle}}|\alpha\rangle, \quad (1.26)$$

in modo che $|\beta\rangle$ è normalizzato e rappresenta lo stesso stato fisico di $|\alpha\rangle$. Notiamo anche che il prodotto scalare tra due vettori $a|\alpha\rangle$ e $b|\beta\rangle$ vale

$$\langle\beta|b^*a|\alpha\rangle = ab^*\langle\beta|\alpha\rangle. \quad (1.27)$$

Due vettori sono *ortogonali* se

$$\langle\alpha|\beta\rangle = 0. \quad (1.28)$$

Un insieme di vettori *ortonormali* è costituito da vettori che sono ortogonali tra loro e hanno norma unitaria. Infine, assumiamo che il prodotto interno si distribuisca sulla somma di vettori in questo modo:

$$\langle\gamma|(|\alpha\rangle + |\beta\rangle)\rangle = \langle\gamma|\alpha\rangle + \langle\gamma|\beta\rangle. \quad (1.29)$$

Formalmente, lo spazio dei *bra* è lo spazio dei funzionali lineari che agiscono sui *ket* in modo che l'azione di un elemento dello spazio dei *bra* su un elemento dello spazio dei *ket* produce un numero complesso. Lo spazio vettoriale lineare dei *ket* con un prodotto interno così definito è uno spazio di Hilbert. Ciò che stiamo dicendo qui, un po' alla volta, è che gli stati di un sistema fisico sono rappresentati dagli elementi di uno spazio di Hilbert.

Un risultato importante che ci servirà in seguito è questo:

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \geq |\langle\alpha|\beta\rangle|^2. \quad (1.30)$$

Si tratta della *disuguaglianza di Schwarz*, che vale per qualsiasi coppia di vettori di stato, analoga alla relazione $|\mathbf{v}|^2|\mathbf{u}|^2 \geq |\mathbf{v} \cdot \mathbf{u}|^2$ nello spazio reale euclideo. La dimostrazione è semplice. Prendiamo il vettore $|\alpha\rangle + \lambda|\beta\rangle$, con λ numero complesso qualsiasi, e calcoliamo il suo prodotto interno con se stesso

$$(\langle\alpha| + \lambda^*\langle\beta|)(|\alpha\rangle + \lambda|\beta\rangle) \geq 0; \quad (1.31)$$

la disuguaglianza vale sempre come conseguenza del postulato della metrica definita positiva. Vale anche per $\lambda = -\langle\beta|\alpha\rangle/\langle\beta|\beta\rangle$. Sostituendo questo valore nell'espressione precedente si trova la disuguaglianza di Schwarz.

1.4.4 Operatori lineari e hermitiani

Torniamo alla questione degli operatori. Sia X un operatore generico, non necessariamente associato ad una osservabile. Come già anticipato, l'azione dell'operatore su un vettore di stato la indichiamo mettendo l'operatore a sinistra del *ket*:

$$X|\alpha\rangle. \quad (1.32)$$

Quello che si ottiene è un nuovo *ket*. Due operatori X e Y sono uguali se

$$X|\alpha\rangle = Y|\alpha\rangle \quad (1.33)$$

per qualsiasi vettore $|\alpha\rangle$. L'operatore X è nullo se

$$X|\alpha\rangle = 0 \quad (1.34)$$

per qualsiasi $|\alpha\rangle$. La somma di operatori va intesa in questo modo

$$(X + Y)|\alpha\rangle = X|\alpha\rangle + Y|\alpha\rangle \quad (1.35)$$

ed è sia *commutativa*

$$X + Y = Y + X \quad (1.36)$$

che *associativa*

$$X + Y + Z = X + (Y + Z) = (X + Y) + Z. \quad (1.37)$$

Noi considereremo solo operatori lineari, che sono quelli per cui vale

$$X(a|\alpha\rangle + b|\beta\rangle) = aX|\alpha\rangle + bX|\beta\rangle. \quad (1.38)$$

Un operatore può anche agire sul vettore duale; in tal caso scriviamo così

$$\langle\alpha|X \quad (1.39)$$

e diciamo che l'operatore agisce da destra sul *bra*. È importante sottolineare che il vettore $\langle\alpha|X$ non è necessariamente il duale di $X|\alpha\rangle$. Per ottenere il duale di $X|\alpha\rangle$, infatti, si deve applicare al duale di $|\alpha\rangle$ un operatore diverso, che indichiamo con $X^\dagger$ e chiamiamo *hermitiano coniugato* (o *aggiunto*) di X , in modo che la corrispondenza duale sia questa

$$\langle\alpha|X^\dagger \longleftrightarrow X|\alpha\rangle. \quad (1.40)$$

Se $X = X^\dagger$ allora diciamo che X è un operatore *hermitiano* (o *autoaggiunto*¹⁴). Gli operatori hermitiani avranno un ruolo fondamentale nella teoria, come si vedrà tra non molto.

Va detto, per correttezza, che qui stiamo sorvolando su alcuni aspetti formali non secondari. Ad esempio, non distinguiamo tra osservabili rappresentate da operatori limitati o non limitati (con autovalori di modulo grande a piacere), né ci occupiamo del loro dominio (ad esempio, se il dominio su cui X è definito sia lo stesso di $X^\dagger$). Questi e altri aspetti formali della teoria potranno essere approfonditi a tempo debito ma, va detto, non hanno grande impatto dal punto di vista della fisica, almeno non al livello che ci riguarda. La definizione che abbiamo dato di operatori hermitiani è quanto ci basta per arrivare a delineare il quadro complessivo senza farci fuorviare dai dettagli.

Possiamo definire i prodotti di operatori. Siano X e Y due operatori. Il loro prodotto XY agisce su un vettore $|\alpha\rangle$ in questo modo: prima si applica Y da sinistra al vettore $|\alpha\rangle$ e poi si applica X da sinistra al vettore $Y|\alpha\rangle$. In simboli, possiamo rappresentare il prodotto in questo modo

$$XY|\alpha\rangle. \quad (1.41)$$

Data questa definizione di prodotto, possiamo anche definire l'inverso di un operatore X , che indichiamo con il simbolo X^{-1} , tale che l'azione successiva dell'operatore e del suo inverso lasci lo stato invariato: $XX^{-1} = X^{-1}X = \mathbb{1}$, dove $\mathbb{1}$ è l'operatore identità.

Il prodotto è sempre associativo

$$XYZ|\alpha\rangle = X(YZ)|\alpha\rangle = (XY)Z|\alpha\rangle, \quad (1.42)$$

ma in generale non è commutativo

$$XY \neq YX \quad (1.43)$$

¹⁴Per essere precisi, i termini hermitiano e autoaggiunto non sono del tutto sinonimi; per essere autoaggiunto un operatore non solo deve coincidere con il suo hermitiano coniugato, ma i due operatori devono anche avere lo stesso dominio, inteso come l'insieme dei *ket* su cui agiscono. Ma si tratta di sottigliezze che non sono rilevanti a questo livello.

e l'ordine dei fattori è rilevante. È utile in proposito definire un nuovo operatore, chiamato *commutatore* di X e Y e indicato con il simbolo $[X, Y]$, tale che

$$[X, Y] = XY - YX. \quad (1.44)$$

Due operatori commutano se $[X, Y] = 0$.

Per definizione, un operatore commuta sempre con se stesso,

$$[X, X] = 0, \quad (1.45)$$

e commuta con un qualsiasi numero k :

$$[X, k] = 0. \quad (1.46)$$

Inoltre vale

$$[X, Y] = -[Y, X]. \quad (1.47)$$

È anche facile verificare che

$$[X, Y + Z] = [X, Y] + [X, Z] \quad (1.48)$$

e

$$[X, YZ] = Y[X, Z] + [X, Y]Z. \quad (1.49)$$

Con un po' di calcoli in più, ma altrettanto facili, si dimostra la cosiddetta *identità di Jacobi*:

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0. \quad (1.50)$$

Notiamo poi che l'operatore prodotto XY ammette un suo operatore hermitiano coniugato, definito da

$$\langle \alpha | (XY)^\dagger \longleftrightarrow XY | \alpha \rangle. \quad (1.51)$$

Ma, avendo definito il prodotto come applicazione di Y e X in sequenza, possiamo anche scrivere

$$\langle \alpha | Y^\dagger X^\dagger = (\langle \alpha | Y^\dagger) X^\dagger \longleftrightarrow X(Y | \alpha) = XY | \alpha \rangle \quad (1.52)$$

e, combinando le due relazioni, abbiamo

$$(XY)^\dagger = Y^\dagger X^\dagger \quad (1.53)$$

dove dobbiamo stare attenti a rispettare l'ordine degli operatori. Da ciò segue anche che l'hermitiano coniugato del commutatore $[X, Y]$ è

$$[X, Y]^\dagger = (XY)^\dagger - (YX)^\dagger = Y^\dagger X^\dagger - X^\dagger Y^\dagger = [Y^\dagger, X^\dagger]. \quad (1.54)$$

Se X e Y sono entrambi hermitiani (cioè $X^\dagger = X$ e $Y^\dagger = Y$), allora

$$[X, Y]^\dagger = [Y^\dagger, X^\dagger] = [Y, X] = -[X, Y] \quad (1.55)$$

e, in questo caso, si dice che il commutatore è un operatore *anti-hermitiano*.

1.4.5 Prodotto esterno

A questo punto siamo pronti per definire anche il *prodotto esterno* di vettori di stato. Lo definiamo così:

$$|\beta\rangle\langle\alpha| \quad (1.56)$$

ed è un operatore. Per convincerci che si tratta di un operatore, basta applicarlo ad un altro vettore

$$(|\beta\rangle\langle\alpha|)|\gamma\rangle \quad (1.57)$$

e usare l'associatività dei prodotti tra vettori (e loro duali)

$$(|\beta\rangle\langle\alpha|)|\gamma\rangle = |\beta\rangle(\langle\alpha|\gamma\rangle) . \quad (1.58)$$

A destra abbiamo un vettore, $|\beta\rangle$, moltiplicato per un prodotto interno di altri due vettori. Il prodotto interno è un numero e quindi il risultato complessivo è un vettore. Come risultato dell'applicazione del prodotto esterno ad un vettore otteniamo un vettore, che significa che il prodotto esterno è un operatore. Prodotto interno ed esterno sono gli analoghi del prodotto scalare e vettoriale di vettori nello spazio delle coordinate euclideo.

L'associatività estesa a tutti i vettori e operatori è un postulato importante della teoria; viene chiamato *assioma associativo della moltiplicazione*. Nella pratica, ci permette di omettere tutte le parentesi nei prodotti. Ad esempio possiamo scrivere

$$(\langle\beta|)(X|\alpha\rangle) = (\langle\beta|X)|\alpha\rangle = \langle\beta|X|\alpha\rangle , \quad (1.59)$$

dove nell'ultima espressione il risultato è lo stesso sia che X venga inteso agire a destra su $|\alpha\rangle$ o a sinistra su $\langle\beta|$. D'altra parte, la prima espressione a sinistra è un prodotto interno al quale possiamo applicare la proprietà (1.22) per scrivere

$$(\langle\beta|)(X|\alpha\rangle) = [(\langle\alpha|X^\dagger)(|\beta\rangle)]^* , \quad (1.60)$$

ovvero, togliendo le parentesi tra i prodotti,

$$\langle\beta|X|\alpha\rangle = \langle\alpha|X^\dagger|\beta\rangle^* . \quad (1.61)$$

Se X è un operatore hermitiano ($X = X^\dagger$) allora

$$\langle\beta|X|\alpha\rangle = \langle\alpha|X|\beta\rangle^* \quad (1.62)$$

e questo implica che la quantità $\langle\alpha|X|\alpha\rangle$ è un numero reale.

Per inciso, si capisce a questo punto l'origine del neologismo *bra* e *ket* introdotto da Dirac: quando si estraggono numeri dai vettori di stato, come nel caso dell'ultima espressione qui sopra, oppure del prodotto interno, si finisce sempre per avere espressioni della forma $\langle\dots\rangle$ e le due parole *bra* e *ket* non fanno altro che suggerire una sorta di parentesi (*bracket*) contenente le informazioni rilevanti sul sistema.

Infine, torniamo alla definizione di prodotto esterno e chiamiamo $X = |\beta\rangle\langle\alpha|$. È facile dimostrare che l'hermitiano coniugato di X è $X^\dagger = |\alpha\rangle\langle\beta|$. Infatti, se applichiamo X ad uno stato generico $|\gamma\rangle$ otteniamo il vettore

$$X|\gamma\rangle = |\beta\rangle\langle\alpha|\gamma\rangle \quad (1.63)$$

e, per definizione, il suo duale è $\langle\gamma|X^\dagger$. Dunque

$$\langle\gamma|X^\dagger = \langle\alpha|\gamma\rangle^*\langle\beta| = \langle\gamma|\alpha\rangle\langle\beta| = \langle\gamma|(|\alpha\rangle\langle\beta|) \quad (1.64)$$

come si voleva dimostrare.

1.5 Stati di base e rappresentazione matriciale

1.5.1 Vettori e basi

Vogliamo che le osservabili fisiche siano rappresentate da operatori lineari che agiscono sui vettori di stato. Vogliamo anche che gli autovalori degli operatori associati alle osservabili siano numeri reali, perché desideriamo identificarli con i possibili esiti della misura. A questo scopo, gli operatori devono essere hermitiani. Infatti, con le definizioni e le proprietà introdotte nella sezione precedente, è possibile dimostrare il seguente

TEOREMA: gli autovalori di un operatore hermitiano sono reali e gli autostati corrispondenti ad autovalori distinti sono ortogonali.

Lo possiamo dimostrare in questo modo¹⁵. Prendiamo un operatore hermitiano A e due suoi autostati generici $|a_1\rangle$ e $|a_2\rangle$ con autovalori rispettivamente a_1 e a_2 e con norma 1. Per definizione

$$A|a_1\rangle = a_1|a_1\rangle \quad (1.65)$$

e

$$A|a_2\rangle = a_2|a_2\rangle. \quad (1.66)$$

Della seconda prendiamo il duale, tenendo conto che A è hermitiano:

$$\langle a_2|A = a_2^*\langle a_2|. \quad (1.67)$$

Poi moltiplichiamo la prima a sinistra per $\langle a_2|$ e l'ultima a destra per $|a_1\rangle$ e ne facciamo la differenza:

$$\langle a_2|A|a_1\rangle - \langle a_2|A|a_1\rangle = a_1\langle a_2|a_1\rangle - a_2^*\langle a_2|a_1\rangle \quad (1.68)$$

e quindi

$$0 = (a_1 - a_2^*)\langle a_2|a_1\rangle. \quad (1.69)$$

Ora, se $a_1 = a_2 = a$, allora $\langle a_2|a_1\rangle = \langle a|a\rangle = 1$ e deve valere $a - a^* = 0$, che implica a reale. Se invece $a_1 \neq a_2$, allora $a_1 - a_2^* \neq 0$ e deve valere $\langle a_2|a_1\rangle = 0$, ovvero gli autostati sono ortogonali. Come si voleva dimostrare.

L'insieme degli autostati di un operatore hermitiano, normalizzati e ortogonali tra loro, costituisce una *base ortonormale e completa* dello spazio dei vettori di stato, e l'insieme degli autovalori è detto *spettro* dell'operatore. La completezza segue dal fatto che la dimensione dello spazio vettoriale coincide, per costruzione, con il numero di esiti che si possono ottenere nella misura dell'osservabile scelta per caratterizzare gli stati fisici del sistema¹⁶. Occorre però prestare attenzione ad un fatto: la dimostrazione precedente non esclude la possibilità che ad uno stesso autovalore di un'osservabile corrispondano più autovettori linearmente indipendenti. In tal caso si dice che lo spettro è *degenere*. Questi vettori che condividono uno stesso autovalore formano un sottospazio dello spazio di Hilbert complessivo, e ogni loro combinazione lineare è ortogonale a tutti gli autovettori corrispondenti ad autovalori diversi. La determinazione di una base ortonormale per un sottospazio degenere qualsiasi è una procedura semplice, che affronteremo se e quando necessario.

¹⁵Qui limitiamo la dimostrazione al caso di un insieme discreto di autovalori. Nel caso di un insieme continuo, le cose sono un po' più complicate; affronteremo questo problema più avanti in questo capitolo.

¹⁶Anche qui stiamo tralasciando alcuni dettagli matematici. In particolare, se è pur vero che per gli spazi di Hilbert di dimensione finita gli autovettori di un operatore hermitiano generano lo spazio stesso (cioè ne costituiscono una base completa), nel caso di spazi di dimensione infinita la cosa è meno ovvia; in alcuni casi è dimostrabile, in altri no. Per garantire la coerenza interna della teoria conviene accettarlo come un assioma.

In uno spazio i cui vettori di base, $\{|a_1\rangle, |a_2\rangle, |a_3\rangle, \dots\}$, autovettori di un operatore hermitiano A , sono numerabili, ogni generico vettore di stato può essere scritto nella forma

$$|\alpha\rangle = \sum_j c_j |a_j\rangle, \quad (1.70)$$

dove l'indice j corre su tutti gli autovettori (finiti o infiniti). I coefficienti della decomposizione possono essere calcolati esplicitamente. Basta moltiplicare la relazione precedente a sinistra per il *bra* $\langle a_i|$, in modo che

$$\langle a_i|\alpha\rangle = \sum_j c_j \langle a_i|a_j\rangle = c_i \quad (1.71)$$

dove abbiamo usato l'ortonormalità della base, per cui $\langle a_i|a_j\rangle = \delta_{ij}$. La relazione che abbiamo trovato vale per tutti i coefficienti della decomposizione di $|\alpha\rangle$. I coefficienti c_j ci permettono di rappresentare univocamente il vettore di stato come un insieme di numeri complessi. Possiamo riscrivere la stessa decomposizione in questo modo:

$$|\alpha\rangle = \sum_j (\langle a_j|\alpha\rangle) |a_j\rangle = \sum_j |a_j\rangle (\langle a_j|\alpha\rangle) = \sum_j |a_j\rangle \langle a_j|\alpha\rangle. \quad (1.72)$$

Questa espressione dovrebbe ricordarci la decomposizione dei vettori nello spazio delle coordinate dove, avendo individuato un insieme di versori ortogonali $\{\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \dots\}$, ogni vettore $\mathbf{V}$ può essere scritto come somma di ciascun versore volte la proiezione del vettore in quella direzione, la proiezione essendo data dal prodotto scalare:

$$\mathbf{V} = \sum_j \hat{\mathbf{u}}_j (\hat{\mathbf{u}}_j \cdot \mathbf{V}). \quad (1.73)$$

Ma l'assioma associativo della moltiplicazione ci permette di interpretare la stessa decomposizione di $|\alpha\rangle$ anche in questo modo equivalente:

$$|\alpha\rangle = \sum_j |a_j\rangle \langle a_j|\alpha\rangle = \left(\sum_j |a_j\rangle \langle a_j| \right) |\alpha\rangle \quad (1.74)$$

ed essendo $|\alpha\rangle$ un vettore del tutto generico, concludiamo che

$$\sum_j |a_j\rangle \langle a_j| = \mathbf{1} \quad (1.75)$$

dove $\mathbf{1}$ è l'operatore identità. Questa espressione è nota come *relazione di completezza* oppure *relazione di chiusura*. Vedremo più avanti come estendere questi concetti al caso di uno spazio infinito dimensionale continuo.

La relazione di completezza è molto comoda. Possiamo inserirla in altre espressioni per ricavare risultati utili. Un esempio è il seguente. Prendiamo la norma quadra di $|\alpha\rangle$ che per definizione è $\langle \alpha|\alpha\rangle$. Poi, tra il *bra* e il *ket* possiamo infilare l'identità (che non cambia nulla) in questo modo:

$$\langle \alpha|\alpha\rangle = \langle \alpha| \left(\sum_j |a_j\rangle \langle a_j| \right) |\alpha\rangle. \quad (1.76)$$

Inserendo il vettore $|\alpha\rangle$ e il suo duale nella sommatoria e togliendo le parentesi, troviamo

$$\langle \alpha|\alpha\rangle = \sum_j \langle \alpha|a_j\rangle \langle a_j|\alpha\rangle = \sum_j c_j^* c_j = \sum_j |c_j|^2. \quad (1.77)$$

Dunque scopriamo che se un vettore di stato ha norma 1 allora

$$\sum_j |c_j|^2 = 1. \quad (1.78)$$

A questo punto vi potrebbero venire in mente le espressioni (1.7) e (1.11) per la decomposizione degli autostati dello spin lungo x e lungo y nella base degli autostati dello spin lungo z . I moduli quadri dei coefficienti in quelle espressioni erano tutti $1/2$ e la somma dava 1, consistentemente con quanto appena trovato. In quel caso, inoltre, riuscivamo ad attribuire un significato fisico al modulo quadro del coefficiente: esprime la frazione di atomi che vengono deflessi nell'una o nell'altra direzione dall'apparato di Stern-Gerlach.

Un'altra relazione interessante è questa:

$$(|a_j\rangle\langle a_j|) |\alpha\rangle = |a_j\rangle\langle a_j|\alpha\rangle = c_j |a_j\rangle. \quad (1.79)$$

Qui vediamo che l'operatore $|a_j\rangle\langle a_j|$ applicato al vettore $|\alpha\rangle$ ne seleziona la componente nella direzione individuata dal vettore di base $|a_j\rangle$. È un operatore di *proiezione*. Lo indichiamo con il simbolo Λ_j :

$$\Lambda_j = |a_j\rangle\langle a_j| \quad (1.80)$$

e la relazione di completezza diventa

$$\sum_j \Lambda_j = \mathbb{1}. \quad (1.81)$$

Se, una volta proiettato un vettore, lo proiettiamo di nuovo nella stessa direzione, il risultato non deve cambiare. Questo equivale alla condizione $\Lambda_j^2 = \Lambda_j$ che deve valere per tutti i proiettori. Si verifica facilmente anche nel nostro caso. Infatti

$$\Lambda_j^2 = \Lambda_j \Lambda_j = |a_j\rangle\langle a_j|a_j\rangle\langle a_j| = |a_j\rangle\langle a_j| = \Lambda_j, \quad (1.82)$$

essendo $\langle a_j|a_j\rangle = 1$.

1.5.2 Operatori e matrici

Abbiamo visto come si possono decomporre i vettori di stato usando la base ortonormale e completa degli autostati di un operatore hermitiano. Ora vediamo come possiamo rappresentare gli operatori usando la stessa base. Prendiamo un generico operatore X e gli mettiamo a sinistra e a destra l'operatore identità espresso tramite la relazione di completezza,

$$X = \left(\sum_i |a_i\rangle\langle a_i| \right) X \left(\sum_j |a_j\rangle\langle a_j| \right), \quad (1.83)$$

poi togliamo le parentesi sfruttando l'assioma associativo delle moltiplicazioni e scriviamo

$$X = \sum_{ij} |a_i\rangle\langle a_i|X|a_j\rangle\langle a_j|. \quad (1.84)$$

Si tratta di una doppia sommatoria su tutti gli stati di base. In ciascun elemento della sommatoria entra un numero complesso, che possiamo indicare con il simbolo X_{ij} e che è dato da $X_{ij} = \langle a_i|X|a_j\rangle$. Si tratta quindi di N^2 numeri complessi, se la dimensione dello spazio (numerabile) è N . Li possiamo mettere in forma di matrice, usando i come indice di

riga e j come indice di colonna. Otteniamo così la rappresentazione matriciale dell'operatore X

$$X = \begin{pmatrix} \langle a_1|X|a_1\rangle & \langle a_1|X|a_2\rangle & \langle a_1|X|a_3\rangle & \dots \\ \langle a_2|X|a_1\rangle & \langle a_2|X|a_2\rangle & \langle a_2|X|a_3\rangle & \dots \\ \langle a_3|X|a_1\rangle & \langle a_3|X|a_2\rangle & \langle a_3|X|a_3\rangle & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (1.85)$$

dove il segno di uguale sta per “rappresentato da”. Notiamo che il simbolo “=” può avere diversi significati a seconda del contesto. In questo caso specifico va sottolineato che la matrice rappresenta l'operatore, non lo definisce. L'operatore viene definito per come agisce sui vettori e uno stesso operatore, definito in un certo modo, può essere rappresentato da matrici diverse a seconda della base che si sceglie di utilizzare. Ciò non toglie che, una volta concordata una base, la matrice identifica univocamente l'operatore.

Avevamo definito l'hermitiano coniugato di un operatore X in modo tale che $\langle \alpha|X^\dagger$ è il duale di $X|\alpha\rangle$. Per questo possiamo scrivere

$$X_{ij}^\dagger = \langle a_i|X^\dagger|a_j\rangle = (\langle a_i|X^\dagger)|a_j\rangle = [\langle a_j|(X|a_i)\rangle]^* = \langle a_j|X|a_i\rangle^* = X_{ji}^*. \quad (1.86)$$

Dunque, la matrice che rappresenta $X^\dagger$ è la trasposta coniugata di quella che rappresenta X . Inoltre, se X è hermitiano, cioè $X = X^\dagger$, allora

$$X_{ij} = X_{ij}^\dagger = X_{ji}^* \quad (1.87)$$

e questo implica che gli elementi diagonali della matrice sono tutti reali.

La rappresentazione dei vettori con n -ple di numeri e degli operatori con matrici quadre ci permette di usare tutte le regole dell'algebra lineare. Ad esempio, l'operatore Z , definito come il prodotto di altri due operatori, $Z = XY$, ha come elementi di matrice

$$Z_{ij} = \langle a_i|Z|a_j\rangle = \langle a_i|XY|a_j\rangle. \quad (1.88)$$

Inserendo la relazione di completezza tra X e Y , otteniamo

$$Z_{ij} = \langle a_i|X \left(\sum_k |a_k\rangle\langle a_k| \right) Y|a_j\rangle = \sum_k \langle a_i|X|a_k\rangle\langle a_k|Y|a_j\rangle, \quad (1.89)$$

ma questa altro non è che la solita regola di moltiplicazione di matrici riga per colonna

$$Z_{ij} = \sum_k X_{ik} Y_{kj}. \quad (1.90)$$

Un altro esempio è l'azione di un operatore X su un vettore di stato $|\alpha\rangle$ che produce un nuovo vettore di stato $|\gamma\rangle$. Possiamo scrivere

$$|\gamma\rangle = X|\alpha\rangle, \quad (1.91)$$

poi possiamo moltiplicare a sinistra per un vettore della base e introdurre la relazione di completezza in questo modo

$$\langle a_i|\gamma\rangle = \langle a_i|X \left(\sum_j |a_j\rangle\langle a_j| \right) |\alpha\rangle = \sum_j \langle a_i|X|a_j\rangle\langle a_j|\alpha\rangle. \quad (1.92)$$

Se rappresentiamo i due vettori $|\alpha\rangle$ e $|\gamma\rangle$ come vettori colonna usando i coefficienti della loro decomposizione sugli stati di base

$$|\alpha\rangle = \begin{pmatrix} \langle a_1|\alpha\rangle \\ \langle a_2|\alpha\rangle \\ \langle a_3|\alpha\rangle \\ \vdots \end{pmatrix} \quad |\gamma\rangle = \begin{pmatrix} \langle a_1|\gamma\rangle \\ \langle a_2|\gamma\rangle \\ \langle a_3|\gamma\rangle \\ \vdots \end{pmatrix} \quad (1.93)$$

allora la relazione precedente diventa l'usuale prodotto tra matrice e vettore, riga per colonna.

Possiamo usare la stessa procedura per scrivere i vettori duali, ma scrivendoli come vettori riga anziché colonna. Ad esempio

$$\langle\beta| = (\langle\beta|a_1\rangle, \langle\beta|a_2\rangle, \langle\beta|a_3\rangle, \dots) = (\langle a_1|\beta\rangle^*, \langle a_2|\beta\rangle^*, \langle a_3|\beta\rangle^*, \dots) , \quad (1.94)$$

allora il prodotto interno $\langle\beta|\alpha\rangle$ diventa

$$\langle\beta|\alpha\rangle = \langle\beta| \left(\sum_j |a_j\rangle \langle a_j| \right) |\alpha\rangle = \sum_j \langle\beta|a_j\rangle \langle a_j|\alpha\rangle \quad (1.95)$$

oppure

$$\langle\beta|\alpha\rangle = \sum_j \langle a_j|\beta\rangle^* \langle a_j|\alpha\rangle \quad (1.96)$$

che è il solito prodotto riga per colonna di due vettori scritti come n -ple di numeri.

È facile verificare che, se A è l'operatore hermitiano i cui autostati costituiscono la base scelta per lo spazio vettoriale, allora la matrice che rappresenta A è diagonale e gli elementi sulla diagonale sono proprio gli autovalori dell'operatore. Infatti

$$A_{ij} = \langle a_i|A|a_j\rangle = \langle a_i|a_j|a_j\rangle = a_j \langle a_i|a_j\rangle = a_j \delta_{ij} \quad (1.97)$$

e quindi tutti gli elementi fuori diagonale sono nulli e quelli sulla diagonale sono gli a_j . Giusto per abituarsi alla notazione, insistiamo sul significato dei simboli. Nella scrittura $\langle a_i|a_j|a_j\rangle$, i due simboli a_i e a_j laterali non sono numeri, sono solo etichette che usiamo per nominare i vettori; potremmo usare al loro posto delle faccine colorate, purché si sappia a quale faccina corrisponde quale autovalore. Il simbolo a_j centrale invece è un numero reale, cioè l'autovalore di A corrispondente al vettore di destra.

Possiamo anche usare la relazione di completezza assieme all'hermitianità di A per scrivere

$$A = \left(\sum_j |a_j\rangle \langle a_j| \right) A = \sum_j |a_j\rangle \langle a_j|A = \sum_j a_j |a_j\rangle \langle a_j|. \quad (1.98)$$

che è un modo diverso di rappresentare l'operatore usando i suoi autostati. Dunque ogni operatore hermitiano può essere scritto come somma dei proiettori corrispondenti ai suoi autostati, con coefficienti dati dagli autovalori. La somma (1.98) è chiamata *decomposizione spettrale* dell'operatore.

1.5.3 L'esempio dello spin

Per abituarci al formalismo appena sviluppato torniamo all'esempio iniziale dello spin dell'atomo di argento. Un'osservabile è la componente z dello spin, S_z . Come sappiamo dagli esperimenti con gli apparati di Stern-Gerlach, la misura di S_z fornisce solo due possibili

valori, $\pm\hbar/2$. Un atomo che venga deflesso verso l'alto da un gradiente di campo magnetico positivo lungo z ha spin $+\hbar/2$ e quello deflesso verso il basso ha spin $-\hbar/2$. Nella sezione 1.3 avevamo scritto i corrispondenti vettori di stato come $|S_z \uparrow\rangle$ e $|S_z \downarrow\rangle$. Assumiamo che abbiano norma 1 e li usiamo come base per lo spazio degli stati di spin. Per semplicità di notazione, omettiamo il simbolo S_z dai *ket* di base, indicandoli semplicemente come $|\uparrow\rangle$ e $|\downarrow\rangle$, con S_z sottinteso. L'operatore S_z è un operatore hermitiano per cui vale la relazione agli autovalori

$$S_z|\uparrow\rangle = \frac{\hbar}{2}|\uparrow\rangle \quad , \quad S_z|\downarrow\rangle = -\frac{\hbar}{2}|\downarrow\rangle . \quad (1.99)$$

La relazione di completezza è

$$|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow| = \mathbf{1} . \quad (1.100)$$

I due vettori di base possono essere scritti come vettori colonna

$$|\uparrow\rangle = \begin{pmatrix} \langle\uparrow|\uparrow\rangle \\ \langle\downarrow|\uparrow\rangle \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (1.101)$$

e

$$|\downarrow\rangle = \begin{pmatrix} \langle\uparrow|\downarrow\rangle \\ \langle\downarrow|\downarrow\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \quad (1.102)$$

La matrice che rappresenta S_z è

$$S_z = \begin{pmatrix} \langle\uparrow|S_z|\uparrow\rangle & \langle\uparrow|S_z|\downarrow\rangle \\ \langle\downarrow|S_z|\uparrow\rangle & \langle\downarrow|S_z|\downarrow\rangle \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \quad (1.103)$$

e le relazioni (1.99) possono essere scritte nella forma

$$\frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \quad (1.104)$$

Altre osservabili sono le due componenti S_x e S_y . Le matrici che rappresentano questi operatori hermitiani le ricaveremo più avanti. Invece mostriamo qui un esempio di operatori non hermitiani. Consideriamo gli operatori

$$S_+ = \hbar|\uparrow\rangle\langle\downarrow| \quad (1.105)$$

e

$$S_- = \hbar|\downarrow\rangle\langle\uparrow| . \quad (1.106)$$

Essi agiscono in questo modo sugli stati di base

$$\begin{aligned} S_+|\uparrow\rangle &= \hbar|\uparrow\rangle\langle\downarrow|\uparrow\rangle = 0 \\ S_+|\downarrow\rangle &= \hbar|\uparrow\rangle\langle\downarrow|\downarrow\rangle = \hbar|\uparrow\rangle \\ S_-|\uparrow\rangle &= \hbar|\downarrow\rangle\langle\uparrow|\uparrow\rangle = \hbar|\downarrow\rangle \\ S_-|\downarrow\rangle &= \hbar|\downarrow\rangle\langle\uparrow|\downarrow\rangle = 0 . \end{aligned} \quad (1.107)$$

In pratica, S_+ alza lo spin se era basso e S_- lo abbassa se era alto, altrimenti danno un vettore nullo. Si chiamano operatori di *spin flip*. La loro rappresentazione matriciale è

$$S_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad , \quad S_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} . \quad (1.108)$$

Essendo matrici reali e l'una la trasposta dell'altra, i due operatori sono anche l'uno hermitiano coniugato dell'altro. Operatori del genere possono sembrare strani, ma chi usa per

mestiere la meccanica quantistica li conosce bene perché capita di incontrarli spesso. Si pensi ad esempio ad un sistema composto da una particella che può stare in due soli stati di energia diversa, come potrebbe essere nel caso della soluzione dell'equazione di Schrödinger in una qualche buca di potenziale. Se la particella si trova nello stato più basso e riceve dall'esterno l'energia giusta, può saltare nello stato più alto; se è già sopra non succede nulla. Invece, se la particella si trova nello stato più alto e riesce a cedere all'esterno l'energia giusta, può cadere nello stato in basso. È ragionevole pensare che i meccanismi fisici responsabili per l'assorbimento e l'emissione di energia in un *sistema a due livelli* di questo tipo trovino una rappresentazione in operatori non hermitiani simili a quelli di *spin flip*. In effetti, uno dei motivi per cui è comodo usare gli stati di spin per iniziare a ragionare sulla meccanica quantistica sta nel fatto che poi lo stesso formalismo lo si ritrova in tutti i sistemi a due livelli, che sono un po' ovunque nella fisica moderna e che stanno alla base del concetto moderno di *qubit*.

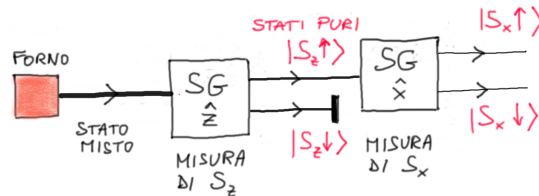
1.6 Definizione di misura

1.6.1 Misura

Tutta la fisica si basa sull'idea che si possano stabilire connessioni logiche tra gli esiti di osservazioni empiriche. Il primo passo, come ci ha insegnato Galileo, consiste nell'estrarre numeri dalle osservazioni usando opportuni strumenti di misura. Gli strumenti di misura sono apparati, più o meno complicati, il cui *output* può essere dato dalla posizione di un ago su un quadrante graduato, o dalle cifre visualizzate su un *display* di un dispositivo elettronico, o da una serie di *bit* memorizzati in un computer, o in altra forma, ma sempre traducibile in un numero reale seguito dalle opportune unità. Le grandezze fisiche da misurare in un dato sistema sono definite tramite le procedure utilizzate per misurarle e coinvolgono il confronto, diretto o indiretto, con campioni di misura convenzionalmente scelti. Per sviluppare una teoria fisica è dunque essenziale chiarire, nel modo più rigoroso possibile, cosa si intenda per *misura* nel contesto formale della teoria stessa. Finora su questo siamo stati ancora troppo vaghi. Ripartiamo dallo spin e proviamo a delineare meglio i nostri ragionamenti.

Gli atomi d'argento hanno una proprietà, che abbiamo chiamato spin, che possiamo misurare con gli apparati di Stern-Gerlach. In particolare, un apparato con un gradiente di campo magnetico lungo z misura la componente S_z dello spin; la misura delle altre componenti si ottiene cambiando la disposizione e l'orientamento dello stesso apparato. Per ogni componente si possono ottenere solo due valori $\pm\hbar/2$ come esiti della misura, corrispondenti a deflessioni opposte della traiettoria degli atomi. Gli atomi escono dal forno con spin casuale. Questo lo vediamo scegliendo tutte le orientazioni possibili del fascio in uscita dal forno e tutte le orientazioni possibili del primo apparato di Stern-Gerlach che gli atomi incontrano: l'apparato produrrà sempre due macchie uguali. Questo implica che la distribuzione statistica dello spin degli atomi è isotropa. Dunque, il primo Stern-Gerlach produce due sole macchie per ragioni legate al comportamento quantistico degli stati di spin, ma le produce uguali per una mera ragione di probabilità classica. Semplicemente il forno produce un *ensemble* statistico di atomi in cui ciascun atomo ha la stessa probabilità di avere lo spin S_z in su o in giù. Potremmo chiamare tali probabilità $w_\uparrow$ e $w_\downarrow$ e, in questo caso, varrebbero entrambe $1/2$. Un ricercatore distratto, che avesse dimenticato qualche magnete vicino al forno, potrebbe indurre una anisotropia nella distribuzione degli spin al momento della produzione del fascio di atomi, variando i valori delle probabilità $w_\uparrow$ e $w_\downarrow$: le macchie sullo schermo avrebbero così intensità diverse. Ma non ci sarebbe nulla di quantistico in questa variazione di intensità, trattandosi puramente di una modifica nella distribuzione statistica degli spin nell'*ensemble*, spiegabile con la fisica classica. Stati fisici come questi

si chiamano *stati misti* o *ensemble misti* e l'esito della misura di un'osservabile eseguita su uno stato misto dipende sia dalla natura quantistica del sistema (quali stati sono ammessi per quell'osservabile in quel sistema), sia dalla distribuzione statistica, in senso classico, che caratterizza l'insieme delle repliche dello stesso sistema (con quale frequenza un dato stato, tra quelli ammessi, compare nell'ensemble).



All'uscita del primo apparato di Stern-Gerlach, incluso lo schermo con il foro, ciascun atomo si trova o nello stato $|S_z \uparrow\rangle$ (se il foro è in alto) o nello stato $|S_z \downarrow\rangle$ (se il foro è in basso). Questi sono *stati puri*, cioè sono stati in cui l'atomo ha una ben definita orientazione dello spin lungo z . Lo sappiamo, perché possiamo esaminare ciascuno dei due fasci con un altro SG $\hat{z}$ e, in questo caso, troviamo che tutti gli atomi nello stato $|S_z \uparrow\rangle$ hanno spin $\hbar/2$ (sono deflessi in su) e tutti quelli nello stato $|S_z \downarrow\rangle$ hanno spin $-\hbar/2$ (sono deflessi in giù). Dunque, il primo apparato, bloccando una delle due uscite, trasforma lo stato *misto* d'ingresso in stati *puri*, autostati di S_z , e il secondo apparato opera una misura di S_z che non altera lo stato di spin in cui l'atomo si trova.

Ora veniamo alla parte più difficile. Se il secondo apparato è uno SG $\hat{x}$ e gli diamo in pasto il fascio di atomi preventivamente selezionati nello stato puro $|S_z \uparrow\rangle$, come in figura, l'esito è la produzione di due macchie di eguale intensità corrispondenti agli stati puri $|S_x \uparrow\rangle$ e $|S_x \downarrow\rangle$. L'analogia con la polarizzazione della luce ci aveva portato ad affermare che gli stati di spin lungo x e y possono essere scritti come combinazioni lineari degli stati $|S_z \uparrow\rangle$ e $|S_z \downarrow\rangle$ e che i coefficienti della combinazione devono contenere informazioni sull'intensità delle macchie. A macchie uguali avevamo fatto corrispondere coefficienti uguali in modulo. Per formalizzare questa idea in modo più rigoroso, dobbiamo chiarirci ancora le idee sul significato degli stati fisici.

Sottolineiamo dapprima il fatto che uno stato $|S_z \uparrow\rangle$ o $|S_z \downarrow\rangle$, o gli altri stati puri di spin, non descrivono lo stato del fascio di atomi, ma quello di ogni singolo atomo. In altri termini, esso non determina le proprietà statistiche dell'*ensemble*; ogni atomo si trova precisamente in quello stato. Potremmo far passare un atomo al minuto e le macchie, con un po' di pazienza, si formerebbero sullo schermo allo stesso modo.

Ora mettiamo in evidenza un possibile bivio logico che si rivelerà cruciale e che sta al cuore stesso della teoria. Da una parte potremmo pensare che gli atomi del fascio in uscita dal primo SG $\hat{z}$ e lasciati andare verso il secondo SG $\hat{x}$ siano sì nello stesso stato di S_z , ma che contestualmente abbiano anche un valore fissato dello spin lungo x , metà di loro positivo e metà negativo. In questo modo, il secondo apparato non fa altro che separare gli atomi sulla base di una proprietà che questi avevano prima di misurare S_x . Dall'altra parte invece potremmo pensare che prima di entrare nel secondo apparato SG $\hat{x}$ ciascuno degli atomi che arriva non abbia alcun valore prefissato di S_x ; semplicemente l'atomo non ha questa proprietà. L'apparato SG $\hat{x}$, esso stesso, determina l'esito della misura separando atomi che, a seguito della misura, si trovano nello stato puro $|S_x \uparrow\rangle$ da quelli che si trovano nello stato puro $|S_x \downarrow\rangle$. La prima strada, che sarebbe la più ovvia da seguire sulla base della nostra esperienza di fisici classici, ci viene sbarrata dall'esito di esperimenti con tre apparati di Stern-Gerlach in sequenza. Infatti, se l'apparato separa gli atomi in base a loro proprietà

pre-esistenti, allora un terzo apparato $SG\hat{z}$ aggiunto ai precedenti non dovrebbe mai produrre un fascio in $|S_z \downarrow\rangle$. Invece lo fa. La seconda strada, che appare più nebbiosa, ci offre un appiglio: sappiamo che lo stato di partenza $|S_z \uparrow\rangle$ può essere scritto come combinazione lineare degli autostati di S_x . Possiamo identificare i moduli quadri dei coefficienti con la probabilità che ciascun atomo, dopo la misura di S_x si trovi nello stato con S_x uguale a $+\hbar/2$ o $-\hbar/2$. Dalle combinazioni lineari estratte per analogia con la polarizzazione della luce abbiamo già visto che funziona. Ora ribaltiamo la logica. Ci sbarazziamo dell'analogia, che ci è solo servita per trovare l'imbocco della strada, e trasformiamo la congettura in postulato:

Supponiamo di voler misurare l'osservabile A in un dato sistema fisico preparato in uno stato *puro* $|\alpha\rangle$. Siano $|a_1\rangle, |a_2\rangle, \dots$, gli autostati normalizzati a 1 dell'operatore hermitiano A , corrispondenti agli autovalori reali $a_1, a_2, \dots$. Gli autostati formano una base su cui si può decomporre $|\alpha\rangle$ in questo modo:

$$|\alpha\rangle = \sum_i c_i |a_i\rangle = \sum_i |a_i\rangle \langle a_i | \alpha \rangle, \quad (1.109)$$

dove $c_i = \langle a_i | \alpha \rangle$ sono numeri complessi. Postuliamo che l'atto della misura consista nel far "collassare" (o "precipitare") lo stato del sistema in uno degli autostati di A e che il valore misurato coincida con l'autovalore di A corrispondente. Ad esempio,

$$|\alpha\rangle \longrightarrow |a_j\rangle, \quad (1.110)$$

quando il valore misurato dell'osservabile A è a_j . Inoltre postuliamo che la probabilità $\mathcal{P}$ di ottenere il valore a_j in una serie di misure effettuate sullo stesso sistema preparato nelle stesse condizioni, sia uguale a

$$\mathcal{P}(a_j) = |c_j|^2 = |\langle a_j | \alpha \rangle|^2. \quad (1.111)$$

L'ultima relazione ha efficacia empirica solo quando la serie comprende un gran numero di misure, sufficientemente grande da estrarre una probabilità da una frequenza di eventi. In altri termini, lo stato di un sistema non può essere determinato da una singola misura. Inoltre, bisogna accertarsi che le copie del sistema, nella serie di misure, si trovino sempre nello stesso stato.

Un caso particolare è quello in cui lo stato $|\alpha\rangle$ prima della misura è già un autostato dell'osservabile. In questo caso, la decomposizione di $|\alpha\rangle$ ha un solo coefficiente uguale a 1 e tutti gli altri sono nulli. Ne segue che la misura non cambia lo stato del sistema, ma si limita a fornire l'autovalore, che è noto a priori per come lo stato iniziale è stato preparato. In tutti gli altri casi la misura cambia lo stato del sistema e, in una singola misura, lo stato cambia in modo imprevedibile: non si può sapere in anticipo quale valore dell'osservabile verrà ottenuto; quello che si sa è la probabilità di ottenerlo.

È importante sottolineare che la trasformazione del vettore di stato $|\alpha\rangle$ indotta dal processo di misura di una generica osservabile A non corrisponde all'azione dell'operatore A sullo stato $|\alpha\rangle$. Infatti, in generale, se si applica A ad un vettore $|\alpha\rangle$ non si ottiene un autostato di A (l'unica eccezione è il caso in cui $|\alpha\rangle$ sia già un autostato), mentre la misura lo fa sempre.

Ovviamente la probabilità di ottenere in una singola misura uno qualsiasi dei valori possibili è pari a 1 (si è certi di ottenere almeno uno dei valori ammessi). Ciò si traduce

nella condizione

$$\sum_j \mathcal{P}(a_j) = \sum_j |c_j|^2 = 1 \quad (1.112)$$

che è quanto avevamo scritto in (1.78). A questo punto si capisce anche la necessità del postulato della metrica definita positiva (1.24).

Supponiamo di eseguire una serie di misure di A ripetute sullo stesso sistema preparato ogni volta nello stesso stato $|\alpha\rangle$ per estrarre la media dei valori misurati. Mostriamo che tale media coincide, nel limite di un grande numero di ripetizioni, con la quantità $\langle\alpha|A|\alpha\rangle$. Per vederlo, basta usare la relazione di completezza e l'ortonormalità della base:

$$\begin{aligned} \langle\alpha|A|\alpha\rangle &= \langle\alpha|A\left(\sum_i |a_i\rangle\langle a_i|\right)|\alpha\rangle \\ &= \sum_i \langle\alpha|A|a_i\rangle\langle a_i|\alpha\rangle = \sum_i a_i \langle\alpha|a_i\rangle\langle a_i|\alpha\rangle \\ &= \sum_i |c_i|^2 a_i = \sum_i \mathcal{P}(a_i) a_i. \end{aligned} \quad (1.113)$$

Si tratta quindi di una media pesata dei valori misurabili di A , usando le loro probabilità come peso statistico: è il loro valor medio. La quantità $\langle\alpha|A|\alpha\rangle$ è detta *valore di aspettazione di A* . Basta prendere il valor medio dei valori ottenuti in una serie di misure; se la serie è grande, esso tenderà al valore di aspettazione. Salvo eccezioni, quando scriviamo il valore di aspettazione assumiamo implicitamente che il vettore su cui si fa la media ha norma 1. Possiamo scrivere

$$\langle A \rangle = \langle\alpha|A|\alpha\rangle \quad (1.114)$$

dove l'uguaglianza può essere intesa sia come definizione del simbolo $\langle A \rangle$, sia come eguaglianza tra $\langle\alpha|A|\alpha\rangle$ e il valor medio delle misure di A definito empiricamente a partire dalla distribuzione statistica dei valori misurati.

La quantità $\langle X \rangle = \langle\alpha|X|\alpha\rangle$ può essere calcolata anche per un operatore generico X , non necessariamente un operatore hermitiano associato ad una osservabile, ma in tal caso è un numero complesso e non vale l'interpretazione in termini di media statistica dei valori misurati. La relazione (1.61) ci dice che

$$\langle X \rangle = \langle X^\dagger \rangle^*$$

e inoltre possiamo dimostrare che vale la disuguaglianza

$$\langle X^\dagger X \rangle \geq \langle X^\dagger \rangle \langle X \rangle. \quad (1.115)$$

Per dimostrarla supponiamo che le medie siano calcolate su uno stato generico $|\gamma\rangle$ di norma unitaria. Poi prendiamo la disuguaglianza di Schwarz (1.30), già dimostrata per i vettori e che qui riportiamo per comodità

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \geq |\langle\beta|\alpha\rangle|^2,$$

e sostituiamo $|\alpha\rangle$ con $X|\gamma\rangle$ e $|\beta\rangle$ con $\langle X \rangle|\gamma\rangle$, tenendo conto che $\langle X \rangle$ è un numero complesso. Otteniamo

$$\langle\gamma|X^\dagger X|\gamma\rangle|\langle X \rangle|^2 \langle\gamma|\gamma\rangle \geq |\langle X \rangle^* \langle\gamma|X|\gamma\rangle|^2 \quad (1.116)$$

ovvero

$$\langle\gamma|X^\dagger X|\gamma\rangle|\langle X \rangle|^2 \geq |\langle X \rangle^* \langle X \rangle|^2 = |\langle X \rangle|^2 |\langle X \rangle|^2 \quad (1.117)$$

da cui

$$\langle\gamma|X^\dagger X|\gamma\rangle \geq |\langle X \rangle|^2 = \langle X \rangle^* \langle X \rangle = \langle X^\dagger \rangle \langle X \rangle, \quad (1.118)$$

che è quanto si voleva. Tra qualche pagina ci tornerà utile.

1.6.2 L'esempio dello spin (ancora una volta)

Ora che abbiamo definito cos'è la misura di un'osservabile in meccanica quantistica, possiamo tornare indietro a rivedere gli stati di spin nell'esperimento di Stern-Gerlach. Stavolta lo facciamo per verificare che possiamo formulare quel problema in modo rigoroso sulla base dei postulati appena enunciati, senza dover più ricorrere all'analogia con la polarizzazione della luce. Dunque, dimentichiamoci le relazioni (1.7) e (1.11) che erano state congetturate in base a quell'analogia, e proviamo a ricavarle soltanto sulla base delle macchie ottenute nelle sequenze di Stern-Gerlach.

Mettiamo come primo apparato un $\text{SG}\hat{x}$, in modo da ottenere in uscita lo stato puro $|S_x \uparrow\rangle$ a partire dallo stato misto in uscita dal forno, o in alternativa lo stato $|S_x \downarrow\rangle$ invertendo il blocco e il foro nello schermo. Prendiamo un secondo apparato, stavolta un $\text{SG}\hat{z}$ per misurare S_z . Anche qui, come avevamo fatto in precedenza, usiamo la notazione sintetica $|\uparrow\rangle$ e $|\downarrow\rangle$ per questi due autostati di S_z . Se facciamo l'esperimento troviamo che per metà atomi si ottiene il collasso nello stato $|\uparrow\rangle$ e metà nello stato $|\downarrow\rangle$. La singola misura è quella sul singolo atomo, mentre la serie di misure identiche corrisponde al comportamento del fascio. Se le macchie sullo schermo sono uguali vuol dire che la probabilità che lo stato $|S_x \uparrow\rangle$ collassi nello stato $|\uparrow\rangle$ è uguale a quella che collassi in $|\downarrow\rangle$. Per il postulato della misura, questo implica che i due coefficienti $\langle \uparrow | S_x \uparrow \rangle$ e $\langle \downarrow | S_x \uparrow \rangle$ devono avere lo stesso modulo quadro, e la somma dei moduli quadri deve dare 1. Questo è vero se i due coefficienti valgono $1/\sqrt{2}$ in modulo e differiscono solo per un fattore di fase. Dunque deve essere

$$|S_x \uparrow\rangle = \frac{1}{\sqrt{2}}|\uparrow\rangle + \frac{1}{\sqrt{2}}e^{i\delta_1}|\downarrow\rangle. \quad (1.119)$$

Lo stesso vale se all'ingresso del secondo apparato prepariamo lo stato $|S_x \downarrow\rangle$: sempre due macchie uguali si ottengono. D'altra parte, questo stato deve essere ortogonale a $|S_x \uparrow\rangle$. L'unica possibilità è che i due stati differiscano per un segno:

$$|S_x \downarrow\rangle = \frac{1}{\sqrt{2}}|\uparrow\rangle - \frac{1}{\sqrt{2}}e^{i\delta_1}|\downarrow\rangle, \quad (1.120)$$

in modo che $\langle S_x \downarrow | S_x \uparrow \rangle = 0$. La scelta di mettere il fattore di fase solo sul secondo addendo è arbitraria e ininfluente, dato che altre scelte differiscono solo per una fase complessiva che non cambia lo stato fisico.

Ora rifacciamo tutto mettendo un $\text{SG}\hat{y}$ come primo apparato. Tutto funziona come prima e troveremo

$$|S_y \uparrow\rangle = \frac{1}{\sqrt{2}}|\uparrow\rangle + \frac{1}{\sqrt{2}}e^{i\delta_2}|\downarrow\rangle \quad (1.121)$$

$$|S_y \downarrow\rangle = \frac{1}{\sqrt{2}}|\uparrow\rangle - \frac{1}{\sqrt{2}}e^{i\delta_2}|\downarrow\rangle, \quad (1.122)$$

con $\delta_2 \neq \delta_1$, da determinare.

Per fissare le fasi ci serve qualche altra informazione. La troviamo usando una sequenza di due Stern-Gerlach $\text{SG}\hat{x}$ e $\text{SG}\hat{y}$. Ad esempio, se usiamo uno $\text{SG}\hat{y}$ per preparare gli atomi nello stato $|S_y \uparrow\rangle$ e li facciamo passare in uno $\text{SG}\hat{x}$, otteniamo le solite due macchie uguali. Dunque, i coefficienti della decomposizione di $|S_y \uparrow\rangle$ sulla base degli autostati di S_x devono avere modulo pari a $1/\sqrt{2}$. Esprimendo gli stessi stati nella base di S_z , la condizione diventa

$$\begin{aligned} \frac{1}{\sqrt{2}} &= |\langle S_x \uparrow | S_y \uparrow \rangle| \\ &= \left| \left(\frac{1}{\sqrt{2}}\langle \uparrow | + \frac{1}{\sqrt{2}}e^{-i\delta_1}\langle \downarrow | \right) \left(\frac{1}{\sqrt{2}}|\uparrow\rangle + \frac{1}{\sqrt{2}}e^{i\delta_2}|\downarrow\rangle \right) \right| \\ &= \frac{1}{2}|1 + e^{i(\delta_2 - \delta_1)}| \end{aligned} \quad (1.123)$$

e

$$\begin{aligned}
\frac{1}{\sqrt{2}} &= |\langle S_x \downarrow | S_y \uparrow \rangle| \\
&= \left| \left(\frac{1}{\sqrt{2}} \langle \uparrow | + \frac{1}{\sqrt{2}} e^{-i\delta_1} \langle \downarrow | \right) \left(\frac{1}{\sqrt{2}} | \uparrow \rangle - \frac{1}{\sqrt{2}} e^{i\delta_2} | \downarrow \rangle \right) \right| \\
&= \frac{1}{2} |1 - e^{i(\delta_2 - \delta_1)}|
\end{aligned} \tag{1.124}$$

da cui

$$\frac{1}{\sqrt{2}} = \frac{1}{2} |1 \pm e^{i(\delta_2 - \delta_1)}| \tag{1.125}$$

che è soddisfatta solo se $\delta_2 - \delta_1 = \pm\pi/2$. Lo stesso risultato si ottiene sostituendo $\langle S_y \uparrow |$ con $\langle S_y \downarrow |$ nelle relazioni precedenti. Dunque, i coefficienti della decomposizione degli autostati di S_x e S_y nella base degli autostati di S_z non possono essere tutti reali. Possiamo scegliere reali quelli degli autostati di S_x , ponendo $\delta_1 = 0$, ma a quel punto spunteranno necessariamente coefficienti immaginari negli autostati di S_y , dato che $\delta_2 = \pm\pi/2$. La scelta tra il $+$ e il $-$ ha a che fare con i sistemi di coordinate destrorsi e sinistrorsi; prendiamo il $+$ e questa è l'unica scelta che qui non giustifichiamo (è quella giusta in un sistema destrorso se si interpreta lo spin come un particolare tipo di momento angolare). Dunque otteniamo

$$\begin{cases} |S_x \uparrow \rangle = \frac{1}{\sqrt{2}} (| \uparrow \rangle + | \downarrow \rangle) \\ |S_x \downarrow \rangle = \frac{1}{\sqrt{2}} (| \uparrow \rangle - | \downarrow \rangle) \end{cases} \tag{1.126}$$

e

$$\begin{cases} |S_y \uparrow \rangle = \frac{1}{\sqrt{2}} (| \uparrow \rangle + i | \downarrow \rangle) \\ |S_y \downarrow \rangle = \frac{1}{\sqrt{2}} (| \uparrow \rangle - i | \downarrow \rangle) \end{cases}, \tag{1.127}$$

che sono le stesse di (1.7) e (1.11)¹⁷.

Possiamo ricavarci l'espressione per gli operatori ricordandoci della (1.98):

$$\begin{aligned}
S_x &= \frac{\hbar}{2} |S_x \uparrow \rangle \langle S_x \uparrow | - \frac{\hbar}{2} |S_x \downarrow \rangle \langle S_x \downarrow | \\
&= \frac{\hbar}{2} \left(\frac{1}{2} (| \uparrow \rangle + | \downarrow \rangle) (\langle \uparrow | + \langle \downarrow |) - \frac{1}{2} (| \uparrow \rangle - | \downarrow \rangle) (\langle \uparrow | - \langle \downarrow |) \right) \\
&= \frac{\hbar}{2} (| \uparrow \rangle \langle \downarrow | + | \downarrow \rangle \langle \uparrow |)
\end{aligned} \tag{1.128}$$

e

$$\begin{aligned}
S_y &= \frac{\hbar}{2} |S_y \uparrow \rangle \langle S_y \uparrow | - \frac{\hbar}{2} |S_y \downarrow \rangle \langle S_y \downarrow | \\
&= \frac{\hbar}{2} \left(\frac{1}{2} (| \uparrow \rangle + i | \downarrow \rangle) (\langle \uparrow | - i \langle \downarrow |) - \frac{1}{2} (| \uparrow \rangle - i | \downarrow \rangle) (\langle \uparrow | + i \langle \downarrow |) \right) \\
&= \frac{\hbar}{2} (-i | \uparrow \rangle \langle \downarrow | + i | \downarrow \rangle \langle \uparrow |) .
\end{aligned} \tag{1.129}$$

E possiamo anche calcolare gli elementi di matrice di S_x e S_y del tipo $\langle \uparrow | S_x | \uparrow \rangle$, $\langle \uparrow | S_x | \downarrow \rangle$, eccetera. Si trova

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.130}$$

¹⁷A meno del segno di $|S_x \downarrow \rangle$, che non cambia la fisica. D'ora in poi teniamo per buono quello ottenuto qui.

dove l'ultima è la solita rappresentazione diagonale per l'osservabile che definisce la base; l'avevamo già vista qualche pagina fa e l'abbiamo aggiunta qui per completezza. Le stesse matrici possono essere scritte nella forma compatta $S_i = (\hbar/2)\sigma_i$, con $i = x, y, z$, e le σ_i sono dette *matrici di Pauli*. Osserviamo che i tre operatori sono hermitiani a vista, dato che le matrici sono uguali a se stesse se trasposte e coniugate. Infine, è facile vedere che gli operatori di *spin flip* introdotti in precedenza, possono essere scritti nella forma

$$S_+ = S_x + iS_y \quad , \quad S_- = S_x - iS_y \quad (1.131)$$

e questi non sono hermitiani.

Le matrici di Pauli sono estremamente utili nella descrizione di qualsiasi sistema a due livelli¹⁸ o, in altri termini, in tutte le situazioni in cui l'informazione sugli stati fisici del sistema è codificabile in *qubit*. Una loro proprietà importante è la regola di commutazione

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \quad (1.132)$$

dove ϵ_{ijk} è il cosiddetto tensore totalmente antisimmetrico, che vale 1 per ϵ_{xyz} e ogni permutazione ciclica dei suoi indici, vale -1 per ϵ_{yxz} e per ogni permutazione ciclica dei suoi indici, vale 0 se due indici sono uguali. La dimostrazione è un semplice calcolo di prodotti di matrici 2×2 . Scritte per gli operatori di spin le regole di commutazione diventano

$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k \quad (1.133)$$

Queste ci torneranno molto utili in seguito.

1.7 Osservabili compatibili e incompatibili

Supponiamo che due osservabili A e B definite per uno stesso sistema fisico ammettano un insieme completo di autostati simultanei. Quando ciò avviene diremo che le osservabili sono *compatibili*, altrimenti sono *incompatibili*. Cosa abbia a che fare la coincidenza degli autostati di due osservabili con la loro compatibilità lo si intuisce da questo fatto: se un autostato di A è anche autostato di B , l'uno con autovalore a e l'altro b nelle rispettive unità di misura allora, per come abbiamo definito la misura, possiamo misurare A , ottenendo un suo autostato, e poi misurare B senza che questa seconda misura cambi lo stato del sistema. In questo modo si ottiene uno stato che corrisponde a valori ben definiti sia dell'osservabile A che dell'osservabile B e l'ordine con cui si eseguono le misure è irrilevante.

Si può dimostrare che due osservabili sono compatibili se e solo se il commutatore degli operatori ad esse associati è nullo

$$[A, B] = AB - BA = 0 \quad (1.134)$$

Per semplicità qui lo dimostriamo nel caso in cui tutti gli autovalori di ciascuna osservabile siano diversi tra loro, in modo che ad ogni autostato corrisponda un valore diverso della misura. In questo caso si dice che gli autovalori sono *non degeneri*, o che lo spettro è *non degenero* (la parola spettro va intesa come *spettro degli autovalori*). Se invece ad un singolo autovalore corrispondono più autostati, come già si è detto qualche pagina fa, si dice che lo spettro è *degenero*. L'estensione della dimostrazione agli spettri degeneri la si trova in quasi tutti i libri di testo e la lasciamo a chi vuole approfondire gli aspetti formali della teoria.

¹⁸Le osservabili in un sistema a due livelli sono rappresentate da matrici hermitiane 2×2 e tutte le matrici hermitiane 2×2 possono essere scritte come combinazioni lineari, con coefficienti reali, delle matrici di Pauli e della matrice identità.

Dunque, prendendo come base quella degli autovettori di un'osservabile A , assumiamo che B commuti con A e scriviamo

$$\langle a_i | [A, B] | a_j \rangle = 0, \quad (1.135)$$

che vale per qualsiasi coppia di autovettori. Questo implica

$$\begin{aligned} 0 &= \langle a_i | (AB - BA) | a_j \rangle \\ &= \langle a_i | AB | a_j \rangle - \langle a_i | BA | a_j \rangle \\ &= a_i \langle a_i | B | a_j \rangle - a_j \langle a_i | B | a_j \rangle \\ &= (a_i - a_j) \langle a_i | B | a_j \rangle. \end{aligned} \quad (1.136)$$

Dato che, per ipotesi, $a_i \neq a_j$ per ogni i diverso da j , allora la relazione precedente implica che $\langle a_i | B | a_j \rangle$ può assumere valori non nulli solo per $i = j$ e possiamo scrivere

$$\langle a_i | B | a_j \rangle = \delta_{ij} \langle a_i | B | a_i \rangle. \quad (1.137)$$

Ne segue che la matrice che rappresenta B è diagonale nella base degli autostati di A . Infatti, usando la (1.84), abbiamo

$$B = \sum_{ij} |a_i\rangle \langle a_i| B | a_j\rangle \langle a_j| = \sum_i |a_i\rangle \langle a_i| B | a_i\rangle \langle a_i|. \quad (1.138)$$

e facendolo agire su un $|a_j\rangle$ qualsiasi troviamo

$$B | a_j \rangle = \sum_i |a_i\rangle \langle a_i| B | a_i\rangle \langle a_i | a_j \rangle = |a_j\rangle \langle a_j | B | a_j \rangle. \quad (1.139)$$

dato che $\langle a_i | a_j \rangle = \delta_{ij}$. A questo punto chiamiamo $b_j = \langle a_j | B | a_j \rangle$ in modo da poter scrivere

$$B | a_j \rangle = b_j | a_j \rangle, \quad (1.140)$$

ma questo significa che gli autostati di A sono anche autostati di B . Possiamo indicare gli autostati comuni, democraticamente, usando entrambe le etichette $|a_j b_j\rangle$ in modo che

$$A | a_j b_j \rangle = a_j | a_j b_j \rangle, \quad (1.141)$$

$$B | a_j b_j \rangle = b_j | a_j b_j \rangle. \quad (1.142)$$

Per gli spettri *non degeneri* si tratta di una scrittura ridondante; infatti, per ogni a_j il valore di b_j è dato in modo univoco e viceversa, e possiamo limitarci alle sole etichette a_j , che sono sufficienti a distinguere tutti gli stati di base.

La dimostrazione inversa è pure più semplice. Se A e B condividono un insieme completo di autostati simultanei del tipo $|ab\rangle$, allora l'applicazione degli operatori prodotto AB e BA dà

$$AB | ab \rangle = bA | ab \rangle = ba | ab \rangle \quad (1.143)$$

$$BA | ab \rangle = aB | ab \rangle = ab | ab \rangle \quad (1.144)$$

da cui

$$(AB - BA) | ab \rangle = (ba - ab) | ab \rangle = 0 \quad (1.145)$$

dove l'ultima uguaglianza segue dal fatto che il prodotto di numeri è commutativo. Questo risultato vale per tutti gli autostati e, dato che essi costituiscono una base completa, la stessa relazione vale per tutti i vettori dello spazio vettoriale. Ne segue che $[A, B] = 0$.

Se lo spettro è *degenere* qualcosa cambia nel modo di etichettare gli stati. Supponiamo ad esempio che n autostati diversi dell'osservabile A corrispondano ad uno stesso autovalore. Indichiamo con a_ξ l'autovalore e con $|a_\xi^{(i)}\rangle$ gli autostati, essendo $i = 1, 2, \dots, n$; gli autostati sono normalizzati e ortogonali tra loro, come al solito, e tali che

$$A|a_\xi^{(i)}\rangle = a_\xi|a_\xi^{(i)}\rangle. \quad (1.146)$$

In tal caso si dice che l'autovalore è n volte degenere, o ha degenerazione n . Ora, la dimensione complessiva dello spazio vettoriale è data dal numero N di autostati ortonormali, ma se n di questi hanno lo stesso autovalore di A , allora l'insieme degli autovalori diversi di A è minore di N e non abbiamo abbastanza etichette per distinguere gli stati di base. In altri termini, non possiamo usare un apparato che misura A per selezionare uno stato del sistema in modo univoco, dato che qualsiasi combinazione lineare degli stati $|a_\xi^{(i)}\rangle$ è un vettore che fornisce lo stesso numero a_ξ come esito della misura. A questo punto ci torna utile cercare un'altra osservabile, B , che sia compatibile con A (cioè tale che $[A, B] = 0$) e che, dunque, ne condivide gli autostati. Se lo spettro di B è non degenere nei sottospazi in cui lo spettro di A è degenere, allora i suoi autovalori b_i potranno essere utilizzati per completare le etichette, dato che

$$A|a_\xi b_i\rangle = a_\xi|a_\xi b_i\rangle \quad (1.147)$$

$$B|a_\xi b_i\rangle = b_i|a_\xi b_i\rangle \quad (1.148)$$

con $i = 1, 2, \dots, n$. Per trovare i valori b_i si può, ad esempio, ricorrere alla rappresentazione matriciale dell'operatore B nella base degli autostati di A . È possibile che questa matrice non sia diagonale nel sottospazio corrispondente all'autovalore a_ξ ; allora basta diagonalizzarla in quel sottospazio per trovare gli autovalori b_i ; i nuovi vettori di base sono automaticamente anche autovettori di A , dato che l'operazione di diagonalizzazione ci mantiene all'interno del sottospazio di partenza. Può capitare che anche B abbia alcuni autovalori che si ripetono sulla diagonale; si dovrà quindi cercare una terza osservabile C compatibile sia con A che con B , in modo che nei sottospazi di vettori degeneri per A e B si possa utilizzare l'autovalore di C come etichetta. Nel caso più generale, possiamo trovare un certo *insieme massimale* di osservabili compatibili, $A, B, C, \dots$ tali che gli autostati etichettati con i loro autovalori

$$|abc\dots\rangle \quad (1.149)$$

godano della proprietà di ortonormalizzazione

$$\langle abc\dots|a'b'c'\dots\rangle = \delta_{aa'}\delta_{bb'}\delta_{cc'}\dots \quad (1.150)$$

e completezza

$$\sum_a \sum_b \sum_c \dots |abc\dots\rangle \langle abc\dots| = \mathbf{1}. \quad (1.151)$$

Fate attenzione al fatto che, dato un sistema e uno spazio di Hilbert di vettori che ne rappresenta gli stati fisici, l'insieme massimale di osservabili compatibili non è l'insieme di *tutte* le osservabili definite in quello spazio che commutano tra di loro, ma è un insieme di esse tale da poter usare i loro autovalori per identificare in modo univoco ciascun elemento della base ortonormale comune. Dato un insieme massimale di osservabili compatibili, ad esempio composto da tre operatori A, B e C , possiamo introdurre nuove osservabili definite come funzioni di questi operatori, che commutano con essi per costruzione¹⁹. Tuttavia,

¹⁹Si può in effetti dimostrare che un qualsiasi operatore che commuta con tutti gli operatori di un insieme massimale di osservabili compatibili è una funzione di questi.

dato uno stato $|abc\rangle$, in cui gli esiti delle misure di A, B e C sono noti e predeterminati, l'eventuale misura di queste ulteriori osservabili non aggiungerebbe alcuna informazione nuova sullo stato del sistema. Per questo motivo, l'insieme degli operatori A, B e C è anche detto “sistema completo di osservabili” e, inoltre, l'informazione codificata nello stato rappresentato dal vettore $|abc\rangle$ è la massima informazione che possiamo avere sul sistema, compatibilmente con le regole della meccanica quantistica²⁰.

Cosa significa in termini di misure? Consideriamo il caso semplice in cui l'insieme massimale di osservabili è costituito solo da A e B . Misuriamo prima A su uno stato rappresentato dal vettore generico $|\alpha\rangle$ e supponiamo di ottenere un certo valore a . Se lo spettro di A è non degenere allora la misura avrà fatto precipitare il vettore $|\alpha\rangle$ nell'autostato di A con autovalore a ; ma questo è anche autostato di B con un autovalore b univocamente fissato. Dunque la misura corrisponde a

$$|\alpha\rangle \xrightarrow{A} |ab\rangle. \quad (1.152)$$

Una successiva misura di B non modifica lo stato fisico del sistema

$$|\alpha\rangle \xrightarrow{A} |ab\rangle \xrightarrow{B} |ab\rangle \quad (1.153)$$

e se poi misurassimo di nuovo A , troveremmo ancora il valore a .

Se invece lo spettro è degenere e a è l'autovalore comune a n autostati diversi di A , allora la prima misura fa precipitare $|\alpha\rangle$ in un qualche vettore appartenente al sottospazio n -dimensionale dei vettori degeneri con autovalore a . Tale vettore è esprimibile come combinazione lineare degli autostati con lo stesso a ma con valori diversi di b . Possiamo quindi scrivere

$$|\alpha\rangle \xrightarrow{A} \sum_{i=1}^n (\langle ab_i | \alpha \rangle) |ab_i\rangle. \quad (1.154)$$

Una successiva misura di B fornirà uno dei valori di b_i , che possiamo chiamare b , e lo stato del sistema precipiterà nell'autostato corrispondente

$$|\alpha\rangle \xrightarrow{A} \sum_{i=1}^n (\langle ab_i | \alpha \rangle) |ab_i\rangle \xrightarrow{B} |ab\rangle. \quad (1.155)$$

Un'ulteriore misura di A darebbe come esito ancora a . Di nuovo, come per lo spettro degenere, la misura intermedia di B non distrugge l'informazione sullo stato del sistema in relazione all'osservabile A . Questo è il senso del concetto di compatibilità.

Tra l'altro, nel caso di spettri degeneri, notiamo che l'atto di misura va inteso come una *proiezione in un sottospazio*, come nella misura di A nella sequenza in (1.155), piuttosto che come collasso in un singolo stato. La definizione di misura come proiezione in sottospazi è più generale di quella data in (1.110), che ne rappresenta un caso particolare, e d'ora in poi diamo per implicito che va intesa in tal senso.

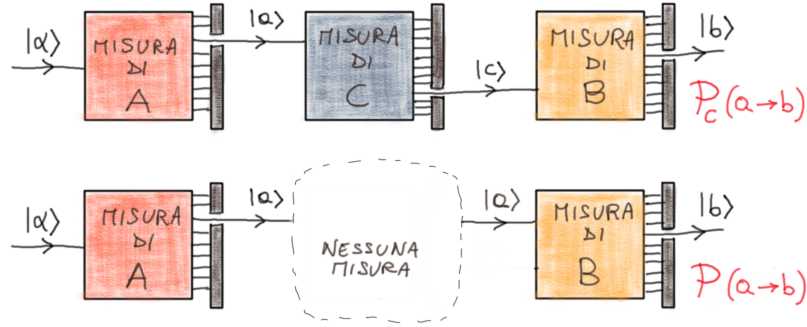
Ora vediamo cosa succede nella misura di osservabili incompatibili. Prendiamo tre osservabili A, B e C , di cui A e B possono essere indifferentemente compatibili o incompatibili (potrebbero anche coincidere), mentre C è incompatibile con entrambe. Eseguiamo misure in sequenza. Ad esempio partiamo da uno stato generico $|\alpha\rangle$ e misuriamo A . Se il valore che otteniamo è a , uno dei possibili autovalori dell'operatore, allora dopo la misura il sistema si trova nello stato $|a\rangle$ che è l'autostato corrispondente. Poi misuriamo B . La misura fa precipitare lo stato $|a\rangle$ in un autostato di B . Dato B è hermitiano possiamo usare i suoi

²⁰Vale la pena insistere sul fatto che l'informazione codificata in un vettore di stato non corrisponde a quanto *sappiamo*, ma a quanto *possiamo sapere in principio* sullo stato del sistema. La differenza è apparentemente sottile, ma ha conseguenze importanti.

autostati come base per decomporre lo stato d'ingresso. Per il postulato della misura, la probabilità di ottenere un certo valore b è data nel postulato

$$\mathcal{P}(a \rightarrow b) = |\langle b|a \rangle|^2. \quad (1.156)$$

Supponiamo di aver ottenuto b ; allora il sistema dopo la misura si trova in $|b\rangle$. Fin qui tutto bene.



I problemi nascono se inseriamo una misura intermedia dell'osservabile C tra quelle di A e di B e ci calcoliamo di nuovo la probabilità di ottenere lo stesso valore b alla fine della sequenza, se la prima misura è a . Misurando C otteniamo uno dei suoi possibili autovalori, ad esempio un certo c , con probabilità $|\langle c|a \rangle|^2$. A quel punto il sistema si trova nello stato $|c\rangle$. Effettuando la misura di B , la probabilità di ottenere il valore b , quello di prima, a partire da $|c\rangle$ è $|\langle b|c \rangle|^2$. Il prodotto

$$|\langle b|c \rangle|^2 |\langle c|a \rangle|^2 \quad (1.157)$$

è dunque la probabilità condizionata di ottenere b nella misura di B , partendo dallo stato $|a\rangle$ e avendo ottenuto c nella misura intermedia. Si tratta di un prodotto di probabilità dato che le misure di B e C sono eventi casuali indipendenti. Il punto cruciale è che il valore b può essere ottenuto anche quando la misura intermedia dà un autovalore diverso di C . La probabilità di ottenere b va dunque calcolata sommando le probabilità su tutti i valori possibili della misura intermedia:

$$\mathcal{P}_C(a \rightarrow b) = \sum_j |\langle b|c_j \rangle|^2 |\langle c_j|a \rangle|^2 = \sum_j \langle b|c_j \rangle \langle c_j|a \rangle \langle a|c_j \rangle \langle c_j|b \rangle. \quad (1.158)$$

Per confrontarla con la probabilità $\mathcal{P}(b)$ nel caso senza misura intermedia, riscriviamo quest'ultima usando la decomposizione $|a\rangle = \sum_j |c_j\rangle \langle c_j|a \rangle$ sulla base degli autostati di C per scrivere

$$\mathcal{P}(a \rightarrow b) = \left| \sum_j \langle b|c_j \rangle \langle c_j|a \rangle \right|^2 = \sum_{ij} \langle b|c_i \rangle \langle c_i|a \rangle \langle a|c_j \rangle \langle c_j|b \rangle \quad (1.159)$$

Le due probabilità risultano essere diverse. Eppure in entrambi i casi operiamo una somma su tutti i possibili valori di C . La differenza, dal punto di vista pratico, sta solo nel fatto che nel caso di $\mathcal{P}_C$ in (1.158) sappiamo, ogni volta, qual era il valore di C prima di fare la misura di B . Nel caso di $\mathcal{P}$ in (1.159) non lo sappiamo.

Se ragionassimo in modo classico, le due probabilità dovrebbero essere uguali, dato che il sistema, tra la misura di A e la misura di B deve comunque trovarsi in una configurazione in cui la grandezza C ha un certo valore definito, che lo si misuri o meno. E se sommiamo

su tutti questi casi “possibili”, poco importa se facciamo veramente la verifica di quali si verificano nelle singole sequenze. Ma il risultato che abbiamo ottenuto appare diverso. In un caso, quello in cui misuriamo C , otteniamo una singola somma sugli stati intermedi; nell’altro, quando non eseguiamo la misura di C , otteniamo una doppia somma, che esprime l’interferenza²¹ tra tutti i possibili stati intermedi, non misurati. Questo è uno dei punti cruciali della teoria. Ci torneremo più avanti.

Quanto detto sopra si applica anche al caso più semplice in cui $A = B \neq C$. In tal caso, la probabilità di riottenere lo stesso valore a (ovvero $b = a$ nelle espressioni precedenti), avendo eseguito una misura intermedia dell’osservabile C diventa

$$\mathcal{P}_C(a \rightarrow a) = \sum_j |\langle c_j | a \rangle|^4, \quad (1.160)$$

ma se la misura intermedia non viene eseguita, allora

$$\mathcal{P}(a \rightarrow a) = \left| \sum_j |\langle c_j | a \rangle|^2 \right|^2 = 1. \quad (1.161)$$

Ad esempio, nel caso di sequenze di Stern-Gerlach in cui $A = B = S_z$ e $C = S_x$, la prima probabilità vale $1/2$, mentre la seconda vale sempre 1. Dunque, misurare o non misurare S_x tra due misure di S_z non è la stessa cosa nemmeno se si somma su tutti gli esiti possibili della misura intermedia. Per questo motivo non è possibile affermare che un atomo con un certo valore di S_z abbia contestualmente anche uno o l’altro dei possibili valori di S_x . Le due grandezze sono semplicemente incompatibili²².

1.8 Relazione di indeterminazione

Una conseguenza importante della compatibilità o incompatibilità delle osservabili è un teorema che riguarda la distribuzione dei risultati di misure di due osservabili misurate sullo stesso stato fisico di un sistema.

Supponiamo che il sistema si trovi in un generico stato rappresentato dal vettore $|\alpha\rangle$. Prendiamo un’osservabile A e il suo valor medio $\langle A \rangle = \langle \alpha | A | \alpha \rangle$, che è un numero reale. Definiamo il nuovo operatore $A - \langle A \rangle$ e calcoliamo il valor medio del suo quadrato:

$$\langle (A - \langle A \rangle)^2 \rangle = \langle (A - \langle A \rangle)(A - \langle A \rangle) \rangle = \langle A^2 - 2A\langle A \rangle + \langle A \rangle^2 \rangle \quad (1.162)$$

ovvero

$$\langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2, \quad (1.163)$$

²¹Il termine “interferenza” usato in questo contesto non è casuale. In effetti, nelle due espressioni (1.158) e (1.159) ogni prodotto interno è un numero complesso con modulo e fase. Nella (1.158) le fasi spariscono nel prodotto $|\langle b | c_j \rangle|^2 |\langle c_j | a \rangle|^2$, mentre in (1.159) si fa il modulo quadro della somma e, dunque, rimarranno termini contenenti funzioni sinusoidali delle differenze di fase tra i vari addendi. Ne segue che la distribuzione statistica degli esiti avrà una modulazione tipica dei fenomeni d’interferenza. Il prototipo di questo comportamento è quello di un fascio coerente di particelle che incidono su uno schermo con doppia fenditura e che poi vengono rilevate su un secondo schermo, dietro al primo ad una certa distanza. Se si introduce un dispositivo che discrimina i casi in cui la particella passa da una o dall’altra fenditura (misura di C nel nostro linguaggio), allora l’immagine prodotta sullo schermo finale è semplicemente la somma degli accumuli di particelle che hanno attraversato ciascuna fenditura; se invece non inseriamo tale dispositivo e lasciamo che le particelle arrivino al secondo schermo senza sapere quale fenditura viene attraversata da ciascuna particella, allora l’immagine finale mostrerà una modulazione d’intensità dovuta all’interferenza tra gli stati che corrispondono ai due possibili passaggi per ciascuna particella. Su questo esempio qui non diciamo altro, perché la corretta interpretazione richiederebbe nozioni sulle osservabili a spettro continuo e le funzioni d’onda, che non abbiamo ancora introdotto, ma già si intuisce dove si va a parare.

²²In proposito, vi consiglio caldamente di leggere i commenti finali della sezione intitolata “Spin $1/2$ ” nel testo di David J. Griffiths, *Introduzione alla Meccanica Quantistica* (pagg.179-180 nella seconda edizione italiana, ed. CEA).

dove abbiamo usato il fatto che il valor medio di un numero è il numero stesso. Ora definiamo lo *scarto quadratico medio* ΔA come

$$\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}. \quad (1.164)$$

Per come l'abbiamo definita, questa quantità dà una stima di quanto la distribuzione statistica degli esiti della misura di A è dispersa rispetto al suo valor medio. Se lo stato $|\alpha\rangle$ è un'autostato di A , allora la misura di A dà sempre, con certezza, lo stesso valore, il valor medio coincide con quel valore e la dispersione è nulla. Se invece $|\alpha\rangle$ non è un autostato di A la dispersione attorno al valor medio è non nulla e l'esito della singola misura è tanto più indeterminato quanto più ΔA è grande.

Fatta questa premessa, il teorema è il seguente:

Relazione di indeterminazione: se A e B sono due osservabili qualsiasi e ΔA e ΔB sono gli scarti quadratici medi dei risultati delle misure di A e B su un generico stato rappresentato dal vettore $|\alpha\rangle$, allora

$$\Delta A \Delta B \geq \frac{1}{2} |\langle \alpha | [A, B] | \alpha \rangle|. \quad (1.165)$$

Per dimostrarla, date le osservabili A e B , introduciamo un operatore X definito come

$$X = A + i\eta B, \quad (1.166)$$

dove η è un numero reale qualsiasi; scriviamo il prodotto $X^\dagger X$

$$X^\dagger X = (A - i\eta B)(A + i\eta B) = A^2 + \eta^2 B^2 + i\eta[A, B] \quad (1.167)$$

e ne calcoliamo il valor medio

$$\langle X^\dagger X \rangle = \langle \alpha | A^2 + \eta^2 B^2 + i\eta[A, B] | \alpha \rangle = \langle A^2 \rangle + \eta^2 \langle B^2 \rangle + \eta \langle i[A, B] \rangle. \quad (1.168)$$

Ora possiamo applicare la disuguaglianza $\langle X^\dagger X \rangle \geq \langle X^\dagger \rangle \langle X \rangle$, che, come abbiamo visto in (1.115), vale per qualsiasi operatore X ed è una diretta conseguenza della disuguaglianza di Schwarz (1.30). Dunque

$$\langle A^2 \rangle + \eta^2 \langle B^2 \rangle + \eta \langle i[A, B] \rangle \geq \langle A - i\eta B \rangle \langle A + i\eta B \rangle, \quad (1.169)$$

da cui

$$\langle A^2 \rangle + \eta^2 \langle B^2 \rangle + \eta \langle i[A, B] \rangle \geq \langle A \rangle^2 + \eta^2 \langle B \rangle^2. \quad (1.170)$$

Portando tutto a sinistra e ricordando la definizione di scarto quadratico medio, la stessa disuguaglianza assume la forma

$$\eta^2 (\Delta B)^2 + \eta \langle i[A, B] \rangle + (\Delta A)^2 \geq 0. \quad (1.171)$$

Si noti che abbiamo tenuto il fattore i all'interno del secondo valor medio. L'abbiamo fatto perché sappiamo che il commutatore di due osservabili è *anti-hermitiano* e dunque l'operatore $i[A, B]$ è hermitiano; infatti

$$(i[A, B])^\dagger = (i)^* [A, B]^\dagger = (-i)(-[A, B]) = i[A, B]. \quad (1.172)$$

Allora $\langle i[A, B] \rangle$ è un numero reale e l'espressione che abbiamo trovato è una forma quadratica in η con coefficienti reali. Il fatto che non possa essere negativa, impone che il discriminante sia minore o uguale di zero:

$$\langle i[A, B] \rangle^2 - 4(\Delta A)^2 (\Delta B)^2 \leq 0, \quad (1.173)$$

da cui segue

$$\Delta A \Delta B \geq \frac{1}{2} |\langle i[A, B] \rangle| = \frac{1}{2} |\langle [A, B] \rangle|, \quad (1.174)$$

che è la relazione di indeterminazione²³.

La relazione di indeterminazione ci dice che, eseguendo la misura di due osservabili per le quali $\langle [A, B] \rangle \neq 0$, non è possibile ottenere due valori con indeterminazione nulla per entrambe. Esiste un valore minimo per il prodotto delle indeterminazioni fissato dal commutatore dei due operatori che rappresentano le osservabili; se lo stato su cui si esegue la misura viene preparato in modo da ridurre il più possibile l'indeterminazione su una delle due osservabili, allora l'indeterminazione sull'altra diventa grande. Per due osservabili compatibili il commutatore è nullo e, in tal caso, possono esistere stati in cui le indeterminazioni ΔA e ΔB sono simultaneamente nulle: questi sono proprio gli autostati comuni di A e B .

Come vedremo nel seguito, nel caso in cui le due osservabili siano la posizione e il momento (quantità di moto) di una particella in una data direzione la relazione diventa il ben noto *principio di indeterminazione di Heisenberg*, $\Delta x \Delta p \geq \hbar/2$. Nella nostra formulazione non si tratta di un principio, ma bensì di un teorema che segue dai postulati della teoria; in particolare, è una conseguenza del fatto che le osservabili in meccanica quantistica, a differenza di quella classica, in generale non costituiscono un'algebra commutativa e, inoltre, dal fatto che i vettori di stato forniscono predizioni di tipo probabilistico sugli esiti delle misure.

1.9 Cambiamenti di base

Come abbiamo già detto, una grandezza fisica è definita tramite le procedure che devono essere seguite per misurarla; queste procedure portano ad un confronto, diretto o indiretto, con campioni di misura e il risultato del confronto è un numero reale. In meccanica quantistica ognuno dei possibili esiti della misura è associato ad un vettore di stato in uno spazio complesso tale che, eseguendo la misura quando il sistema si trova in quello stato, l'esito è proprio quel valore e lo stato fisico non cambia. L'insieme dei vettori corrispondenti ai possibili esiti della misura forma una base ortonormale completa per lo spazio vettoriale, in modo da poter rappresentare tutti gli altri vettori di stato come n -ple di numeri (vettori colonna per i *ket* e riga per i duali *bra*) e tutti gli operatori come matrici quadre. Naturalmente, queste rappresentazioni non sono univoche. Data la definizione delle grandezze fisiche e dei loro autostati, la rappresentazione dei vettori e delle matrici (cioè i numeri che compaiono nelle righe e nelle colonne) dipende da quale base si è scelta. È quindi importante sapere come si passa da una rappresentazione all'altra per effetto di un cambiamento di base.

Data una base ortonormale costituita dai vettori $\{|a_i\rangle\}$ autostati di un'osservabile A vogliamo trovare un operatore U tale che, quando viene applicato ad uno qualsiasi degli autostati $|a_i\rangle$, produca l'autostato $|b_i\rangle$ di un'altra osservabile B . Questo si ottiene se

$$U = \sum_k |b_k\rangle \langle a_k|. \quad (1.175)$$

Infatti, quando applichiamo questo operatore a $|a_i\rangle$ otteniamo

$$U|a_i\rangle = \sum_k |b_k\rangle \langle a_k|a_i\rangle = |b_i\rangle, \quad (1.176)$$

essendo $\langle a_k|a_i\rangle = \delta_{ik}$. Ora, se prendiamo un generico vettore $|\alpha\rangle$, che può essere scritto come combinazione lineare degli autovettori di A con coefficienti $c_i^{(A)} = \langle a_i|\alpha\rangle$, e vogliamo invece

²³Di passaggio notiamo che qui sia la derivazione che la notazione sono diverse da quelle del Sakurai, così come del Griffiths. Non è mai un male che esista una varietà di dimostrazioni diverse di uno stesso risultato.

conoscere i coefficienti della decomposizione dello stesso stato sulla base degli autovettori di B , $c_i^{(B)} = \langle b_i | \alpha \rangle$, possiamo usare il duale della relazione appena trovata per scrivere

$$c_i^{(B)} = \langle b_i | \alpha \rangle = \langle a_i | U^\dagger | \alpha \rangle = \sum_j \langle a_i | U^\dagger | a_j \rangle \langle a_j | \alpha \rangle, \quad (1.177)$$

dove nell'ultimo passaggio abbiamo inserito la relazione di completezza. Dunque possiamo scrivere

$$c_i^{(B)} = \sum_j U_{ij}^\dagger c_j^{(A)}, \quad (1.178)$$

dove $U_{ij}^\dagger$ sono gli elementi della matrice che rappresenta $U^\dagger$ nella base di A . In pratica, si tratta della solita moltiplicazione di una matrice per un vettore colonna, come si fa usualmente in algebra lineare.

Per sapere com'è fatta la matrice, basta applicare la definizione di U data sopra, ricordando che l'hermitiano coniugato di un prodotto esterno è lo stesso prodotto ma scambiando l'ordine dei vettori:

$$U_{ij}^\dagger = \langle a_i | U^\dagger | a_j \rangle = \sum_k \langle a_i | a_k \rangle \langle b_k | a_j \rangle = \sum_k \delta_{ik} \langle b_k | a_j \rangle = \langle b_i | a_j \rangle. \quad (1.179)$$

La matrice che rappresenta U è la sua trasposta coniugata e dunque

$$U_{ij} = \langle b_j | a_i \rangle^* = \langle a_i | b_j \rangle. \quad (1.180)$$

La matrice così ottenuta ha come elementi di matrice i coefficienti degli autovettori di B nella base degli autovettori di A , con un autovettore per colonna. Nulla di nuovo: lo stesso si aveva nell'algebra lineare, con i vettori definiti nello spazio euclideo. Questa matrice è anche chiamata *matrice di trasformazione*.

Una proprietà molto importante dell'operatore U è questa:

$$UU^\dagger = U^\dagger U = \mathbb{1}. \quad (1.181)$$

Infatti

$$\begin{aligned} U^\dagger U &= \left(\sum_k |a_k\rangle \langle b_k| \right) \left(\sum_{k'} |b_{k'}\rangle \langle a_{k'}| \right) = \sum_{kk'} |a_k\rangle \langle b_k | b_{k'} \rangle \langle a_{k'}| \\ &= \sum_{kk'} |a_k\rangle \delta_{kk'} \langle a_{k'}| = \sum_k |a_k\rangle \langle a_k| = \mathbb{1}, \end{aligned} \quad (1.182)$$

e altrettanto per $UU^\dagger$. Un operatore che soddisfa la condizione $UU^\dagger = U^\dagger U = \mathbb{1}$ è detto *unitario*²⁴. Un operatore unitario non cambia la norma dello stato su cui agisce. Infatti, dato un vettore generico $|\alpha\rangle$, l'applicazione di $U^\dagger$ produce un nuovo vettore $|\gamma\rangle = U^\dagger |\alpha\rangle$ la cui norma è

$$\langle \gamma | \gamma \rangle = \langle \alpha | UU^\dagger | \alpha \rangle = \langle \alpha | \mathbb{1} | \alpha \rangle = \langle \alpha | \alpha \rangle. \quad (1.183)$$

L'operatore U permette anche di ricavare la relazione tra gli elementi di matrice di un operatore qualsiasi X nella vecchia e nuova base:

$$\begin{aligned} X_{ij}^{(B)} &= \langle b_i | X | b_j \rangle = \sum_{kk'} \langle b_i | a_k \rangle \langle a_k | X | a_{k'} \rangle \langle a_{k'} | b_j \rangle \\ &= \sum_{kk'} U_{ik}^\dagger X_{kk'}^{(A)} U_{k'j} \end{aligned} \quad (1.184)$$

²⁴Dato un operatore generico X , si può definire il suo *inverso*, indicato con il simbolo X^{-1} , tale che $X^{-1}X = \mathbb{1}$. Allora, un operatore X è unitario se il suo hermitiano coniugato è uguale all'inverso: $X^\dagger = X^{-1}$.

che è l'usuale prodotto tra matrici $X^{(B)} = U^\dagger X^{(A)} U$. Usando come base gli autostati di A la matrice che rappresenta B è diagonale solo se B e A sono osservabili compatibili. In generale la matrice che rappresenta B non è diagonale. Diagonalizzare la matrice corrisponde a trovare la matrice di trasformazione U tale che $U^\dagger B U$ è diagonale. Vale la pena notare che ogni cambio di base preserva la traccia della matrice, definita come la somma dei suoi elementi diagonali. Infatti

$$\begin{aligned}
 \sum_i X_{ii}^{(B)} &= \sum_i \sum_{kk'} \langle b_i | a_k \rangle \langle a_k | X | a_{k'} \rangle \langle a_{k'} | b_i \rangle \\
 &= \sum_{kk'} \langle a_{k'} | \left(\sum_i | b_i \rangle \langle b_i | \right) | a_k \rangle \langle a_k | X | a_{k'} \rangle \\
 &= \sum_{kk'} \langle a_{k'} | a_k \rangle \langle a_k | X | a_{k'} \rangle \\
 &= \sum_k \langle a_k | X | a_k \rangle = \sum_k X_{kk}^{(A)}.
 \end{aligned} \tag{1.185}$$

e si vede che la traccia è la stessa.

Per spazi numerabili di dimensione N tutto il formalismo che serve è quello dell'algebra lineare imparata nei corsi introduttivi di matematica, estesa al campo dei numeri complessi. Ora vediamo come estendere il formalismo anche agli spazi infiniti e continui.

1.10 Spettri continui, posizione

Finora abbiamo lavorato con osservabili aventi uno *spettro discreto* di autovalori. Lo spunto iniziale c'era venuto parlando di spin negli esperimenti di Stern-Gerlach, ma poi abbiamo supposto che il formalismo si applicasse a qualsiasi osservabile. Ci sono però grandezze fisiche i cui valori misurabili formano un continuo. Ad esempio, se vogliamo descrivere lo stato di una particella possiamo ricorrere a strumenti che ci dicono dove la particella si trova e qual è il suo momento (tra parentesi, d'ora in poi useremo sempre la parola momento come sinonimo di quantità di moto). Le misure di *posizione* e di *momento* forniscono numeri reali in un insieme continuo. Molte delle definizioni e relazioni discusse finora si applicano tali e quali anche a questi casi semplicemente sostituendo le somme con integrali. Altre richiedono cautela e assunzioni aggiuntive. Per fare solo un esempio, ci possiamo chiedere come si possa esprimere la condizione di ortonormalità, che nel caso discreto è

$$\langle a_i | a_j \rangle = \delta_{ij}, \tag{1.186}$$

passando da indici discreti a variabili continue.

Come prototipo di variabile continua consideriamo la posizione di una particella in una direzione data, ad esempio x . L'osservabile è x , i suoi autovalori sono tutti i valori di x misurabili e gli autostati corrispondenti sono i vettori $|x\rangle$. La misura della posizione della particella fa precipitare il suo stato generico in uno degli autostati. Ad esempio, se la particella viene trovata in x' , dopo la misura il suo stato è $|x'\rangle$. Possiamo scrivere

$$x|x'\rangle = x'|x'\rangle. \tag{1.187}$$

Ora, bisogna ammettere che questa equazione può sembrare strana. Il punto è che i vari x e x' che vi compaiono hanno significati diversi. Il primo è un operatore. Per evitare confusioni spesso lo si denota con il cappello sopra: $\hat{x}$, anche se così sembra un versore nello spazio euclideo; oppure si può usare x_{op} . Il simbolo x' nei due *ket* è semplicemente l'etichetta

dell'autostato; non è importante, purché sia chiara la relazione tra etichetta e autovalore. Il primo simbolo x' a destra dell'uguale è un numero reale, che nella misura concreta sarà espresso in metri. Dunque, potremmo riscriverla così, per evitare troppe ambiguità:

$$x_{\text{op}}|x'\rangle = x'|x'\rangle. \quad (1.188)$$

L'insieme di tutti gli stati $\{|x'\rangle\}$ costituisce una base per tutti gli stati che rappresentano la posizione di una particella. Questo significa che uno stato generico $|\alpha\rangle$ può essere scritto come combinazione lineare degli autostati, come facevamo nel caso di spettri discreti, ma stavolta usando integrali al posto delle somme:

$$|\alpha\rangle = \int_{-\infty}^{\infty} dx' |x'\rangle \langle x'|\alpha\rangle. \quad (1.189)$$

Per dare un significato concreto a questa espressione dobbiamo stabilire un legame tra i coefficienti della decomposizione, $\langle x'|\alpha\rangle$, e le probabilità di ottenere x' come esito della misura della posizione quando la particella si trova nello stato $|\alpha\rangle$. Per gli spettri discreti avevamo postulato che tale probabilità fosse $\mathcal{P}(a_i) = |\langle a_i|\alpha\rangle|^2$. Possiamo dire lo stesso per $\mathcal{P}(x') = |\langle x'|\alpha\rangle|^2$? No, non si può: la probabilità di trovare la particella da qualche parte può essere definita solo per intervalli spaziali, non per singoli punti. Il quesito corretto è qual è la probabilità di trovare la particella entro un certo intervallo $(x', x' + \Delta x)$. Ci aspettiamo che la probabilità sia proporzionale a Δx quando Δx è piccolo. Questo, oltre ad avere senso dal punto di vista pratico (si pensi ad esempio ai rivelatori di particelle con la loro risoluzione spaziale finita), permette anche di utilizzare il formalismo dei differenziali quando si fa tendere Δx a zero. Possiamo così scrivere

$$\mathcal{P}(x') = dx' |\langle x'|\alpha\rangle|^2. \quad (1.190)$$

Dunque la quantità $|\langle x'|\alpha\rangle|^2$ ha il significato di probabilità per unità di lunghezza, o densità di probabilità, ed è una funzione della posizione la cui forma dipende dallo stato $|\alpha\rangle$ in cui la particella si trova. Prendiamo questo come un postulato (o meglio, come una integrazione al postulato della misura già enunciato, in cui includiamo anche le osservabili con spettro continuo).

Per un intervallo finito, compreso tra due valori x_1 e x_2 , la probabilità diventa

$$\mathcal{P} = \int_{x_1}^{x_2} dx' |\langle x'|\alpha\rangle|^2 \quad (1.191)$$

e sullo spazio intero

$$1 = \int_{-\infty}^{\infty} dx' |\langle x'|\alpha\rangle|^2, \quad (1.192)$$

dato che, una volta identificato il sistema fisico con la particella, la probabilità di trovarla in una qualsiasi delle posizioni possibili è pari alla certezza. Se $|\alpha\rangle$ ha norma 1, la relazione precedente permette di scrivere anche

$$\begin{aligned} 1 = \langle \alpha|\alpha \rangle &= \int_{-\infty}^{\infty} dx' |\langle x'|\alpha\rangle|^2 \\ &= \int_{-\infty}^{\infty} dx' \langle \alpha|x' \rangle \langle x'|\alpha \rangle \\ &= \langle \alpha| \left(\int_{-\infty}^{\infty} dx' |x' \rangle \langle x'| \right) |\alpha \rangle. \end{aligned} \quad (1.193)$$

da cui segue la relazione di completezza

$$\int_{-\infty}^{\infty} dx' |x'\rangle\langle x'| = \mathbb{1}. \quad (1.194)$$

analoga alla (1.75) per gli spettri discreti.

Ora vediamo cosa succede alla definizione di ortogonalità. Nel caso discreto si aveva $\langle a_i | a_j \rangle = \delta_{ij}$, dove δ_{ij} è la delta di Kronecker. Questo simbolo non va bene per gli spettri continui. Va sostituito con la *delta di Dirac*

$$\langle x | x' \rangle = \delta(x - x'). \quad (1.195)$$

Dirac la introdusse apposta per estendere la formulazione della meccanica quantistica agli spettri continui. Nel linguaggio matematico, non si tratta di una funzione, ma di una *distribuzione*, ovvero un funzionale lineare continuo. La teoria delle distribuzioni viene normalmente introdotta nei corsi di analisi matematica o di metodi matematici per la fisica. Qui diamo per assodato che gli studenti la conoscano. Data una funzione $f(x)$ qualsiasi, continua in un intorno di $x = x'$, vale

$$\int_{-\infty}^{\infty} dx f(x) \delta(x - x') = f(x'). \quad (1.196)$$

Questo non è un integrale nel senso usuale, ma è da intendersi come l'applicazione del funzionale δ alla funzione f ; il funzionale trasforma la funzione $f(x)$ nel numero $f(x')$. Nel caso particolare della funzione $f(x) = 1$ si ha

$$\int_{-\infty}^{\infty} dx \delta(x) = 1. \quad (1.197)$$

Sotto ipotesi appropriate, la $\delta(x)$ può essere ottenuta come limite di una successione di funzioni che hanno un picco stretto attorno a $x = 0$, con altezza che tende a infinito e larghezza che tende a zero in modo che l'integrale rimanga sempre uguale a 1.

Una volta introdotta la delta di Dirac, possiamo usare la tavola di corrispondenza scritta qui. Questa tabella si applica a qualsiasi osservabile il cui spettro sia costituito dall'insieme continuo dei numeri reali. La definizione di densità di probabilità rimane la stessa, anche se l'osservabile non è la posizione, ma la probabilità va riferita allo spazio degli autovalori dell'osservabile che si considera, anziché allo spazio coordinate x . A parte questo, null'altro cambia nella trattazione matematica.

È opportuno segnalare a questo punto che per arrivare fin qua abbiamo nascosto sotto il tappeto alcune questioni delicate. Ad esempio, abbiamo detto che la misura della posizione della particella non può essere mai, in pratica, un punto di dimensione nulla: questo

Spettri discreti	Spettri continui
$\langle a_i a_j \rangle = \delta_{ij}$	$\langle x x' \rangle = \delta(x - x')$
$\sum_i a_i\rangle\langle a_i = \mathbb{1}$	$\int dx x\rangle\langle x = \mathbb{1}$
$ \alpha\rangle = \sum_i a_i\rangle\langle a_i \alpha\rangle$	$ \alpha\rangle = \int dx x\rangle\langle x \alpha\rangle$
$\sum_i \langle a_i \alpha\rangle ^2 = 1$	$\int dx \langle x \alpha\rangle ^2 = 1$
$\langle \beta \alpha\rangle = \sum_i \langle \beta a_i\rangle\langle a_i \alpha\rangle$	$\langle \beta \alpha\rangle = \int dx \langle \beta x\rangle\langle x \alpha\rangle$
$\langle a_i A a_j\rangle = a_i\delta_{ij}$	$\langle x x_{\text{op}} x'\rangle = x\delta(x - x')$

comporta che la quantità $|\langle x'|\alpha\rangle|^2$ è una densità di probabilità e non una probabilità (e questo l'abbiamo già detto), ma comporta anche che $|x'\rangle$ non è un autovettore dell'operatore posizione nel senso attribuito a questo concetto nel caso di spettri discreti. Nel caso di spettri continui, come quello della posizione, si parla di *vettori impropri* e la misura dell'osservabile stessa dovrebbe essere intesa come il limite di una misura eseguita per intervalli finiti, con intervalli sempre più piccoli. Possiamo dire che $|x'\rangle$ è un autovettore *improprio* dell'operatore x_{op} con autovalore *improprio* x' . L'introduzione della distribuzione (o funzione *impropria*) delta di Dirac serve proprio a rendere gli autovettori degli spettri continui equivalenti a quelli degli spettri discreti a tutti i fini pratici. Continueremo perciò a dire che $|x'\rangle$ è un autovettore dell'osservabile con spettro continuo x_{op} , lasciando le delicatezze sotto il tappeto. Chi è interessato a questi aspetti può trovare più soddisfazione leggendo il capitolo 6 del testo di Luigi Picasso citato in premessa.

Misurare la posizione di una particella che sta nello spazio tridimensionale corrisponde a misurare tre coordinate spaziali. Possiamo definire la posizione come un'osservabile a tre componenti, $\mathbf{r} = (x, y, z)$, dove le posizioni x , y e z sono loro stesse delle osservabili. Naturalmente, affinché sia possibile assegnare ad esse simultaneamente i tre valori corrispondenti alla posizione $\mathbf{r}$, è necessario assumere che x , y e z siano osservabili compatibili, ovvero che i loro commutatori siano tutti nulli. Questa assunzione è ragionevole se si pensa alla misura di posizione come l'attivazione di opportuni dispositivi rivelatori disposti nello spazio e che fanno *click* quando la particella si trova nella regione occupata dal dispositivo; in tal modo si possono assegnare contemporaneamente le tre coordinate della particella, identificate come le coordinate del rivelatore che ha fatto *click*, con precisione fissata solo dalla dimensione del rivelatore stesso. Se il rivelatore ha dimensioni molto piccole rispetto a tutte le scale di lunghezza rilevanti per un dato sistema, allora possiamo considerarlo come un volume infinitesimo e definire la densità di probabilità come funzione di tre coordinate. Possiamo così etichettare gli autostati simultanei di x_{op} , y_{op} e z_{op} in questo modo:

$$|\mathbf{r}\rangle = |x, y, z\rangle \quad (1.198)$$

e gli operatori agiscono così

$$x_{\text{op}}|\mathbf{r}\rangle = x|\mathbf{r}\rangle, \quad y_{\text{op}}|\mathbf{r}\rangle = y|\mathbf{r}\rangle, \quad z_{\text{op}}|\mathbf{r}\rangle = z|\mathbf{r}\rangle. \quad (1.199)$$

Le stesse relazioni possono essere scritte in forma più compatta come

$$\mathbf{r}_{\text{op}}|\mathbf{r}\rangle = \mathbf{r}|\mathbf{r}\rangle, \quad (1.200)$$

dove il primo $\mathbf{r}$ a destra dell'uguale è il vettore reale avente come tre componenti gli esiti delle misure x , y e z .

1.11 Traslazioni spaziali e momento

È molto interessante vedere cosa succede quando si compie una traslazione spaziale del sistema fisico o, in alternativa, del sistema di coordinate. Notiamo che la scelta del sistema di coordinate fa parte delle procedure operative che definiscono l'osservabile posizione, esattamente come in fisica classica. Traslare il sistema in una direzione, tenendo le coordinate fisse, equivale a tenere fisso il sistema ma traslare le coordinate in verso opposto. Quindi possiamo scegliere arbitrariamente una delle due opzioni e scegliamo la prima. Consideriamo un autostato qualsiasi $|x\rangle$ della posizione di una particella lungo x e definiamo un *operatore di traslazione infinitesima*, $\tau_x(dx)$, tale che

$$\tau_x(dx)|x\rangle = |x + dx\rangle. \quad (1.201)$$

Se la particella si trova in uno stato rappresentato dal vettore $|\alpha\rangle$ allora l'azione di $\tau_x(dx)$ su tale stato può essere scritta nella forma

$$\begin{aligned}\tau_x(dx)|\alpha\rangle &= \tau_x(dx) \int dx' |x'\rangle \langle x'|\alpha\rangle \\ &= \int dx' \tau_x(dx)|x'\rangle \langle x'|\alpha\rangle \\ &= \int dx' |x' + dx\rangle \langle x'|\alpha\rangle ,\end{aligned}\tag{1.202}$$

dove nella prima uguaglianza abbiamo usato la decomposizione di $|\alpha\rangle$ nella base degli autostati della posizione, nella seconda abbiamo applicato la linearità dell'operatore traslazione, nella terza abbiamo applicato l'operatore a ciascun autostato. A questo punto possiamo eseguire un cambio di variabile, sostituendo $x' + dx$ con un nuovo x' . L'integrale è esteso a tutto lo spazio e il cambio di variabile non modifica il valore dell'integrale²⁵. Dunque

$$\tau_x(dx)|\alpha\rangle = \int dx' |x'\rangle \langle x' - dx|\alpha\rangle ,\tag{1.203}$$

e se moltiplichiamo a sinistra per $\langle x|$ otteniamo

$$\langle x|\tau_x(dx)|\alpha\rangle = \int dx' \langle x|x'\rangle \langle x' - dx|\alpha\rangle = \langle x - dx|\alpha\rangle .\tag{1.204}$$

Questa relazione ci dice che la densità di probabilità di trovare la particella in x dopo che ha agito l'operatore $\tau_x(dx)$ è la stessa che si aveva in $x - dx$ prima, consistentemente con l'idea di traslazione spaziale degli stati fisici del sistema.

L'operatore $\tau_x(dx)$, essendo una pura traslazione in uno spazio continuo, deve soddisfare alcune importanti proprietà. Ad esempio deve conservare la norma dei vettori,

$$\langle \alpha|\tau_x^\dagger(dx)\tau_x(dx)|\alpha\rangle = \langle \alpha|\alpha\rangle\tag{1.205}$$

e dunque deve essere un operatore unitario:

$$\tau_x^\dagger(dx)\tau_x(dx) = \tau_x(dx)\tau_x^\dagger(dx) = \mathbb{1} .\tag{1.206}$$

Deve obbedire ad una regola di composizione: due traslazioni successive di dx_1 e dx_2 sono equivalenti ad una traslazione di $dx_1 + dx_2$:

$$\tau_x(dx_2)\tau_x(dx_1) = \tau_x(dx_1 + dx_2) .\tag{1.207}$$

Una traslazione nel verso opposto, di una distanza $-dx$, corrisponde all'inverso della traslazione di dx :

$$\tau_x(-dx) = \tau_x^{-1}(dx) .\tag{1.208}$$

Infine, una traslazione si sviluppa in modo continuo dall'identità, la differenza tra τ_x e l'identità essendo del primo ordine in dx . È possibile riscrivere l'operatore in una forma che soddisfa tutti quattro i requisiti:

$$\tau_x(dx) = \mathbb{1} - iK_x dx ,\tag{1.209}$$

²⁵Può sembrare strano applicare il concetto di cambio di variabile ai *ket*, $|x' + dx\rangle \rightarrow |x'\rangle$, ma ciò si giustifica con il fatto che l'etichetta del *ket* è essa stessa una variabile continua, in relazione biunivoca con l'insieme dei numeri reali. Analogamente, il prodotto interno $\langle x'|\alpha\rangle$ è una funzione della variabile continua x' e il cambiamento di variabile ha il significato usuale.

dove K_x è un operatore hermitiano. Infatti, l'unitarietà è soddisfatta in quanto $\tau_x^\dagger(dx)\tau_x(dx) = (\mathbb{1} + iK_x dx)(\mathbb{1} - iK_x dx) = \mathbb{1}$ (i termini al primo ordine si cancellano e quelli al secondo ordine vanno ignorati per definizione). Per la regola di composizione delle traslazioni successive si ha $\tau_x(dx_2)\tau_x(dx_1) = (\mathbb{1} - iK_x dx_2)(\mathbb{1} - iK_x dx_1) = \mathbb{1} - iK_x(dx_1 + dx_2)$. Per la traslazione inversa si ha $\tau_x(-dx) = \mathbb{1} + iK_x dx = \tau_x^\dagger(dx)$ ma, essendo τ_x unitario, si ha anche $\tau_x^\dagger(dx) = \tau_x^{-1}(dx)$. La quarta condizione è soddisfatta per costruzione.

Ripetendo lo stesso ragionamento per traslazioni nelle altre direzioni, possiamo definire un operatore di traslazione infinitesima $\boldsymbol{\tau}(d\mathbf{r})$ come un operatore a tre componenti τ_x , τ_y e τ_z , tale che

$$\boldsymbol{\tau}(d\mathbf{r})|\mathbf{r}\rangle = |\mathbf{r} + d\mathbf{r}\rangle \quad (1.210)$$

e avente la forma

$$\boldsymbol{\tau}(d\mathbf{r}) = \mathbb{1} - i\mathbf{K} \cdot d\mathbf{r}, \quad (1.211)$$

dove il prodotto a destra va inteso nel comune significato di somma dei prodotti delle singole componenti degli operatori, in analogia al prodotto scalare tra vettori euclidei.

Per capire il significato fisico dell'operatore $\mathbf{K}$ conviene calcolarne il commutatore con l'operatore posizione. Consideriamo ad esempio la misura di x eseguita sullo stato $|\mathbf{r}\rangle = |x, y, z\rangle$ seguita da una traslazione lungo x di dx e calcoliamo la differenza di esito nel caso in cui si scambi l'ordine degli operatori:

$$\begin{aligned} [x_{\text{op}}, \tau_x(dx)]|\mathbf{r}\rangle &= x_{\text{op}}\tau_x(dx)|\mathbf{r}\rangle - \tau_x(dx)x_{\text{op}}|\mathbf{r}\rangle \\ &= x_{\text{op}}|x + dx, y, z\rangle - \tau_x(dx)x|x, y, z\rangle \\ &= (x + dx)|x + dx, y, z\rangle - x|x + dx, y, z\rangle \\ &= dx|x + dx, y, z\rangle, \end{aligned} \quad (1.212)$$

che al primo ordine dà

$$[x_{\text{op}}, \tau_x(dx)]|x, y, z\rangle = dx|x, y, z\rangle. \quad (1.213)$$

Sostituendo τ_x con la sua espressione (1.209) e notando che l'operatore identità commuta con l'operatore posizione, si ottiene

$$-idx[x_{\text{op}}, K_x]|x, y, z\rangle = dx|x, y, z\rangle \quad (1.214)$$

ovvero

$$[x_{\text{op}}, K_x] = i\mathbb{1}. \quad (1.215)$$

Facendo lo stesso calcolo per le altre direzioni, otterremo lo stesso risultato:

$$[y_{\text{op}}, K_y] = i\mathbb{1}, \quad [z_{\text{op}}, K_z] = i\mathbb{1}. \quad (1.216)$$

Cosa succede se, invece, misuriamo la posizione in una direzione e operiamo la traslazione in una direzione ortogonale? Rifacciamo come prima e prendiamo, ad esempio, la misura di posizione lungo x ma operando una traslazione lungo y . Abbiamo

$$\begin{aligned} [x_{\text{op}}, \tau_y(dy)]|\mathbf{r}\rangle &= x_{\text{op}}\tau_y(dy)|\mathbf{r}\rangle - \tau_y(dy)x_{\text{op}}|\mathbf{r}\rangle \\ &= x_{\text{op}}|x, y + dy, z\rangle - \tau_y(dy)x|x, y, z\rangle \\ &= x|x, y + dy, z\rangle - x|x, y + dy, z\rangle = 0, \end{aligned} \quad (1.217)$$

che in termini dell'operatore K_y implica

$$[x_{\text{op}}, K_y] = 0. \quad (1.218)$$

Lo stesso risultato per le altre combinazioni di coordinate traslazioni. In sintesi, possiamo esprimere questi risultati nella relazione

$$[r_{i,\text{op}}, K_j] = i\delta_{ij}\mathbb{1}, \quad (1.219)$$

dove ij indica le coppie di indici x, y, z e dove r_i sono le tre componenti x, y, z dell'operatore posizione.

A questo punto è utile ricordarsi un paio di nozioni di fisica classica. La prima è che in meccanica classica una traslazione infinitesima può essere considerata come una trasformazione canonica che coinvolge il momento $\mathbf{p}$. La seconda è che, sempre in meccanica classica, assegnate le s coordinate generalizzate q di un sistema e i loro momenti coniugati p , prese due funzioni qualsiasi $A(q, p)$ e $B(q, p)$, è possibile definire la parentesi di Poisson

$$[A(q, p), B(q, p)]_{\text{Poisson}} = \sum_{i=1}^s \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right), \quad (1.220)$$

e vale

$$[q_i, p_j]_{\text{Poisson}} = \delta_{ij}. \quad (1.221)$$

Questa è davvero molto simile alla relazione di commutazione ricavata prima. Per analogia, ci viene naturale identificare l'operatore $\mathbf{K}$ (che è hermitiano) con un'osservabile che chiamiamo momento, e che indichiamo con la lettera $\mathbf{p}$. In realtà, se vogliamo che il $\mathbf{p}$ che introduciamo nella nostra teoria abbia le stesse unità del $\mathbf{p}$ classico, dobbiamo moltiplicare $\mathbf{K}$ (che ha le dimensioni dell'inverso di una lunghezza) per una costante che ha le dimensioni di energia per tempo, le stesse dimensioni che ha la costante $\hbar$. Dunque facciamo questa assunzione:

$$\mathbf{p} = \hbar \mathbf{K}, \quad (1.222)$$

dove $\mathbf{p}$ corrisponde alla terna di operatori p_x, p_y e p_z , che rappresentano le componenti x, y e z dell'osservabile momento. In questo modo si ottiene la corrispondenza

$$[\ , \]_{\text{Poisson}} \leftrightarrow \frac{1}{i\hbar} [\ , \] \quad (1.223)$$

e le relazioni (1.219) diventano

$$[r_i, p_j] = i\hbar\delta_{ij}\mathbb{1}, \quad (1.224)$$

a cui si aggiungono le

$$[r_i, r_j] = 0, \quad [p_i, p_j] = 0, \quad (1.225)$$

dove stavolta abbiamo tolto il suffisso “op” dagli operatori per alleggerire la notazione. Queste si chiamano *relazioni di commutazione canonica*.

Per inciso, notiamo anche che le parentesi di Poisson godono delle stesse proprietà (1.45)-(1.50) dei commutatori, il che rafforza l'analogia.

Vale la pena insistere qui sulla logica che abbiamo seguito: abbiamo definito un nuovo operatore $\mathbf{p}$ tramite la relazione (1.222) che lo lega all'operatore $\mathbf{K}$ associato alla traslazioni spaziali infinitesime. Abbiamo usato la lettera $\mathbf{p}$ e il nome “momento” per via di una certa analogia con la fisica classica. Che poi questo operatore, in un appropriato limite classico della teoria, vada a coincidere con la definizione $\mathbf{p} = m\mathbf{v}$ rimane tutto da dimostrare, anche perché finora non abbiamo definito alcuna osservabile velocità²⁶.

²⁶Non solo non abbiamo definito la velocità, ma nemmeno la massa. Diamo per scontato che una qualsiasi particella possieda una massa che ne quantifica l'inerzia ai cambiamenti di moto (nel senso newtoniano del termine), ma in realtà finora non abbiamo mai dovuto introdurre tale concetto. Quando parleremo di dinamica quantistica la massa m comparirà nell'energia cinetica di una particella, laddove era prevista dalla meccanica classica, ma in un'equazione diversa da quella di Newton.

Una conseguenza importante della non commutatività delle osservabili posizione e momento è la relazione nota come “principio di indeterminazione di Heisenberg”. Per come abbiamo formulato la teoria, non si tratta di un principio, ma di una conseguenza diretta della relazione di indeterminazione (1.165) applicata agli operatori $\mathbf{r}$ e $\mathbf{p}$,

$$\Delta r_i \Delta p_j \geq \frac{1}{2} |\langle [r_i, p_j] \rangle| = \frac{\hbar}{2} \delta_{ij}, \quad (1.226)$$

che nel caso di coppie di componenti nella stessa direzione diventa

$$\boxed{\Delta x \Delta p_x \geq \frac{\hbar}{2}}. \quad (1.227)$$

Si tratta di una delle leggi più famose della fisica moderna. Il suo significato apparirà chiaro mano a mano che procederemo con l'esposizione della teoria²⁷.

In sintesi, in questa sezione abbiamo identificato il momento $\mathbf{p}$ come il generatore delle traslazioni spaziali. Ad esempio, per una traslazione infinitesima nella direzione x possiamo scrivere

$$\left(\mathbb{1} - \frac{i}{\hbar} p_x dx \right) |x\rangle = |x + dx\rangle. \quad (1.228)$$

L'operatore a sinistra, opera la traslazione del vettore lasciando invariata la norma. Per traslazioni finite, di una distanza Δx , è sufficiente dividere questa distanza in intervalli più piccoli $\Delta x/N$ con N tendente all'infinito e applicare una sequenza di traslazioni infinitesime in questo modo:

$$|x + \Delta x\rangle = \lim_{N \rightarrow \infty} \left(\mathbb{1} - \frac{i}{\hbar} p_x \frac{\Delta x}{N} \right)^N |x\rangle = \exp \left(-\frac{i}{\hbar} p_x \Delta x \right) |x\rangle, \quad (1.229)$$

dove l'ultima uguaglianza si trova confrontando lo sviluppo dei prodotti degli operatori a sinistra con lo sviluppo in serie di potenze dell'esponenziale.

La relazione (1.228) ci dice che il vettore $(\mathbb{1} - (i/\hbar)p_x dx)|x\rangle$ coincide con l'autostato dell'operatore posizione x_{op} con autovalore $x + dx$. Dunque, se applichiamo l'operatore x_{op} a questo vettore dobbiamo trovare

$$x_{\text{op}} \left(\mathbb{1} - \frac{i}{\hbar} p_x dx \right) |x\rangle = (x + dx) \left(\mathbb{1} - \frac{i}{\hbar} p_x dx \right) |x\rangle. \quad (1.230)$$

Per dimostrare che questo è vero, basta applicare ad entrambi i membri l'operatore $\mathbb{1} + (i/\hbar)p_x dx$, ricordando che $x + dx$ è un numero e può essere spostato davanti all'operatore,

$$\left(\mathbb{1} + \frac{i}{\hbar} p_x dx \right) x_{\text{op}} \left(\mathbb{1} - \frac{i}{\hbar} p_x dx \right) |x\rangle = (x + dx) \left(\mathbb{1} + \frac{i}{\hbar} p_x dx \right) \left(\mathbb{1} - \frac{i}{\hbar} p_x dx \right) |x\rangle \quad (1.231)$$

²⁷Heisenberg e Bohr spesero molto tempo e molti sforzi per spiegare il principio di indeterminazione (o di incertezza) in termini di relazione tra la precisione di una misura e il disturbo che essa crea al sistema misurato. Il prototipo di tale ragionamento è il tentativo di misurare otticamente la posizione di un elettrone; per localizzarlo entro un intervallo piccolo $\epsilon(x)$ è necessario usare fotoni di lunghezza d'onda sufficientemente corta e questi, interagendo con l'elettrone, gli impartiscono un momento tale da indurre una indeterminazione $\eta(p)$ consistente con la relazione $\epsilon(x)\eta(p) \geq \hbar/2$. Questo tipo di ragionamenti hanno dominato la scena per decenni ed erano uno degli ingredienti principali delle famose discussioni tra Bohr e Einstein. Non sorprende quindi che molti libri di testo li usino ancora per “spiegare” il principio. Va detto, tuttavia, che si tratta di un approccio non rigoroso e potenzialmente fuorviante. Come abbiamo visto qui, la relazione d'indeterminazione è una conseguenza diretta dei postulati della teoria, che, a loro volta, trovano conferma nel successo predittivo della teoria nel suo complesso. Si tratta di una relazione che esprime una proprietà “intrinseca” dei sistemi fisici, valida per tutti i sistemi e tutte le coppie di osservabili incompatibili, indipendentemente dal fatto che si usino o meno apparati di misura tali da produrre disturbi. Al contrario, i concetti di precisione e disturbo sono meno rigorosi e la relazione $\epsilon(A)\eta(B) \geq (1/2)|\langle [A, B] \rangle|$, dove $\epsilon(A)$ è la precisione richiesta nella misura di A e $\eta(B)$ è il disturbo nella successiva determinazione di B , non è rigorosa, né esatta; esistono esperimenti che la violano, senza tuttavia violare la relazione esatta (1.174).

e poi svolgere i prodotti a sinistra e a destra, tenendo i termini solo al primo ordine:

$$\left(x_{\text{op}} - \frac{i}{\hbar}[x_{\text{op}}, p_x]dx\right)|x\rangle = (x + dx)|x\rangle. \quad (1.232)$$

Questa è vera se e solo se $[x_{\text{op}}, p_x] = i\hbar$, che è la relazione di commutazione canonica.

1.12 Funzioni d'onda

Precedentemente, nella discussione attorno alla definizione (1.190), avevamo definito la probabilità di misurare la posizione x di una particella entro un intervallo di larghezza dx' centrato in x' , se la particella si trova nello stato rappresentato dal vettore $|\alpha\rangle$, come la quantità $dx' |\langle x'|\alpha\rangle|^2$. Ora notiamo che $\langle x'|\alpha\rangle$ è una funzione complessa della variabile continua x' . D'ora in poi la chiameremo *funzione d'onda* e la indicheremo così

$$\langle x'|\alpha\rangle = \psi_\alpha(x') \quad (1.233)$$

e $|\psi|^2$ è una densità di probabilità.

Possiamo esprimere i prodotti interni, gli elementi di matrice di un operatore, il valor medio e altre grandezze in termini della funzione d'onda. Per il prodotto interno vale

$$\langle \beta|\alpha\rangle = \int dx' \langle \beta|x'\rangle \langle x'|\alpha\rangle = \int dx' \psi_\beta^*(x') \psi_\alpha(x'). \quad (1.234)$$

Se un vettore di stato viene decomposto nella base di un osservabile A al solito modo $|\alpha\rangle = \sum_i |a_i\rangle \langle a_i|\alpha\rangle$, possiamo moltiplicare a sinistra per $\langle x'|$

$$\langle x'|\alpha\rangle = \sum_i \langle x'|a_i\rangle \langle a_i|\alpha\rangle, \quad (1.235)$$

e riscrivere la decomposizione nella forma

$$\psi_\alpha(x') = \sum_i c_i \varphi_i(x'), \quad (1.236)$$

dove $c_i = \langle a_i|\alpha\rangle$, mentre $\varphi_i(x')$ è la funzione d'onda che rappresenta l'autovettore $|a_i\rangle$ di A . Diremo che $\varphi_i(x')$ è l'*autofunzione* di A con autovalore a_i . Per gli elementi di matrice di A possiamo usare la relazione di completezza e scrivere

$$\begin{aligned} \langle \beta|A|\alpha\rangle &= \iint dx' dx'' \langle \beta|x'\rangle \langle x'|A|x''\rangle \langle x''|\alpha\rangle \\ &= \iint dx' dx'' \psi_\beta^*(x') \langle x'|A|x''\rangle \psi_\alpha(x''). \end{aligned} \quad (1.237)$$

dove $\langle x'|A|x''\rangle$ è una funzione complessa delle due variabili reali continue x' e x'' . Se A è una funzione della posizione, $A = f(x_{\text{op}})$, allora

$$\langle \beta|f(x_{\text{op}})|\alpha\rangle = \iint dx' dx'' \psi_\beta^*(x') \langle x'|f(x_{\text{op}})|x''\rangle \psi_\alpha(x''), \quad (1.238)$$

da cui, estraendo gli autovalori e usando l'ortogonalità, si ottiene

$$\begin{aligned} \langle \beta|f(x_{\text{op}})|\alpha\rangle &= \iint dx' dx'' f(x'') \psi_\beta^*(x') \langle x'|x''\rangle \psi_\alpha(x'') \\ &= \iint dx' dx'' f(x'') \psi_\beta^*(x') \delta(x' - x'') \psi_\alpha(x'') \\ &= \int dx' f(x') \psi_\beta^*(x') \psi_\alpha(x') \end{aligned} \quad (1.239)$$

e il valore di aspettazione dell'osservabile f diventa

$$\langle f(x) \rangle = \langle \alpha | f(x_{\text{op}}) | \alpha \rangle = \int dx' f(x') |\psi_\alpha(x')|^2, \quad (1.240)$$

dove conviene sottolineare che la funzione f nella prima uguaglianza è un operatore, funzione di un operatore, mentre $f(x')$ nell'integrale è una funzione di variabile reale. Ad esempio, se $f = x_{\text{op}}^2$, allora

$$\langle x^2 \rangle = \langle \alpha | x_{\text{op}}^2 | \alpha \rangle = \int dx' x'^2 |\psi_\alpha(x')|^2. \quad (1.241)$$

Tutto ciò si applica pari pari anche nelle direzioni y e z in modo, per un dato vettore di stato $|\alpha\rangle$, possiamo definire la funzione d'onda $\psi_\alpha(\mathbf{r})$ e generalizzare i risultati precedenti sostituendo gli integrali unidimensionali con integrali in tre dimensioni.

Come si comporta l'operatore momento, $\mathbf{p}$, quando agisce sulle funzioni d'onda? Per vederlo ci mettiamo in una dimensione e riprendiamo la definizione dell'operatore di traslazione infinitesima

$$\tau_x(dx) = \mathbb{1} - \frac{i}{\hbar} p_x dx. \quad (1.242)$$

Applichiamolo ad un generico stato rappresentato dal vettore $|\alpha\rangle$. L'azione di $\tau(dx)$ è data dalla relazione (1.203). La riscriviamo, ma stavolta usando l'espressione di $\tau(dx)$ in termini di p_x

$$\left(\mathbb{1} - \frac{i}{\hbar} p_x dx \right) |\alpha\rangle = \int dx' |x'\rangle \langle x' - dx | \alpha \rangle \quad (1.243)$$

e poi sviluppiamo la funzione $\langle x' - dx | \alpha \rangle$ al primo ordine in dx :

$$\left(\mathbb{1} - \frac{i}{\hbar} p_x dx \right) |\alpha\rangle = \int dx' |x'\rangle \left(\langle x' | \alpha \rangle - dx \frac{d}{dx'} \langle x' | \alpha \rangle \right). \quad (1.244)$$

Ne segue che

$$p_x |\alpha\rangle = \int dx' |x'\rangle \left(\frac{\hbar}{i} \frac{d}{dx'} \right) \langle x' | \alpha \rangle \quad (1.245)$$

e anche

$$\begin{aligned} \langle \beta | p_x | \alpha \rangle &= \int dx' \langle \beta | x' \rangle \left(\frac{\hbar}{i} \frac{d}{dx'} \right) \langle x' | \alpha \rangle \\ &= \int dx' \psi_\beta^*(x') \left(\frac{\hbar}{i} \frac{d}{dx'} \right) \psi_\alpha(x'), \end{aligned} \quad (1.246)$$

da cui, se $|\beta\rangle = |\alpha\rangle$,

$$\langle p_x \rangle = \int dx' \psi_\alpha^*(x') \left(\frac{\hbar}{i} \frac{d}{dx'} \right) \psi_\alpha(x'). \quad (1.247)$$

Gli stessi ragionamenti applicati a $p_x^2 = p_x p_x$ portano anche a

$$p_x^2 |\alpha\rangle = \int dx' |x'\rangle \left(-\hbar^2 \frac{d^2}{dx'^2} \right) \langle x' | \alpha \rangle \quad (1.248)$$

e

$$\langle p_x^2 \rangle = \int dx' \psi_\alpha^*(x') \left(-\hbar^2 \frac{d^2}{dx'^2} \right) \psi_\alpha(x'), \quad (1.249)$$

o più in generale

$$\langle f(p_x) \rangle = \int dx' \psi_\alpha^*(x') f \left(\frac{\hbar}{i} \frac{d}{dx'} \right) \psi_\alpha(x'), \quad (1.250)$$

dove si sottintende che tutte le funzioni che ci interessano possano essere scritte come serie di potenze, in modo da avere termini del tipo p_x^n a cui corrispondono derivate di ordine n rispetto alla coordinata x' .

Questo risultato, combinato con l'analogo (1.240), ci permette di calcolare valori medi di osservabili che siano funzioni della coordinata e del momento, $f(x, p_x)$, eseguendo integrali che coinvolgono le funzione d'onda. È un fatto importante, perché ci permette di ridurre il livello di astrazione della teoria, riconducendo il calcolo di quantità che hanno un ben definito significato fisico al calcolo integrale con funzioni complesse. Ciò è possibile quando le grandezze che prendiamo in esame sono funzioni delle posizioni e dei momenti, potendo in tal caso utilizzare come base gli autostati della posizione.

Naturalmente l'operatore hermitiano p_x avrà pure i suoi autostati (*impropri*, come quelli della posizione), che soddisfano l'equazione agli autovalori

$$p_x |p'\rangle = p' |p'\rangle, \quad (1.251)$$

dove $|p'\rangle$ è un generico autovettore e p' il corrispondente autovalore. Gli auto-vettori formeranno una base ortonormale, tale che

$$\langle p' | p'' \rangle = \delta(p' - p''). \quad (1.252)$$

e

$$\int dp' |p'\rangle \langle p'| = \mathbb{1}, \quad (1.253)$$

e ogni vettore generico $|\alpha\rangle$ potrà essere decomposto in questo modo

$$|\alpha\rangle = \int dp' |p'\rangle \langle p' | \alpha \rangle. \quad (1.254)$$

Possiamo definire la *funzione d'onda nello spazio dei momenti*

$$\tilde{\psi}_\alpha(p) = \langle p | \alpha \rangle, \quad (1.255)$$

il cui modulo quadro è la densità di probabilità nello stesso spazio, cioè $|\tilde{\psi}(p)|^2 dp$ è la probabilità che il momento misurato stia in un intervallo di ampiezza dp attorno a p . Il valor medio di p sarà esprimibile nella forma

$$\langle p_x \rangle = \langle \alpha | p_x | \alpha \rangle = \int dp' \tilde{\psi}_\alpha^*(p') p' \tilde{\psi}_\alpha(p') = \int dp' p' |\tilde{\psi}_\alpha(p')|^2. \quad (1.256)$$

Ci interessa ora capire come si passa dalla base degli autostati della posizione a quella degli autostati dei momenti. Nel caso degli spettri discreti, il cambiamento di base si attuava tramite una matrice di trasformazione unitaria U i cui elementi di matrice erano i prodotti interni tra gli elementi delle rispettive basi, $\langle a_i | b_j \rangle$. Qui ci aspettiamo di avere una relazione che coinvolge le funzioni $\langle x | p \rangle$, che possono essere intese sia come funzioni di x , a p fissato, che come funzioni di p , a x fissato.

Cominciamo con prendere la (1.245) e applicarla al caso $|\alpha\rangle = |p'\rangle$:

$$p_x |p'\rangle = \int dx |x\rangle \left(\frac{\hbar}{i} \frac{d}{dx} \right) \langle x | p' \rangle; \quad (1.257)$$

poi moltiplichiamo a sinistra per $\langle x' |$ e ricordiamo che $\langle x' | x \rangle = \delta(x' - x)$:

$$\langle x' | p_x | p' \rangle = \int dx \langle x' | x \rangle \left(\frac{\hbar}{i} \frac{d}{dx} \right) \langle x | p' \rangle = \frac{\hbar}{i} \frac{d}{dx'} \langle x' | p' \rangle. \quad (1.258)$$

Il membro di sinistra può essere riscritto sapendo che $|p'\rangle$ è autostato di p_x con autovalore p' :

$$p'\langle x'|p'\rangle = \frac{\hbar}{i} \frac{d}{dx'} \langle x'|p'\rangle. \quad (1.259)$$

Abbiamo così ottenuto un'equazione differenziale per la funzione $\langle x'|p'\rangle$, intesa come funzione di x' a p' fissato. La soluzione è

$$\langle x'|p'\rangle = C \exp\left(\frac{ip'x'}{\hbar}\right), \quad (1.260)$$

dove C è una costante di normalizzazione. Questa può essere fissata ricorrendo all'ortonormalità $\langle x'|x''\rangle = \delta(x' - x'')$ che, usando l'ultima espressione, possiamo riscrivere così

$$\begin{aligned} \delta(x' - x'') &= \langle x'|x''\rangle = \int dp' \langle x'|p'\rangle \langle p'|x''\rangle \\ &= |C|^2 \int dp' \exp\left(\frac{ip'(x' - x'')}{\hbar}\right) \\ &= 2\pi\hbar |C|^2 \delta(x' - x''), \end{aligned} \quad (1.261)$$

dove gli integrali si estendono da $-\infty$ a $+\infty$ e abbiamo usato la rappresentazione di Fourier della delta: $2\pi\delta(x) = \int dk \exp(ikx)$, che diamo per nota dai corsi di matematica. Dunque concludiamo che $2\pi\hbar|C|^2 = 1$ e, se prendiamo C reale, questo corrisponde a scrivere

$$\langle x'|p'\rangle = \frac{1}{\sqrt{2\pi\hbar}} \exp\left(\frac{ip'x'}{\hbar}\right). \quad (1.262)$$

Dunque, la funzione d'onda che rappresenta un autostato del momento nello spazio delle coordinate è un'onda piana. Prendendo il coniugato complesso della stessa espressione troviamo

$$\langle p'|x'\rangle = \frac{1}{\sqrt{2\pi\hbar}} \exp\left(-\frac{ip'x'}{\hbar}\right), \quad (1.263)$$

e vediamo che la funzione d'onda che rappresenta un autostato della posizione nello spazio dei momenti è anch'essa un'onda piana. Ciò è consistente con la relazione d'indeterminazione posizione-momento, secondo la quale, se l'indeterminazione di una delle due osservabili tende a zero, come nel caso in cui lo stato del sistema sia un autostato di quell'osservabile, allora l'indeterminazione dell'altra è infinita e la funzione d'onda ha modulo costante ovunque nello spazio corrispondente.

Per le funzioni d'onda associate ad uno stato generico basta usare la completezza delle basi nei due spazi per scrivere

$$\psi_\alpha(x) = \langle x|\alpha\rangle = \int dp \langle x|p\rangle \langle p|\alpha\rangle \quad (1.264)$$

e

$$\tilde{\psi}_\alpha(p) = \langle p|\alpha\rangle = \int dx \langle p|x\rangle \langle x|\alpha\rangle \quad (1.265)$$

e inserire le espressioni precedenti per $\langle x|p\rangle$ e $\langle p|x\rangle$ come onde piane. Otteniamo

$$\psi_\alpha(x) = \frac{1}{\sqrt{2\pi\hbar}} \int dp \tilde{\psi}_\alpha(p) \exp\left(\frac{ipx}{\hbar}\right) \quad (1.266)$$

e

$$\tilde{\psi}_\alpha(p) = \frac{1}{\sqrt{2\pi\hbar}} \int dx \psi_\alpha(x) \exp\left(-\frac{ipx}{\hbar}\right). \quad (1.267)$$

Queste due relazioni corrispondono alle trasformate di Fourier. Queste trasformate svolgono per gli spettri continui il ruolo che le matrici di trasformazione svolgevano per gli spettri discreti.

Tutto questo può essere generalizzato facilmente al caso tridimensionale, dove si avranno funzioni d'onda $\psi(\mathbf{r})$ e le loro trasformate $\tilde{\psi}(\mathbf{p})$, il cui modulo quadro sarà la densità di probabilità nello spazio delle posizioni e dei momenti; l'operatore momento nello spazio delle posizioni sarà identificato da $(\hbar/i)\nabla$ anziché dalla semplice derivata, le trasformate di Fourier saranno integrali di volume, e così via.

Dunque, nel formalismo delle funzioni d'onda il cambiamento di base coincide con la matematica delle trasformate di Fourier. Inoltre, i due spazi che chiamiamo “delle posizioni” e “dei momenti” non sono altro che lo stesso spazio di Hilbert che descrive gli stati fisici di una particella, ma rappresentati usando due basi diverse, quella degli autostati della posizione e quella degli autostati del momento.

Una funzione d'onda nello spazio delle posizioni può essere vista come sovrapposizione di onde piane con vettore d'onda $\mathbf{k} = \mathbf{p}/\hbar$. Si ritrova quindi la vecchia idea di onda di materia di de Broglie, da cui aveva preso spunto la meccanica ondulatoria, ma in questo nuovo contesto il significato della funzione d'onda appare più chiaro e privo delle ambiguità tipiche delle prime discussioni sul dualismo onda-particella.

La sovrapposizione di onde piane può essere tale da produrre una funzione d'onda localizzata in una certa regione limitata. Ad esempio, in una dimensione, sia Δx l'ampiezza dell'intervallo entro cui la funzione è significativamente diversa da zero. L'ampiezza dell'intervallo dei valori di k , diciamo Δk , necessari per rappresentare tale funzione d'onda non è indipendente da Δx ; dalla teoria delle trasformate di Fourier si sa, ad esempio, che per una funzione gaussiana il prodotto $\Delta x \Delta k$ è una costante di ordine 1. Il risultato è consistente con la relazione di indeterminazione di Heisenberg, già ricavata alla fine della sezione precedente. La funzione gaussiana, intesa come funzione d'onda, è anche chiamata *pacchetto di minima indeterminazione*. Più in generale si parla di *pacchetti d'onda* quando le funzioni d'onda delle particelle sono localizzate entro distanze piccole rispetto a quelle rilevanti per la descrizione di un sistema complessivo. Ma non si faccia l'errore di identificare le particelle con i pacchetti d'onda. Sarebbe una scorciatoia rovinosa verso la fisica classica, che porta a risultati in contraddizione con le osservazioni, un po' come certe scorciatoie dei sentieri di montagna, apparentemente ovvie, ma che una volta imboccate richiedono ore di cammino per tornare indietro dopo aver perso la strada.

Abbiamo così concluso la descrizione dei concetti alla base della teoria, dando forma all'idea di stato fisico di un sistema e definendo cosa intendiamo per misura di una grandezza fisica (osservabile). Di tutto quanto scritto in questo capitolo, le idee forse più importanti e attorno alle quali tutto ruota, sono due: gli stati fisici obbediscono ad un principio di sovrapposizione e la misura di un'osservabile fa collassare lo stato di un sistema in uno degli autostati dell'osservabile, quello identificato dall'esito della misura (o meglio, nel caso di spettro degenere, nel sottospazio degli autostati che condividono lo stesso esito della misura). L'algebra dei vettori e degli operatori l'abbiamo costruita a partire da queste ipotesi nate dal tentativo di interpretare gli esiti di esperimenti con sequenze di apparati di Stern-Gerlach. Tali esperimenti hanno a che fare con una grandezza, lo spin, che non ha equivalente in fisica classica, malgrado assomigli a un momento angolare per il suo legame con il momento magnetico, e la cui origine non si trova nemmeno nell'ambito della meccanica quantistica non-relativistica. Occorrerà rendere la teoria compatibile con la relatività speciale di Einstein per capire cos'è lo spin. Ma poco importa; qui ci interessava prendere uno spunto da un esperimento paradigmatico per ricavarne suggerimenti su come muoverci, avendo davanti a noi una mappa vuota. Siamo partiti apposta dal punto più lontano dalla fisica classica che potevamo immaginare proprio per non essere fuorviati da pregiudizi che ci portiamo dietro da Newton o da Maxwell. Introducendo le osservabili posizione e momento ci stiamo

avvicinando verso regioni più familiari. Procediamo su questa strada. La prossima tappa è la formulazione del problema dinamico: introdurre il tempo e studiare come evolvono gli stati fisici.

Capitolo 2

Dinamica quantistica

2.1 Evoluzione temporale

Per introdurre il tempo è necessario per prima cosa decidere se la dinamica che vogliamo sviluppare debba essere invariante per trasformazioni di Galileo, come nella meccanica newtoniana, o per trasformazioni di Lorentz, come prescritto dalla relatività speciale di Einstein. Possiamo procedere per gradi. In questo corso cominciamo con l'introdurre una dinamica non-relativistica, per particelle che viaggiano a velocità molto minori della velocità della luce. Poi si potrà riformulare il tutto in modo che sia invariante per trasformazioni di Lorentz e questo lo lasciamo per i corsi più avanzati.

Dunque, prendiamo tempo e spazio come grandezze indipendenti. Lo spazio è definito tramite la misura della posizione. All'osservabile posizione si può associare un operatore hermitiano (tre operatori nello spazio tridimensionale) e la procedura di misura è quella descritta nel capitolo precedente. In particolare, data una particella che si trova in uno certo stato fisico, ha significato chiedersi quale sia la probabilità di ottenere come esito della misura della posizione un valore compreso in un dato intervallo. Nel caso del tempo questo non ha senso. Il tempo *non* è un'osservabile a cui si possa assegnare un operatore hermitiano t_{op} i cui autovalori sono i possibili valori misurati di t . In particolare non ha senso chiedersi quale sia la probabilità di ottenere un certo valore del tempo facendo una misura sugli stati fisici di un sistema. Il tempo è un parametro, che varia in modo autonomo rispetto alle configurazioni dei sistemi e che si misura con gli orologi. In un mondo in cui il tempo scorre è necessario piuttosto stabilire quale sia lo stato di un sistema *in un certo istante*. Nel capitolo precedente abbiamo implicitamente assunto che gli stati fossero assegnati in un certo istante, ma quale istante fosse non era rilevante. Ora lo diventa.

Supponiamo che al tempo t_0 un sistema si trovi nello stato generico rappresentato dal vettore $|\alpha\rangle$. Per introdurre nella notazione anche l'informazione sul tempo, usiamo il simbolo

$$|\alpha, t_0\rangle. \quad (2.1)$$

In generale ci aspettiamo che lo stato del sistema in un istante successivo t sia diverso. Indichiamo l'evoluzione in questo modo

$$|\alpha, t_0\rangle \longrightarrow |\alpha, t_0; t\rangle. \quad (2.2)$$

La trasformazione di un vettore in un altro può essere rappresentata tramite un operatore. Nel nostro caso lo indichiamo con $U(t, t_0)$ e lo chiamiamo *operatore di evoluzione temporale*; per definizione, esso agisce in modo che

$$|\alpha, t_0; t\rangle = U(t, t_0)|\alpha, t_0\rangle. \quad (2.3)$$

Il significato di questa espressione è che U agisce sullo stato in cui il sistema si trovava al tempo t_0 per trasformarlo nello stato in cui si trova al tempo $t > t_0$.

Se vogliamo che l'interpretazione in termini di probabilità degli esiti delle misure si conservi nel tempo, la norma dei vettori deve rimanere costante nel tempo, e questo implica che U deve essere un operatore unitario ($U^\dagger U = U U^\dagger = \mathbb{1}$). Inoltre, l'evoluzione temporale è equivalente ad una traslazione nel tempo e deve essere sviluppabile con continuità dall'identità, in modo che ad evoluzioni per un tempo infinitesimo corrispondano variazioni infinitesime degli stati. Infine, deve essere componibile per intervalli temporali successivi, cioè $U(t_2, t_0) = U(t_2, t_1)U(t_1, t_0)$, se $t_0 < t_1 < t_2$, tenendo conto dell'ordine del prodotto, ovvero il primo operatore che si applica è quello a destra. Queste proprietà sono identiche a quelle soddisfatte dall'operatore di traslazione spaziale $\tau_x(dx)$. In modo del tutto analogo a quel caso, anche qui possiamo scrivere l'operatore di traslazione temporale infinitesima nella forma

$$U(t + dt, t) = \mathbb{1} - i\Omega dt. \quad (2.4)$$

dove Ω è un operatore hermitiano con le dimensioni dell'inverso di un tempo. Questo operatore soddisfa i requisiti appena menzionati, esattamente come l'operatore

$$\tau_x(dx) = \mathbb{1} - \frac{i}{\hbar} p_x dx, \quad (2.5)$$

ampiamente discusso nel capitolo precedente.

A questo punto notiamo che in fisica classica le traslazioni temporali sono associate alla grandezza energia; ad esempio, se nella descrizione di un sistema la scelta dell'origine del tempo ($t = 0$) è irrilevante o, in altri termini, se il sistema è invariante per traslazioni temporali, allora l'energia si conserva. Analogamente, in sistemi invarianti per traslazione spaziale, il momento si conserva. Volendo mantenere nella nuova teoria quantistica le stesse simmetrie della fisica classica, siamo quindi incoraggiati ad associare all'operatore Ω la grandezza energia, così come avevamo fatto nel rimpiazzare l'operatore K con il momento p_x in (1.222). Dato che Ω ha le dimensioni dell'inverso di un tempo, possiamo scrivere

$$H = \hbar\Omega, \quad (2.6)$$

dove H è un nuovo operatore, che chiamiamo *operatore di Hamilton* (o, in breve, *hamiltoniano*) avente le dimensioni dell'energia. Se un sistema fisico, in qualche limite opportuno, ammette una descrizione classica, allora assumiamo anche che l'operatore H si possa ottenere dalla funzione di Hamilton $H(q, p)$ della meccanica classica sostituendo le coordinate generalizzate e i momenti coniugati con i corrispondenti operatori posizione e momento definiti nel capitolo precedente; la funzione di Hamilton in fisica classica rappresenta proprio l'energia del sistema in funzione di posizioni e momenti¹.

È interessante notare che le due relazioni tra operatori $p_x = \hbar K$ e $H = \hbar\Omega$, hanno la stessa forma delle relazioni derivate da L. de Broglie nella sua descrizione ondulatoria delle particelle, $p = \hbar k$ e $E = \hbar\omega$, ma qui il significato è diverso e più generale. L'analogia, comunque, sembra giustificare l'uso di $\hbar$ come fattore moltiplicativo nella definizione del generatore delle traslazioni, sia spaziali che temporali.

Dunque l'operatore di traslazione temporale infinitesima diventa

$$U(t + dt, t) = \mathbb{1} - \frac{i}{\hbar} H dt, \quad (2.7)$$

¹Inoltre, l'hamiltoniana classica è anche la funzione generatrice dell'evoluzione temporale, intesa come trasformazione canonica delle variabili dinamiche di un sistema (coordinate e momenti coniugati, presi come funzioni del tempo). In effetti, ricordarsi la meccanica analitica aiuta molto a cogliere il senso delle congettture utilizzate nel formulare la meccanica quantistica.

e H assume il ruolo di generatore delle traslazioni temporali.

A questo livello, la relazione tra la funzione di Hamilton classica e l'operatore di Hamilton quantistico è solo una congettura. Non siamo tenuti ora a giustificarla al di là di questi argomenti euristici; vediamo piuttosto quali sono le conseguenze.

2.2 Equazione di Schrödinger

Ci poniamo il problema di determinare lo stato di un sistema al tempo $t > t_0$, essendo noto lo stato al tempo t_0 . Un'idea che ci può venire in mente, ragionando in modo analogo a quanto fatto per la traslazioni spaziali in (1.229), è quella di suddividere l'intervallo temporale $t - t_0$ in N intervalli, con N che tende all'infinito, applicando a ciascuno l'evoluzione temporale infinitesima appena definita:

$$|\alpha, t_0; t\rangle = \lim_{N \rightarrow \infty} \left(\mathbb{1} - \frac{i}{\hbar} H \frac{t - t_0}{N} \right)^N |\alpha, t_0\rangle = \exp \left(-\frac{i}{\hbar} H(t - t_0) \right) |\alpha, t_0\rangle, \quad (2.8)$$

ovvero

$$U(t, t_0) = \exp \left[-\frac{iH(t - t_0)}{\hbar} \right]. \quad (2.9)$$

Tuttavia, dobbiamo stare ben attenti a quello che abbiamo fatto, perché c'è un'assunzione implicita in questi passaggi: valgono solo se l'operatore H non dipende dal tempo. In generale H può dipendere dal tempo²; in tal caso nella sequenza delle evoluzioni infinitesime per ogni intervallino dt' compreso tra t_0 e t , l'operatore $\mathbb{1} - \frac{i}{\hbar} H(t')dt'$ contiene un hamiltoniano diverso e non possiamo scrivere la sequenza intera come una potenza N -esima di uno stesso operatore. Ne parleremo più avanti; qui piuttosto conviene affrontare il problema in modo diverso.

La cosa migliore è calcolarsi la derivata di $U(t, t_0)$ a t_0 fissato. La derivata va intesa in questo senso: U è un operatore che cambia al variare di due variabili continue e si può definire il differenziale dell'operatore, dU , rispetto ad una delle due variabili, esattamente come si fa con le funzioni, ovvero

$$dU(t, t_0) = U(t + dt, t_0) - U(t, t_0) = \left(\frac{\partial U}{\partial t} \right) dt. \quad (2.10)$$

Poi usiamo il fatto che l'evoluzione da t_0 a $t + dt$ può essere ottenuta combinando un'evoluzione da t_0 a t , seguita da un'evoluzione infinitesima da t a $t + dt$:

$$\begin{aligned} U(t + dt, t_0) &= U(t + dt, t)U(t, t_0) \\ &= \left(\mathbb{1} - \frac{i}{\hbar} H dt \right) U(t, t_0) \\ &= U(t, t_0) - \frac{i}{\hbar} H U(t, t_0) dt, \end{aligned} \quad (2.11)$$

dove abbiamo tenuto conto che il tempo cresce da destra a sinistra nei prodotti degli operatori e che l'operatore di evoluzione infinitesima è dato dall'espressione (2.7). Mettendo assieme queste informazioni, otteniamo

$$\boxed{i\hbar \frac{\partial}{\partial t} U(t, t_0) = H U(t, t_0)}. \quad (2.12)$$

²Si pensi ad esempio al caso di un hamiltoniano di una particella carica che interagisce con un campo elettrico variabile nel tempo. Un altro esempio, che citeremo alla fine di questa sezione, è quello di uno spin in un campo magnetico variabile.

L'abbiamo incorniciata perché è un'equazione fondamentale. Tutta la dinamica quantistica sta qui dentro.

Se applichiamo i due operatori a sinistra e a destra dell'uguale in (2.12) ad uno stesso vettore generico $|\alpha, t_0\rangle$, otteniamo l'equazione

$$i\hbar \left(\frac{\partial}{\partial t} U(t, t_0) \right) |\alpha, t_0\rangle = H U(t, t_0) |\alpha, t_0\rangle, \quad (2.13)$$

ma $|\alpha, t_0\rangle$ è lo stato iniziale e non dipende da t ; quindi può essere inserito nella derivata in modo da ottenere

$$i\hbar \frac{\partial}{\partial t} (U(t, t_0) |\alpha, t_0\rangle) = H U(t, t_0) |\alpha, t_0\rangle, \quad (2.14)$$

ovvero

$$\boxed{i\hbar \frac{\partial}{\partial t} |\alpha, t_0; t\rangle = H |\alpha, t_0; t\rangle}. \quad (2.15)$$

Questa equazione differenziale fornisce il vettore di stato a qualsiasi tempo t maggiore di t_0 ; per essere risolta richiede la conoscenza del vettore di stato al tempo t_0 , che dipende da come il sistema è stato preparato, e della forma esplicita dell'hamiltoniano, che dipende dalla natura stessa del sistema che si vuole indagare³.

Questa equazione e la (2.12) sono generali ma astratte. Per tradurle in numeri ci serve scegliere una rappresentazione, ovvero scegliere una base ortonormale e completa per lo spazio dei vettori di stato.

Facciamo un esempio particolarmente significativo, ma non esaustivo. Prendiamo una particella che si muove in una dimensione e prendiamo come base l'insieme degli autostati della posizione $|x\rangle$; questa base non dipende dal tempo, per definizione, essendo costruita a partire dall'insieme di tutti i possibili esiti della misura di posizione. Introduciamo la funzione d'onda, stavolta dipendente anche dal tempo, come

$$\langle x | \alpha, t_0; t \rangle = \psi_\alpha(x, t). \quad (2.16)$$

Poi prendiamo la funzione di Hamilton classica $H = p^2/(2m) + V(x)$, dove m è la massa della particella⁴ e $V(x)$ è l'energia potenziale della particella, supponendo che questa sia soggetta ad un campo di forze conservative e indipendenti dal tempo⁵. Secondo quanto detto in precedenza, dobbiamo prendere la stessa grandezza H , ma stavolta considerata come un operatore, funzione degli operatori p e x (anche qui tralasciamo i suffissi “op” per semplicità, ma ricordiamoci che le osservabili sono operatori). L'operatore H è hermitiano, perché è la somma di due funzioni reali del momento p e della posizione x , che sono essi stessi hermitiani. Dunque, moltiplicando a sinistra per $\langle x |$ entrambi i membri della (2.15), otteniamo

$$i\hbar \frac{\partial}{\partial t} \langle x | \alpha, t_0; t \rangle = \langle x | \left(\frac{p^2}{2m} + V(x) \right) | \alpha, t_0; t \rangle, \quad (2.17)$$

³Seguendo una logica alternativa, il fatto che l'evoluzione temporale sia data dall'equazione (2.15) può essere visto come una conseguenza di alcuni requisiti che potremmo imporre alla nostra teoria, in particolare, quello di essere deterministica (l'equazione per gli stati deve essere del primo ordine in t per poter essere risolta ad ogni t se è noto lo stato ad un tempo t_0 assegnato), di preservare la linearità (la sovrapposizione di stati) e la norma (l'interpretazione probabilistica degli esiti delle misure), e di essere riducibile all'evoluzione classica nei casi in cui quest'ultima è applicabile (l'effetto dell'operatore H deve ricondursi a quello della funzione di Hamilton classica, generatrice delle traslazioni temporali). Si veda ad esempio il cap.3.1 del libro *Quantum Mechanics* di Auletta, Fortunato e Parisi (Cambridge Univ. Press, 2009).

⁴Ecco che qui compare per la prima volta la massa. Scrivendo H in questo modo diamo per assodato che il sistema è costituito da una particella la cui energia cinetica è $p^2/2m$. In questo contesto m è un parametro che si assume noto, nelle appropriate unità di misura, e mantiene lo stesso significato anche quando H diventa l'operatore hamiltoniano nell'ambito della meccanica quantistica.

⁵D'ora in poi useremo semplicemente la parola “potenziale” invece di energia potenziale, senza timore di fare confusione.

da cui

$$i\hbar \frac{\partial}{\partial t} \psi_\alpha(x, t) = \langle x | \frac{p^2}{2m} | \alpha, t_0; t \rangle + \langle x | V(x) | \alpha, t_0; t \rangle. \quad (2.18)$$

Nel primo termine a destra dell'uguale possiamo usare la (1.248), in modo che

$$\langle x | \frac{p^2}{2m} | \alpha, t_0; t \rangle = \int dx' \langle x | x' \rangle \left(-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x'^2} \right) \langle x' | \alpha, t_0; t \rangle, \quad (2.19)$$

dove la derivata al secondo membro stavolta è una derivata parziale perché la funzione d'onda dipende anche dal tempo. Usando l'ortogonalità della base e la definizione della funzione d'onda, si ottiene

$$\langle x | \frac{p^2}{2m} | \alpha, t_0; t \rangle = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi_\alpha(x, t). \quad (2.20)$$

Per il termine con il potenziale, basta notare che il *bra* a sinistra dell'operatore $V(x)$ è un autostato della posizione e, agendo su quello, l'operatore produce la funzione reale $V(x)$, dove qui la x indica gli autovalori e non più gli operatori. Dunque

$$\langle x | V(x) | \alpha, t_0; t \rangle = V(x) \langle x | \alpha, t_0; t \rangle = V(x) \psi_\alpha(x, t). \quad (2.21)$$

Mettendo assieme i pezzi arriviamo a

$$i\hbar \frac{\partial}{\partial t} \psi_\alpha(x, t) = \left(-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right) \psi_\alpha(x, t), \quad (2.22)$$

la cui estensione al caso tridimensionale è

$$\boxed{i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = \left(-\frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r}) \right) \psi(\mathbf{r}, t)}, \quad (2.23)$$

dove abbiamo anche tolto l'indice α , per compattezza di notazione, dato che l'equazione vale per qualsiasi stato generico. Questa è l'equazione di Schrödinger che gli studenti avranno già visto nei corsi introduttivi alla fisica moderna. Venne ricavata da Schrödinger sulla base di congetture diverse da quelle usate nella nostra formulazione della meccanica quantistica, in particolare da analogie con l'ottica geometrica. Già sappiamo, perché qualcuno ce l'ha raccontato, che risolvendo questa equazione si trovano i corretti stati stazionari di un elettrone nell'atomo di idrogeno⁶, e molto altro. Siamo quindi sulla strada giusta.

Il fatto di aver ottenuto la stessa equazione, così come Schrödinger l'aveva scritta, corrobora a posteriori la nostra scelta di identificare la grandezza $\mathbf{p}$, definita nel nostro caso come generatore delle traslazioni spaziali, con il momento di una particella analogo a quello classico e anche la scelta di identificare l'hamiltoniano H con il generatore delle traslazioni temporali⁷. Tuttavia vale la pena sottolineare che, per come ci siamo arrivati, l'equazione

⁶Questo lo vedremo più avanti in dettaglio, ma dovete avere pazienza, perché prima dovremo definire il momento angolare.

⁷Questa derivazione è consistente anche con la scelta di operare nel contesto della relatività galileiana. Si può infatti dimostrare, con pochi passaggi, che l'informazione fisica che si estrae dall'equazione di Schrödinger è invariante per trasformazioni di Galileo. Infatti, per due osservatori le cui coordinate sono legate da $t = t'$ e $\mathbf{r} = \mathbf{r}' + \mathbf{v}t$, dove $\mathbf{v}$ è la velocità relativa (intesa in senso classico) dei sistemi di riferimento, il momento e l'energia di una particella obbediscono alle relazioni $\mathbf{p} = \mathbf{p}' + m\mathbf{v}$ e $E = E' + \mathbf{v} \cdot \mathbf{p}' + mv^2/2$ e si può vedere facilmente che la funzione d'onda $\psi(\mathbf{r}, t)$, trovata da un osservatore come soluzione dell'equazione di Schrödinger, si trasforma nella funzione d'onda $\psi'(\mathbf{r}', t') = \psi(\mathbf{r} - \mathbf{v}t, t) \exp(i\phi/\hbar)$, con $\phi = m\mathbf{v} \cdot \mathbf{r} - (1/2)mv^2t$. Dato che un qualsiasi fattore di fase globale non cambia le informazioni fisiche sullo stato del sistema, possiamo concludere che l'invarianza per trasformazioni di Galileo è rispettata.

di Schrödinger nella forma (2.23) rappresenta solo un caso particolare dell'equazione (2.15), che pure possiamo chiamare equazione di Schrödinger per estensione, che ha validità più generale e si applica anche a stati fisici caratterizzati da osservabili che non hanno un analogo classico. Abbiamo quindi a disposizione un panorama più ampio.

A questo punto, ci fidiamo delle equazioni trovate e le trasformiamo in un postulato:

l'evoluzione temporale degli stati fisici è data dall'equazione di Schrödinger

$$i\hbar \frac{\partial}{\partial t} |\alpha, t_0; t\rangle = H |\alpha, t_0; t\rangle ,$$

dove H è l'hamiltoniano del sistema e $|\alpha, t_0; t\rangle$ è il vettore di stato al tempo generico t a partire da un vettore iniziale $|\alpha, t_0\rangle$.

Aggiungiamo questo postulato al principio di sovrapposizione per gli stati fisici e al postulato della misura delle osservabili discussi al capitolo precedente, in modo da dare corpo alla nuova teoria.

Notiamo che l'equazione di Schrödinger è un'equazione differenziale al primo ordine in t la cui soluzione è univoca, data la condizione iniziale: assegnato il vettore di stato in un istante generico, il vettore in un qualsiasi istante successivo è noto in modo deterministico. Questo garantisce che la teoria soddisfi un principio di *causalità*, che è un requisito implicito in ogni teoria fisica. Ma la causalità si applica ai vettori di stato e non agli esiti delle misure di osservabili, per i quali invece vale una regola di tipo probabilistico. In questo senso possiamo dire che la meccanica quantistica è una teoria causale delle probabilità.

Prima di procedere, torniamo un attimo all'equazione di Schrödinger (2.12) per l'operatore di evoluzione temporale. Avevamo visto che, se l'hamiltoniano H non dipende dal tempo, allora la soluzione è l'esponenziale (2.9). Tuttavia, come già avevamo detto, la soluzione generale non è sempre semplice. Se al tempo t_1 l'hamiltoniano del sistema è $H(t_1)$ e al tempo t_2 successivo è $H(t_2)$, non è nemmeno detto che i due operatori commutino. Prendiamo il solito esempio dell'atomo d'argento in un campo magnetico inizialmente diretto lungo z . L'energia dell'atomo, e quindi anche l'operatore hamiltoniano associato, contiene un termine proporzionale alla componente z del momento magnetico, a sua volta proporzionale alla componente z dello spin. Supponiamo che, mentre l'atomo passa tra i poli del magnete, lo sperimentatore aumenti o diminuisca l'intensità del campo senza cambiarne la direzione, allora l'hamiltoniano cambierà nel tempo solo tramite una funzione reale di t che moltiplica l'operatore S_z ; ne segue che l'hamiltoniano in t_2 commuta con l'hamiltoniano in t_1 , dato che S_z commuta con se stesso. Se al contrario lo sperimentatore, invece di cambiare l'intensità del campo ne cambia l'orientazione, allora l'hamiltoniano al tempo t_2 potrebbe essere proporzionale a S_y , e dunque non commuta con l'hamiltoniano al tempo t_1 , proporzionale a S_z . Se gli operatori H in istanti diversi commutano, allora si può seguire la procedura delle sequenze di N trasformazioni infinitesime, e si arriva al risultato

$$U(t, t_0) = \exp \left[- \left(\frac{i}{\hbar} \right) \int_{t_0}^t dt' H(t') \right] , \quad (2.24)$$

Se non commutano, allora l'ordine degli operatori diventa cruciale e le cose si complicano parecchio. Quello che si ottiene è noto come *serie di Dyson*. Qui per il momento sorvoliamo la questione. Ci torneremo quando parleremo dell'evoluzione temporale nella rappresentazione d'interazione. Nel seguito assumiamo che H sia indipendente dal tempo, salvo diversa indicazione.

2.3 Energia e costanti del moto

2.3.1 Stati stazionari e costanti del moto

L'operatore hamiltoniano H , essendo il generatore delle traslazioni temporali, è hermitiano per definizione. Dunque i suoi autovalori sono reali e i corrispondenti autovettori formano una base ortonormale completa dello spazio di Hilbert associato al sistema. Supponiamo per semplicità che gli autovalori siano discreti, ma quello che diremo qui si applica anche al caso di uno spettro continuo. Scriviamo l'equazione agli autovalori

$$H|E_i\rangle = E_i|E_i\rangle, \quad (2.25)$$

dove l'indice i corre su tutti gli autostati dello spettro. Il numero E_i lo identifichiamo con l'energia del sistema quando questo si trova nello stato rappresentato dal vettore $|E_i\rangle$.

Riscriviamo l'operatore di evoluzione temporale

$$U(t, 0) = \exp\left(-\frac{iHt}{\hbar}\right), \quad (2.26)$$

avendo scelto $t_0 = 0$ per comodità. Poi ci mettiamo nel caso in cui lo stato iniziale $|\alpha, 0\rangle$ coincide con uno degli autostati $|E_i\rangle$. Allora la sua evoluzione temporale diventa

$$|\alpha, 0; t\rangle = U(t, 0)|\alpha, 0\rangle = U(t, 0)|E_i\rangle = \exp\left(-\frac{iHt}{\hbar}\right)|E_i\rangle. \quad (2.27)$$

L'esponenziale va identificato con la corrispondente serie di potenze, in cui entrano potenze di H ; ogni potenza di ordine n agisce come n applicazioni in sequenza dello stesso H e, ad ogni applicazione, si ottiene l'autovalore associato all'autostato su cui H agisce. Alla fine, la serie di potenze rimane la stessa, ma con l'autovalore E_i al posto di ogni H . Il risultato è

$$|\alpha, 0; t\rangle = \exp\left(-\frac{iE_it}{\hbar}\right)|E_i\rangle. \quad (2.28)$$

Tutta l'evoluzione si riduce così ad una semplice moltiplicazione per un fattore di fase, $\exp(-i\omega_it)$, con frequenza fissata dall'energia di quello stato tramite la relazione $E_i = \hbar\omega_i$. Il vettore di stato continua a essere normalizzato a 1 e lo stato fisico del sistema non cambia.

Se invece lo stato iniziale è uno stato generico, lo possiamo comunque scrivere come combinazione lineare di autostati di H

$$|\alpha, 0\rangle = \sum_j |E_j\rangle \langle E_j|\alpha, 0\rangle = \sum_j c_j(0)|E_j\rangle, \quad (2.29)$$

dove $\langle E_j|\alpha, 0\rangle = c_j(0)$ è il coefficiente della decomposizione al tempo $t = 0$. Poi possiamo applicare l'operatore di evoluzione temporale come prima,

$$|\alpha, 0; t\rangle = \exp\left(-\frac{iHt}{\hbar}\right) \sum_j c_j(0)|E_j\rangle = \sum_j c_j(0) \exp\left(-\frac{iHt}{\hbar}\right)|E_j\rangle. \quad (2.30)$$

Agendo sugli autostati, l'esponenziale dell'hamiltoniano estrae gli autovalori. Stavolta ce ne sono tanti e bisogna sommare:

$$|\alpha, 0; t\rangle = \sum_j c_j(0) \exp\left(-\frac{iE_jt}{\hbar}\right)|E_j\rangle = \sum_j c_j(t)|E_j\rangle. \quad (2.31)$$

Come si vede, lo stato evoluto nel tempo è ancora una combinazione lineare degli autostati, ma ora i coefficienti sono questi:

$$c_j(t) = c_j(0) \exp\left(-\frac{iE_j t}{\hbar}\right). \quad (2.32)$$

Dunque, l'evoluzione temporale dello stato $|\alpha\rangle$ corrisponde a moltiplicare ciascun coefficiente della decomposizione per un fattore di fase, con frequenza fissata dall'autovalore dell'energia⁸. Questo implica che la probabilità di trovare un certo valore dell'energia non dipende da quando si esegue la misura; infatti, dato che la probabilità è fissata dal modulo quadro dei coefficienti, il fattore di fase non conta. Inoltre, se la distribuzione di probabilità non cambia nel tempo, ne segue che anche il valor medio dell'energia, $\langle H \rangle$ non dipende dal tempo. Questo è consistente con l'ipotesi che l'hamiltoniano H , come operatore, sia indipendente dal tempo o, in altri termini, il sistema sia invariante per traslazioni temporali. Come in fisica classica, questa invarianza implica la conservazione dell'energia.

Possiamo prendere una generica osservabile A , associata ad un operatore hermitiano che agisce nello stesso spazio vettoriale, e calcolarne il valore d'aspettazione in un generico istante t ,

$$\langle A \rangle_t = \langle \alpha, 0; t | A | \alpha, 0; t \rangle = \langle \alpha, 0 | U^\dagger(t, 0) A U(t, 0) | \alpha, 0 \rangle; \quad (2.33)$$

nel caso particolare in cui il sistema sia stato preparato al tempo $t = 0$ in uno degli autostati di H , ad esempio $|\alpha, 0\rangle = |E_j\rangle$, si ottiene

$$\langle A \rangle_t = \langle E_j | U^\dagger(t, 0) A U(t, 0) | E_j \rangle. \quad (2.34)$$

Poi possiamo usare l'espressione (2.26) per scrivere

$$\langle A \rangle_t = \langle E_j | \exp\left(\frac{iHt}{\hbar}\right) A \exp\left(-\frac{iHt}{\hbar}\right) | E_j \rangle \quad (2.35)$$

e applicando H ai suoi autostati si ottiene

$$\begin{aligned} \langle A \rangle_t &= \langle E_j | \exp\left(\frac{iE_j t}{\hbar}\right) A \exp\left(-\frac{iE_j t}{\hbar}\right) | E_j \rangle \\ &= \exp\left(\frac{iE_j t}{\hbar}\right) \exp\left(-\frac{iE_j t}{\hbar}\right) \langle E_j | A | E_j \rangle \\ &= \langle E_j | A | E_j \rangle = \langle A \rangle_0, \end{aligned} \quad (2.36)$$

e così troviamo che il valor medio di A al tempo t è lo stesso che al tempo iniziale. Concludiamo che il valore di aspettazione di una *qualsiasi* osservabile non dipende dal tempo se lo stato del sistema viene preparato in un autostato dell'hamiltoniano. Per questo motivo diciamo che qualsiasi autostato di H è uno *stato stazionario* del sistema.

Supponiamo di avere a disposizione un insieme massimale di osservabili compatibili tra loro e con l'hamiltoniano,

$$[A, B] = [B, C] = [A, C] = \dots = 0 \quad (2.37)$$

$$[A, H] = [B, H] = [C, H] = \dots = 0. \quad (2.38)$$

⁸Se rappresentiamo il coefficiente della decomposizione nel piano in cui la parte reale è l'ascissa e la parte immaginaria è l'ordinata, vediamo che l'evoluzione temporale è una semplice rotazione con velocità angolare ω .

Gli operatori $A, B, C, \dots$ ammettono una base comune di autovettori $\{|abc\dots\rangle\}$, la cui completezza può essere usata per esprimere l'operatore di evoluzione temporale (2.26) in questo modo:

$$U(t, 0) = \sum_{a,b,c,\dots} \exp\left(-\frac{iHt}{\hbar}\right) |abc\dots\rangle \langle abc\dots|. \quad (2.39)$$

Ma anche H ammette lo stesso spettro di autovettori, a ciascuno dei quali corrisponde un ben preciso valore di energia, $E_{a,b,c,\dots}$. Dunque l'operatore esponenziale trova alla sua destra suoi autovettori e agendo su di essi estrae i relativi autovalori:

$$U(t, 0) = \sum_{a,b,c,\dots} \exp\left(-\frac{iE_{a,b,c,\dots}t}{\hbar}\right) |abc\dots\rangle \langle abc\dots|. \quad (2.40)$$

Ne segue anche che l'evoluzione temporale di un vettore, inizialmente preparato in uno qualsiasi degli autovettori $|abc\dots\rangle$, si riduce a

$$|\alpha, 0; t\rangle = U(t, 0)|abc\dots\rangle = \exp\left(-\frac{iE_{a,b,c,\dots}t}{\hbar}\right) |abc\dots\rangle. \quad (2.41)$$

che è un altro modo di scrivere la (2.28) per gli stati stazionari di un sistema, usando come loro etichette, o come *numeri quantici*, gli autovalori di un insieme massimale di osservabili compatibili.

Per una particella che si muove nello spazio, possiamo usare al solito la rappresentazione in termini di funzioni d'onda. Basta prendere il risultato precedente e moltiplicare entrambi i membri per il generico *bra* $\langle \mathbf{r}|$. Definendo le funzioni d'onda $\psi(\mathbf{r}, 0) = \langle \mathbf{r}|abc\dots\rangle$ e $\psi(\mathbf{r}, t) = \langle \mathbf{r}|\alpha, 0; t\rangle$, otteniamo

$$\psi(\mathbf{r}, t) = \psi(\mathbf{r}, 0) \exp\left(-\frac{iEt}{\hbar}\right). \quad (2.42)$$

Questa espressione può essere inserita nell'equazione di Schrödinger (2.23) che diventa

$$\left(-\frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r})\right) \psi(\mathbf{r}) = E\psi(\mathbf{r}), \quad (2.43)$$

dove abbiamo ommesso il riferimento tempo arbitrario $t = 0$. In questa forma si chiama *equazione di Schrödinger stazionaria*. Il calcolo dello spettro dell'hamiltoniano $H = p^2/(2m) + V(\mathbf{r})$ e delle corrispondenti autofunzioni si riduce quindi al problema di risolvere un'equazione differenziale del secondo ordine nelle coordinate spaziali, soggetta ad appropriate condizioni al contorno⁹. Una volta trovate le soluzioni $\psi(\mathbf{r})$, la quantità $|\psi(\mathbf{r})|^2 d^3r$ fornisce, al solito, la probabilità di trovare la particella in un volumetto d^3r intorno alla posizione $\mathbf{r}$. Essendo lo stato stazionario, tale probabilità non dipende dal tempo.

Uno stato generico del sistema può essere sempre scritto come combinazione degli stati stazionari $|E_j\rangle$ come avevamo fatto nella (2.29). Allora avevamo mostrato che il valor medio dell'energia non cambia comunque nel tempo, anche in uno stato non stazionario; ma questa è una peculiarità dell'hamiltoniano, che ha un ruolo molto speciale nella dinamica. Per altre osservabili, invece, nulla ci garantisce che la distribuzione statistica degli esiti e il valore d'aspettazione non dipendano dal tempo. È ancora vero se il sistema si trova in uno stato

⁹Lo spettro dell'operatore hamiltoniano che descrive sistemi fisici ha anche l'importante proprietà di essere sempre limitato inferiormente: esiste sempre un autovalore minimo, finito, equivalente alla minima energia che il sistema può possedere. Lo stato di energia minima è detto *stato fondamentale*, o *ground state* del sistema. Lavorare con operatori che sono limitati dal basso evita anche alcuni problemi formali nella trattazione degli operatori hermitiani negli spazi di Hilbert, semplificando la vita dei fisici (almeno di quelli meno appassionati ai dettagli dell'analisi funzionale).

stazionario, ma non lo è più se lo stato è generico. Infatti, dato un operatore A qualsiasi, possiamo scrivere

$$\begin{aligned}
\langle A \rangle_t &= \langle \alpha, 0; t | A | \alpha, 0; t \rangle \\
&= \sum_{ij} c_j^*(0) \langle E_j | U^\dagger(t, 0) A U(t, 0) | E_i \rangle c_i(0) \\
&= \sum_{ij} c_j^*(0) c_i(0) \langle E_j | \exp\left(\frac{iHt}{\hbar}\right) A \exp\left(-\frac{iHt}{\hbar}\right) | E_i \rangle \\
&= \sum_{ij} c_j^*(0) c_i(0) \langle E_j | \exp\left(\frac{iE_j t}{\hbar}\right) A \exp\left(-\frac{iE_i t}{\hbar}\right) | E_i \rangle \\
&= \sum_{ij} c_j^*(0) c_i(0) \exp\left(-\frac{i(E_i - E_j)t}{\hbar}\right) \langle E_j | A | E_i \rangle.
\end{aligned} \tag{2.44}$$

Il risultato lo possiamo scrivere in questo modo

$$\langle A \rangle_t = \sum_{ij} c_j^*(0) c_i(0) A_{ji} e^{-i\omega_{ij}t}, \tag{2.45}$$

dove A_{ji} sono gli elementi della matrice che rappresenta A nella base degli autostati di H , i coefficienti c_i e c_j sono quelli al tempo zero, il valor medio $\langle A \rangle_t$ è calcolato al tempo t e, infine, la frequenza ω_{ij} è definita da

$$\hbar\omega_{ij} = E_i - E_j. \tag{2.46}$$

Questa potrebbe far pensare alla regola di quantizzazione di Bohr per l'emissione e l'assorbimento di radiazione da parte dell'atomo di idrogeno. Ci assomiglia, ma non è quella. Qui non ci sono scambi di energia tra il sistema e il resto del mondo. La regola di Bohr la ritroveremo quando parleremo di transizioni tra gli stati di un sistema indotte da perturbazioni esterne. Ma il fatto di poter associare una frequenza all'energia è comunque interessante e ha varie implicazioni. Un esempio semplice è la precessione dello spin in un campo magnetico, che vedremo tra un attimo.

Notiamo che, nel caso in cui A sia compatibile con H , cioè $[A, H] = 0$, allora la matrice che rappresenta A nella base comune di autovettori è diagonale, ovvero

$$\langle E_j | A | E_i \rangle = a_i \delta_{ij} \tag{2.47}$$

e questo implica

$$\langle A \rangle_t = \sum_i a_i |c_i(0)|^2 = \langle A \rangle \tag{2.48}$$

e di nuovo il valore d'aspettazione non dipende dal tempo. In conclusione, se A è un'osservabile compatibile con l'hamiltoniano, cioè $[A, H] = 0$, allora la distribuzione statistica degli esiti delle misure di A rimane la stessa, così come il suo valore d'aspettazione, indipendentemente da quale istante scegliamo per eseguire le misure. Si dice allora che A è una *costante del moto*:

$$[A, H] = 0 \implies A \text{ costante del moto.} \tag{2.49}$$

2.3.2 Precessione dello spin

In fisica classica una particella con momento magnetico $\boldsymbol{\mu}$, che si trova in un campo magnetico $\mathbf{B}$, ha un'energia $-\boldsymbol{\mu} \cdot \mathbf{B}$. Se il campo è diretto lungo z , allora l'energia è $-\mu_z B$ e il vettore

$\boldsymbol{\mu}$ precede attorno all'asse z , essendo sottoposto ad un momento di forze $\boldsymbol{\mu} \times \mathbf{B}$. La rotazione nel piano $x-y$ avviene con una velocità angolare che, nel sistema di misura CGS, è data da $\omega = |q|B/2mc$, dove q è la carica della particella a cui è associato il momento magnetico, m è la sua massa e c è la velocità della luce nel vuoto. La quantità ω è detta *frequenza di Larmor*. Se la particella carica ha un momento angolare orbitale $\mathbf{L}$, allora il suo momento magnetico è $\boldsymbol{\mu} = [q/(2mc)]\mathbf{L}$. Questo lo diamo per noto dai corsi di elettromagnetismo.

Ora prendiamo i soliti atomi di argento. Dalle misure con apparati di Stern-Gerlach sappiamo che la componente del loro momento magnetico lungo una qualsiasi direzione può assumere solo due valori, in contrasto con le predizioni classiche. Dato che il momento magnetico può sempre essere scritto come una costante per una grandezza che ha le dimensioni di un momento angolare, avevamo introdotto un momento angolare di spin $\mathbf{S}$ tale da ammettere soli i valori $\pm\hbar/2$, per le sue componenti. Abbiamo anche trattato le sue componenti come operatori hermitiani nella nuova teoria che stiamo formulando, ma non ci siamo ancora posti il problema di quale sia la costante di proporzionalità tra $\mathbf{S}$ e $\boldsymbol{\mu}$. Ora facciamo un passo in più. Per prima cosa, come fecero Stern e Gerlach, assumiamo che il momento magnetico dell'atomo di argento sia interamente determinato da un singolo elettrone. Poi, per analogia con il caso classico, assumiamo che la relazione tra momento magnetico e spin sia $\boldsymbol{\mu} = g[-e/(2m_e c)]\mathbf{S}$, dove e è la carica dell'elettrone¹⁰ e m_e è la sua massa, e dove abbiamo aggiunto un fattore numerico g (rapporto giromagnetico) che può dipendere dalla natura dell'osservabile spin. Con queste assunzioni, l'energia (in senso classico) dell'atomo nel campo magnetico uniforme B diretto lungo z è data da $-\mu_z B = [geB/(2m_e c)]S_z$ e il suo corrispondente operatore quantistico è

$$H = \frac{geB}{2m_e c} S_z. \quad (2.50)$$

Notiamo che S_z commuta con l'hamiltoniano, dato che commuta con se stesso. Dunque gli autostati di S_z , $\{|\uparrow\rangle, |\downarrow\rangle\}$ con autovalori $\pm\hbar/2$, definiti nelle (1.17)-(1.18), sono anche autostati di H . Gli autovalori corrispondenti di H sono quindi

$$E_{\pm} = \pm \frac{geB}{2m_e c} \frac{\hbar}{2}. \quad (2.51)$$

Ispirati dalla (2.46), definiamo una frequenza ω in modo che $\hbar\omega$ sia la differenza di energia tra i due autostati. Allora

$$\omega = \frac{geB}{2m_e c}. \quad (2.52)$$

Questa, a meno del fattore incognito g , ha la stessa forma della frequenza di Larmor classica per un elettrone. In questo modo possiamo riscrivere l'hamiltoniano nella forma

$$H = \omega S_z. \quad (2.53)$$

Fatte queste premesse, ora prepariamo lo stato iniziale dell'atomo di argento (ad esempio tramite un apparato di Stern-Gerlach con lo schermo forato) con lo spin orientato lungo x , cioè nello stato $|S_x \uparrow\rangle$. Sappiamo che questo vettore può essere decomposto così

$$|\alpha, 0\rangle = |S_x \uparrow\rangle = \frac{1}{\sqrt{2}}|\uparrow\rangle + \frac{1}{\sqrt{2}}|\downarrow\rangle. \quad (2.54)$$

¹⁰Qui consideriamo e positivo, così che l'elettrone ha carica $-e$.

Se lasciamo evolvere il sistema nel campo magnetico uniforme, dalla (2.32) sappiamo che lo stato dopo un tempo generico t sarà

$$\begin{aligned} |\alpha, 0; t\rangle &= \frac{1}{\sqrt{2}} e^{-iE_+ t/\hbar} |\uparrow\rangle + \frac{1}{\sqrt{2}} e^{-iE_- t/\hbar} |\downarrow\rangle \\ &= \frac{1}{\sqrt{2}} e^{-i\omega t/2} |\uparrow\rangle + \frac{1}{\sqrt{2}} e^{i\omega t/2} |\downarrow\rangle. \end{aligned} \quad (2.55)$$

Cosa succede se misuriamo S_x al tempo t ? Calcoliamoci la probabilità di trovare uno dei due autovalori:

$$\begin{aligned} \mathcal{P}_\uparrow &= |\langle S_x \uparrow | \alpha, 0; t \rangle|^2 \\ &= \left| \left(\frac{1}{\sqrt{2}} \langle \uparrow | + \frac{1}{\sqrt{2}} \langle \downarrow | \right) \left(\frac{1}{\sqrt{2}} e^{-i\omega t/2} |\uparrow\rangle + \frac{1}{\sqrt{2}} e^{i\omega t/2} |\downarrow\rangle \right) \right|^2 \\ &= \left| \frac{1}{2} (e^{-i\omega t/2} + e^{i\omega t/2}) \right|^2 = \cos^2 \frac{\omega t}{2} \end{aligned} \quad (2.56)$$

e similmente

$$\mathcal{P}_\downarrow = |\langle S_x \downarrow | \alpha, 0; t \rangle|^2 = \sin^2 \frac{\omega t}{2}, \quad (2.57)$$

dato che $|S_x \downarrow\rangle$ differisce solo per il segno tra i due addendi della decomposizione. Da queste espressioni possiamo ricavare il valor medio di S_x

$$\langle S_x \rangle_t = \frac{\hbar}{2} \mathcal{P}_\uparrow - \frac{\hbar}{2} \mathcal{P}_\downarrow = \frac{\hbar}{2} \left(\cos^2 \frac{\omega t}{2} - \sin^2 \frac{\omega t}{2} \right) = \frac{\hbar}{2} \cos \omega t. \quad (2.58)$$

Procedendo allo stesso modo possiamo calcolare anche il valore medio di S_y , per il quale si trova

$$\langle S_y \rangle_t = \frac{\hbar}{2} \sin \omega t. \quad (2.59)$$

Entrambi i valori medi oscillano periodicamente nel tempo con frequenza fissata dalla differenza di energia tra i due autostati. Il risultato è analogo a quello della precessione classica di un momento magnetico che ruota nel piano x - y , perpendicolarmente al campo magnetico. Ma attenzione, perché qui non c'è alcun vettore che ruota nello spazio euclideo; le quantità che oscillano nel tempo sono le probabilità $\mathcal{P}_\uparrow$ e $\mathcal{P}_\downarrow$ e la precessione avviene nello spazio astratto dei vettori di stato. In pratica, prepariamo gli atomi di argento del fascio tutti allo stesso modo in $|S_x \uparrow\rangle$; li facciamo passare per un certo tempo t in un campo magnetico uniforme B diretto lungo z e, al tempo t , usiamo uno Stern-Gerlach lungo x per vedere quanti atomi vanno in una direzione o nell'altra, corrispondenti agli stati $|S_x \uparrow\rangle$ e $|S_x \downarrow\rangle$. Possiamo ripetere le misure anche con lo Stern-Gerlach lungo y . Nella misura, ogni atomo verrà trovato in una o nell'altra direzione. La probabilità associata a ciascuna scelta la otteniamo misurando l'intensità delle macchie sullo schermo in uscita dallo Stern-Gerlach. Ripetiamo l'esperimento per diversi valori di t e troveremo che l'intensità delle macchie varia in modo periodico, l'una come $\cos^2(\omega t/2)$ e l'altra come $\sin^2(\omega t/2)$. Questo è ciò che succede. Non c'è alcun vettore che ruota. Per avere un vettore euclideo che ruota dovrebbe essere possibile in ogni istante conoscere le tre componenti cartesiane del vettore. Ma le tre componenti dello spin sono osservabili incompatibili. Su questo torneremo più oltre, quando parleremo del momento angolare e del suo legame con le rotazioni.

Dalla misura delle componenti dello spin a tempi diversi, possiamo ricavare empiricamente il rapporto giromagnetico g . Basta eseguire un *fit* ai dati con una curva sinusoidale per estrarre frequenza ω e, da questa, g . Gli esperimenti mostrano che $g = 2$, entro un margine dell'un per mille¹¹. I due valori ammessi per μ_z risultano quindi essere

¹¹Il rapporto giromagnetico può essere ottenuto anche da misure di spettroscopia, dato che lo spin entra in gioco nella struttura degli spettri degli atomi e delle molecole.

$\mu_z = \pm e\hbar/(2m_e c) = \pm\mu_B$, dove μ_B è detto *magnetone di Bohr*. Questo lo prendiamo come un fatto empirico. Per dare conto del valore $g = 2$ dal punto di vista della teoria sarà necessario ricorrere alla versione relativistica della meccanica quantistica.

Vale la pena sottolineare che gli stessi risultati (2.58) e (2.59) potevano essere ottenuti direttamente dalla relazione (2.45). Si tratta di un semplice esercizio. Basta usare come base quella degli autostati $|\uparrow\rangle$ e $|\downarrow\rangle$ di S_z e ricordarsi la rappresentazione matriciale degli operatori S_x e S_y di cui vogliamo calcolare il valore di aspettazione:

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (2.60)$$

Per il valor medio di S_x possiamo scrivere

$$\langle S_x \rangle_t = c_{\uparrow}^* c_{\downarrow} S_{x,\uparrow\downarrow} e^{-i\omega_{\downarrow\uparrow}t} + c_{\downarrow}^* c_{\uparrow} S_{x,\downarrow\uparrow} e^{-i\omega_{\uparrow\downarrow}t}, \quad (2.61)$$

dove $c_{\uparrow} = c_{\downarrow} = 1/\sqrt{2}$ e dove abbiamo tenuto conto del fatto che solo due elementi di matrice sono non nulli, quelli fuori diagonale, che valgono entrambi $\hbar/2$. E dunque

$$\langle S_x \rangle_t = \frac{\hbar}{4} (e^{i\omega t} + e^{-i\omega t}) = \frac{\hbar}{2} \cos \omega t, \quad (2.62)$$

con la frequenza di Larmor

$$\omega_{\uparrow\downarrow} = \omega = \frac{geB}{2m_e c} \quad (2.63)$$

e $\omega_{\downarrow\uparrow} = -\omega_{\uparrow\downarrow}$. Analogamente per il valor medio di S_y , abbiamo

$$\langle S_y \rangle_t = c_{\uparrow}^* c_{\downarrow} S_{y,\uparrow\downarrow} e^{-i\omega_{\downarrow\uparrow}t} + c_{\downarrow}^* c_{\uparrow} S_{y,\downarrow\uparrow} e^{-i\omega_{\uparrow\downarrow}t}, \quad (2.64)$$

con gli stessi coefficienti $c_{\uparrow} = c_{\downarrow} = 1/\sqrt{2}$ e con gli elementi di matrice non diagonali che valgono $\mp i\hbar/2$. E dunque

$$\langle S_y \rangle_t = \frac{i\hbar}{4} (-e^{i\omega t} + e^{-i\omega t}) = \frac{\hbar}{2} \sin \omega t. \quad (2.65)$$

2.4 Rappresentazione di Heisenberg

Torniamo al problema generale dell'evoluzione temporale degli stati fisici. Se al tempo iniziale il sistema è rappresentato da un vettore di stato $|\alpha\rangle$, allora lo stato del sistema dopo un certo tempo è dato da $U|\alpha\rangle$ dove U è un operatore unitario

$$|\alpha\rangle \longrightarrow U|\alpha\rangle. \quad (2.66)$$

Prendiamo un operatore generico A (non necessariamente un'osservabile), che agisce nello stesso spazio vettoriale, e calcoliamo l'elemento di matrice di A tra due stati qualsiasi

$$\langle \alpha | A | \beta \rangle. \quad (2.67)$$

Secondo la logica seguita fin qui, l'operatore è qualcosa di definito indipendentemente dal tempo, mentre gli stati evolvono nel tempo. Perciò, l'evoluzione temporale dell'elemento di matrice è

$$\langle \alpha | A | \beta \rangle \longrightarrow (\langle \alpha | U^\dagger) A (U | \beta \rangle) = \langle \alpha | U^\dagger A U | \beta \rangle. \quad (2.68)$$

Senza modificare in alcun modo il risultato, questa stessa scrittura può essere letta in un modo alternativo: i vettori di stato rimangono invariati nel tempo, mentre l'operatore A evolve così

$$A \longrightarrow U^\dagger A U. \quad (2.69)$$

La scelta (2.66), con operatori che rimangono invariati nel tempo (e di conseguenza anche i loro autovettori e autovalori), è detta *rappresentazione di Schrödinger* dell'evoluzione temporale, mentre la scelta (2.69), con i vettori di stato che rimangono invariati nel tempo, è detta *rappresentazione di Heisenberg*. Per non fare confusione, qui di seguito usiamo i simboli $A^{(S)}$ e $A^{(H)}$ per indicare l'operatore nella rappresentazione rispettivamente di Schrödinger e di Heisenberg. Per l'operatore U usiamo la definizione (2.7), in modo che obbedisca all'equazione (2.12), che è indipendente dalla rappresentazione.

Se l'hamiltoniano del sistema H non dipende dal tempo, allora possiamo scrivere

$$U(t) = \exp\left(-\frac{iHt}{\hbar}\right). \quad (2.70)$$

dove abbiamo preso $t_0 = 0$ nella (2.9) e alleggerito la notazione. Possiamo definire un'operatore $A^{(H)}$ a partire dal suo corrispondente $A^{(S)}$ in questo modo

$$A^{(H)}(t) = U^\dagger(t)A^{(S)}U(t). \quad (2.71)$$

Al tempo iniziale $t = 0$ i due operatori coincidono, $A^{(H)}(t) = A^{(S)}$, e anche i vettori di stato coincidono, $|\alpha, 0\rangle_{(H)} = |\alpha, 0\rangle_{(S)} = |\alpha, 0\rangle$. Negli istanti successivi, lo stato nella rappresentazione di Schrödinger evolve, mentre quello nella rappresentazione di Heisenberg no, ma il valore di aspettazione dell'osservabile A rimane lo stesso nelle due rappresentazioni¹². Infatti

$$\begin{aligned} \langle A \rangle^{(S)} &= {}_{(S)}\langle \alpha, 0; t | A^{(S)} | \alpha, 0; t \rangle_{(S)} \\ &= \langle \alpha, 0 | U^\dagger(t) A^{(S)} U(t) | \alpha, 0 \rangle \\ &= \langle \alpha, 0 | A^{(H)}(t) | \alpha, 0 \rangle = \langle A \rangle^{(H)}. \end{aligned}$$

Abbiamo già visto che l'evoluzione temporale dei vettori di stato nella rappresentazione di Schrödinger è governata dall'equazione (2.15). Ora ci chiediamo quale sia l'equazione che governa l'evoluzione degli operatori nella rappresentazione di Heisenberg. Per arrivare alla risposta, torniamo alla definizione (2.71) e deriviamola rispetto a t :

$$\begin{aligned} \frac{d}{dt}A^{(H)}(t) &= \frac{d}{dt}[U^\dagger(t)A^{(S)}U(t)] \\ &= \left(\frac{\partial U^\dagger(t)}{\partial t}\right)A^{(S)}U(t) + U^\dagger(t)A^{(S)}\left(\frac{\partial U(t)}{\partial t}\right). \end{aligned} \quad (2.72)$$

D'altra parte, per le derivate di U vale l'equazione (2.12), da cui

$$\frac{\partial U}{\partial t} = \frac{1}{i\hbar}HU, \quad \frac{\partial U^\dagger}{\partial t} = -\frac{1}{i\hbar}U^\dagger H, \quad (2.73)$$

che inserite nella precedente

$$\begin{aligned} \frac{d}{dt}A^{(H)}(t) &= -\frac{1}{i\hbar}U^\dagger(t)HA^{(S)}U(t) + \frac{1}{i\hbar}U^\dagger(t)A^{(S)}HU(t) \\ &= -\frac{1}{i\hbar}HU^\dagger(t)A^{(S)}U(t) + \frac{1}{i\hbar}U^\dagger(t)A^{(S)}U(t)H, \end{aligned} \quad (2.74)$$

¹²Qui ci concentriamo sui valori di aspettazione, ma si potrebbe dire di più, volendo approfondire. In particolare, sebbene gli autovettori di $A^{(H)}(t)$ cambino nel tempo, i suoi autovalori restano gli stessi, indipendenti dal tempo, come conseguenza del fatto che la trasformazione che definisce $A^{(H)}(t)$ è unitaria. Inoltre si può dimostrare che la distribuzione di probabilità degli esiti delle misure di A è la stessa nelle due rappresentazioni.

dove nella seconda uguaglianza abbiamo usato il fatto che U e H commutano. Ora, usando la definizione di $A^{(H)}$, troviamo

$$\begin{aligned}\frac{d}{dt}A^{(H)}(t) &= -\frac{1}{i\hbar}HA^{(H)}(t) + \frac{1}{i\hbar}A^{(H)}(t)H \\ &= \frac{1}{i\hbar}[A^{(H)}(t), H].\end{aligned}\quad (2.75)$$

Dunque, se H non dipende dal tempo, l'equazione per l'evoluzione temporale degli operatori è

$$\boxed{\frac{d}{dt}A^{(H)} = \frac{1}{i\hbar}[A^{(H)}, H]}.\quad (2.76)$$

Questa si chiama *equazione del moto di Heisenberg* e merita di essere incorniciata.

Usare questa per gli operatori, tenendo i vettori di stato costanti, oppure l'equazione di Schrödinger per i vettori di stato, con gli operatori costanti, porta agli stessi risultati fisici¹³. Scegliere una rappresentazione o l'altra è una questione di convenienza, ma il fatto di avere l'alternativa è comunque un vantaggio: ci sono aspetti della teoria che si chiariscono meglio usando una delle due rappresentazioni, o l'altra¹⁴. Ad esempio, l'equazione del moto di Heisenberg si presta ad un parallelo importante con la fisica classica. Infatti, dalla meccanica analitica sappiamo che, se H è la funzione di Hamilton e A è una funzione delle coordinate generalizzate e dei momenti coniugati che non dipenda esplicitamente dal tempo, vale

$$\frac{dA}{dt} = [A, H]_{\text{Poisson}}.\quad (2.77)$$

Come conseguenza, se la parentesi di Poisson è nulla, allora $dA/dt = 0$ e quindi A è una grandezza conservata, una costante del moto. In meccanica quantistica abbiamo la relazione analoga per i commutatori. Se A è un osservabile che commuta con l'hamiltoniano, allora è una costante del moto. L'analogia formale tra i due risultati si ottiene se

$$[\ , \]_{\text{Poisson}} \leftrightarrow \frac{1}{i\hbar}[\ , \]\quad (2.78)$$

ma questa l'avevamo già scritta in (1.223). Il risultato era lo stesso, ma in quel caso veniva dall'aver identificato l'operatore $\mathbf{K}$ delle traslazioni spaziali infinitesime con $\mathbf{p}/\hbar$, mentre in questo caso ci siamo arrivati identificando l'operatore Ω per l'evoluzione temporale infinitesima con $H/\hbar$. Nessuno ci garantiva a priori che la costante di proporzionalità fosse la stessa nei due casi. A posteriori si vede che quella scelta porta alla stessa relazione tra commutatori e parentesi di Poisson. È un buon indizio che siamo nel giusto. Dirac la prese come regola per quantizzare la meccanica classica, partendo dal presupposto che la fisica classica e quella quantistica dovessero condividere le stesse strutture di simmetria (stesso legame tra costanza del momento e invarianza traslazionale, costanza dell'energia e invarianza temporale, ecc.).

¹³Per chi volesse approfondire, suggerisco la lettura del capitolo 5.3 del Forte-Rottoli, dove il formalismo è più generale, in quanto include il caso in cui l'hamiltoniano stesso contiene una dipendenza (parametrica) dal tempo. Si noti poi che anche l'operatore $A^{(S)}$ potrebbe contenere una dipendenza esplicita dal tempo, non contemplata nella derivazione in (2.72); questo accade quando si quantizza una grandezza fisica classica che dipende dal tempo sia implicitamente, tramite la dipendenza da t delle variabili $\mathbf{r}$ e $\mathbf{p}$, sia esplicitamente. In tal caso, dovremmo aggiungere la derivata $\partial A^{(H)}/\partial t$ al membro di destra dell'equazione del moto di Heisenberg.

¹⁴Il fatto che si possa spostare la dipendenza temporale tra stati e operatori ci dice che gli uni e gli altri entrano nella teoria con ruoli complementari e di uguale rilevanza. L'importante è che questo spostamento (il cambio di rappresentazione) avvenga tramite trasformazioni unitarie, che sono quelle che garantiscono l'invarianza dei prodotti interni tra vettori e degli elementi di matrice degli operatori, e, di conseguenza, anche l'invarianza delle distribuzioni di probabilità, ovvero di tutto il contenuto fisico della teoria. Vedremo più avanti anche un terzo tipo di rappresentazione, quella di interazione, o di Dirac.

In questo senso si può dire che la procedura di quantizzazione canonica di Dirac è un modo elegante di applicare il principio di corrispondenza di Bohr, laddove la corrispondenza viene identificata con l'invarianza della struttura di simmetria dei sistemi meccanici. Ma è bene ricordare che la meccanica quantistica non si esaurisce con la quantizzazione della fisica classica, dato che essa comprende anche osservabili che non hanno analogo classico. A tali osservabili possiamo ancora applicare l'equazione del moto di Heisenberg, senza che per esse esista un'analogia parentesi di Poisson.

Per dare un'idea concreta di come la rappresentazione di Heisenberg può portare a risultati interessanti in modo efficace possiamo tornare all'esempio della precessione dello spin descritta nella sezione precedente. Se il nostro scopo è ricavare le relazioni (2.58) e (2.59) per i valori medi dello spin lungo x e lungo y in funzione del tempo, allora basta partire dall'hamiltoniano (2.53) e scrivere le equazioni del moto di Heisenberg per gli operatori S_x e S_y :

$$\begin{aligned}\frac{dS_x}{dt} &= \frac{1}{i\hbar}[S_x, H] = \frac{\omega}{i\hbar}[S_x, S_z] \\ \frac{dS_y}{dt} &= \frac{1}{i\hbar}[S_y, H] = \frac{\omega}{i\hbar}[S_y, S_z].\end{aligned}\quad (2.79)$$

Ora possiamo ricordare le regole di commutazione (1.133) per gli operatori di spin, in particolare $[S_x, S_z] = -i\hbar S_y$ e $[S_y, S_z] = i\hbar S_x$. Con queste, le due equazioni del moto diventano

$$\begin{aligned}\frac{dS_x}{dt} &= -\omega S_y \\ \frac{dS_y}{dt} &= \omega S_x.\end{aligned}\quad (2.80)$$

Derivando di nuovo rispetto al tempo ciascuna delle due si ottengono le due equazioni

$$\begin{aligned}\frac{d^2 S_x}{dt^2} &= -\omega^2 S_x \\ \frac{d^2 S_y}{dt^2} &= -\omega^2 S_y.\end{aligned}\quad (2.81)$$

Queste sono equazioni per gli operatori nella rappresentazione di Heisenberg. Possiamo prendere il valore di aspettazione del membro di sinistra e del membro di destra, calcolati su uno stesso stato generico:

$$\begin{aligned}\frac{d^2 \langle S_x \rangle_t}{dt^2} &= -\omega^2 \langle S_x \rangle_t \\ \frac{d^2 \langle S_y \rangle_t}{dt^2} &= -\omega^2 \langle S_y \rangle_t.\end{aligned}$$

Qui non serve più specificare la rappresentazione, perché i valori di aspettazione sono uguali nelle due rappresentazioni di Heisenberg e Schrödinger. Inoltre, queste due equazioni sono equazioni differenziali per funzioni reali di variabile reale (il tempo); in particolare, sono semplicemente due equazioni armoniche disaccoppiate, le cui soluzioni sono $\langle S_x \rangle_t = \langle S_x \rangle_{\max} \cos(\omega t + \delta_x)$ e $\langle S_y \rangle_t = \langle S_y \rangle_{\max} \cos(\omega t + \delta_y)$. Le ampiezze e le fasi possono essere fissate conoscendo le condizioni iniziali. Ad esempio, se al tempo 0 lo spin era diretto lungo x , allora le variabili d'integrazione sono $\langle S_x \rangle_{\max} = \hbar/2$ e $\delta_x = 0$ per lo spin lungo x , e $\langle S_y \rangle_{\max} = \hbar/2$ e $\delta_y = -\pi/2$ per lo spin lungo y , in modo che

$$\begin{aligned}\langle S_x \rangle_t &= (\hbar/2) \cos \omega t \\ \langle S_y \rangle_t &= (\hbar/2) \sin \omega t.\end{aligned}\quad (2.82)$$

Queste coincidono con le (2.58) - (2.59) calcolate nella rappresentazione di Schrödinger.

2.5 Teorema di Ehrenfest e limite classico

Consideriamo una particella libera nello spazio tridimensionale. Il suo hamiltoniano è

$$H = \frac{p^2}{2m} = \frac{p_x^2 + p_y^2 + p_z^2}{2m} \quad (2.83)$$

dove ciascuno dei p in questa espressione va inteso come operatore.

Se vogliamo usare la rappresentazione di Heisenberg ci serve calcolare qualche commutatore, usando le regole (1.45)-(1.49). È immediato vedere che nel caso della particella libera si ha

$$[p_i, H] = \frac{1}{2m}[p_i, p^2] = 0, \quad (2.84)$$

per $i = x, y, z$, e dunque gli autostati dell'hamiltoniano sono anche autostati del momento $\mathbf{p}$ e il momento è una costante del moto, come in fisica classica.

Poi possiamo scrivere l'equazione del moto di Heisenberg per l'operatore posizione. Ad esempio, per la componente x , usando le commutazioni canoniche (1.224)-(1.225) si ha

$$\frac{dx}{dt} = \frac{1}{i\hbar}[x, H] = \frac{1}{2im\hbar}[x, p^2] = \frac{1}{2im\hbar}[x, p_x^2] = \frac{p_x}{m} \quad (2.85)$$

L'ultimo commutatore può essere calcolato riscrivendolo nella forma $[x, p_x p_x]$ e applicando la regola (1.49) valida per qualsiasi terna di operatori. Si trova

$$\frac{dx}{dt} = \frac{p_x}{m} \quad (2.86)$$

e altrettanto per le altre componenti. Dunque l'equazione del moto è

$$\frac{d\mathbf{r}}{dt} = \frac{\mathbf{p}}{m} \quad (2.87)$$

che integrata nel tempo dà

$$\mathbf{r}(t) = \mathbf{r}(0) + \left(\frac{\mathbf{p}(0)}{m} \right) t. \quad (2.88)$$

In fisica classica, questo corrisponderebbe al moto uniforme con velocità $\mathbf{v} = \mathbf{p}/m$. Qui però $\mathbf{r}$ e $\mathbf{p}$ sono operatori e non abbiamo ancora definito la grandezza *velocità*. La difficoltà di definire una velocità in meccanica quantistica si capisce immediatamente se si applica la relazione di indeterminazione (1.174) all'operatore posizione a tempi diversi. Nella rappresentazione di Heisenberg gli operatori cambiano nel tempo e non è detto che, se commutano al tempo 0, commutino anche la tempo $t > 0$. Nel caso della posizione x infatti abbiamo

$$[x(t), x(0)] = \left[\left(x(0) + \frac{p_x(0)}{m} t \right), x(0) \right] = \frac{t}{m} [p_x(0), x(0)] = -\frac{i\hbar t}{m}. \quad (2.89)$$

e la relazione d'indeterminazione diventa

$$(\Delta x)_t (\Delta x)_0 \geq \frac{\hbar t}{2m}. \quad (2.90)$$

Il significato è il seguente. Se nell'istante $t = 0$ una particella si trova in uno stato tale per cui una misura della sua posizione darebbe un certo valore della coordinata x con un'indeterminazione $(\Delta x)_0$, allora una misura di posizione effettuata in un tempo successivo t avrebbe un'indeterminazione $(\Delta x)_t$ maggiore o uguale di $\hbar t/[2m(\Delta x)_0]$. Aumentare la precisione della prima misura, diminuendo $(\Delta x)_0$, fa automaticamente crescere l'indeterminazione sulla seconda. Ciò impedisce di definire in modo rigoroso un'osservabile che abbia il significato

di velocità in termini di distanza percorsa in un certo intervallo di tempo, ottenuta misurando la posizione di una particella in due istanti successivi. Ne segue anche, di conseguenza, l'impossibilità di determinare la *traiettoria* di una particella, come luogo dei punti percorsi in istanti successivi: ogni volta che si volesse verificare la traiettoria con una misura di posizione, si introdurrebbe una nuova indeterminazione aggiuntiva nelle misure successive. La traiettoria riacquista un senso solo quando, nel problema studiato, l'indeterminazione $(\Delta x)_t$ rimane comunque molto inferiore alle scale di lunghezza utilizzate per descrivere la dinamica del sistema.

Gli autostati dell'hamiltoniano della particella libera coincidono con gli autostati del momento $\mathbf{p}$. Come abbiamo visto in sezione 1.12, nel formalismo delle funzioni d'onda nello spazio delle coordinate questi autostati sono rappresentati da onde piane

$$\psi(\mathbf{r}) = \langle \mathbf{r} | \mathbf{p} \rangle = \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(\frac{i\mathbf{p} \cdot \mathbf{r}}{\hbar}\right) \quad (2.91)$$

che sono anche le soluzioni dell'equazione di Schrödinger stazionaria

$$-\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}) = E \psi(\mathbf{r}), \quad (2.92)$$

con

$$E = \frac{\mathbf{p}^2}{2m}. \quad (2.93)$$

Lo stato generico che rappresenta una particella che, in un certo istante, si trova in una certa regione limitata di spazio sarà scrivibile come combinazione lineare (un integrale) di onde piane con diversi valori di momento. Un modo alternativo per esprimere quanto detto al paragrafo precedente sta nel considerare la limitazione imposta dalla relazione di indeterminazione di Heisenberg $\Delta x \Delta p_x \geq \hbar/2$ (o dalle trasformate di Fourier, che la traducono nel linguaggio delle funzioni complesse): se la combinazione di onde piane è tale da produrre una funzione d'onda confinata in una regione di spazio molto piccola allora, nello stesso istante, si avrà un'indeterminazione molto grande sul valore del momento. Più ci si incaponisce a conoscere con precisione dove una particella si trova e meno potremo sapere su dove andrà in seguito.

Possiamo riscrivere la relazione (2.87) per i valori medi delle osservabili supponendo che la particella si trovi in un generico stato $|\alpha\rangle$. Dato che i vettori di stato nella rappresentazione di Heisenberg non dipendono dal tempo, possiamo identificare il valore medio della derivata di $\mathbf{r}$ con la derivata del valor medio. Dunque

$$\frac{d\langle \mathbf{r} \rangle}{dt} = \left\langle \frac{\mathbf{p}}{m} \right\rangle. \quad (2.94)$$

Anche qui possiamo interpretare il membro di sinistra come una velocità, ma stavolta si tratta di una velocità associata alla posizione media ed è opportuno insistere sul fatto che non si tratta di una media della posizione di una stessa particella misurata più volte, ma di una media delle posizioni di un insieme di particelle identiche preparate allo stesso modo.

Ora supponiamo che la particella non sia libera, ma agiscano su di essa forze conservative a cui corrisponda un potenziale scalare $V(\mathbf{r})$. In meccanica quantistica questo potenziale va inteso come funzione degli operatori x, y, z . Nella rappresentazione di Schrödinger il problema si affronta risolvendo l'equazione (2.23). Nella rappresentazione di Heisenberg invece possiamo scrivere

$$\frac{dp_x}{dt} = \frac{1}{i\hbar} [p_x, H] = \frac{1}{i\hbar} \left[p_x, \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) \right] = \frac{1}{i\hbar} [p_x, V(\mathbf{r})]. \quad (2.95)$$

Per calcolare il commutatore a destra dell'uguale, ricordiamo che $V(\mathbf{r})$ è un operatore, funzione degli operatori posizione x, y, z . Prendiamo il caso semplice di una funzione generica, $V(x)$, della sola posizione lungo x . La funzione di operatori è definita dalla sua serie di Taylor; nella serie che definisce la generica $V(x)$ entrano termini con potenze x^n . Per ciascuno di questi termini dimostriamo che vale

$$[p_x, x^n] = \frac{\hbar}{i} n x^{n-1} = \frac{\hbar}{i} \frac{dx^n}{dx}. \quad (2.96)$$

Si dimostra per induzione. Infatti, per $n = 1$ la regola coincide con la commutazione canonica tra p_x e x . Inoltre, supponiamo che valga per un generico n e dimostriamo che allora vale per $n + 1$:

$$[p_x, x^{n+1}] = [p_x, x x^n] = x [p_x, x^n] + [p_x, x] x^n = \frac{\hbar}{i} n x^n + \frac{\hbar}{i} x^n = \frac{\hbar}{i} (n+1) x^n. \quad (2.97)$$

Dunque la (2.96) vale per tutti i termini della serie di potenze che definisce la funzione V . Ma questo implica che

$$[p_x, V(x)] = \frac{\hbar}{i} \frac{dV(x)}{dx}. \quad (2.98)$$

Applicando lo stesso ragionamento a funzioni che dipendono anche da y e z si vede facilmente che il commutatore non deriva in quelle direzioni (il commutatore di p_x con y è nullo, e lo stesso con z). Dunque si ha

$$[p_x, V(\mathbf{r})] = \frac{\hbar}{i} \frac{\partial V(\mathbf{r})}{\partial x}, \quad (2.99)$$

che inserito nell'equazione (2.95) dà

$$\frac{dp_x}{dt} = - \frac{\partial V(\mathbf{r})}{\partial x}. \quad (2.100)$$

In termini di valori medi degli operatori questo implica

$$\frac{d\langle p_x \rangle}{dt} = \langle - \frac{\partial V(\mathbf{r})}{\partial x} \rangle. \quad (2.101)$$

Questa ha la forma della seconda legge di Newton, dove il membro di destra gioca il ruolo di una forza che agisce sulla particella. Ma dobbiamo sempre fare attenzione al fatto che si tratta di un'equazione per i valori medi di misure effettuate su repliche del sistema, non su misure di momento eseguite sulla stessa particella in istanti diversi.

Possiamo anche riscriverla usando solo l'osservabile posizione. Infatti, per quanto riguarda la derivata prima dell'operatore x nulla cambia rispetto al caso della particella libera, dato che $[x, V(\mathbf{r})] = 0$ e dunque vale ancora la relazione

$$\frac{dx}{dt} = \frac{p_x}{m}.$$

Per l'operatore dx/dt possiamo usare di nuovo l'equazione del moto di Heisenberg

$$\frac{d^2x}{dt^2} = \frac{d}{dt} \left(\frac{dx}{dt} \right) = \frac{1}{i\hbar} \left[\frac{dx}{dt}, H \right] = \frac{1}{i\hbar} \left[\frac{p_x}{m}, H \right] = \frac{1}{i\hbar m} [p_x, V(\mathbf{r})], \quad (2.102)$$

da cui

$$\frac{d^2x}{dt^2} = - \frac{1}{m} \frac{\partial V(\mathbf{r})}{\partial x}, \quad (2.103)$$

che per i valori medi diventa

$$\frac{d^2\langle x \rangle}{dt^2} = -\frac{1}{m} \left\langle \frac{\partial V(\mathbf{r})}{\partial x} \right\rangle. \quad (2.104)$$

Ripetendo il calcolo anche per le derivate nelle altre direzioni otteniamo il risultato finale

$$m \frac{d^2\langle \mathbf{r} \rangle}{dt^2} = \langle -\nabla V(\mathbf{r}) \rangle. \quad (2.105)$$

Abbiamo quindi ritrovato l'analogo quantistico della seconda legge di Newton, stavolta espresso in termini dell'operatore posizione. Questo risultato è noto come *teorema di Ehrenfest*. È importante sottolineare che si tratta di un risultato indipendente dalla rappresentazione, dato che i valori medi sono gli stessi se calcolati nella rappresentazione di Heisenberg o di Schrödinger, e non dipende nemmeno dalla base degli autostati utilizzata per esprimere i vettori e gli operatori.

Ripetiamo il senso fisico di quanto abbiamo trovato: misurando la posizione al tempo t di particelle tutte preparate in uno stesso stato puro al tempo $t = 0$ e soggette ad un potenziale statico $V(\mathbf{r})$, si otterranno valori della posizione distribuiti statisticamente attorno ad un valor medio $\langle \mathbf{r} \rangle$ che obbedisce ad un'equazione tipo Newton. Se la distribuzione dei valori è larga sulla scala delle lunghezze tipiche del problema, allora questa relazione per i valori medi, pur rimanendo valida, ha poca utilità pratica. Se, al contrario, la distribuzione è stretta, cioè lo stato della particella è localizzato entro regioni molto piccole rispetto alle lunghezze tipiche (la risoluzione ottica degli strumenti, le distanze percorse, le distanze tra le particelle di un fascio, la larghezza della buca di potenziale in cui la particella si trova, ecc.) allora i valori medi sono una buona rappresentazione dei valori misurati e l'equazione (2.105) va a coincidere con l'equazione del moto di Newton per la posizione di una particella in senso classico. In questo limite vediamo anche che la costante di Planck sparisce dalla dinamica. Sappiamo quindi come arrivare alla fisica classica partendo dalla meccanica quantistica, in modo rigoroso, almeno per il moto di una particella in un potenziale $V(\mathbf{r})$.

2.6 L'oscillatore armonico

2.6.1 Lo spettro dell'oscillatore

Un esempio importante di potenziale scalare è il potenziale armonico unidimensionale $V(x) = (1/2)m\omega^2 x^2$. È importante per almeno due motivi. Il primo è che si tratta di un potenziale semplice da trattare e che si presta ad essere usato come archetipo di problema quantistico con varie alternative eleganti per la sua soluzione. Il secondo è che gli oscillatori armonici, come gli studenti sanno già dalla fisica classica, sono ovunque, dalla meccanica dei mezzi elastici macroscopici, fino alle vibrazioni degli atomi nelle molecole, e in ogni sistema dove la dinamica si riduce a piccole oscillazioni attorno ad un minimo di potenziale.

Nel seguito deriviamo le proprietà dell'oscillatore armonico con un approccio puramente algebrico, che fa uso delle proprietà degli operatori posizione e momento. Lo stesso problema potrebbe essere affrontato risolvendo l'equazione di Schrödinger stazionaria per le funzioni d'onda, autofunzioni dell'hamiltoniano, ma su questo daremo solo qualche breve accenno.

Per sfruttare le proprietà algebriche degli operatori nel calcolo dello spettro di H , conviene definire due nuovi operatori a e $a^\dagger$ a partire dagli operatori posizione, x , e momento, p , in questo modo:

$$a = \sqrt{\frac{m\omega}{2\hbar}} \left(x + \frac{ip}{m\omega} \right), \quad a^\dagger = \sqrt{\frac{m\omega}{2\hbar}} \left(x - \frac{ip}{m\omega} \right). \quad (2.106)$$

Questi operatori sono l'uno l'hermitiano coniugato dell'altro e godono della seguente regola di commutazione:

$$[a, a^\dagger] = \frac{i}{2\hbar}([p, x] - [x, p]) = 1. \quad (2.107)$$

Poi definiamo l'operatore

$$N = a^\dagger a, \quad (2.108)$$

che è hermitiano. Possiamo scrivere

$$N = a^\dagger a = \frac{m\omega}{2\hbar} \left(x^2 + \frac{p^2}{m^2\omega^2} \right) + \frac{i}{2\hbar} [x, p], \quad (2.109)$$

D'altra parte l'hamiltoniano è

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \quad (2.110)$$

e dunque

$$N = \frac{H}{\hbar\omega} - \frac{1}{2}, \quad (2.111)$$

ovvero

$$H = \hbar\omega \left(N + \frac{1}{2} \right). \quad (2.112)$$

Ne segue che $[N, H] = 0$ e gli autovettori di H sono anche autovettori di N . Indichiamo con n il generico autovalore di N e con $|n\rangle$ il corrispondente autovettore normalizzato:

$$N|n\rangle = n|n\rangle, \quad (2.113)$$

da cui

$$H|n\rangle = \hbar\omega \left(n + \frac{1}{2} \right) |n\rangle \quad (2.114)$$

e gli autovalori dell'energia sono

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right). \quad (2.115)$$

Il problema dello spettro di H si riduce quindi alla determinazione dei numeri n . In effetti, la scelta dei coefficienti delle combinazioni lineari di x e p che definiscono a e $a^\dagger$ in (2.106) è stata fatta proprio per garantire la regola di commutazione (2.107) e per ottenere una relazione semplice tra H e N . Ora però resta aperto il problema di stabilire quali sono i valori ammessi per l'autovalore n .

A questo scopo vediamo qual è l'azione degli operatori a e $a^\dagger$ sugli autovettori di N . Per prima cosa notiamo che

$$[N, a] = [a^\dagger a, a] = a^\dagger [a, a] + [a^\dagger, a]a = -a \quad (2.116)$$

e analogamente

$$[N, a^\dagger] = [a^\dagger a, a^\dagger] = a^\dagger [a, a^\dagger] + [a^\dagger, a^\dagger]a = a^\dagger. \quad (2.117)$$

Usiamo questi risultati per scrivere

$$Na|n\rangle = (aN + [N, a])|n\rangle = a(N - 1)|n\rangle = (n - 1)a|n\rangle \quad (2.118)$$

e

$$Na^\dagger|n\rangle = (a^\dagger N + [N, a^\dagger])|n\rangle = a^\dagger(N + 1)|n\rangle = (n + 1)a^\dagger|n\rangle. \quad (2.119)$$

Questo ci dice che $a|n\rangle$ è un autovettore di N con autovalore $n-1$, mentre $a^\dagger|n\rangle$ è un autovettore di N con autovalore $n+1$. Nessuno ci assicura però che siano normalizzati a 1, come invece lo sono $|n-1\rangle$ e $|n+1\rangle$, da cui possono differire per un semplice fattore numerico moltiplicativo. Se

$$a|n\rangle = c|n-1\rangle, \quad (2.120)$$

allora si ha

$$\langle n|a^\dagger a|n\rangle = (\langle n|a^\dagger)(a|n\rangle) = |c|^2 \langle n-1|n-1\rangle = |c|^2. \quad (2.121)$$

D'altra parte, lo stesso valore d'aspettazione lo potevamo calcolare in questo modo

$$\langle n|a^\dagger a|n\rangle = (\langle n|)(a^\dagger a|n\rangle) = (\langle n|)(N|n\rangle) = n\langle n|n\rangle = n \quad (2.122)$$

e dunque deve valere

$$|c|^2 = n. \quad (2.123)$$

A meno di un fattore di fase irrilevante, possiamo scegliere c reale ottenendo $c = \sqrt{n}$:

$$a|n\rangle = \sqrt{n} |n-1\rangle. \quad (2.124)$$

Procedendo allo stesso modo per $a^\dagger|n\rangle$ si trova

$$a^\dagger|n\rangle = \sqrt{n+1} |n+1\rangle. \quad (2.125)$$

Questi operatori modificano lo stato della particella togliendo o aggiungendo un *quanto* di energia $\hbar\omega$. Per questo si chiamano operatori di *distruzione* e di *creazione*. Vengono detti anche operatori *a scala*, perchè permettono di salire o scendere la scaletta degli autovalori¹⁵. Qui li stiamo usando come operatori ausiliari per calcolarci lo spettro, ma si può intuire che essi potranno avere un significato fisico rilevante nel descrivere le transizioni di una particella tra i suoi stati stazionari.

Ora supponiamo che la particella si trovi in uno stato generico $|n\rangle$ e applichiamo ripetutamente l'operatore di distruzione a . Ogni volta, si otterrà lo stato con autovalore diminuito di una unità. Se n è un numero reale qualsiasi, non intero, allora la sequenza prima o poi genera autovalori negativi. Ma il postulato della metrica definita positiva impone che n sia non negativo. Infatti, possiamo scrivere

$$n = \langle n|N|n\rangle = \langle n|a^\dagger a|n\rangle = (\langle n|a^\dagger)(a|n\rangle), \quad (2.126)$$

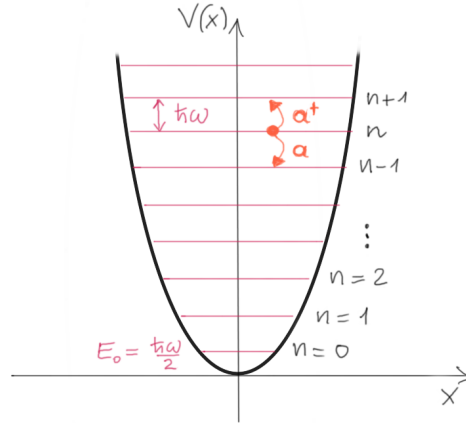
ma l'ultimo termine è la norma quadra del vettore $a|n\rangle$ e deve essere maggiore o uguale di zero. Dunque

$$n \geq 0 \quad (2.127)$$

e la sequenza deve interrompersi prima di ottenere n negativi. Questo è possibile solo se tutti gli n sono interi. Infatti, partendo da un n intero, dopo n applicazioni di a si arriva allo stato $|0\rangle$, e le successive applicazioni producono un vettore nullo. Concludiamo quindi che lo spettro degli stati stazionari, autostati di H , è quello scritto in (2.115) con n interi non negativi. Lo riscriviamo qui per completezza e lo incorniciamo per la sua rilevanza in molti problemi fisici:

$$\boxed{E_n = \hbar\omega \left(n + \frac{1}{2} \right) \quad \text{con } n = 0, 1, 2, \dots} \quad (2.128)$$

¹⁵La forma degli operatori a e $a^\dagger$ nelle espressioni (2.106) è stata scelta proprio in modo tale da ottenere operatori che trasformano ogni autovettore di H in un altro autovettore di H , a meno della normalizzazione; la norma non è conservata dai due operatori, perché non sono unitari ($a^\dagger a \neq 1$). La scelta di scrivere a e $a^\dagger$ come combinazioni lineari di x e p è naturale, dato che queste sono le uniche osservabili che entrano nell'hamiltoniano.



Lo stato con $n = 0$ è quello di minima energia

$$E_0 = \frac{\hbar\omega}{2} \quad (2.129)$$

e può essere definito come lo stato per cui $a|0\rangle = 0$. Ricordando la definizione di a , possiamo scrivere

$$a|0\rangle = \sqrt{\frac{m\omega}{2\hbar}} \left(x + \frac{ip}{m\omega} \right) |0\rangle = 0, \quad (2.130)$$

nella rappresentazione degli autostati di posizione possiamo scrivere

$$\langle x|a|0\rangle = \sqrt{\frac{m\omega}{2\hbar}} \langle x| \left(x + \frac{ip}{m\omega} \right) |0\rangle = 0, \quad (2.131)$$

ovvero, ricordando la (1.245),

$$\sqrt{\frac{m\omega}{2\hbar}} \left(x + \frac{\hbar}{m\omega} \frac{d}{dx} \right) \psi_0(x) = 0, \quad (2.132)$$

che si può scrivere nella forma

$$\left(x + a_{\text{ho}}^2 \frac{d}{dx} \right) \psi_0(x) = 0, \quad (2.133)$$

dove

$$a_{\text{ho}} = \sqrt{\frac{\hbar}{m\omega}} \quad (2.134)$$

è una lunghezza tipica dell'oscillatore ("ho" sta per *harmonic oscillator*). La soluzione dell'equazione differenziale è la gaussiana

$$\psi_0(x) = \frac{1}{\pi^{1/4} \sqrt{a_{\text{ho}}}} \exp\left(-\frac{x^2}{2a_{\text{ho}}^2}\right), \quad (2.135)$$

dove il fattore moltiplicativo è stato scelto in modo da garantire la normalizzazione $\int dx |\psi_0(x)|^2 = 1$. La particella è spazialmente delocalizzata attorno al minimo del potenziale armonico. Misure di posizione della particella ottenute in repliche dello stesso esperimento sullo stesso

sistema producono valori di x statisticamente distribuiti in modo gaussiano con media nulla, $\langle x \rangle = 0$, e con scarto quadratico medio

$$\Delta x = \sqrt{\langle x^2 \rangle} = \left(\int dx x^2 |\psi_0(x)|^2 \right)^{1/2} = \frac{a_{\text{ho}}}{\sqrt{2}}. \quad (2.136)$$

La funzione d'onda nello spazio dei momenti è la trasformata di Fourier della gaussiana, che è ancora una gaussiana

$$\tilde{\psi}_0(p) = \frac{1}{\pi^{1/4}} \sqrt{\frac{a_{\text{ho}}}{\hbar}} \exp\left(-\frac{a_{\text{ho}}^2 p^2}{2\hbar^2}\right). \quad (2.137)$$

con valor medio nullo, $\langle p \rangle = 0$, e con scarto quadratico medio

$$\Delta p = \sqrt{\langle p^2 \rangle} = \frac{\hbar}{\sqrt{2}a_{\text{ho}}}. \quad (2.138)$$

Il prodotto delle indeterminazioni delle misure è

$$\Delta x \Delta p = \frac{\hbar}{2} \quad (2.139)$$

e corrisponde alla minima indeterminazione ammessa dalla relazione di indeterminazione di Heisenberg, come ci si aspetta per un pacchetto gaussiano. La particella non può localizzarsi più di quanto concesso dalla relazione di indeterminazione. Ciò si riflette nel valore minimo di energia $\hbar\omega/2$, che è detta *energia di punto zero* ed è l'energia dello *stato fondamentale* dell'oscillatore.

Gli autostati con $n > 0$ si ottengono per applicazioni successive dell'operatore di creazione $a^\dagger$. Per il primo, ad esempio, possiamo scrivere $|1\rangle = a^\dagger|0\rangle$, ovvero

$$\psi_1(x) = \langle x|a^\dagger|0\rangle = \sqrt{\frac{m\omega}{2\hbar}} \langle x| \left(x - \frac{ip}{m\omega} \right) |0\rangle, \quad (2.140)$$

da cui

$$\psi_1(x) = \frac{1}{\sqrt{2}a_{\text{ho}}} \left(x - a_{\text{ho}}^2 \frac{d}{dx} \right) \psi_0(x), \quad (2.141)$$

che dà una funzione del tipo $\psi_1(x) \propto x \exp[-x^2/(2a_{\text{ho}}^2)]$. L'applicazione ulteriore di $a^\dagger$ darà ancora la gaussiana, ma moltiplicata per un polinomio di secondo grado, e così via. I polinomi che si ottengono in questo modo, per applicazioni successive di $a^\dagger$, sono detti *polinomi di Hermite*. In questo modo possiamo risolvere completamente il problema della determinazione degli stati stazionari dell'oscillatore armonico unidimensionale.

Come detto all'inizio di questa sezione, lo stesso spettro può essere calcolato a partire dall'equazione di Schrödinger stazionaria (2.43) che per l'oscillatore unidimensionale è

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m \omega^2 x^2 \right) \psi(x) = E \psi(x). \quad (2.142)$$

Possiamo riscalarle le distanze con la lunghezza tipica a_{ho} , introducendo la nuova variabile adimensionale $\tilde{x} = x/a_{\text{ho}}$ e possiamo anche riscalarle l'energia introducendo la grandezza $\epsilon = 2E/(\hbar\omega)$. Allora la stessa equazione diventa

$$\frac{d^2}{d\tilde{x}^2} \psi(\tilde{x}) + (\epsilon - \tilde{x}^2) \psi(\tilde{x}) = 0. \quad (2.143)$$

Le soluzioni devono essere normalizzabili. Bisogna quindi eliminare le soluzioni che divergono all'infinito. Un'equazione del tipo $\psi''(\tilde{x}) - \tilde{x}^2\psi(\tilde{x}) = 0$ ha soluzioni del tipo $\psi(\tilde{x}) \propto \exp(\pm\tilde{x}^2/2)$. Quindi le soluzioni buone avranno la forma

$$\psi(\tilde{x}) = \phi(\tilde{x}) \exp(-\tilde{x}^2/2). \quad (2.144)$$

con la nuova funzione $\phi(\tilde{x})$ da determinare risolvendo l'equazione

$$\frac{d^2\phi}{d\tilde{x}^2} - 2\tilde{x}\frac{d\phi}{d\tilde{x}} + (\epsilon - 1)\phi = 0. \quad (2.145)$$

A questo punto, se vogliamo ottenere risultati analitici, dobbiamo ricorrere alla teoria delle equazioni differenziali, o a un buon manuale (gli studenti trovano la soluzione completa anche nel Sakurai). Si trovano soluzioni solo per $\epsilon - 1 = 2n$ con $n = 0, 1, 2, \dots$, che corrisponde allo spettro $E_n = \hbar\omega(n + 1/2)$ già trovato in precedenza. Per ogni n la soluzione è

$$\psi_n(x) = c_n H_n(\tilde{x}) \exp(-\tilde{x}^2/2), \quad (2.146)$$

dove c_n è una costante numerica che può essere fissata normalizzando la funzione a 1 e le funzioni $H_n(x)$ sono i *polinomi di Hermite*. La scelta di affrontare il problema in questo modo, utilizzando le equazioni differenziali, oppure con l'algebra degli operatori di creazione e distruzione è una questione di gusti. Tutti i risultati, in termini di stati fisici e osservabili, sono gli stessi.

2.6.2 Evoluzione temporale e stati coerenti

Ora che abbiamo sistemato il problema degli stati stazionari, rimettiamoci nella rappresentazione di Heisenberg e studiamo l'evoluzione temporale degli operatori. Le equazioni del moto per x e p sono

$$\frac{dx}{dt} = \frac{1}{i\hbar}[x, H] = \frac{1}{2mi\hbar}[x, p^2] = \frac{p}{m} \quad (2.147)$$

$$\frac{dp}{dt} = \frac{1}{i\hbar}[p, H] = \frac{m\omega^2}{2i\hbar}[p, x^2] = -m\omega^2 x. \quad (2.148)$$

Sono equazioni accoppiate, ma possiamo disaccoppiarle derivando la prima di nuovo rispetto al tempo. Otteniamo

$$m \frac{d^2x}{dt^2} = -m\omega^2 x, \quad (2.149)$$

che ha forma identica alla seconda legge di Newton, $m\ddot{x} = F$, per una particella classica soggetta ad una forza di richiamo elastica. La soluzione è

$$x(t) = x(0) \cos \omega t + \frac{p(0)}{m\omega} \sin \omega t, \quad (2.150)$$

dove $x(0)$ e $p(0)$ sono gli operatori posizione e momento nella rappresentazione di Heisenberg, definiti al tempo $t = 0$. Dato $x(t)$ si può anche calcolare $p(t)$ usando la (2.147):

$$p(t) = -m\omega x(0) \sin \omega t + p(0) \cos \omega t. \quad (2.151)$$

Per esercizio, verifichiamo che le stesse soluzioni possono essere ricavate dalle equazioni del moto per gli operatori a e $a^\dagger$:

$$\frac{da}{dt} = \frac{1}{i\hbar}[a, H] = \frac{\omega}{i}[a, N] = -i\omega a \quad (2.152)$$

$$\frac{da^\dagger}{dt} = \frac{1}{i\hbar}[a^\dagger, H] = \frac{\omega}{i}[a^\dagger, N] = i\omega a^\dagger. \quad (2.153)$$

Notiamo che, mentre le equazioni per x e p sono accoppiate, quelle per a e $a^\dagger$ non lo sono. Le loro soluzioni sono

$$a(t) = a(0) \exp(-i\omega t) \quad , \quad a^\dagger(t) = a^\dagger(0) \exp(i\omega t). \quad (2.154)$$

Una conseguenza è che l'operatore $N = a^\dagger a$ è indipendente dal tempo, perché i fattori di fase si cancellano nel prodotto, consistentemente con l'indipendenza dal tempo di H . Poi, possiamo riscrivere le stesse soluzioni ricordando le definizioni

$$a = \frac{1}{a_{\text{ho}}\sqrt{2}} \left(x + \frac{ip}{m\omega} \right) \quad , \quad a^\dagger = \frac{1}{a_{\text{ho}}\sqrt{2}} \left(x - \frac{ip}{m\omega} \right) \quad (2.155)$$

e le loro inverse

$$x = \frac{a_{\text{ho}}}{\sqrt{2}} (a + a^\dagger) \quad , \quad p = \frac{\hbar}{ia_{\text{ho}}\sqrt{2}} (a - a^\dagger). \quad (2.156)$$

Da queste segue che

$$\begin{aligned} x(t) &= \frac{a_{\text{ho}}}{\sqrt{2}} (a(t) + a^\dagger(t)) \\ &= \frac{a_{\text{ho}}}{\sqrt{2}} (a(0) \exp(-i\omega t) + a^\dagger(0) \exp(i\omega t)) \\ &= \frac{1}{2} \left[\left(x(0) + \frac{ip(0)}{m\omega} \right) \exp(-i\omega t) + \left(x(0) - \frac{ip(0)}{m\omega} \right) \exp(i\omega t) \right] \\ &= x(0) \cos \omega t + \frac{p(0)}{m\omega} \sin \omega t, \end{aligned} \quad (2.157)$$

che è lo stesso risultato di prima, e lo stesso si troverebbe per $p(t)$.

Le soluzioni hanno la stessa forma di quelle per la coordinata e il momento di un oscillatore classico, ma dobbiamo sempre ricordarci che qui abbiamo a che fare con operatori e nessuno ci autorizza a concludere che la posizione di una particella oscilla nel tempo: l'operatore posizione è una cosa, l'esito della misura della posizione è un'altra. Ad esempio, se calcoliamo il valor medio della posizione quando la particella è in uno stato stazionario, $\langle n|x|n\rangle$, vediamo subito che, mentre l'operatore x varia periodicamente come scritto qui sopra, il valor medio rimane costantemente nullo. Per vederlo basta usare l'espressione di x in termini di $(a + a^\dagger)$ e notare che, quando applichiamo x a $|n\rangle$ otteniamo una somma di due termini con $|n-1\rangle$ e $|n+1\rangle$, che sono entrambi ortogonali a $|n\rangle$. Peraltro, sapevamo già dalla sezione precedente che il valore di aspettazione di un'osservabile per un sistema in uno stato stazionario è costante. Dunque, il fatto che gli operatori x e p obbediscano a leggi armoniche non implica che il moto della singola particella sia armonico, in senso classico.

Qualcosa che assomigli di più ad un oscillatore classico, almeno per il comportamento dei valori medi, lo si trova preparando la particella in uno stato $|\alpha\rangle$ che sia una sovrapposizione lineare di stati stazionari di diversa energia. Ad esempio prendiamo lo stato normalizzato

$$|\alpha\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad (2.158)$$

allora

$$\begin{aligned}
\langle x(t) \rangle &= \frac{1}{2} (\langle 0| + \langle 1|) x(t) (|0\rangle + |1\rangle) \\
&= \frac{a_{ho}}{2\sqrt{2}} (\langle 0| + \langle 1|) (a(t) + a^\dagger(t)) (|0\rangle + |1\rangle) \\
&= \frac{a_{ho}}{2\sqrt{2}} (\langle 0|a(t)|1\rangle + \langle 1|a^\dagger(t)|0\rangle) \\
&= \frac{a_{ho}}{2\sqrt{2}} (\langle 0|a(0)|1\rangle \exp(-i\omega t) + \langle 1|a^\dagger(0)|0\rangle \exp(i\omega t)) \\
&= (a_{ho}/\sqrt{2}) \cos \omega t
\end{aligned} \tag{2.159}$$

e il valor medio oscilla con periodo $2\pi/\omega$ con ampiezza $a_{ho}/\sqrt{2}$. Ma siamo ancora ben lontani dal caso classico: la funzione d'onda rimane delocalizzata su tutto lo spazio disponibile alla particella e la forma della funzione d'onda non assomiglia affatto a quella di una particella che va avanti e indietro nel potenziale armonico. Per ottenere questo, occorre combinare autostati con n grande, entro un intervallo sufficientemente ristretto di energia, in modo che la funzione d'onda abbia la forma di un pacchetto d'onde localizzato su una scala molto minore dell'ampiezza di oscillazione. In altri termini, se lo studio del moto di un corpo soggetto ad una forza elastica implica scale di energia tali che $E \gg \hbar\omega$ e l'incertezza ammissibile per l'energia ΔE sia anch'essa tale che $\Delta E \gg \hbar\omega$, allora lo spettro dell'oscillatore sarà trattabile di fatto come un continuo, la sua discretizzazione essendo del tutto irrilevante; inoltre, al corpo che si muove sarà associabile uno stato rappresentato da una combinazione lineare di molte autofunzioni di H con energia prossima a E , ciascuna diversa dall'altra per il numero di nodi, analogamente a quanto si fa nei pacchetti d'onda costruiti con onde piane. Questo è il limite classico dell'oscillatore, in cui tutto l'apparato formale della meccanica quantistica, pur rimanendo valido e generale, diventa ridondante e può essere abbandonato a favore delle leggi di Newton.

Un problema con i pacchetti d'onda è che, se non si scelgono in modo appropriato le fasi delle onde, il pacchetto si disperde, aumentando la sua larghezza nel tempo. È tuttavia possibile realizzare combinazioni lineari di autostati dell'oscillatore corrispondenti a pacchetti gaussiani che oscillano come una particella classica senza dispersione. In questo caso si parla di *stati coerenti*. Si dimostra (ma qui non lo facciamo) che un autostato dell'operatore di distruzione

$$a|\lambda\rangle = \lambda|\lambda\rangle, \tag{2.160}$$

dove λ è in generale un numero complesso, è uno *stato coerente*. Il modo più semplice per ottenerlo, nella rappresentazione delle coordinate, consiste nel traslare lo stato fondamentale dell'oscillatore armonico (con funzione d'onda gaussiana di minima indeterminazione) di una distanza finita dall'origine al tempo $t = 0$, lasciando poi evolvere il sistema. Se la distanza iniziale del pacchetto dall'origine è molto maggiore di a_{ho} , allora l'energia media del pacchetto sarà molto maggiore di $\hbar\omega$ e il pacchetto si comporterà come una particella classica che oscilla nello stesso potenziale armonico. Una trattazione chiara e semplice di questi stati coerenti si trova nel testo di Forte e Rottoli, al capitolo 8.5, a cui rimando per approfondimenti.

2.7 Equazione di continuità

Per concludere questo capitolo sulla dinamica, torniamo al formalismo generale nella rappresentazione di Schrödinger e, specificamente, all'equazione di Schrödinger (2.23). La

riscriviamo in questo modo

$$i\hbar \frac{\partial}{\partial t} \psi = \left(-\frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r}) \right) \psi$$

e ne prendiamo anche la complessa coniugata

$$-i\hbar \frac{\partial}{\partial t} \psi^* = \left(-\frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r}) \right) \psi^* .$$

Moltiplichiamo a sinistra entrambi i membri della prima per ψ^* e della seconda per ψ e poi sottraiamo l'una dall'altra. Otteniamo

$$i\hbar \left(\psi^* \frac{\partial}{\partial t} \psi + \psi \frac{\partial}{\partial t} \psi^* \right) = -\frac{\hbar^2}{2m} (\psi^* \nabla^2 \psi - \psi \nabla^2 \psi^*) . \quad (2.161)$$

Il membro di sinistra lo possiamo scrivere così

$$i\hbar \left(\psi^* \frac{\partial}{\partial t} \psi + \psi \frac{\partial}{\partial t} \psi^* \right) = i\hbar \frac{\partial}{\partial t} |\psi|^2 = i\hbar \frac{\partial n}{\partial t} , \quad (2.162)$$

dove $n(\mathbf{r}, t)$ è la densità di probabilità di trovare la particella in $\mathbf{r}$ al tempo t . Poi definiamo la quantità

$$\mathbf{j} = \frac{\hbar}{2mi} (\psi^* \nabla \psi - \psi \nabla \psi^*) = \frac{\hbar}{m} \text{Im}(\psi^* \nabla \psi) \quad (2.163)$$

in modo che

$$\nabla \cdot \mathbf{j} = \frac{\hbar}{2mi} (\psi^* \nabla^2 \psi - \psi \nabla^2 \psi^*) . \quad (2.164)$$

Con queste definizioni di n e $\mathbf{j}$ l'equazione (2.161) diventa

$$\frac{\partial}{\partial t} n(\mathbf{r}, t) + \nabla \cdot \mathbf{j}(\mathbf{r}, t) = 0 . \quad (2.165)$$

Questa ha la forma di un'equazione di continuità, in cui n è la densità e $\mathbf{j}$ è la densità di corrente.

Un'equazione della stessa forma sta alla base di qualsiasi teoria idrodinamica che descriva un gas, un fluido, un plasma, ecc., per cui valga una legge di conservazione per il numero totale di particelle, cioè un fluido che si muove in una regione priva di pozzi e sorgenti. Se le particelle non si distruggono e non si creano, allora ogni variazione locale della densità in un punto $\mathbf{r}$ deve essere accompagnata da un flusso di particelle uscente dal volumetto d^3r intorno ad $\mathbf{r}$, o entrante in esso, e l'equazione di continuità è l'espressione formale di questo fatto. Qui però la particella è una sola e ciò che varia è la densità di probabilità definita tramite la funzione d'onda. Di conseguenza, la $\mathbf{j}$ va intesa come *densità di corrente di probabilità*, oppure *flusso di probabilità*, e l'equazione di continuità esprime la conservazione della norma della funzione d'onda che, a sua volta, rende conto del fatto che la particella non si crea e non si distrugge, ma semplicemente si muove nello spazio. L'equazione di Schrödinger garantisce la conservazione della norma in quanto rappresenta, in termini di funzioni d'onda, l'azione dell'operatore di traslazione temporale U , che è un operatore unitario. Così come n , anche $\mathbf{j}$ assume significato solo in senso statistico, riferita a sequenze di misure di posizione eseguite su copie dello stesso sistema preparato nelle stesse identiche condizioni. Notiamo anche che la densità di corrente è identicamente nulla quando la funzione d'onda è reale oppure quando è il prodotto di un numero complesso per una funzione reale.

Il significato di n e $\mathbf{j}$ può essere apprezzato scrivendo la generica funzione d'onda $\psi(\mathbf{r}, t)$ nella forma

$$\psi(\mathbf{r}, t) = \sqrt{n(\mathbf{r}, t)} \exp \left(\frac{iS(\mathbf{r}, t)}{\hbar} \right) \quad (2.166)$$

con S reale e $n \geq 0$. Ogni funzione complessa può essere scritta in questo modo. Nel caso specifico delle funzioni d'onda, la funzione $n(\mathbf{r}, t)$, che ne fissa il modulo quadro, è proprio la densità di probabilità, per costruzione. Per capire il significato della fase $S(\mathbf{r}, t)$ basta ricorrere alla definizione di $\mathbf{j}$. Infatti

$$\nabla\psi = [\nabla\sqrt{n} + (i/\hbar)\sqrt{n}\nabla S] \exp\left(\frac{iS}{\hbar}\right) \quad (2.167)$$

da cui

$$\psi^*\nabla\psi = \sqrt{n}\nabla\sqrt{n} + (i/\hbar)n\nabla S \quad (2.168)$$

e infine

$$\mathbf{j} = \frac{\hbar}{m} \text{Im}(\psi^*\nabla\psi) = \frac{n\nabla S}{m}. \quad (2.169)$$

Qui si vede che un gradiente della fase S determina un flusso di probabilità. Tanto più è grande il gradiente di S tanto più lo è il flusso $\mathbf{j}$, che è diretto in direzione ortogonale alla superficie di fase costante. In un fluido classico normalmente si scrive la densità di corrente $\mathbf{j}$ come il prodotto della densità per la velocità locale del fluido stesso: $\mathbf{j} = n\mathbf{v}$. Questo ci suggerisce di identificare la quantità $(1/m)\nabla S$ con una qualche velocità $\mathbf{v}$. Così facendo l'equazione di continuità scritta sopra diventerebbe

$$\frac{\partial}{\partial t}n + \nabla \cdot (n\mathbf{v}) = 0 \quad (2.170)$$

che è la forma più conosciuta dell'equazione di continuità per fluidi classici. Inoltre, la stessa definizione si applica in modo semplice alle funzioni d'onda che rappresentano le autofunzioni dell'operatore momento, che sono onde piane del tipo

$$\psi(\mathbf{r}, t) \propto \exp\left[\frac{i}{\hbar}(\mathbf{p} \cdot \mathbf{r} - Et)\right] \quad (2.171)$$

e quindi

$$\nabla S = \nabla(\mathbf{p} \cdot \mathbf{r} - Et) = \mathbf{p} \quad (2.172)$$

che è consistente con la definizione di $\mathbf{v}$ appena data.

Non è la prima volta che tentiamo di definire una velocità in meccanica quantistica e tutte le volte insistiamo sulla necessità di evitare fraintendimenti. La velocità di cui parliamo qui ha un significato intrinsecamente statistico e non può essere identificata con la definizione classica di $\mathbf{v}$ in termini di rapporto incrementale della posizione di una stessa particella, misurata in tempi successivi. Tuttavia, fatto questo richiamo alla cautela, non è affatto detto che la velocità definita come $(1/m)\nabla S$ sia inutile, anzi. Se in laboratorio produciamo un fascio di particelle collimate e preparato in modo che la velocità (intesa in senso classico) sia la stessa a meno di piccole fluttuazioni attorno alla media, allora dal punto di vista quantistico è come aver prodotto molte repliche di uno stesso stato di singola particella, descrivibile in buona approssimazione in termini di autostati del momento. In tal caso la densità di corrente di probabilità è effettivamente interpretabile come un flusso di particelle di velocità nota entro certi margini di errore. Per questo motivo la densità di corrente di probabilità è particolarmente utile nel risolvere problemi che coinvolgono gli stati liberi di particelle soggette localmente a barriere o buche di potenziale. Gli studenti avranno probabilmente già visto come si utilizza nei problemi unidimensionali con potenziali a gradino o a barriera quadra, ad esempio, dove la densità di corrente serve per calcolare i coefficienti di riflessione e trasmissione. Ma la velocità $(1/m)\nabla S$ è utile anche quando un grande numero di particelle identiche e non interagenti tra loro si trovano in uno stato quantistico descritto da un'unica funzione d'onda. Realizzare uno stato del genere non è

banale e richiede condizioni molto particolari, ma non è impossibile. In uno stato del genere, la replica di molte misure di posizione su stati identici di singola particella preparati allo stesso modo è equivalente ad una singola misura della densità del sistema e, analogamente, il gradiente della fase S fissa il campo di velocità locale delle particelle del sistema. Dato che finora abbiamo trattato esplicitamente solo sistemi composti da una singola particella, lasciamo in sospenso le implicazioni di questo argomento, che potranno essere apprezzate pienamente solo dopo aver studiato la meccanica statistica quantistica.

Capitolo 3

Momento angolare

3.1 Definizione, commutatori e spettro

Come in meccanica classica anche in meccanica quantistica si può introdurre il momento angolare e associarlo alle rotazioni nello spazio delle coordinate. In questo capitolo ne diamo la definizione in termini di operatori posizione e momento partendo dalla definizione classica. Poi mostreremo quale sia il suo legame con le rotazioni. Qui ci discosteremo parecchio dalla trattazione del Sakurai, che è più completa e si basa sul formalismo della teoria dei gruppi. Gli studenti interessati sono incoraggiati a leggerlo.

Dunque partiamo dalla definizione classica

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}, \quad (3.1)$$

dove le tre grandezze sono vettori nello spazio euclideo e il prodotto è un prodotto vettoriale. Teniamo la stessa definizione, ma sostituiamo tutte le grandezze classiche con operatori. Per componenti abbiamo

$$\begin{aligned} L_x &= yp_z - zp_y \\ L_y &= zp_x - xp_z \\ L_z &= xp_y - yp_x, \end{aligned} \quad (3.2)$$

dove questi sono tutti operatori. Calcoliamo l'hermitiano coniugato di L_x

$$L_x^\dagger = (yp_z - zp_y)^\dagger = (yp_z)^\dagger - (zp_y)^\dagger = p_z^\dagger y^\dagger - p_y^\dagger z^\dagger, \quad (3.3)$$

e, usando l'hermitianità degli operatori posizione e momento e le loro regole di commutazione canonica, scriviamo

$$L_x^\dagger = p_z y - p_y z = yp_z - zp_y = L_x, \quad (3.4)$$

Dunque L_x è un operatore hermitiano e altrettanto vale, con dimostrazione simile, per L_y e L_z . Quindi $\mathbf{L}$ può essere definito come l'operatore le cui tre componenti sono tre operatori hermitiani che rappresentano le componenti dell'osservabile momento angolare, analogo al momento angolare classico.

Ora vediamo quali sono le loro regole di commutazione. Il calcolo è semplice. Cominciamo da

$$\begin{aligned} [L_x, L_y] &= [yp_z - zp_y, zp_x - xp_z] \\ &= [yp_z, zp_x] - [yp_z, xp_z] - [zp_y, zp_x] + [zp_y, xp_z] \\ &= y[p_z, z]p_x + x[z, p_z]p_y = i\hbar(xp_y - yp_x) = i\hbar L_z. \end{aligned} \quad (3.5)$$

Analogamente

$$\begin{aligned}[L_y, L_z] &= [zp_x - xp_z, xp_y - yp_x] \\ &= z[p_x, x]p_y + y[x, p_x]p_z = i\hbar(yp_z - zp_y) = i\hbar L_x\end{aligned}\quad (3.6)$$

e

$$\begin{aligned}[L_z, L_x] &= [xp_y - yp_x, yp_z - zp_y] \\ &= x[p_y, y]p_z + z[y, p_y]p_x = i\hbar(zp_x - xp_z) = i\hbar L_y.\end{aligned}\quad (3.7)$$

In sintesi:

$$[L_x, L_y] = i\hbar L_z, \quad [L_y, L_z] = i\hbar L_x, \quad [L_z, L_x] = i\hbar L_y. \quad (3.8)$$

Queste relazioni di commutazione possono essere espresse in forma più compatta in questo modo

$$\boxed{[L_i, L_j] = i\hbar \epsilon_{ijk} L_k} \quad (3.9)$$

dove ϵ_{ijk} è il tensore totalmente antisimmetrico¹.

Poi introduciamo l'operatore hermitiano

$$L^2 = L_x^2 + L_y^2 + L_z^2. \quad (3.10)$$

Possiamo vedere facilmente che L^2 commuta con ciascuna delle componenti di $\mathbf{L}$. Ad esempio

$$\begin{aligned}[L^2, L_x] &= [L_x^2 + L_y^2 + L_z^2, L_x] \\ &= [L_x^2, L_x] + [L_y^2, L_x] \\ &= L_y[L_y, L_x] + [L_y, L_x]L_y + L_z[L_z, L_x] + [L_z, L_x]L_z \\ &= i\hbar(-L_y L_z - L_z L_y + L_z L_y + L_y L_z) = 0.\end{aligned}\quad (3.11)$$

e analogamente per le altre componenti. Dunque l'operatore hermitiano L^2 è un'osservabile compatibile con ciascuna delle componenti del momento angolare ed è possibile trovare un insieme completo di autostati simultanei di L^2 e di uno qualsiasi dei tre operatori L_x, L_y, L_z , ma non più di uno, dato che i tre non sono tra loro compatibili². La scelta è arbitraria. Scegliamo L_z e cerchiamo la base di autostati comuni a L^2 e L_z .

Come primo passo osserviamo che gli autovalori di L^2 non possono essere negativi. Per dimostrarlo supponiamo che un autovalore generico di L^2 sia un numero reale λ e l'autostato corrispondente sia $|\lambda\rangle$, di norma unitaria, in modo che $L^2|\lambda\rangle = \lambda|\lambda\rangle$. Poi consideriamo un qualsiasi stato di un sistema generico, rappresentato dal vettore $|\alpha\rangle$, e calcoliamo il valore di aspettazione di L^2 :

$$\langle\alpha|L^2|\alpha\rangle = \sum_{i=x,y,z} \langle\alpha|L_i^2|\alpha\rangle = \sum_{i=x,y,z} (\langle\alpha|L_i)(L_i|\alpha\rangle). \quad (3.12)$$

Ma questa è la somma delle norme quadre dei vettori $L_i|\alpha\rangle$, che non può essere negativa. Dunque $\langle\alpha|L^2|\alpha\rangle \geq 0$. Nel caso particolare in cui $|\alpha\rangle = |\lambda\rangle$, la stessa condizione vale ancora e diventa

$$\langle\lambda|L^2|\lambda\rangle = \lambda\langle\lambda|\lambda\rangle = \lambda \geq 0. \quad (3.13)$$

¹Queste regole di commutazione sono identiche a quelle già trovate per le tre componenti dello spin in (1.133). Ciò non implica, tuttavia, che lo spin e il momento angolare siano la stessa grandezza fisica, come vedremo tra un po'.

²Una conseguenza importante della incompatibilità delle componenti x, y e z del momento angolare è l'impossibilità (di principio, non solo pratica) di fissare contemporaneamente i tre valori delle componenti di $\mathbf{L}$ di una particella. Rappresentare $\mathbf{L}$ con un segmento orientato, una freccia, come si fa in fisica classica, non ha più senso, dato che la conoscenza di una componente di tale freccia in una direzione generica esclude che si possano conoscere con certezza i valori delle componenti nelle due direzioni ortogonali ad essa. Quindi niente freccia!

Dunque λ è non negativo e dobbiamo anche ricordarci che non è un numero puro, ma ha le dimensioni di un momento angolare al quadrato. Se vogliamo scriverlo come numero puro, allora ci conviene, ad esempio, scrivere l'autovalore di L^2 come $\hbar^2\lambda$, sfruttando il fatto che $\hbar$ ha le dimensioni del momento angolare.

Gli autostati di L^2 sono anche autostati di L_z . Scriviamo l'autovalore di L_z nella forma $\hbar m$ in modo che m sia un numero puro, reale e incognito. Usiamo i valori λ e m per etichettare gli stati $\{|\lambda m\rangle\}$ che costituiscono la base ortonormale simultanea dei due operatori, tali che

$$\begin{aligned} L^2|\lambda m\rangle &= \hbar^2\lambda |\lambda m\rangle \\ L_z|\lambda m\rangle &= \hbar m |\lambda m\rangle. \end{aligned} \quad (3.14)$$

Per arrivare fin qui abbiamo usato solo il fatto che il momento angolare ha le stesse dimensioni di $\hbar$, che L^2 e L_z sono operatori hermitiani e sono anche osservabili compatibili e che L^2 è definito positivo (ha autovalori positivi o al più nulli). Questo è solo l'aperitivo. Ora dobbiamo risolvere il vero problema: determinare i valori λ e m .

Per procedere usiamo un metodo algebrico simile a quello utilizzato per ricavare lo spettro dell'oscillatore armonico. A questo scopo introduciamo i seguenti operatori

$$L_+ = L_x + iL_y, \quad L_- = L_x - iL_y, \quad (3.15)$$

che sono anche analoghi agli operatori di *spin flip* definiti nel capitolo 1. Questi operatori sono l'uno l'hermitiano coniugato dell'altro. Inoltre, assieme a L_z sono sufficienti a definire completamente il momento angolare, in alternativa alle tre componenti L_x, L_y, L_z . Le regole di commutazione per gli operatori L_+ e L_- si possono calcolare facilmente:

$$[L^2, L_{\pm}] = [L^2, L_x \pm iL_y] = [L^2, L_x] \pm i[L^2, L_y] = 0, \quad (3.16)$$

$$\begin{aligned} [L_z, L_+] &= [L_z, L_x + iL_y] = [L_z, L_x] + i[L_z, L_y] \\ &= i\hbar L_y + \hbar L_x = \hbar L_+, \end{aligned} \quad (3.17)$$

$$\begin{aligned} [L_z, L_-] &= [L_z, L_x - iL_y] = [L_z, L_x] - i[L_z, L_y] \\ &= i\hbar L_y - \hbar L_x = -\hbar L_- \end{aligned} \quad (3.18)$$

e infine

$$\begin{aligned} [L_+, L_-] &= [L_x + iL_y, L_x - iL_y] = -i[L_x, L_y] + i[L_y, L_x] \\ &= -2i[L_x, L_y] = 2\hbar L_z. \end{aligned} \quad (3.19)$$

Poi notiamo che i due valori di aspettazione

$$\langle \lambda m | L_+ L_- | \lambda m \rangle, \quad \langle \lambda m | L_- L_+ | \lambda m \rangle \quad (3.20)$$

devono essere entrambi non negativi, dato che coincidono rispettivamente con la norma quadra dei vettori $L_-|lm\rangle$ e $L_+|lm\rangle$. Ma dalle definizioni di L_+ e L_- segue che

$$L_+ L_- = L_x^2 + L_y^2 - i[L_x, L_y] = L_x^2 + L_y^2 + \hbar L_z = L^2 - L_z^2 + \hbar L_z \quad (3.21)$$

$$L_- L_+ = L_x^2 + L_y^2 + i[L_x, L_y] = L_x^2 + L_y^2 - \hbar L_z = L^2 - L_z^2 - \hbar L_z \quad (3.22)$$

e quindi possiamo scrivere

$$0 \leq \langle \lambda m | L_+ L_- | \lambda m \rangle = \langle \lambda m | L^2 - L_z^2 + \hbar L_z | \lambda m \rangle = \hbar^2[\lambda - m(m-1)] \quad (3.23)$$

$$0 \leq \langle \lambda m | L_- L_+ | \lambda m \rangle = \langle \lambda m | L^2 - L_z^2 - \hbar L_z | \lambda m \rangle = \hbar^2 [\lambda - m(m+1)] \quad (3.24)$$

ovvero

$$\lambda - m(m-1) \geq 0 \quad (3.25)$$

$$\lambda - m(m+1) \geq 0. \quad (3.26)$$

Dunque, dato un valore di λ , l'autovalore m non può assumere valori grandi o piccoli a piacere. Ora, nulla ci vieta di adottare una nuova notazione per l'autovalore di L^2 , introducendo un numero l non negativo, tale che $\lambda = l(l+1)$. Il numero l è incognito tanto quanto lo era λ . La convenienza sta nel fatto che, riscrivendo le due condizioni precedenti in termini di l ,

$$l(l+1) - m(m-1) \geq 0 \quad (3.27)$$

$$l(l+1) - m(m+1) \geq 0, \quad (3.28)$$

si vede chiaramente che m deve essere confinato nell'intervallo compreso tra $-l$ e l . Infatti, se m fosse negativo e minore di $-l$ la prima sarebbe violata, se fosse positivo e maggiore di l sarebbe violata la seconda. La prima conclusione a cui arriviamo, dunque, è che lo spettro deve essere rappresentato dalle due equazioni agli autovalori

$$\begin{aligned} L^2 |lm\rangle &= \hbar^2 l(l+1) |lm\rangle \\ L_z |lm\rangle &= \hbar m |lm\rangle, \end{aligned} \quad (3.29)$$

con l non negativo e m compreso tra $-l$ e l , estremi inclusi.

Ora prendiamo un generico autovettore della base ortonormale $\{|lm\rangle\}$ e da questo, applicando L_- , generiamo il nuovo vettore $L_- |lm\rangle$. Verifichiamo se questo vettore è ancora autostato di L_z o no. Abbiamo

$$\begin{aligned} L_z (L_- |lm\rangle) &= L_z L_- |lm\rangle \\ &= (L_- L_z - [L_-, L_z]) |lm\rangle \\ &= (L_- L_z - \hbar L_-) |lm\rangle \\ &= L_- (L_z - \hbar) |lm\rangle \\ &= L_- (\hbar m - \hbar) |lm\rangle \\ &= \hbar(m-1) L_- |lm\rangle \end{aligned} \quad (3.30)$$

e dunque $L_- |lm\rangle$ è effettivamente un autostato di L_z , con autovalore $\hbar(m-1)$. L'applicazione di L_- ha semplicemente abbassato di una unità il valore di m . In realtà, nessuno ci assicura che la norma del vettore $L_- |lm\rangle$ sia ancora 1. Il vettore $L_- |lm\rangle$ e l'autovettore normalizzato $|l, m-1\rangle$ possono differire per una costante moltiplicativa c ,

$$L_- |lm\rangle = c |l, m-1\rangle. \quad (3.31)$$

Ne segue che la norma quadra del vettore $L_- |lm\rangle$ è

$$\langle lm | L_-^\dagger L_- | lm \rangle = |c|^2 \langle l, m-1 | l, m-1 \rangle = |c|^2. \quad (3.32)$$

Ma la stessa quantità può anche essere scritta sapendo che $L_-^\dagger = L_+$ e usando l'espressione (3.22) per l'operatore $L_- L_+$:

$$\langle lm | L_+ L_- | lm \rangle = \langle lm | L^2 - L_z^2 - \hbar L_z | lm \rangle = \hbar^2 [l(l+1) - m(m-1)] \quad (3.33)$$

e quindi

$$c = \hbar \sqrt{l(l+1) - m(m-1)}, \quad (3.34)$$

dove abbiamo scelto convenzionalmente la fase in modo da ottenere un numero reale, che dalla (3.27) sappiamo essere non negativo. In conclusione

$$L_-|lm\rangle = \hbar\sqrt{l(l+1) - m(m-1)} |l, m-1\rangle. \quad (3.35)$$

L'intera procedura può essere ripetuta con l'operatore L_+ al posto di L_- . L'unica differenza sta nel segno opposto del commutatore con L_z . Si ottiene

$$L_+|lm\rangle = \hbar\sqrt{l(l+1) - m(m+1)} |l, m+1\rangle. \quad (3.36)$$

Naturalmente l'applicazione di L_+ e L_- non cambia il valore di l , dato che entrambi gli operatori commutano con L^2 .

Dalle ultime due relazioni vediamo che, se m è un autovalore di L_z , anche $m \pm 1$, $m \pm 2$, $m \pm 3 \dots$ lo sono. Per passare da un valore all'altro basta applicare in sequenza gli operatori L_+ e L_- ; l'uno aumenta m di un'unità, l'altro lo diminuisce. In questo senso possiamo dire che si tratta di operatori *a scala*, analoghi a quelli già usati per lo spettro dell'oscillatore armonico.

C'è un vincolo di cui dobbiamo tenere conto: m non può violare le condizioni (3.27)-(3.28). In altre parole, affinché m rimanga in un intervallo limitato, pur potendo applicare L_+ o L_- infinite volte in sequenza, è necessario che, ad un certo punto della sequenza il coefficiente $\sqrt{l(l+1) - m(m \pm 1)}$ si annulli in modo da ottenere un vettore nullo, bloccando così la scaletta. Nel caso di L_+ il coefficiente si annulla quando $m = l$. La sequenza di applicazioni di L_+ si interrompe, quindi, se dopo un certo numero di passi si arriva ad applicare L_+ all'autovettore con $m = l - 1$, ottenendo lo stato $|ll\rangle$. Nel verso opposto, una sequenza di applicazioni di L_- , a partire dallo stesso valore m , si interrompe se in un certo numero di passi si arriva $m = -l$. Quindi, se vogliamo rimanere nei limiti richiesti, la distanza di m sia da l che da $-l$ deve essere un numero intero positivo o al più nullo (è il numero di passi permessi). Ciò implica che $2l$ sia esso stesso un numero intero non negativo. In conclusione, lo spettro è quello dato in (3.29) con le condizioni

$$l \text{ intero o semintero} \geq 0 \quad \text{e} \quad m = -l, -l+1, \dots, l-1, l. \quad (3.37)$$

Ma non è ancora finita. Fin qui abbiamo usato come ipotesi solo il fatto che $\mathbf{L}$ è un operatore hermitiano, con le dimensioni di un momento angolare, le cui tre componenti L_x , L_y e L_z sono tali che $[L_i, L_j] = i\hbar\epsilon_{ijk}L_k$. Ogni operatore che soddisfi queste condizioni avrebbe lo stesso spettro che abbiamo appena trovato. Tuttavia non abbiamo ancora usato, in alcun punto della derivazione, il fatto che $\mathbf{L}$ è il momento angolare $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, ottenuto sostituendo le grandezze classiche con operatori quantistici. Nulla vieta *a priori* che questa ulteriore specificazione riduca lo spettro ad un sottoinsieme di quello appena calcolato. Per dare una risposta occorre studiare il legame tra momento angolare e rotazioni.

3.2 Momento angolare orbitale e rotazioni

Qui potremmo partire dal caso generale delle rotazioni attorno ad un asse generico usando gli angoli di Eulero, ecc., come viene fatto in molti libri di testo più esaustivi di questi appunti. Tuttavia, per capire il meccanismo di base è sufficiente iniziare dal caso di una rotazione attorno all'asse z , che è relativamente semplice e coinvolge quanto basta per capire la fisica che ci serve. L'estensione al caso generale, anche se formalmente più complicata, non aggiunge nulla di sostanziale.

Ora seguiamo la stessa logica che avevamo adottato per le traslazioni spaziali nella sezione 1.11. In particolare partiamo da un autostato generico dell'operatore posizione $\mathbf{r}_{\text{op}}$ e

operiamo una rotazione di un angolo infinitesimo $d\phi$ attorno all'asse z . Per capire di quale rotazione si tratta conviene scrivere la relazione agli autovalori³

$$\mathbf{r}_{\text{op}}|\mathbf{r}\rangle = \mathbf{r}|\mathbf{r}\rangle, \quad (3.38)$$

dove l'autovalore $\mathbf{r}$ a destra dell'uguale è un vettore nello spazio euclideo delle coordinate, ovvero una terna di numeri. La terna può essere espressa come (x, y, z) oppure come terna di coordinate polari (r, θ, ϕ) , legate dalle relazioni

$$\begin{cases} x = r \cos \phi \sin \theta \\ y = r \sin \phi \sin \theta \\ z = r \cos \theta \end{cases} \quad (3.39)$$

Quello che vogliamo è trovare l'operatore che trasforma l'autostato di $\mathbf{r}_{\text{op}}$ con autovalore (r, θ, ϕ) nell'autostato, sempre di $\mathbf{r}_{\text{op}}$, ma con autovalore ruotato in $(r, \theta, \phi + d\phi)$. Vogliamo che lo faccia mantenendo la norma, ammettendo uno sviluppo continuo dall'identità e garantendo una regola di composizione in sequenze di rotazioni, esattamente com'avevamo fatto per le traslazioni spaziali e temporali. Per soddisfare questi requisiti, l'operatore di rotazione infinitesima attorno all'asse z deve avere la forma

$$\mathbb{1} - \frac{i}{\hbar} \mathcal{L} d\phi \quad (3.40)$$

dove $\mathcal{L}$ è un operatore hermitiano, da determinare, e la sua azione è espressa dalla relazione

$$\left(\mathbb{1} - \frac{i}{\hbar} \mathcal{L} d\phi \right) |r, \theta, \phi\rangle = |r, \theta, \phi + d\phi\rangle, \quad (3.41)$$

analogamente alla

$$\left(\mathbb{1} - \frac{i}{\hbar} p_x dx \right) |x, y, z\rangle = |x + dx, y, z\rangle, \quad (3.42)$$

per le traslazioni spaziali. Avendo scelto di inserire $\hbar$ a denominatore ed essendo $d\phi$ adimensionale, l'operatore $\mathcal{L}$ deve avere le stesse dimensioni di $\hbar$, che sono quelle di un momento angolare.

L'espressione (3.41) ci dice che il vettore

$$\left(\mathbb{1} - \frac{i}{\hbar} \mathcal{L} d\phi \right) |r, \theta, \phi\rangle \quad (3.43)$$

è ancora autovettore di $\mathbf{r}_{\text{op}}$, ma con autovalore ruotato, $\mathbf{r}' = (r, \theta, \phi + d\phi)$. Questa informazione può essere tradotta nell'equazione agli autovalori

$$\mathbf{r}_{\text{op}} \left(\mathbb{1} - \frac{i}{\hbar} \mathcal{L} d\phi \right) |r, \theta, \phi\rangle = \mathbf{r}' \left(\mathbb{1} - \frac{i}{\hbar} \mathcal{L} d\phi \right) |r, \theta, \phi\rangle. \quad (3.44)$$

Ora applichiamo a sinistra e a destra lo stesso operatore $\mathbb{1} + (i/\hbar)\mathcal{L}d\phi$ e svolgiamo i prodotti tenendo i termini fino al primo ordine in $d\phi$:

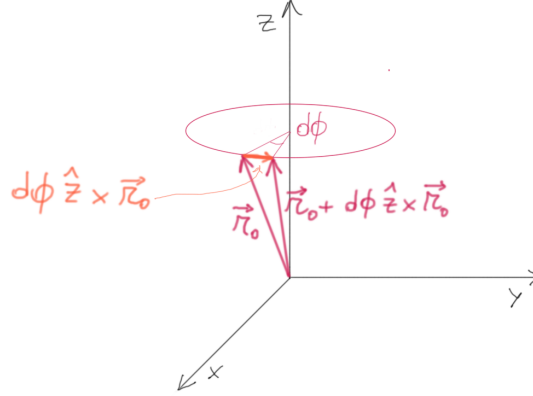
$$\left(\mathbf{r}_{\text{op}} + \frac{i}{\hbar} [\mathcal{L}, \mathbf{r}_{\text{op}}] d\phi \right) |r, \theta, \phi\rangle = \mathbf{r}' |r, \theta, \phi\rangle. \quad (3.45)$$

³Si noti che in questa sezione uso il suffisso "op" per l'operatore posizione, $\mathbf{r}_{\text{op}}$, per non confonderlo con i vettori euclidei, ma non lo uso per il momento $\mathbf{p}$ e per il momento angolare $\mathbf{L}$, per i quali non c'è rischio di confusione.

Poi notiamo che

$$\mathbf{r}' = \mathbf{r} + \hat{\mathbf{z}} \times \mathbf{r} d\phi . \quad (3.46)$$

Quest'ultima uguaglianza viene semplicemente dalla geometria. Basta guardare la figura qui sotto, dove un generico vettore euclideo $\mathbf{r}_0$ viene ruotato di un angolo $d\phi$ e questo corrisponde a sommare a $\mathbf{r}_0$, nell'usuale significato di somma di vettori euclidei, un vettore infinitesimo le cui componenti coincidono con quelle del prodotto vettoriale $\hat{\mathbf{z}} \times \mathbf{r}_0$ volte l'ampiezza della rotazione $d\phi$.



Dunque, possiamo scrivere

$$\left(\mathbf{r}_{\text{op}} + \frac{i}{\hbar} [\mathcal{L}, \mathbf{r}_{\text{op}}] d\phi \right) |\mathbf{r}\rangle = (\mathbf{r} + \hat{\mathbf{z}} \times \mathbf{r} d\phi) |\mathbf{r}\rangle . \quad (3.47)$$

Questa è vera se

$$\mathbf{r}_{\text{op}} + \frac{i}{\hbar} [\mathcal{L}, \mathbf{r}_{\text{op}}] d\phi = \mathbf{r}_{\text{op}} + \hat{\mathbf{z}} \times \mathbf{r}_{\text{op}} d\phi , \quad (3.48)$$

in modo che, agendo su $|\mathbf{r}\rangle$, questo operatore dia proprio l'autovalore $\mathbf{r}' = \mathbf{r} + \hat{\mathbf{z}} \times \mathbf{r} d\phi$. Ne segue che

$$[\mathcal{L}, \mathbf{r}_{\text{op}}] = -i\hbar \hat{\mathbf{z}} \times \mathbf{r}_{\text{op}} . \quad (3.49)$$

L'operatore $\mathcal{L}$ che soddisfa questa relazione è la componente z del momento angolare $\mathbf{L} = \mathbf{r}_{\text{op}} \times \mathbf{p}$. Infatti,

$$[L_z, x_{\text{op}}] = [x_{\text{op}} p_y - y_{\text{op}} p_x, x_{\text{op}}] = -y_{\text{op}} [p_x, x_{\text{op}}] = i\hbar y_{\text{op}} \quad (3.50)$$

$$[L_z, y_{\text{op}}] = [x_{\text{op}} p_y - y_{\text{op}} p_x, y_{\text{op}}] = x_{\text{op}} [p_y, y_{\text{op}}] = -i\hbar x_{\text{op}} \quad (3.51)$$

$$[L_z, z_{\text{op}}] = [x_{\text{op}} p_y - y_{\text{op}} p_x, z_{\text{op}}] = 0 \quad (3.52)$$

e

$$(\hat{\mathbf{z}} \times \mathbf{r}_{\text{op}})_x = -y_{\text{op}} , \quad (\hat{\mathbf{z}} \times \mathbf{r}_{\text{op}})_y = x_{\text{op}} , \quad (\hat{\mathbf{z}} \times \mathbf{r}_{\text{op}})_z = 0 , \quad (3.53)$$

da cui appunto segue che

$$[L_z, \mathbf{r}_{\text{op}}] = -i\hbar \hat{\mathbf{z}} \times \mathbf{r}_{\text{op}} . \quad (3.54)$$

In conclusione, abbiamo dimostrato che L_z è il generatore delle rotazioni attorno all'asse z . La rotazione infinitesima corrisponde all'azione dell'operatore

$$1 - \frac{i}{\hbar} L_z d\phi , \quad (3.55)$$

in modo che

$$\left(\mathbb{1} - \frac{i}{\hbar} L_z d\phi \right) |r, \theta, \phi\rangle = |r, \theta, \phi + d\phi\rangle. \quad (3.56)$$

Per le rotazioni di un angolo finito a partire da quelle infinitesime basta applicare lo stesso operatore in sequenza, dividendo l'angolo in intervalli infinitesimi in modo che

$$|r, \theta, \phi + \Delta\phi\rangle = \lim_{N \rightarrow \infty} \left(\mathbb{1} - \frac{i}{\hbar} L_z \frac{\Delta\phi}{N} \right)^N |r, \theta, \phi\rangle = \exp \left(-\frac{i}{\hbar} L_z \Delta\phi \right) |r, \theta, \phi\rangle. \quad (3.57)$$

La scelta iniziale dell'asse z è del tutto arbitraria e potevamo partire da un asse qualsiasi, individuato da un versore $\hat{\mathbf{u}}$ generico. Nulla cambia nella procedura per determinare l'operatore associato alle rotazioni attorno a questo asse. Nella figura precedente il versore $\hat{\mathbf{z}}$ va sostituito con $\hat{\mathbf{u}}$ e, usando le regole di commutazione tra le componenti di $\mathbf{L}$, si trova che la rotazione infinitesima di un angolo $d\phi_u$ attorno a $\hat{\mathbf{u}}$ è prodotta dall'operatore

$$\mathbb{1} - \frac{i}{\hbar} \mathbf{L} \cdot \hat{\mathbf{u}} d\phi. \quad (3.58)$$

Il generatore della rotazione è la componente di $\mathbf{L}$ nella direzione $\hat{\mathbf{u}}$. Per rotazioni di un angolo $\Delta\phi_u$ finito, l'operatore è

$$\exp \left(-\frac{i}{\hbar} \mathbf{L} \cdot \hat{\mathbf{u}} \Delta\phi_u \right). \quad (3.59)$$

Le componenti L_x e L_y del momento angolare corrispondono ai casi particolari delle rotazioni attorno agli assi x e y .

A questo punto si capisce anche meglio la ragione per cui le componenti di $\mathbf{L}$ non commutano tra loro. Un diverso ordine delle rotazioni applicate ad un vettore porta a diversi vettori finali. Ad esempio, se il vettore di partenza sta nella direzione z e applichiamo prima una rotazione di $\pi/2$ attorno all'asse y e poi un'altra, sempre di $\pi/2$, attorno all'asse z otteniamo un vettore lungo y ; invece, se allo stesso vettore applichiamo prima la rotazione attorno a z e poi quella attorno a y , otteniamo un vettore lungo x . Le regole di commutazione (3.9) esprimono formalmente la non commutatività delle rotazioni spaziali. Il gruppo delle rotazioni di cui L_x , L_y e L_z sono i generatori è detto *non abeliano*, proprio perché i generatori non commutano. La cosa è diversa nel caso delle traslazioni spaziali, dove i generatori p_x , p_y e p_z commutano tra loro e, per questo, tali trasformazioni formano un gruppo *abeliano*.

Il fatto che l'operatore $\mathbf{L} = \mathbf{r}_{\text{op}} \times \mathbf{p}$ è associato alle rotazioni nello spazio euclideo, ha una conseguenza molto importante sulla struttura dello spettro degli operatori L^2 e L_z , già introdotto nella sezione precedente. Avevamo visto che, usando le sole regole di commutazione degli operatori, avevamo dedotto che equazioni agli autovalori dovevano avere la forma (3.29). Le riscriviamo qui per comodità

$$\begin{aligned} L^2 |lm\rangle &= \hbar^2 l(l+1) |lm\rangle \\ L_z |lm\rangle &= \hbar m |lm\rangle. \end{aligned}$$

Avevamo anche visto che l e m potevano essere numeri interi o seminteri, con l non negativo e m compreso tra $-l$ e l , estremi inclusi. Ora prendiamo il duale della relazione (3.57),

$$\langle r, \theta, \phi + \Delta\phi | = \langle r, \theta, \phi | \exp \left(\frac{i}{\hbar} L_z \Delta\phi \right) \quad (3.60)$$

e moltiplichiamo a destra entrambi i membri per un generico autovettore $|lm\rangle$:

$$\langle r, \theta, \phi + \Delta\phi | lm\rangle = \langle r, \theta, \phi | \exp \left(\frac{i}{\hbar} L_z \Delta\phi \right) | lm\rangle. \quad (3.61)$$

Poi sfruttiamo il fatto che l'operatore esponenziale nel membro di destra si trova ad agire su un suo autovettore e può essere sostituito con il corrispondente autovalore:

$$\langle r, \theta, \phi + \Delta\phi | lm \rangle = e^{im\Delta\phi} \langle r, \theta, \phi | lm \rangle. \quad (3.62)$$

I due prodotti interni a sinistra e a destra, non sono altro che le componenti del vettore $|lm\rangle$ sulla base degli autostati della posizione e dunque, per definizione, corrispondono all'autofunzione $\psi_{lm}(\mathbf{r}) = \langle \mathbf{r} | lm \rangle$ calcolata per due angoli diversi:

$$\psi_{lm}(r, \theta, \phi + \Delta\phi) = e^{im\Delta\phi} \psi_{lm}(r, \theta, \phi). \quad (3.63)$$

E qui arriviamo ad un punto cruciale: in coordinate polari l'angolo ϕ è definito “modulo 2π ”, nel senso che l'angolo $\phi = 2\pi$ è identico all'angolo $\phi = 0$ dal punto di vista delle coordinate spaziali; in altri termini, due punti nello spazio che differiscono solo per una variazione di 2π dell'angolo ϕ sono lo stesso punto! Dato che la funzione d'onda contiene le informazioni sullo stato di una particella espresse nello spazio delle coordinate, e queste informazioni devono essere note in modo univoco, allora la funzione d'onda deve essere una funzione monodroma, cioè deve ammettere un solo valore in ciascun punto. Compiere un giro completo di 2π attorno all'asse z porta allo stesso punto e, dunque, allo stesso valore della funzione d'onda. La funzione d'onda ruotata

$$\psi_{lm}(r, \theta, \phi + 2\pi) = e^{im2\pi} \psi_{lm}(r, \theta, \phi) \quad (3.64)$$

deve essere uguale alla $\psi_{lm}(r, \theta, \phi)$, ma questo è vero se e solo se

$$e^{im2\pi} = 1, \quad (3.65)$$

e quindi solo se m è intero. Abbiamo quindi trovato un vincolo ulteriore sugli autovalori: se $\mathbf{L}$ è il momento angolare che si ottiene sostituendo nella definizione classica $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ i corrispondenti operatori quantistici, allora m può essere solo intero e, di conseguenza, anche l deve essere intero. Questo segue dalle proprietà delle rotazioni nello spazio delle coordinate, di cui le tre componenti di $\mathbf{L}$ sono i generatori. Questo momento angolare, associato a funzioni monodrome nella base delle coordinate, è comunemente detto *momento angolare orbitale*, proprio per il suo legame con le rotazioni nello spazio tridimensionale euclideo⁴. Il suo spettro è

$$\begin{array}{l} \left\{ \begin{array}{l} L^2 |lm\rangle = \hbar^2 l(l+1) |lm\rangle \\ L_z |lm\rangle = \hbar m |lm\rangle \end{array} \right. \\ \text{con} \\ l \geq 0 \text{ intero} \\ m = -l, -l+1, \dots, l-1, l. \end{array}$$

Ora che abbiamo sistemato la questione degli autovalori l e m , possiamo determinare la forma delle autofunzioni di L^2 e L_z nella rappresentazione delle coordinate. Per procedere, andiamo dapprima a determinare l'azione degli operatori associati al momento angolare quando agiscono sulle generiche funzioni d'onda di una particella. Partiamo da L_z e

⁴Un'applicazione diretta di questi concetti è il caso dello spettro dei livelli energetici di un rotore rigido. Consideriamo un'asta rigida con due masse agli estremi e sia I il momento d'inerzia riferito ad un asse principale di rotazione, perpendicolare all'asta. L'energia cinetica classica associata ad una rotazione con momento angolare $\mathbf{L}$ è $E = L^2/2I$. Nel caso quantistico possiamo convertire l'energia classica in un operatore hamiltoniano avente la stessa forma: $H = L^2/2I$. Il calcolo dello spettro di H è immediato, essendo $E_l = \hbar^2 l(l+1)/2I$ con $l = 0, 1, 2, \dots$. Per molecole biatomiche, questo significa che nello spettro di emissione o assorbimento si troveranno frequenze del tipo $\nu_{l,l+1} = \Delta E_{l,l+1}/h = 2B(l+1)$, dove $B = h/(8\pi^2 I)$ è detta costante rotazionale. Si tratta di serie di righe spettrali equispaziate. La spaziatura ci dà una stima diretta del momento d'inerzia della molecola.

prendiamo la relazione (3.56); ne scriviamo il duale

$$\langle r, \theta, \phi + d\phi | = \langle r, \theta, \phi | \left(\mathbb{1} + \frac{i}{\hbar} L_z d\phi \right) \quad (3.66)$$

e moltiplichiamo entrambi i membri per un generico vettore $|\alpha\rangle$:

$$\langle r, \theta, \phi + d\phi | \alpha \rangle = \langle r, \theta, \phi | \left(\mathbb{1} + \frac{i}{\hbar} L_z d\phi \right) | \alpha \rangle. \quad (3.67)$$

A sinistra abbiamo una funzione complessa di r, θ e ϕ che possiamo sviluppare al primo ordine in ϕ

$$\langle r, \theta, \phi + d\phi | \alpha \rangle = \langle r, \theta, \phi | \alpha \rangle + d\phi \frac{\partial}{\partial \phi} \langle r, \theta, \phi | \alpha \rangle \quad (3.68)$$

in modo che la relazione precedente diventa

$$\frac{\partial}{\partial \phi} \langle r, \theta, \phi | \alpha \rangle = \frac{i}{\hbar} \langle r, \theta, \phi | L_z | \alpha \rangle, \quad (3.69)$$

ovvero

$$\langle r, \theta, \phi | L_z | \alpha \rangle = \frac{\hbar}{i} \frac{\partial}{\partial \phi} \psi_\alpha(r, \theta, \phi). \quad (3.70)$$

Dunque, quando agisce sulle funzioni d'onda l'operatore L_z è rappresentato dall'operatore differenziale

$$L_z \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial \phi}. \quad (3.71)$$

Ad esempio, se vogliamo calcolare il suo valore di aspettazione usando le funzioni d'onda, abbiamo

$$\begin{aligned} \langle \alpha | L_z | \alpha \rangle &= \int d^3r \langle \alpha | r, \theta, \phi \rangle \langle r, \theta, \phi | L_z | \alpha \rangle \\ &= \int d^3r \langle \alpha | r, \theta, \phi \rangle \frac{\hbar}{i} \frac{\partial}{\partial \phi} \langle r, \theta, \phi | \alpha \rangle \\ &= \int d^3r \psi_\alpha^*(r, \theta, \phi) \frac{\hbar}{i} \frac{\partial}{\partial \phi} \psi_\alpha(r, \theta, \phi). \end{aligned} \quad (3.72)$$

Questa è analoga alla (1.247) già scritta in sezione 1.12 per l'operatore momento e che riscriviamo qui per comodità nella forma

$$\langle \alpha | p_x | \alpha \rangle = \int d^3r \psi_\alpha^*(x, y, z) \frac{\hbar}{i} \frac{\partial}{\partial x} \psi_\alpha(x, y, z). \quad (3.73)$$

Nel caso del momento, anche le altre componenti p_y e p_z , sono rappresentate allo stesso modo dalle rispettive derivate parziali. Nel caso del momento angolare non è altrettanto semplice e gli operatori differenziali che rappresentano L_x e L_y sono più complicati. Il motivo della complicazione sta nel fatto che, mentre l'angolo di rotazione attorno all'asse z coincide con ϕ , che è una delle coordinate polari, l'angolo di rotazione attorno ad assi diversi implica combinazioni di diverse coordinate polari. Il calcolo è un po' tedioso, ma lo facciamo almeno per uno dei due.

Consideriamo una rotazione di un angolo $d\phi_x$ attorno all'asse x . Per un vettore generico $\mathbf{r}$ tale rotazione implica una variazione $d\mathbf{r} = \hat{\mathbf{x}} \times \mathbf{r} d\phi_x$. In termini di coordinate questo si traduce in $dx = 0$, $dy = -z d\phi_x$ e $dz = y d\phi_x$. Possiamo ricordare le definizioni delle coordinate polari in (3.39) e, in particolare, la relazione $z = r \cos \theta$, da cui segue che $dz =$

$-r \sin \theta d\theta$. Confrontando le due espressioni di dz e usando anche la $y = r \sin \phi \sin \theta$, vediamo che

$$d\theta = -\sin \phi d\phi_x. \quad (3.74)$$

Per sapere quanto varia l'angolo ϕ nella stessa rotazione basta ricordare che $dx = 0$ e che $x = r \cos \phi \sin \theta$, da cui $d(\cos \phi \sin \theta) = 0$, che implica

$$-\sin \phi \sin \theta d\phi + \cos \phi \cos \theta d\theta = 0. \quad (3.75)$$

Inserendo in questa il risultato precedente per $d\theta$ otteniamo

$$-\sin \phi \sin \theta d\phi - \cos \phi \cos \theta \sin \phi d\phi_x = 0, \quad (3.76)$$

da cui

$$d\phi = -\cot \theta \cos \phi d\phi_x. \quad (3.77)$$

Ora possiamo scrivere la relazione analoga alla (3.66), ma per la rotazione attorno all'asse x , il cui generatore è l'operatore L_x ,

$$\langle r, \theta + d\theta, \phi + d\phi | = \langle r, \theta, \phi | \left(\mathbb{1} + \frac{i}{\hbar} L_x d\phi_x \right), \quad (3.78)$$

e usare le espressioni di $d\theta$ e $d\phi$ appena ricavate, in modo che

$$\langle r, \theta - \sin \phi d\phi_x, \phi - \cot \theta \cos \phi d\phi_x | = \langle r, \theta, \phi | \left(\mathbb{1} + \frac{i}{\hbar} L_x d\phi_x \right). \quad (3.79)$$

Come avevamo fatto prima per L_z , moltiplichiamo a destra per un generico vettore di stato $|\alpha\rangle$

$$\langle r, \theta - \sin \phi d\phi_x, \phi - \cot \theta \cos \phi d\phi_x | \alpha \rangle = \langle r, \theta, \phi | \left(\mathbb{1} + \frac{i}{\hbar} L_x d\phi_x \right) | \alpha \rangle \quad (3.80)$$

e sviluppiamo la funzione a sinistra al primo ordine

$$\langle r, \theta, \phi | \alpha \rangle - \sin \phi d\phi_x \frac{\partial}{\partial \theta} \langle r, \theta, \phi | \alpha \rangle - \cot \theta \cos \phi d\phi_x \frac{\partial}{\partial \phi} \langle r, \theta, \phi | \alpha \rangle \quad (3.81)$$

da cui segue che

$$\langle r, \theta, \phi | L_x | \alpha \rangle = \frac{\hbar}{i} \left(-\sin \phi \frac{\partial}{\partial \theta} - \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) \langle r, \theta, \phi | \alpha \rangle. \quad (3.82)$$

Dunque, quando agisce sulle funzioni d'onda l'operatore L_x è rappresentato dall'operatore differenziale

$$L_x \rightarrow \frac{\hbar}{i} \left(-\sin \phi \frac{\partial}{\partial \theta} - \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right). \quad (3.83)$$

che è effettivamente più complicato che nel caso di L_z in (3.71). Seguendo la stessa procedura è facile vedere che

$$L_y \rightarrow \frac{\hbar}{i} \left(\cos \phi \frac{\partial}{\partial \theta} - \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right). \quad (3.84)$$

Gli operatori *a scala* diventano

$$L_{\pm} = L_x \pm iL_y \rightarrow \frac{\hbar}{i} e^{\pm i\phi} \left(\pm i \frac{\partial}{\partial \theta} - \cot \theta \frac{\partial}{\partial \phi} \right) \quad (3.85)$$

e l'operatore L^2 è rappresentato da

$$L^2 = L_x^2 + L_y^2 + L_z^2 \rightarrow -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right]. \quad (3.86)$$

È importante notare che questo operatore differenziale non è nuovo; lo si incontra anche in fisica classica quando si ha a che fare con il laplaciano ∇^2 . Infatti, se espresso in coordinate polari, il laplaciano è

$$\nabla^2 = \frac{1}{r} \frac{\partial^2}{\partial r^2} r + \frac{1}{r^2} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right], \quad (3.87)$$

e l'espressione nella parentesi quadra è la stessa.

Ora torniamo alla questione dello spettro di L^2 e L_z . Prendiamo le loro equazioni agli autovalori e le moltiplichiamo a sinistra per il generico *bra* $\langle r, \theta, \phi |$:

$$\begin{cases} \langle r, \theta, \phi | L^2 | lm \rangle = \hbar^2 l(l+1) \langle r, \theta, \phi | lm \rangle \\ \langle r, \theta, \phi | L_z | lm \rangle = \hbar m \langle r, \theta, \phi | lm \rangle. \end{cases} \quad (3.88)$$

A questo punto però sappiamo anche come scrivere i termini a sinistra come operatori differenziali che agiscono sulle funzioni d'onda. Infatti, usando le espressioni (3.70) e (3.86) possiamo riscrivere le equazioni agli autovalori nella forma

$$\begin{cases} - \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] \psi_{lm}(r, \theta, \phi) = l(l+1) \psi_{lm}(r, \theta, \phi) \\ -i \frac{\partial}{\partial \phi} \psi_{lm}(r, \theta, \phi) = m \psi_{lm}(r, \theta, \phi). \end{cases} \quad (3.89)$$

Poi notiamo che a sinistra abbiamo operatori differenziali che agiscono solo sulle coordinate angolari θ e ϕ , ma non sulla coordinata radiale r . Ciò risulta particolarmente utile nei casi in cui la funzione d'onda sia fattorizzabile nella forma

$$\psi_{lm}(r, \theta, \phi) = R(r) Y_{lm}(\theta, \phi), \quad (3.90)$$

dove $R(r)$ è la parte radiale e $Y_{lm}(\theta, \phi)$ è la parte angolare. In tal caso, nelle equazioni precedenti la parte radiale si semplifica e si ottengono le due equazioni seguenti:

$$\begin{cases} - \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] Y_{lm}(\theta, \phi) = l(l+1) Y_{lm}(\theta, \phi) \\ -i \frac{\partial}{\partial \phi} Y_{lm}(\theta, \phi) = m Y_{lm}(\theta, \phi). \end{cases} \quad (3.91)$$

In queste equazioni $\hbar$ non c'è più e tutto è adimensionale. Si tratta di un problema puramente matematico. Le funzioni $Y_{lm}(\theta, \phi)$ che soddisfano queste equazioni sono funzioni speciali dette *armoniche sferiche*, note, tabulate e normalmente disponibili nei pacchetti software per il calcolo scientifico.

La seconda equazione è semplice e ammette soluzioni del tipo

$$Y_{lm}(\theta, \phi) = e^{im\phi} F_{lm}(\theta). \quad (3.92)$$

Per determinare le funzioni $F_{lm}(\theta)$ possiamo ricordare che agendo con l'operatore L_+ sullo stato con $m = l$ si ottiene zero. L'espressione dell'operatore L_+ in termini di derivate l'avevamo scritta in (3.85) e la condizione $L_+ Y_{ll}(\theta, \phi) = 0$ diventa

$$\left(i e^{il\phi} \frac{\partial}{\partial \theta} F_{ll}(\theta) - \cot \theta F_{ll}(\theta) \frac{\partial}{\partial \phi} e^{il\phi} \right) = 0 \quad (3.93)$$

ovvero

$$\frac{\partial}{\partial \theta} F_{ll}(\theta) - l \cot \theta F_{ll}(\theta) = 0 \quad (3.94)$$

la cui soluzione è

$$F_{ll}(\theta) = c \sin^l \theta \quad (3.95)$$

dove la costante di proporzionalità c possiamo sceglierla normalizzando le funzioni $Y_{lm}(\theta, \phi)$ come vogliamo. Dunque

$$Y_{ll}(\theta, \phi) = c e^{il\phi} \sin^l \theta \quad (3.96)$$

e tutte le altre possono essere ricavate applicando in sequenza l'operatore L_- a partire da ciascuna delle Y_{ll} . Il problema dello spettro di L^2 e L_z , sia per quanto riguarda la determinazione degli autovalori che delle autofunzioni, può dirsi risolto.

Le armoniche sferiche per $l = 0$ e $l = 1$ hanno la forma

$$Y_{00}(\theta, \phi) = \sqrt{\frac{1}{4\pi}} \quad (3.97)$$

$$Y_{11}(\theta, \phi) = -\sqrt{\frac{3}{8\pi}} \sin \theta e^{i\phi} \quad (3.98)$$

$$Y_{10}(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \cos \theta \quad (3.99)$$

$$Y_{1,-1}(\theta, \phi) = \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\phi} \quad (3.100)$$

dove la normalizzazione è fissata da

$$\int d\Omega |Y_{lm}(\theta, \phi)|^2 = 1 \quad (3.101)$$

essendo $d\Omega = \sin \theta d\theta d\phi$ l'angolo solido infinitesimo. Si può mostrare che le armoniche sferiche soddisfano la relazione di ortonormalità

$$\int d\Omega Y_{lm}^*(\theta, \phi) Y_{l'm'}(\theta, \phi) = \delta_{ll'} \delta_{mm'} \quad (3.102)$$

e che possono essere usate per esprimere una qualsiasi funzione degli angoli⁵:

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_{lm} Y_{lm}(\theta, \phi). \quad (3.103)$$

Per altre proprietà interessanti di queste funzioni rimandiamo ai libri di testo, più esaurienti di questi appunti.

3.3 Particella in un campo centrale

Consideriamo una particella di massa m che si muove soggetta a forze centrali. La funzione di Hamilton classica avrà la forma $H(\mathbf{p}, \mathbf{r}) = p^2/(2m) + V(r)$, dove V è un'energia potenziale che dipende solo dal r . Tale funzione è invariante per rotazioni spaziali e, per

⁵Usando il nostro gergo da fisici, il fatto che le armoniche sferiche costituiscano una base ortonormale completa per le funzioni di θ e ϕ segue in modo diretto dal loro essere autofunzioni dei due operatori L^2, L_z che costituiscono l'insieme massimale di osservabili compatibili per lo spazio di Hilbert associato alle rotazioni nello spazio euclideo.

questo motivo, il momento angolare si conserva. Tradotto nel formalismo della meccanica quantistica il moto della particella è determinato dall'hamiltoniano $H = p^2/(2m) + V(r)$ che si ottiene sostituendo le grandezze classiche con operatori. Possiamo utilizzare l'equazione di Schrödinger (2.23), che diventa

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = \left(-\frac{\hbar^2 \nabla^2}{2m} + V(r) \right) \psi(\mathbf{r}, t), \quad (3.104)$$

e possiamo anche utilizzare l'operatore L^2 , tramite le (3.86) e (3.87), per scrivere la stessa equazione così:

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = \left(-\frac{\hbar^2}{2mr} \frac{\partial^2}{\partial r^2} r + \frac{L^2}{2mr^2} + V(r) \right) \psi(\mathbf{r}, t). \quad (3.105)$$

A questo punto si vede che l'operatore L^2 commuta con l'hamiltoniano, dato che agisce solo sulle coordinate angolari e quindi commuta con qualsiasi funzione di r e con qualsiasi derivata rispetto a r . Lo stesso vale per l'operatore L_z che agisce solo sull'angolo ϕ . Dunque L^2 e L_z sono costanti del moto, come nel caso classico, come conseguenza dell'isotropia del potenziale. Questo ci permette di calcolare gli stati stazionari della particella risolvendo l'equazione di Schrödinger indipendente dal tempo

$$\left(-\frac{\hbar^2}{2mr} \frac{\partial^2}{\partial r^2} r + \frac{L^2}{2mr^2} + V(r) \right) \psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (3.106)$$

potendo contare sul fatto che le autofunzioni ψ che risolvono questo problema agli autovalori per l'hamiltoniano sono anche, contemporaneamente, autofunzioni di L^2 e L_z . Possiamo quindi cercare soluzioni di questa equazione corrispondenti ad un dato valore del momento angolare o, meglio, corrispondenti ad una particella la cui funzione d'onda, per la parte angolare, è un'autofunzione di L^2 e L_z , in questo modo:

$$\psi(r, \theta, \phi) = R(r)Y_{lm}(\theta, \phi). \quad (3.107)$$

Inserendo questa espressione nell'equazione di Schrödinger troviamo

$$-\frac{\hbar^2}{2mr} \frac{d^2}{dr^2} r R(r) + \left(\frac{\hbar^2 l(l+1)}{2mr^2} + V(r) \right) R(r) = ER(r), \quad (3.108)$$

che è un'equazione puramente radiale. Conviene definire una nuova funzione

$$u(r) = rR(r) \quad (3.109)$$

che inserita nell'equazione precedente dà

$$-\frac{\hbar^2}{2m} \frac{d^2 u(r)}{dr^2} + \left(\frac{\hbar^2 l(l+1)}{2mr^2} + V(r) \right) u(r) = Eu(r), \quad (3.110)$$

Questa è l'equazione di Schrödinger per una particella che si muove in una dimensione, soggetta ad un potenziale efficace; quest'ultimo include, in aggiunta al potenziale vero $V(r)$, un termine analogo alla *barriera centrifuga* classica. In effetti, la riduzione del problema da tre dimensioni a una, sfruttando la conservazione del momento angolare, è del tutto analoga al caso newtoniano, salvo che qui le soluzioni non sono leggi orarie e traiettorie, ma bensì funzioni d'onda. L'equazione può essere usata per studiare sia gli stati *liberi* della particella, in cui la densità di probabilità di trovare la particella in un dato punto dello spazio rimane finita anche a distanze infinite dalla sorgente, sia nel caso di stati *legati*, in cui

la densità di probabilità ha valori non nulli solo in una regione limitata attorno alla sorgente. Gli stati liberi sono essenziali per descrivere i fenomeni di diffusione (*scattering*); gli stati legati sono essenziali per comprendere la spettroscopia degli atomi e delle molecole, ma non solo. Esempi relativamente semplici di potenziali centrali sono la buca sferica con pareti infinite (analogo in tre dimensioni della buca quadra che si studia come primo esempio di soluzioni dell'equazione di Schrödinger nei corsi introduttivi alla meccanica ondulatoria) e il potenziale armonico in tre dimensioni. Esercizi come questi, o simili, vengono svolti nel corso di Complementi matematici della meccanica quantistica, e se ne trovano parecchi di interessanti nel testo *Practical Quantum Mechanics* di S. Flügge. Qui ci limiteremo a trattare l'esempio più significativo, l'atomo di idrogeno.

3.4 Atomo d'idrogeno

L'idrogeno è l'atomo più semplice: un protone al centro e un elettrone che gli si muove attorno. L'elettrone è circa 2000 volte più leggero e, per questo, conviene trascurare il moto del protone. Se il protone se ne sta fermo nell'origine del sistema di coordinate, allora l'elettrone si muove soggetto ad un potenziale centrale coulombiano $V(r) = -e^2/r$, dove e è la carica elementare, uguale per le due particelle a meno del segno (che qui però non conta), e dove abbiamo usato ancora una volta il sistema di misura CGS. Gli effetti della massa finita del protone, così come altri effetti correttivi al modello potranno essere studiati in seguito, quando serve (ad oggi le misure spettroscopiche sono talmente accurate che diventano misurabili anche correzioni estremamente piccole, come quelle dovute al raggio finito del protone ad esempio).

Dunque, l'equazione di Schrödinger da risolvere è

$$-\frac{\hbar^2}{2m_e} \frac{d^2 u(r)}{dr^2} + \left(\frac{\hbar^2 l(l+1)}{2m_e r^2} - \frac{e^2}{r} \right) u(r) = E u(r), \quad (3.111)$$

e dovrà essere risolta per ogni l intero non negativo. Una volta trovate le soluzioni $u(r)$, le funzioni d'onda della particella saranno della forma

$$\psi_{lm}(r, \theta, \phi) = \frac{1}{r} u(r) Y_{lm}(\theta, \phi). \quad (3.112)$$

Di passaggio notiamo che abbiamo usato il simbolo m_e per la massa dell'elettrone per non confonderlo con l'autovalore di L_z . Siccome stiamo studiando l'atomo e non il problema dello *scattering* elettrone-protone, consideriamo solo gli stati legati, che sono quelli con energia E negativa. Possiamo anche moltiplicare tutto per $-2m_e/\hbar^2$ e definire le grandezze $\epsilon = \sqrt{2m_e|E|}/\hbar^2$ e $\eta = 2m_e e^2/\hbar^2$, in modo che l'equazione diventa

$$\frac{d^2 u(r)}{dr^2} - \left(\frac{l(l+1)}{r^2} - \frac{\eta}{r} + \epsilon^2 \right) u(r) = 0. \quad (3.113)$$

Ora, per qualsiasi valore di ϵ non nullo, la soluzione asintotica per $r \rightarrow \infty$ è facilmente calcolabile notando che i due termini in $1/r$ e $1/r^2$ diventano trascurabili rispetto a ϵ e l'equazione, in questo limite, si riduce a

$$\frac{d^2 u(r)}{dr^2} = \epsilon^2 u(r), \quad (3.114)$$

la cui soluzione decrescente è

$$u(r) \propto e^{-\epsilon r}. \quad (3.115)$$

Dunque per soddisfare questa condizione asintotica, cerchiamo soluzioni del tipo

$$u(r) = f(r)e^{-\epsilon r}, \quad (3.116)$$

dove la funzione $f(r)$ incognita non deve crescere più rapidamente di $e^{\epsilon r}$ per $r \rightarrow \infty$. Inserendo questa espressione nell'equazione di Schrödinger si ottiene

$$\frac{d^2 f}{dr^2} - 2\epsilon \frac{df}{dr} - \left(\frac{l(l+1)}{r^2} - \frac{\eta}{r} \right) f = 0. \quad (3.117)$$

Possiamo anche ricavare l'andamento della $f(r)$ nell'intorno dell'origine. Infatti, se $r \rightarrow 0$, allora il termine centrifugo domina su quello coulombiano e l'equazione diventa

$$\frac{d^2 f}{dr^2} - 2\epsilon \frac{df}{dr} - \frac{l(l+1)}{r^2} f = 0, \quad (3.118)$$

e possiamo cercare una soluzione del tipo $f(r) \propto r^\alpha$. Sostituendo si ottiene

$$\alpha(\alpha-1)r^{\alpha-2} - 2\epsilon\alpha r^{\alpha-1} - l(l+1)r^{\alpha-2} = 0, \quad (3.119)$$

ovvero

$$\alpha(\alpha-1) = l(l+1) + 2\epsilon\alpha r. \quad (3.120)$$

L'ultimo termine è trascurabile nel limite di $r \rightarrow 0$; perciò

$$\alpha(\alpha-1) = l(l+1), \quad (3.121)$$

e quindi $\alpha = l+1$ (ci sarebbe anche la soluzione $\alpha = -l$, ma non è accettabile perché darebbe funzioni d'onda che divergono all'origine). Dunque la soluzione che cerchiamo deve avere la forma r^{l+1} per $r \rightarrow 0$.

Dato che la funzione d'onda deve essere una funzione continua e derivabile, possiamo scrivere la funzione incognita $f(r)$ come una serie di potenze

$$f(r) = r^{l+1} \sum_{j=0}^{\infty} a_j r^j = \sum_{j=0}^{\infty} a_j r^{l+j+1}, \quad (3.122)$$

con $a_0 \neq 0$. In questo modo l'andamento per $r \rightarrow 0$ è quello giusto. Per trovare i coefficienti incogniti basta inserire quest'ultima serie nell'equazione (3.117):

$$\sum_{j=0}^{\infty} [(l+j+1)(l+j)a_j r^{l+j-1} - 2\epsilon(l+1+j)a_j r^{l+j} - l(l+1)a_j r^{l+j-1} + \eta a_j r^{l+j}] = 0. \quad (3.123)$$

Possiamo ridistribuire gli addendi in modo da scrivere la stessa espressione come somma di quattro sommatorie; la prima e la terza contengono termini in r^{l+j-1} , la seconda e la quarta in r^{l+j} . Se nella prima e la terza introduciamo un nuovo indice $j' = j-1$, allora le potenze di r saranno della forma $r^{l+j'}$. Facendo così, le due sommatorie cominciano con $j' = -1$; ma è immediato vedere che i termini con $j' = -1$ nelle due sommatorie sono uguali ma di segno opposto, $\pm a_0 l(l+1)r^l$, e dunque si elidono. Rimangono così quattro somme con indici interi che corrono da 0 a infinito. Nulla ci vieta di togliere l'apice a j' e chiamarlo nuovamente j , in modo tale da poter raccogliere di nuovo il tutto in un'unica sommatoria:

$$\sum_{j=0}^{\infty} [(l+j+2)(l+j+1)a_{j+1} - 2\epsilon(l+1+j)a_j - l(l+1)a_{j+1} + \eta a_j] r^{l+j} = 0, \quad (3.124)$$

ovvero

$$a_{j+1} = \frac{2\epsilon(l+1+j) - \eta}{(l+j+2)(l+j+1) - l(l+1)} a_j. \quad (3.125)$$

Questa è una *formula di ricorrenza*. Dato il coefficiente a_0 (che può essere scelto arbitrariamente, essendo l'equazione di Schrödinger omogenea), gli altri si calcolano applicando questa formula. Ma c'è una condizione da rispettare: la funzione $f(r)$ non può crescere più rapidamente di $e^{\epsilon r}$. Per vedere come questa condizione si traduce in termini dei coefficienti della formula di ricorrenza, ne analizziamo il comportamento a grandi valori dell'indice j , quando prende la forma

$$\lim_{j \gg 1} \frac{a_{j+1}}{a_j} = \frac{2\epsilon}{j}. \quad (3.126)$$

Ma questa relazione coincide con quella che si ottiene per l'espansione in serie dell'esponenziale $e^{2\epsilon r}$. Infatti la serie dell'esponenziale è

$$e^{2\epsilon r} = \sum_{j=0}^{\infty} \frac{(2\epsilon)^j}{j!} r^j, \quad (3.127)$$

e soddisfa anch'essa la relazione (3.126). Ne segue che, se nella serie di potenze che rappresenta la $f(r)$ sono non nulli anche i coefficienti per j che tende all'infinito, allora la soluzione necessariamente cresce come $e^{2\epsilon r}$, cioè più rapidamente di $e^{\epsilon r}$, violando la condizione di annullamento asintotico della funzione d'onda degli stati legati.

Per evitare la contraddizione, l'unica possibilità è che la serie sia troncata ad un certo valore di j , diciamo $\bar{j}$, in modo che tutti gli a_j per $j > \bar{j}$ siano nulli. Questo succede se per $j = \bar{j}$ il coefficiente davanti ad a_j nella formula di ricorrenza (3.125) si annulla:

$$2\epsilon(l+1+\bar{j}) = \eta. \quad (3.128)$$

Il numero in parentesi è un intero; lo chiamiamo n . Inoltre ricordiamo le definizioni di η e ϵ in modo da poter scrivere

$$n = \frac{\eta}{2\epsilon} = \sqrt{\frac{m_e e^4}{2|E|\hbar^2}}, \quad (3.129)$$

ovvero

$$E_n = -\frac{m_e e^4}{2n^2 \hbar^2}. \quad (3.130)$$

con $n \geq 1$. Quindi l'energia degli stati legati assume valori discreti. L'intero n è detto *numero quantico principale*.

Lo stato di energia più bassa, detto *stato fondamentale*, corrisponde a caso $n = 1$. Dato che $n = l + 1 + \bar{j}$ e che l e $\bar{j}$ sono interi non negativi, allora $n = 1$ implica $l = \bar{j} = 0$. L'energia vale

$$E_1 = -\frac{m_e e^4}{2\hbar^2} = -13.6 \text{ eV}, \quad (3.131)$$

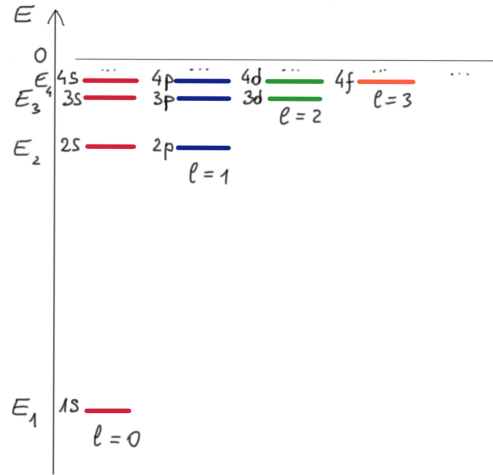
dove ricordiamo che $1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}$. Il modulo di E_1 rappresenta l'energia minima necessaria a liberare l'elettrone dall'atomo e, per questo, è detta *energia di ionizzazione*. Il primo livello sopra lo stato fondamentale corrisponde a $n = 2$. Lo stesso valore di n può essere ottenuto sia con $l = 0$ che con $l = 1$, rispettivamente per $\bar{j} = 1$ e 0 . Allo stesso modo si vede che ad ogni n corrispondono i valori

$$l = 0, 1, 2, \dots, n-1 \quad (3.132)$$

corrispondenti a $j = n - 1, n - 2, \dots, 0$. Inoltre, per ogni l , ci sono $2l + 1$ valori possibili dell'autovalore m della componente L_z del momento angolare. Dunque, la degenerazione totale di ogni livello energetico E_n è

$$\sum_{l=0}^{n-1} (2l + 1) = n^2 \quad (3.133)$$

Questa degenerazione è tipica del potenziale $1/r$. Chi fosse curioso di capirne il perché, vada a vedersi la definizione classica di *vettore di Lenz* e il suo corrispettivo operatore quantistico (ad esempio, leggendo il capitolo 11.3 del Forte e Rottoli).



Lo spettro (3.130) è in accordo con le serie spettrali osservate negli esperimenti di emissione e assorbimento della luce e dà la corretta predizione per l'energia di ionizzazione. È lo stesso ricavato da Bohr nella vecchia teoria dei quanti, ma qui non c'è traccia di orbite degli elettroni né di regole di quantizzazione introdotte *ad hoc*. Inoltre nel modello di Bohr c'era una corrispondenza uno a uno tra il numero quantico principale n e il valore del momento angolare e lo stato fondamentale era caratterizzato da $n = l = 1$. Qui invece al valore $n = 1$ corrisponde $l = 0$, inoltre, diversi valori di l sono ammessi per i livelli con $n \geq 2$. Nel gergo della spettroscopia, gli stati con $l = 0$ sono chiamati orbitali *s*, quelli con $l = 1, 2, 3$ rispettivamente orbitali *p*, *d* e *f*. Gli stati dell'elettrone sono univocamente determinati da tre numeri quantici, n, l e m , che fissano gli autovalori delle tre osservabili H, L^2 e L_z . Queste costituiscono l'insieme massimale di osservabili compatibili per l'atomo d'idrogeno (ma senza tener conto dello spin dell'elettrone, di cui parleremo dopo).

Per quanto riguarda le funzioni d'onda, mettendo assieme i risultati precedenti possiamo scrivere

$$\begin{aligned} \psi_{nlm}(r, \theta, \phi) &= R_{nl}(r) Y_{lm}(\theta, \phi) = \frac{u_{nl}(r)}{r} Y_{lm}(\theta, \phi) \\ &= r^{-1} f_{nl}(r) e^{-r/na_B} Y_{lm}(\theta, \phi) \end{aligned} \quad (3.134)$$

dove $a_B = \hbar^2/(m_e e^2)$ è il cosiddetto *raggio di Bohr* e

$$f_{nl}(r) = r^{l+1} (a_0 + a_1 r + \dots + a_{n-l-1} r^{n-l-1}). \quad (3.135)$$

I polinomi di grado $n - l - 1$ che soddisfano la formula di ricorrenza (3.125) sono noti come *polinomi di Laguerre*, che sono casi particolare di una funzione nota come *funzione*

ipergeometrica confluyente. Per i tre livelli di energia più bassa, la dipendenza da r è data da

$$R_{10}(r) = 2 a_B^{-3/2} e^{-r/a_B} \quad (3.136)$$

$$R_{20}(r) = (2a_B)^{-3/2} \left(2 - \frac{r}{a_B}\right) e^{-r/2a_B} \quad (3.137)$$

$$R_{21}(r) = (2a_B)^{-3/2} \frac{r}{\sqrt{3}a_B} e^{-r/2a_B}, \quad (3.138)$$

Per effetto del potenziale efficace centrifugo, solo gli stati con $l = 0$ hanno funzioni d'onda che non si annullano in $r = 0$. Il numero quantico n determina univocamente il numero di nodi della funzione d'onda degli orbitali s nella direzione radiale, che è pari a $n - 1$. I nodi nelle direzioni angolari sono fissati dalla forma delle armoniche sferiche.

Può sembrare strano che la funzione d'onda dell'elettrone possa essere non nulla in $r = 0$ dove si trova il protone, ma dobbiamo ricordarci che la probabilità di trovare l'elettrone in un volumetto d^3r è data da $|\psi|^2 d^3r$ e in coordinate polari l'elemento di volume è $r^2 dr d\Omega$. Dunque la probabilità di trovare l'elettrone ad una distanza compresa tra r e $r + dr$ è

$$r^2 |R_{nl}(r)|^2 dr \int d\Omega |Y_{lm}(\theta, \phi)|^2 \quad (3.139)$$

e si annulla sempre per $r = 0$. Invece il raggio di Bohr, r_B , fissa la scala radiale su cui la funzione d'onda si annulla a grandi distanze, per effetto dell'esponenziale e^{-r/na_B} nella (3.134). Nel modello di Bohr $n^2 r_B$ era il raggio dell'orbita circolare dell'elettrone nel livello n . Qui invece non ha alcun senso parlare di raggi e di orbite in senso classico, ma solo di densità di probabilità.

Ovviamente, quello proposto qui non è l'unico modo per risolvere l'equazione di Schrödinger con il potenziale coulombiano; gli studenti interessati troveranno derivazioni diverse in altri testi.

3.5 Spin

Nel primo capitolo avevamo già definito gli operatori di spin S_x , S_y e S_z associati agli stati degli atomi nell'esperimento di Stern-Gerlach. La loro rappresentazione matriciale l'avevamo data in sezione 1.6.2 e può essere sintetizzata dall'espressione $S_i = (\hbar/2)\sigma_i$, con $i = x, y, z$ e dove le σ_i sono le matrici di Pauli. Si tratta di tre operatori hermitiani, aventi le stesse dimensioni del momento angolare, che possono essere visti come le tre componenti di un operatore di spin $\mathbf{S}$. Avevamo anche verificato che le loro regole di commutazione sono

$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k. \quad (3.140)$$

Inoltre possiamo anche definire l'operatore $S^2 = S_x^2 + S_y^2 + S_z^2$. Questo commuta con gli operatori S_i e possiamo ricavare gli autostati simultanei di S^2 e S_z seguendo esattamente quanto già fatto in sezione 3.1 per il momento angolare orbitale. Dunque lo spettro sarà dato da

$$\begin{aligned} S^2 |sm_s\rangle &= \hbar^2 s(s+1) |sm_s\rangle \\ S_z |sm_s\rangle &= \hbar m_s |sm_s\rangle. \end{aligned} \quad (3.141)$$

con

$$s \text{ intero o semintero } \geq 0 \quad (3.142)$$

$$m_s = -s, -s+1, \dots, s-1, s. \quad (3.143)$$

La differenza cruciale tra lo spin e il momento angolare orbitale sta nel fatto che quest'ultimo è il generatore delle rotazioni nello spazio euclideo e, per questo, abbiamo visto che vanno esclusi dal suo spettro i valori seminteri di l . Per lo spin non esiste un argomento analogo che ci faccia escludere i valori seminteri di s . Anzi, se scriviamo la matrice che rappresenta S^2

$$\begin{aligned}
 S^2 &= S_x^2 + S_y^2 + S_z^2 = \\
 &= \left(\frac{\hbar}{2}\right)^2 \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right. \\
 &\quad \left. + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right] \\
 &= \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
 \end{aligned} \tag{3.144}$$

vediamo che è diagonale nella base degli autostati di S_z , avendo come autovalore $s(s+1)\hbar^2 = (3/4)\hbar^2$, ovvero $s = 1/2$, a cui corrispondono i due autovalori $m_s = \pm 1/2$ di S_z . Dunque gli stati della base $|\uparrow\rangle$ e $|\downarrow\rangle$ corrispondono ai due autostati $|\frac{1}{2}, \frac{1}{2}\rangle$ e $|\frac{1}{2}, -\frac{1}{2}\rangle$, nella notazione $|sm_s\rangle$ per una particella di spin $1/2$. Da informazioni indipendenti si sa che lo spin dell'atomo di argento coincide con lo spin di un solo elettrone, il 47-esimo, che occupa da solo l'orbitale $5s$, dato che tutti gli altri elettroni dello stesso atomo non contribuiscono globalmente al momento magnetico. L'esperimento quindi è una prova che l'elettrone è una particella con spin $1/2$. Fra parentesi, a questo punto si potrebbe pensare che sarebbe stato più semplice misurare lo spin degli elettroni facendo passare nel campo magnetico dello Stern-Gerlach direttamente un fascio di raggi catodici, che sono appunto elettroni; ma in tal caso l'effetto dominante sarebbe stata una deviazione indotta dalla forza di Lorentz, $-e\mathbf{v} \times \mathbf{B}$, tale da mascherare completamente l'effetto dello spin.

A differenza del momento angolare orbitale $\mathbf{L}$ che ammette valori interi positivi l qualsiasi per una particella che si muove nello spazio, lo spin di una particella ha un valore s fissato, caratteristico della particella stessa, che può essere sia intero che semintero. Un'altra differenza è che gli autostati dello spin non sono rappresentabili da funzioni degli angoli, com'erano le armoniche sferiche $Y_{lm}(\theta, \phi)$, perchè $\mathbf{S}$ non è associato alla rotazione delle coordinate spaziali. Lo spin viene anche chiamato momento angolare *intrinseco* e questo induce a pensare che si tratti di una rotazione della particella attorno al proprio asse. Ma questa idea non è corretta. Una rotazione attorno al proprio asse ha significato solo se la particella è un corpo esteso, ma l'elettrone, per quel che ne sappiamo a tutt'oggi, è una particella puntiforme, o comunque talmente piccola da rendere incompatibile il valore misurato dello spin con qualsiasi rotazione di una sferetta con quel raggio. Inoltre se anche fosse una sferetta che ruota su se stessa il suo momento angolare avrebbe le proprietà di $\mathbf{L}$, essendo associato ad una vera rotazione nello spazio delle coordinate, e non quelle di $\mathbf{S}$ e pertanto i valori seminteri di s non sarebbero ammessi.

In realtà, anche $\mathbf{S}$ è associato a rotazioni, ma non a quelle dei vettori nello spazio usuale euclideo infinito-dimensionale. Lo spin è il generatore di rotazioni in uno spazio astratto finito-dimensionale. Il caso $s = 1/2$ corrisponde ad uno spazio bidimensionale, che è lo spazio di dimensione minima in cui si possono implementare le regole di commutazione del momento angolare. Questo può essere compreso meglio se si conosce la teoria dei gruppi e, in particolare, le proprietà delle trasformazioni unitarie del gruppo $SU(2)$. Per rimanere su un piano semplice, immaginiamo di avere due versori nello spazio delle coordinate $\hat{\mathbf{e}}_1$ e $\hat{\mathbf{e}}_2$, il secondo dei quali essendo ottenuto dal primo per effetto di una rotazione di un angolo ϕ_u attorno ad un asse individuato dalla direzione di un versore $\hat{\mathbf{u}}$. Chiamiamo $|\pm_{(1)}\rangle$ i due autostati dello spin di una particella con spin $1/2$, che si ottengono se si misura la

componente di $\mathbf{S}$ nella direzione individuata dal versore $\hat{\mathbf{e}}_1$. Allora, i due vettori ottenuti con la seguente operazione

$$|\pm_{(2)}\rangle = \exp\left(-\frac{i}{\hbar}\mathbf{S} \cdot \hat{\mathbf{u}} \phi_u\right) |\pm_{(1)}\rangle \quad (3.145)$$

rappresentano una base di autostati della componente di $\mathbf{S}$ nella direzione $\hat{\mathbf{e}}_2$. Con un po' di pazienza, potete verificare che le combinazioni lineari (1.126) e (1.127) per gli autostati delle componenti dello spin sono consistenti con la rotazione degli assi secondo la formula appena scritta. Ad esempio, se applichiamo una rotazione $\phi = \pi/2$ attorno all'asse z agli autostati di S_x , $|S_x \uparrow\rangle$ e $|S_x \downarrow\rangle$, questi si trasformano negli autostati di S_y , $|S_y \uparrow\rangle$ e $|S_y \downarrow\rangle$, a meno di un fattore di fase globale inessenziale. Altrettanto succede per le altre rotazioni attorno ai tre assi ortogonali. La forma dell'operatore che compare nella relazione precedente è uguale a quella in (3.59) con il momento angolare orbitale. In quel caso però, la rotazione agiva nello spazio infinito-dimensionale degli autovettori della posizione, con autovalori rappresentati da un continuo di vettori reali in uno spazio euclideo, mentre qui la rotazione riguarda uno spazio finito-dimensionale complesso, con autovalori discreti. Un'altra differenza è che nel caso del momento angolare orbitale $\mathbf{L}$ avevamo imposto che la rappresentazione dei vettori di stato nello spazio delle coordinate, cioè la funzione d'onda, fosse monodroma, così limitando i valori di l ai soli interi. Nel caso dello spin accade invece che una rotazione di 2π può produrre un cambio di segno del vettore di stato. Infatti, nel caso di spin $s = 1/2$, applicando la trasformazione scritta sopra ad una rotazione di 2π , ad esempio attorno all'asse z , si ottiene

$$\begin{aligned} \exp\left(-\frac{i}{\hbar}S_z 2\pi\right) |S_x \uparrow\rangle &= \exp\left(-\frac{i}{\hbar}S_z 2\pi\right) \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \\ &= \frac{1}{\sqrt{2}}(e^{-i\pi}|\uparrow\rangle + e^{i\pi}|\downarrow\rangle) \\ &= -\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) = -|S_x \uparrow\rangle. \end{aligned} \quad (3.146)$$

Il cambiamento di segno di uno stato di spin per una rotazione di 360° avviene per tutti i casi in cui lo spin s è semintero. A differenza del momento angolare orbitale, che coinvolge operatori che hanno un analogo classico come la posizione e il momento e dove è del tutto naturale imporre che le funzioni d'onda siano monodrome, qui non abbiamo alcuna ragione fisica per escludere il cambiamento di segno. Semplicemente abbiamo una grandezza fisica che non ha analogo classico, che obbedisce all'algebra dei commutatori del momento angolare, ma che ammette anche autovalori s seminteri. Gli esperimenti, tipo quello di Stern e Gerlach, ci portano ad accettare questa idea e a includerla nella teoria⁶.

L'espressione (3.145) ci aiuta anche a capire meglio la precessione di spin che avevamo trattato nella sezione 2.3.2. Infatti, in quel caso avevamo un campo magnetico uniforme lungo z che induceva un'evoluzione temporale degli stati di spin analoga alla precessione classica di un momento magnetico. Avevamo visto che l'hamiltoniano poteva essere scritto come $H = \omega S_z$, dove $\omega = geB/(2m_e c)$ era la frequenza di Larmor. Supponiamo che lo stato iniziale al tempo 0 sia lo stato $|\alpha, 0\rangle = |S_x \uparrow\rangle$. L'hamiltoniano non dipende dal tempo e dunque l'evoluzione temporale è data da

$$|\alpha, 0; t\rangle = \exp\left(-\frac{i}{\hbar}Ht\right) |S_x \uparrow\rangle = \exp\left(-\frac{i}{\hbar}S_z \omega t\right) |S_x \uparrow\rangle \quad (3.147)$$

⁶Un esempio di esperimento in cui si dimostra empiricamente il cambio di segno degli stati di spin per particelle di spin $1/2$ soggetti a rotazioni di 360° è riportato nel capitolo 3.2.3 del Sakurai. Per osservare gli effetti del cambio di segno è necessario progettare gli esperimenti in modo da rilevare l'interferenza tra uno stato ruotato e uno non ruotato. Non è facile.

ma, alla luce dell'espressione precedente, è facile accorgersi che questa evoluzione temporale è equivalente ad una rotazione di un angolo ωt del versore lungo il quale misurare lo spin. In particolare, se a $t = 0$ si otteneva con certezza il valore $+\hbar/2$ misurando lo spin lungo l'asse individuato dal versore $\hat{\mathbf{e}}_1 = \hat{\mathbf{x}}$, al tempo t si ottiene con certezza lo stesso autovalore ma misurando lo spin lungo un asse $\hat{\mathbf{e}}_2$ ruotato di un angolo ωt attorno all'asse z .

Lo spin della particella è un grado di libertà addizionale rispetto ai gradi di libertà spaziali. Gli stati di spin e gli stati associati al moto nello spazio euclideo appartengono a spazi di vettori diversi, in modo tale che lo stato complessivo di una particella (o di un sistema generico) può essere autostato simultaneo di una delle componenti di $\mathbf{S}$ e di un'osservabile come la posizione $\mathbf{r}$ o il momento $\mathbf{p}$ o il momento angolare $\mathbf{L}$ dato che, agendo in spazi vettoriali diversi, commuta con ciascuno di essi. In una situazione come questa si introduce la nozione di *spazio prodotto diretto* (o anche *tensoriale*), intendendo con ciò lo spazio dei vettori

$$|\mathbf{r}, \pm\rangle = |\mathbf{r}\rangle \otimes |\pm\rangle \quad (3.148)$$

dove il simbolo $\otimes$ esprime il prodotto di due spazi distinti di vettori, quello infinito-dimensionale generato dagli autostati della posizione e quello bidimensionale generato dagli autostati dello spin (spin $1/2$ in questo caso). Gli operatori che agiscono nello spazio generato dagli stati $|\mathbf{r}\rangle$ commutano con quelli che agiscono nello spazio generato da $|\pm\rangle$ e viceversa.

Questo implica che, volendo specificare in modo completo lo stato di una particella ad esempio utilizzando le funzioni d'onda, queste ultime dovranno includere anche l'informazione sullo stato di spin. Nel caso di particelle con spin $1/2$, un modo efficace di procedere consiste nel sostituire l'usuale funzione d'onda $\psi(\mathbf{r})$ con una coppia di funzioni, una per ciascuno dei due stati possibili di S_z ,

$$\psi_+(\mathbf{r}) = \langle \mathbf{r}, + | \alpha \rangle \quad (3.149)$$

$$\psi_-(\mathbf{r}) = \langle \mathbf{r}, - | \alpha \rangle, \quad (3.150)$$

che rappresentano i coefficienti della decomposizione dello stato generico $|\alpha\rangle$ sulla base comune della posizione e dello spin, $|\mathbf{r}, \pm\rangle$. Possiamo usarle per formare un oggetto a due componenti

$$\psi(\mathbf{r}) = \begin{pmatrix} \psi_+(\mathbf{r}) \\ \psi_-(\mathbf{r}) \end{pmatrix} \quad (3.151)$$

che chiamiamo *spinore*. Il modulo quadro $|\psi_{\pm}(\mathbf{r})|^2$ corrisponde alla densità di probabilità di trovare la particella in $\mathbf{r}$ con spin $\pm\hbar/2$ e la normalizzazione diventa

$$\int d^3\mathbf{r} (|\psi_+(\mathbf{r})|^2 + |\psi_-(\mathbf{r})|^2) = 1. \quad (3.152)$$

Per tener conto delle proprietà di rotazione di questi stati conviene introdurre un nuovo operatore

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \quad (3.153)$$

che possiamo chiamare momento angolare totale. Dal punto di vista formale, sarebbe più corretto scrivere

$$\mathbf{J} = \mathbf{L} \otimes \mathbf{1} + \mathbf{1} \otimes \mathbf{S} \quad (3.154)$$

dove il primo operatore di ciascun prodotto agisce solo nello spazio degli autostati $|\mathbf{r}\rangle$ e il secondo nello spazio degli autostati $|\pm\rangle$. Possiamo mantenere la notazione precedente, più semplice, a patto di ricordarci il significato della somma.

Per costruzione, l'operatore $\mathbf{J}$ obbedisce alle stesse regole di commutazione di $\mathbf{L}$ e $\mathbf{S}$:

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k. \quad (3.155)$$

Per quanto riguarda le rotazioni, vediamo cosa fa l'operatore $\exp[-(i/\hbar)\mathbf{J} \cdot \hat{\mathbf{u}}\phi_u]$ quando viene applicato ad un autostato della posizione con autovalore $\mathbf{r}_0$ e dello spin lungo una direzione $\hat{\mathbf{e}}_1$ con autovalore $+\hbar/2$:

$$\exp\left(-\frac{i}{\hbar}\mathbf{J} \cdot \hat{\mathbf{u}}\phi_u\right)|\mathbf{r}_0, \pm_{(1)}\rangle = \exp\left[-\frac{i}{\hbar}(\mathbf{L} \cdot \hat{\mathbf{u}} + \mathbf{S} \cdot \hat{\mathbf{u}})\phi_u\right]|\mathbf{r}_0, \pm_{(1)}\rangle. \quad (3.156)$$

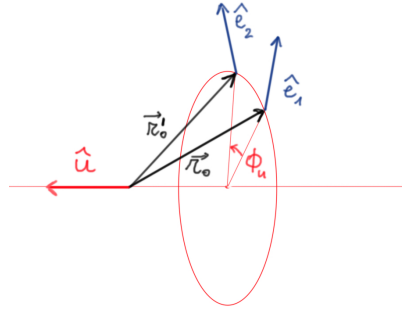
Dato che $\mathbf{L}$ e $\mathbf{S}$ commutano, l'esponenziale può essere separato in due fattori

$$\exp\left(-\frac{i}{\hbar}\mathbf{J} \cdot \hat{\mathbf{u}}\phi_u\right)|\mathbf{r}_0, \pm_{(1)}\rangle = \exp\left(-\frac{i}{\hbar}\mathbf{L} \cdot \hat{\mathbf{u}}\phi_u\right)\exp\left(-\frac{i}{\hbar}\mathbf{S} \cdot \hat{\mathbf{u}}\phi_u\right)|\mathbf{r}_0, \pm_{(1)}\rangle. \quad (3.157)$$

Avevamo già visto prima che i due esponenziali, separatamente, producono una rotazione di un angolo ϕ_u attorno all'asse individuato da $\hat{\mathbf{u}}$, l'uno dell'autovalore della posizione, da $\mathbf{r}_0$ a $\mathbf{r}'_0$, l'altro del versore lungo cui misurare lo spin, da $\hat{\mathbf{e}}_1$ a $\hat{\mathbf{e}}_2$. Dunque

$$\exp\left(-\frac{i}{\hbar}\mathbf{J} \cdot \hat{\mathbf{u}}\phi_u\right)|\mathbf{r}_0, \pm_{(1)}\rangle = |\mathbf{r}'_0, \pm_{(2)}\rangle \quad (3.158)$$

e l'operatore $\mathbf{J}$ si comporta come ci si aspetta per una rotazione complessiva dei vettori di stato, incluso lo spin⁷.



Ma qual è lo spettro del momento angolare totale $\mathbf{J}$? In fisica classica, dati due momenti angolari, la loro somma è una banale somma di vettori ordinari. Qui però abbiamo operatori e la cosa non è altrettanto semplice.

3.6 Addizione di momenti angolari

Innanzitutto ricordiamo che l'importanza del momento angolare risiede nel fatto che in molti casi si tratta di una costante del moto, come succede ad esempio ad una particella che si muove in un campo centrale. In molte situazioni può capitare che il momento angolare orbitale e lo spin commutino con l'hamiltoniano e siano quindi, separatamente costanti del moto; allora anche la loro somma lo sarà. In altre situazioni, invece, una delle due, o entrambe, non sono costanti del moto, ma la loro somma $\mathbf{J}$ sì. Questo capita, ad esempio, quando l'hamiltoniano della particella contiene un termine, detto di *spin-orbit* della forma

$$H_{\text{so}} = f_{\text{so}}(r)\mathbf{L} \cdot \mathbf{S}. \quad (3.159)$$

⁷Per inciso, la commutatività di $\mathbf{L}$ e $\mathbf{S}$ era essenziale a questo scopo perché l'esponenziale e^{A+B} è diverso da $e^A e^B$ se A e B non commutano, come può essere facilmente dimostrato confrontando gli sviluppi in serie di potenze.

In tal caso

$$\begin{aligned}
 [L_z, H_{\text{so}}] &= f_{\text{so}}(r)[L_z, \mathbf{L} \cdot \mathbf{S}] \\
 &= f_{\text{so}}(r)([L_z, L_x]S_x + [L_z, L_y]S_y) \\
 &= i\hbar f_{\text{so}}(r)(L_y S_x - L_x S_y) \neq 0
 \end{aligned} \tag{3.160}$$

e analogamente

$$\begin{aligned}
 [S_z, H_{\text{so}}] &= f_{\text{so}}(r)[S_z, \mathbf{L} \cdot \mathbf{S}] \\
 &= f_{\text{so}}(r)(L_x[S_z, S_x] + L_y[S_z, S_y]) \\
 &= i\hbar f_{\text{so}}(r)(L_x S_y - L_y S_x) \neq 0
 \end{aligned} \tag{3.161}$$

ma, combinando i due risultati si ottiene

$$[J_z, H_{\text{so}}] = f_{\text{so}}(r)[L_z + S_z, \mathbf{L} \cdot \mathbf{S}] = 0. \tag{3.162}$$

Interazioni di spin-orbita sono presenti anche nel moto degli elettroni degli atomi, come effetto di piccole correzioni relativistiche.

Lo stesso problema si pone anche quando $\mathbf{J}$ è la somma di più momenti angolari orbitali o più spin. Prendiamo ad esempio un sistema composto da due particelle identiche. Siano $\mathbf{L}_1$ e $\mathbf{L}_2$ i rispettivi momenti angolari orbitali, a cui corrispondono gli insiemi di autostati $\{|l_1, m_1\rangle\}$ e $\{|l_2, m_2\rangle\}$. Poi supponiamo che le due particelle siano soggette solo ad un campo esterno centrale, in modo che l'hamiltoniano possa essere scritto in questo modo

$$H = \left(\frac{p_1^2}{2m} + V(r_1) \right) + \left(\frac{p_2^2}{2m} + V(r_2) \right) = H_1 + H_2. \tag{3.163}$$

In tal caso, $\mathbf{L}_1$ e $\mathbf{L}_2$ commutano separatamente con H e sono costanti del moto. Se invece le particelle interagiscono anche tra di loro con un potenziale del tipo $v(|\mathbf{r}_1 - \mathbf{r}_2|)$, allora l'hamiltoniano diventa

$$H = H_1 + H_2 + v(|\mathbf{r}_1 - \mathbf{r}_2|). \tag{3.164}$$

In termini di commutatori questo implica che $\mathbf{L}_1$ e $\mathbf{L}_2$ non commutano più con H . Infatti, prendiamo ad esempio la componente z di $\mathbf{L}_1$ e la scriviamo come $L_{z,1} = x_1 p_{y,1} - y_1 p_{x,1}$. Poi ci ricordiamo che, come nella (2.98), vale $[p_x, f(\mathbf{r})] = -i\hbar(\partial f/\partial x)$, e altrettanto per le componenti y e z . Allora possiamo scrivere

$$\begin{aligned}
 [L_{z,1}, H] &= [L_{z,1}, v] \\
 &= [x_1 p_{y,1} - y_1 p_{x,1}, v] \\
 &= -i\hbar \left(x_1 \frac{\partial v}{\partial y_1} - y_1 \frac{\partial v}{\partial x_1} \right)
 \end{aligned} \tag{3.165}$$

e analogamente

$$[L_{z,2}, H] = -i\hbar \left(x_2 \frac{\partial v}{\partial y_2} - y_2 \frac{\partial v}{\partial x_2} \right). \tag{3.166}$$

Per le derivate parziali valgono le relazioni

$$\begin{aligned}
 \frac{\partial v}{\partial x_1} &= v' \frac{\partial |\mathbf{r}_1 - \mathbf{r}_2|}{\partial x_1} = v' \frac{x_1 - x_2}{|\mathbf{r}_1 - \mathbf{r}_2|} \\
 \frac{\partial v}{\partial x_2} &= v' \frac{\partial |\mathbf{r}_1 - \mathbf{r}_2|}{\partial x_2} = -v' \frac{x_1 - x_2}{|\mathbf{r}_1 - \mathbf{r}_2|}
 \end{aligned} \tag{3.167}$$

dove v' è la derivata di v rispetto alla sua variabile $|\mathbf{r}_1 - \mathbf{r}_2|$. Analogamente per le altre due

$$\begin{aligned}\frac{\partial v}{\partial y_1} &= v' \frac{y_1 - y_2}{|\mathbf{r}_1 - \mathbf{r}_2|} \\ \frac{\partial v}{\partial y_2} &= -v' \frac{y_1 - y_2}{|\mathbf{r}_1 - \mathbf{r}_2|}.\end{aligned}\quad (3.168)$$

Se inseriamo queste derivate nei commutatori precedenti, troviamo

$$[L_{z,1}, H] = \frac{i\hbar v'}{|\mathbf{r}_1 - \mathbf{r}_2|} (x_1 y_2 - x_2 y_1) \quad (3.169)$$

e

$$[L_{z,2}, H] = \frac{i\hbar v'}{|\mathbf{r}_1 - \mathbf{r}_2|} (-x_1 y_2 + x_2 y_1). \quad (3.170)$$

Queste espressioni sono generalmente diverse da zero e, dunque, i due commutatori sono operatori non nulli. Tuttavia, se definiamo il momento angolare totale come $\mathbf{J} = \mathbf{L}_1 + \mathbf{L}_2$, abbiamo

$$[J_z, H] = [L_{z,1} + L_{z,2}, H] = [L_{z,1}, H] + [L_{z,2}, H] = 0 \quad (3.171)$$

e J_z commuta con H . In modo simile possiamo dimostrare che J^2 commuta con H , mentre L_1^2 e L_2^2 non commutano. Questo significa che possiamo etichettare gli autostati di H usando gli autovalori di J^2 e J_z , ma non quelli dei momenti angolari di ciascuna particella.

Da quanto detto si capisce l'importanza di saper determinare lo spettro di un momento angolare totale definito come somma di momenti angolari. Cominciamo a mostrare come si fa nel caso dello spin totale di due particelle con spin $s_1 = s_2 = 1/2$. Sia $\mathbf{J} = \mathbf{S}_1 + \mathbf{S}_2$ lo spin totale. Consideriamo i vettori normalizzati

$$|m_{s1} m_{s2}\rangle = |s_1, m_{s1}\rangle \otimes |s_2, m_{s2}\rangle \quad (3.172)$$

che soddisfano le seguenti equazioni agli autovalori:

$$\begin{aligned}S_1^2 |m_{s1} m_{s2}\rangle &= \hbar^2 s_1(s_1 + 1) |m_{s1} m_{s2}\rangle = (3/4)\hbar^2 |m_{s1} m_{s2}\rangle \\ S_2^2 |m_{s1} m_{s2}\rangle &= \hbar^2 s_2(s_2 + 1) |m_{s1} m_{s2}\rangle = (3/4)\hbar^2 |m_{s1} m_{s2}\rangle \\ S_{z1} |m_{s1} m_{s2}\rangle &= \hbar m_{s1} |m_{s1} m_{s2}\rangle \\ S_{z2} |m_{s1} m_{s2}\rangle &= \hbar m_{s2} |m_{s1} m_{s2}\rangle\end{aligned}\quad (3.173)$$

con $m_{s1} = \pm 1/2$, $m_{s2} = \pm 1/2$. Dunque la base degli autostati dei singoli spin è data dai quattro vettori

$$\{|m_{s1} m_{s2}\rangle\} = \{|\frac{1}{2}, \frac{1}{2}\rangle, |\frac{1}{2}, -\frac{1}{2}\rangle, |-\frac{1}{2}, \frac{1}{2}\rangle, |-\frac{1}{2}, -\frac{1}{2}\rangle\}, \quad (3.174)$$

che possiamo anche riscrivere nella solita notazione con le frecce

$$\{|m_{s1} m_{s2}\rangle\} = \{|\uparrow, \uparrow\rangle, |\uparrow, \downarrow\rangle, |\downarrow, \uparrow\rangle, |\downarrow, \downarrow\rangle\}. \quad (3.175)$$

Ora vogliamo usare questi vettori per rappresentare gli autostati dello spin totale. Per prima cosa osserviamo che

$$J_z |m_{s1} m_{s2}\rangle = (S_{z1} + S_{z2}) |m_{s1} m_{s2}\rangle = \hbar(m_{s1} + m_{s2}) |m_{s1} m_{s2}\rangle, \quad (3.176)$$

da cui si deduce che gli autovalori possibili della componente z dello spin totale sono solo questi

$$m = 1, 0, -1 \quad (3.177)$$

dove il valore 1 si ottiene per lo stato $|\uparrow, \uparrow\rangle$, il valore -1 si ottiene per lo stato $|\downarrow, \downarrow\rangle$ e il valore 0 per combinazioni lineari degli altri due. A quali autovalori di J^2 corrispondono questi autovalori di J_z ? Per saperlo costruiamo la matrice che rappresenta J^2 nella base assegnata sopra. Per fare i calcoli conviene riscrivere J^2 in questo modo

$$\begin{aligned} J^2 &= S_1^2 + S_2^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2 \\ &= S_1^2 + S_2^2 + 2S_{z1}S_{z2} + 2(S_{x1}S_{x2} + S_{y1}S_{y2}) \\ &= S_1^2 + S_2^2 + 2S_{z1}S_{z2} + S_{+,1}S_{-,2} + S_{-,1}S_{+,2}, \end{aligned} \quad (3.178)$$

dove al solito $S_{\pm} = S_x \pm iS_y$. In questo modo possiamo scrivere

$$J^2|\uparrow, \uparrow\rangle = \hbar^2 \left(\frac{3}{4} + \frac{3}{4} + \frac{1}{2} \right) |\uparrow, \uparrow\rangle = 2\hbar^2 |\uparrow, \uparrow\rangle \quad (3.179)$$

$$\begin{aligned} J^2|\uparrow, \downarrow\rangle &= \hbar^2 \left(\frac{3}{4} + \frac{3}{4} - \frac{1}{2} \right) |\uparrow, \downarrow\rangle + \hbar^2 |\downarrow, \uparrow\rangle \\ &= \hbar^2 (|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle) \end{aligned} \quad (3.180)$$

$$\begin{aligned} J^2|\downarrow, \uparrow\rangle &= \hbar^2 \left(\frac{3}{4} + \frac{3}{4} - \frac{1}{2} \right) |\downarrow, \uparrow\rangle + \hbar^2 |\uparrow, \downarrow\rangle \\ &= \hbar^2 (|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle) \end{aligned} \quad (3.181)$$

$$J^2|\downarrow, \downarrow\rangle = \hbar^2 \left(\frac{3}{4} + \frac{3}{4} + \frac{1}{2} \right) |\downarrow, \downarrow\rangle = 2\hbar^2 |\downarrow, \downarrow\rangle \quad (3.182)$$

e la matrice che rappresenta J^2 diventa

$$J^2 = \hbar^2 \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad (3.183)$$

Ora la diagonalizziamo per trovare la base dei suoi autostati. È sufficiente diagonalizzare la sottomatrice

$$M = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (3.184)$$

la cui equazione caratteristica $\det(M - \lambda \mathbb{1}) = 0$ è

$$(1 - \lambda)^2 - 1 = 0 \quad (3.185)$$

che ha soluzioni $\lambda = 0, 2$ ed è facile verificare che i corrispondenti autostati sono $(1/\sqrt{2})(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle)$ e $(1/\sqrt{2})(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle)$. Conviene anche riscrivere l'autovalore nella forma $\lambda = j(j+1)$ in modo che i due autovalori corrispondono rispettivamente a $j = 0$ e $j = 1$.

In conclusione la base di autostati $\{|j, m\rangle\}$ delle osservabili J^2 e J_z , scritti come combinazioni lineari degli stati $|m_{s1}m_{s2}\rangle$, è formata dai quattro vettori

$$\begin{aligned} |1, 1\rangle &= |\uparrow, \uparrow\rangle \\ |1, 0\rangle &= \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle) \\ |1, -1\rangle &= |\downarrow, \downarrow\rangle \\ |0, 0\rangle &= \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle). \end{aligned} \quad (3.186)$$

I primi tre stati formano il cosiddetto *tripletto* di spin ($j = 1$), mentre il quarto è detto *singoletto* ($j = 0$). In forma più compatta possiamo scrivere

$$|jm\rangle = \sum_{m_{s1}, m_{s2}} (s_1 m_{s1} s_2 m_{s2} | jm) |m_{s1} m_{s2}\rangle \quad (3.187)$$

dove i simboli $(s_1 m_{s1} s_2 m_{s2} | jm)$ rappresentano i coefficienti numerici

$$\begin{aligned} \left(\frac{1}{2} \frac{1}{2} \frac{1}{2} \frac{1}{2} | 11\right) &= 1 & \left(\frac{1}{2} \frac{1}{2} \frac{1}{2} -\frac{1}{2} | 1-1\right) &= 1 & \left(\frac{1}{2} \frac{1}{2} \frac{1}{2} -\frac{1}{2} | 10\right) &= \frac{1}{\sqrt{2}} \\ \left(\frac{1}{2} \frac{1}{2} \frac{1}{2} -\frac{1}{2} | 00\right) &= \frac{1}{\sqrt{2}} & \left(\frac{1}{2} -\frac{1}{2} \frac{1}{2} \frac{1}{2} | 10\right) &= \frac{1}{\sqrt{2}} & \left(\frac{1}{2} -\frac{1}{2} \frac{1}{2} \frac{1}{2} | 00\right) &= -\frac{1}{\sqrt{2}} \end{aligned}$$

e tutte le altre combinazioni di indici danno coefficienti nulli.

Quanto abbiamo fatto per lo spin $1/2$ può essere generalizzato all'addizione di due momenti angolari qualsiasi

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2. \quad (3.188)$$

Volendo scrivere gli autostati di J^2 , J_z , J_1^2 , J_2^2 in termini degli autostati di J_1^2 , J_2^2 , J_{z1} e J_{z2} , si può utilizzare l'espressione

$$|j_1 j_2 j m\rangle = \sum_{m_1, m_2} (j_1 m_1 j_2 m_2 | jm) |j_1 j_2 m_1 m_2\rangle \quad (3.189)$$

dove i numeri $(j_1 m_1 j_2 m_2 | jm)$ sono detti *coefficienti di Clebsch-Gordan* e possono essere intesi come gli elementi di matrice che connettono le due basi. Quelli scritti sopra corrispondono al caso particolare $j_1 = j_2 = 1/2$. I coefficienti di Clebsch-Gordan sono noti e si trovano tabulati in molti testi. Una loro proprietà fondamentale è che essi non dipendono da eventuali altri numeri quantici (autovalori di altre osservabili) che possono caratterizzare il sistema fisico, ma solo dagli autovalori del momento angolare (orbitale o di spin). Nella definizione dei coefficienti di Clebsch-Gordan c'è una fase arbitraria che viene fissata convenzionalmente in modo che il coefficiente $(j_1 m_1 j_2 m_2 | jj)$ sia sempre reale e non negativo. Nel calcolo dei coefficienti si usano regole di ricorrenza che seguono dalle proprietà algebriche del momento angolare. Un risultato cruciale è la seguente regola di selezione: i coefficienti $(j_1 m_1 j_2 m_2 | jm)$ sono non nulli solo se si verificano simultaneamente le due condizioni

$$m = m_1 + m_2 \quad (3.190)$$

$$|j_1 - j_2| \leq j \leq j_1 + j_2. \quad (3.191)$$

Il caso $j_1 = j_2 = 1/2$ l'abbiamo già visto prima. Invece per un elettrone nell'atomo d'idrogeno conviene scrivere $\mathbf{J} = \mathbf{L} + \mathbf{S}$, dove $\mathbf{L}$ è il momento angolare orbitale e $\mathbf{S}$ è lo spin, con $s = 1/2$. Per la regola precedente, i coefficienti $(s m_s l m_l | jm)$ sono non nulli solo se $j = l \pm 1/2$.

Si faccia attenzione al fatto che la notazione usata per i coefficienti di Clebsch-Gordan può variare da testo a testo. Ad esempio, si usa spesso la scrittura

$$C_{m_1 m_2 m}^{j_1 j_2 j} \quad (3.192)$$

al posto di

$$(j_1 m_1 j_2 m_2 | jm), \quad (3.193)$$

oppure anche

$$\langle j_1 j_2; m_1 m_2 | j_1 j_2; jm \rangle, \quad (3.194)$$

ma sono sempre gli stessi numeri. Per altre proprietà dell'addizione di momenti angolari suggeriamo di leggere ad esempio il capitolo 3.8 del Sakurai⁸.

⁸Oppure il capitolo 4 del Griffiths, dove si trova anche il seguente commento: "Se pensate che questo

assomigli a una sorta di numerologia mistica, vi capisco. Non useremo le tabelle di Clebsch-Gordan nel resto del libro, ma ho voluto che voi vedeste dove si collocano nello schema teorico, nel caso vi capiti di incontrarle in seguito. In senso matematico, abbiamo a che fare con un'applicazione della teoria dei gruppi: in realtà ciò di cui stiamo parlando è la decomposizione del prodotto diretto di due rappresentazioni irriducibili del gruppo delle rotazioni in una somma diretta di rappresentazioni irriducibili (potete dirla in questo modo, se volete impressionare gli amici)".

Capitolo 4

Metodi approssimati

4.1 Soluzioni esatte o approssimate

La teoria che abbiamo introdotto nei primi due capitoli si basa su pochi principi fisici (il principio di sovrapposizione applicato agli stati di un sistema, il concetto di misura di una osservabile e l'evoluzione temporale degli stati) ed ha la pretesa di avere validità generale. La teoria è formulata in termini di equazioni che permettono di ottenere soluzioni anche analitiche per gli stati di un sistema e per la loro evoluzione nel tempo. Abbiamo già visto esempi di soluzioni analitiche per gli stati stazionari dell'oscillatore armonico unidimensionale, per l'elettrone nell'atomo d'idrogeno, per l'evoluzione degli stati di spin in un campo magnetico uniforme. In realtà, già nel caso della dinamica di una singola particella, i casi in cui si possono determinare soluzioni analitiche non sono poi molti e, se le particelle sono più di una, la ricerca di soluzioni analitiche si fa rapidamente proibitiva. Avere a disposizione soluzioni analitiche, benché poche, è comunque molto importante, perché esse rispondono alla nostra esigenza di rendere semplici e comprensibili le leggi di natura e, inoltre, perché costituiscono una sorta di punti di appoggio per l'esplorazione di soluzioni incognite per sistemi più complessi.

Un approccio alternativo, facilitato dalla disponibilità di *computer* sempre più potenti, consiste nel cercare soluzioni delle stesse equazioni ma di tipo numerico. Si tratta di formulare il problema tramite rappresentazioni degli stati e degli operatori compatibili con il linguaggio proprio del calcolo numerico. Il computer è in grado di svolgere il calcolo differenziale e integrale necessario per arrivare a soluzioni esatte, a meno degli errori impliciti nel calcolo (arrotondamenti, fluttuazioni statistiche, troncamenti di serie numeriche, ecc..) che si spera di poter ridurre a piacere, avendo tempo e risorse sufficienti per farlo. Negli ultimi decenni sono state sviluppate tecniche assai potenti per risolvere in questo modo problemi anche molto complessi. I progressi in questa direzione sono molto rapidi e permettono, ad esempio, di applicare con successo la meccanica quantistica alla chimica, ad esempio, o allo studio di fasi liquide o solide della materia. Tuttavia, ad un certo livello di complessità la soluzione esatta (*da principi primi*, come si dice in gergo, oppure *ab initio*) diventa prima o poi impraticabile. Inoltre, anche in sistemi relativamente semplici, avere a disposizione la soluzione numerica non implica necessariamente avere una comprensione fisica adeguata: i numeri generati dal computer sono equiparabili agli esiti di esperimenti, salvo che, nel caso di simulazioni numeriche, si ha di solito maggiore flessibilità nel variare i parametri del sistema e vederne le conseguenze. I numeri da soli non bastano; per “capire” cosa avviene, quali sono i processi fisici rilevanti e come si comporta la natura, occorre anche avere a disposizione un qualche punto di riferimento concettuale che ci guidi nell'interpretazione dei risultati numerici.

Un ruolo cruciale nello sviluppo e nell'applicazione di una qualsiasi teoria fisica è giocato dalle approssimazioni. Ogni modello fisico può essere visto come un'approssimazione della realtà. Dato un sistema da studiare, il fisico, fin dai tempi di Galileo, è abituato a scartare i dettagli che ritiene essere irrilevanti per concentrarsi sulle caratteristiche che invece ritiene essenziali; per queste ultime usa le teorie a sua disposizione al fine di formulare la descrizione del sistema in termini di equazioni matematiche, possibilmente risolubili in modo analitico o numerico. Ne esce un modello del sistema, valido entro i limiti di applicabilità delle teorie utilizzate ed entro i limiti delle approssimazioni compiute. Se le approssimazioni iniziali sono troppo drastiche e, per eccesso di semplificazione, scartano elementi del sistema che sono essenziali anche per una sua descrizione minimale, chiaramente la procedura fallisce (sulla tendenza dei fisici a lavorare con modelli semplificati girano varie barzellette, come quella del fisico a cui viene chiesto come si possa aumentare la produzione di latte delle mucche e che torna dopo alcuni giorni con la soluzione esatta per mucche sferiche). Ma se, al contrario, il modello include quasi tutte le caratteristiche essenziali, a meno di piccole correzioni il cui effetto è in qualche misura controllabile, allora il metodo approssimato può diventare veramente efficace. In questo capitolo vediamo alcuni esempi di metodi approssimati della meccanica quantistica.

4.2 Perturbazioni indipendenti dal tempo

4.2.1 Spettro non degenere

Supponiamo che un sistema sia descritto da un hamiltoniano indipendente dal tempo separabile in due parti

$$H = H_0 + H_1 \quad (4.1)$$

dove H_0 è un hamiltoniano di cui si conoscono autostati e autovalori in modo esatto,

$$H_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle, \quad (4.2)$$

e H_1 sia invece un termine piccolo, in qualche senso da precisare, rispetto ad H_0 . Diciamo che H_0 è l'hamiltoniano *imperturbato*, mentre H_1 è la *perturbazione*. Per rendere evidente che si tratta di un termine piccolo possiamo anche scrivere H_1 come λV , dove λ è un parametro numerico, reale e continuo, che fissa l'intensità della perturbazione e V è un operatore hermitiano (non necessariamente un potenziale scalare) che rappresenta il tipo specifico di perturbazione:

$$H = H_0 + \lambda V. \quad (4.3)$$

Ora facciamo l'ipotesi che gli autostati e gli autovalori dell'hamiltoniano totale H differiscano di poco da quelli di H_0 , in modo da poterli ottenere tramite uno *sviluppo perturbativo* di questo tipo

$$|n\rangle = |n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \dots \quad (4.4)$$

con

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \quad (4.5)$$

Meglio sottolineare che questa è un'ipotesi forte. Non è affatto detto che si possano scrivere sviluppi in serie di questo tipo. Ma se il sistema ammette uno sviluppo perturbativo, allora questo diventa uno strumento efficace ed elegante che spesso dà soluzioni vicine alle soluzioni esatte anche quando ci si limita ai primi termini dello sviluppo.

Vediamo dunque come si possono calcolare le correzioni perturbative, cominciando dal caso di spettri non degeneri. Sia dunque $\{|n^{(0)}\rangle\}$ la base degli autostati di un generico hamiltoniano imperturbato H_0 e siano $E_n^{(0)}$ le corrispondenti energie, tutte diverse tra loro.

Sia $H = H_0 + \lambda V$ l'hamiltoniano perturbato di cui si vogliono ottenere autostati e autovalori con uno sviluppo in λ . L'equazione da risolvere per trovare gli stati stazionari di H è l'equazione di Schrödinger

$$(H_0 + \lambda V)|n\rangle = E_n|n\rangle \quad (4.6)$$

dove gli stati $|n\rangle$ e le energia E_n sono le nostre incognite. Per i vettori di stato conviene rinunciare alla solita condizione di normalizzazione $\langle n|n\rangle = 1$ in favore della nuova condizione

$$\langle n^{(0)}|n\rangle = 1 \quad (4.7)$$

tenendo vera la normalizzazione degli stati imperturbati $\langle n^{(0)}|n^{(0)}\rangle = 1$. Il fatto che i vettori incogniti $|n\rangle$ non siano normalizzati a 1 non è un problema e non dà noia ai calcoli, dato che la normalizzazione non modifica l'equazione da risolvere e può sempre essere cambiata a posteriori. Invece, la nostra scelta di fissare a 1 il prodotto interno tra ciascun $|n\rangle$ e il corrispondente stato imperturbato, ci torna utile. Infatti, se prendiamo lo sviluppo (4.4) e moltiplichiamo a sinistra per $\langle n^{(0)}|$ otteniamo

$$\langle n^{(0)}|n\rangle = \langle n^{(0)}|n^{(0)}\rangle + \lambda \langle n^{(0)}|n^{(1)}\rangle + \lambda^2 \langle n^{(0)}|n^{(2)}\rangle + \dots ; \quad (4.8)$$

ma se ora vogliamo che $\langle n^{(0)}|n\rangle = \langle n^{(0)}|n^{(0)}\rangle = 1$, ne segue che tutti i coefficienti di λ , λ^2 , ..., devono annullarsi, e dunque

$$\langle n^{(0)}|n^{(i)}\rangle = 0 \quad \text{per } i = 1, 2, 3, \dots \quad (4.9)$$

Dunque la nostra scelta per la normalizzazione implica che tutte le correzioni perturbative allo stato n -esimo sono ortogonali allo stato imperturbato corrispondente¹. Fatto questo, torniamo all'equazione di Schrödinger (4.6), inseriamo lo sviluppo (4.4) e eguagliamo i termini allo stesso ordine in λ .

All'ordine zero (termini senza λ) ritroviamo lo spettro dell'hamiltoniano imperturbato (4.2). Al primo ordine (termini proporzionali a λ) invece otteniamo

$$H_0|n^{(1)}\rangle + V|n^{(0)}\rangle = E_n^{(1)}|n^{(0)}\rangle + E_n^{(0)}|n^{(1)}\rangle . \quad (4.10)$$

Possiamo moltiplicare a sinistra tutti i termini per $\langle n^{(0)}|$

$$\langle n^{(0)}|H_0|n^{(1)}\rangle + \langle n^{(0)}|V|n^{(0)}\rangle = E_n^{(1)}\langle n^{(0)}|n^{(0)}\rangle + E_n^{(0)}\langle n^{(0)}|n^{(1)}\rangle . \quad (4.11)$$

Il primo termine vale $E_n^{(0)}\langle n^{(0)}|n^{(1)}\rangle$ e si elimina con l'ultimo a destra, ma sarebbero comunque entrambi nulli per la regola (4.9). Dato che gli autostati imperturbati hanno norma 1, i due termini che rimangono danno

$$E_n^{(1)} = \langle n^{(0)}|V|n^{(0)}\rangle = V_{nn} , \quad (4.12)$$

dove V_{nn} sono gli elementi diagonali della matrice V nella base imperturbata. L'energia dell'autostato n , al primo ordine nello sviluppo in λ , è dunque

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} = E_n^{(0)} + \lambda V_{nn} \quad (4.13)$$

¹I vettori dello spettro imperturbato formano una base completa per lo spazio di Hilbert del sistema. Dunque anche il vettore $|n\rangle$, autostato di H , può essere scritto come una loro combinazione lineare. La regola che abbiamo scelto, secondo cui $\langle n^{(0)}|n\rangle = 1$, equivale a fissare uno dei coefficienti di tale decomposizione, quello corrispondente al vettore $|n^{(0)}\rangle$ di cui si calcola la correzione. Dato che i risultati fisici non dipendono dalla norma dei vettori, uno dei coefficienti può sempre essere fissato a piacere. Quello che conta è il peso relativo degli altri termini rispetto a quello liberamente fissato, dato che quest'ultimo può sempre essere raccolto a fattore comune e inglobato nella norma. Così, senza perdere di generalità, ci ritroveremo con espressioni più semplici da manipolare e calcolare.

e possiamo definire una correzione all'energia al primo ordine

$$\Delta_n^{(1)} = E_n - E_n^{(0)} = \lambda E_n^{(1)} = \lambda V_{nn} . \quad (4.14)$$

Ovviamente gli elementi di matrice V_{nn} sono noti in partenza, dato che sono noti sia H_0 , con i suoi autostati, che l'operatore V . La correzione all'energia al primo ordine è dunque calcolabile in modo semplice.

All'ordine k generico (termini in λ^k , con k intero e positivo) lo sviluppo dell'equazione di Schrödinger (4.6) dà

$$H_0|n^{(k)}\rangle + V|n^{(k-1)}\rangle = E_n^{(k)}|n^{(0)}\rangle + E_n^{(k-1)}|n^{(1)}\rangle + \dots + E_n^{(0)}|n^{(k)}\rangle . \quad (4.15)$$

Moltiplicando tutti i termini per $\langle n^{(0)}|$ e tenendo conto della solita regola (4.9), l'espressione precedente si riduce a

$$\langle n^{(0)}|V|n^{(k-1)}\rangle = E_n^{(k)} . \quad (4.16)$$

Quindi, per conoscere la correzione all'energia di ogni autostato all'ordine k basta conoscere l'autostato all'ordine $k-1$. Il caso $k=1$ l'abbiamo appena visto sopra nella (4.13). L'energia al secondo ordine ($k=2$) la calcoleremo tra un attimo, ma ci serve calcolare prima la correzione agli autostati al primo ordine. E così via, per ricorrenza. Tra l'altro, la relazione precedente può essere riscritta in forma diversa, moltiplicando entrambi i membri per λ^k e sommando sui valori di k da 1 a ∞ ,

$$\begin{aligned} \sum_{k=1}^{\infty} \lambda^k E_n^{(k)} &= \sum_{k=1}^{\infty} \lambda^k \langle n^{(0)}|V|n^{(k-1)}\rangle \\ &= \langle n^{(0)}|\lambda V \left(\sum_{k=1}^{\infty} \lambda^{k-1} |n^{(k-1)}\rangle \right) , \end{aligned} \quad (4.17)$$

e poi possiamo usare i due sviluppi (4.4) e (4.5) per riscriverla in questo modo:

$$E_n - E_n^{(0)} = \langle n^{(0)}|\lambda V|n\rangle \quad (4.18)$$

ovvero

$$E_n = E_n^{(0)} + \lambda \langle n^{(0)}|V|n\rangle . \quad (4.19)$$

Ora procediamo calcolando la correzione al primo ordine per gli autostati. Per farlo riprendiamo la (4.15), portiamo a destra l'elemento di matrice di V e poi moltiplichiamo tutti i termini per il duale $\langle m^{(0)}|$ di un autostato generico della base imperturbata, con $m \neq n$:

$$\begin{aligned} \langle m^{(0)}|H_0|n^{(k)}\rangle &= -\langle m^{(0)}|V|n^{(k-1)}\rangle + E_n^{(k)}\langle m^{(0)}|n^{(0)}\rangle + E_n^{(k-1)}\langle m^{(0)}|n^{(1)}\rangle \\ &+ \dots + E_n^{(1)}\langle m^{(0)}|n^{(k-1)}\rangle + E_n^{(0)}\langle m^{(0)}|n^{(k)}\rangle . \end{aligned} \quad (4.20)$$

Il secondo termine a destra dell'uguale è nullo per l'ortogonalità della base. Il termine a sinistra può essere riscritto come $E_m^{(0)}\langle m^{(0)}|n^{(k)}\rangle$ e possiamo raccogliere questo prodotto interno tra questo termine e l'ultimo a destra. Otteniamo

$$\begin{aligned} \langle m^{(0)}|n^{(k)}\rangle &= \frac{1}{E_n^{(0)} - E_m^{(0)}} [\langle m^{(0)}|V|n^{(k-1)}\rangle - E_n^{(k-1)}\langle m^{(0)}|n^{(1)}\rangle \\ &- \dots - E_n^{(2)}\langle m^{(0)}|n^{(k-2)}\rangle - E_n^{(1)}\langle m^{(0)}|n^{(k-1)}\rangle] , \end{aligned} \quad (4.21)$$

ma questi non sono altro che i coefficienti della decomposizione del vettore incognito $|n^{(k)}\rangle$ sulla base degli autostati imperturbati. Dall'espressione che abbiamo ottenuto si vede

che, per calcolare la correzione perturbativa $|n^{(k)}\rangle$ basta conoscere le correzioni agli ordini precedenti, fino a $k-1$.

Vediamo subito cosa succede al primo ordine ($k=1$). La relazione precedente dà

$$\langle m^{(0)} | n^{(1)} \rangle = \frac{\langle m^{(0)} | V | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} = \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} . \quad (4.22)$$

Possiamo scrivere $|n^{(1)}\rangle$ come combinazione degli stati della base imperturbata

$$|n^{(1)}\rangle = \sum_m |m^{(0)}\rangle \langle m^{(0)} | n^{(1)} \rangle . \quad (4.23)$$

Il termine con $m=n$ si annulla per la regola (4.9) e gli altri li abbiamo appena calcolati. Dunque

$$|n^{(1)}\rangle = \sum_{m \neq n} |m^{(0)}\rangle \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \quad (4.24)$$

e l'espansione dell'autovettore n -esimo troncata al primo ordine diventa

$$|n\rangle = |n^{(0)}\rangle + \lambda \sum_{m \neq n} |m^{(0)}\rangle \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \quad (4.25)$$

a cui corrisponde l'energia

$$E_n = E_n^{(0)} + \lambda V_{nn} . \quad (4.26)$$

La correzione al secondo ordine è un po' più complicata da calcolare, ma non troppo. Per l'energia basta prendere la (4.16) con $k=2$,

$$E_n^{(2)} = \langle n^{(0)} | V | n^{(1)} \rangle , \quad (4.27)$$

e usare il risultato già ottenuto per la correzione (4.24) al primo ordine degli autostati:

$$E_n^{(2)} = \langle n^{(0)} | V \left(\sum_{m \neq n} |m^{(0)}\rangle \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \right) , \quad (4.28)$$

ovvero

$$E_n^{(2)} = \sum_{m \neq n} \frac{V_{nm} V_{mn}}{E_n^{(0)} - E_m^{(0)}} = \sum_{m \neq n} \frac{|V_{nm}|^2}{E_n^{(0)} - E_m^{(0)}} , \quad (4.29)$$

dove abbiamo usato il fatto che V è un operatore hermitiano. Dunque, l'energia di ciascun autostato sviluppata al secondo ordine diventa

$$E_n = E_n^{(0)} + \lambda V_{nn} + \lambda^2 \sum_{m \neq n} \frac{|V_{nm}|^2}{E_n^{(0)} - E_m^{(0)}} \quad (4.30)$$

e la correzione rispetto allo stato imperturbato è

$$\Delta_n^{(2)} = \lambda V_{nn} + \lambda^2 \sum_{m \neq n} \frac{|V_{nm}|^2}{E_n^{(0)} - E_m^{(0)}} . \quad (4.31)$$

Per quanto riguarda gli autostati basta usare l'espressione (4.21) con $k=2$:

$$\langle m^{(0)} | n^{(2)} \rangle = \frac{1}{E_n^{(0)} - E_m^{(0)}} [\langle m^{(0)} | V | n^{(1)} \rangle - E_n^{(1)} \langle m^{(0)} | n^{(1)} \rangle] , \quad (4.32)$$

dove $E_n^{(1)}$ e $|n^{(1)}\rangle$ sono quelli già trovati al primo ordine. Dunque

$$\langle m^{(0)} | V | n^{(1)} \rangle = \langle m^{(0)} | V \left(\sum_{l \neq n} |l^{(0)}\rangle \frac{V_{ln}}{E_n^{(0)} - E_l^{(0)}} \right) = \sum_{l \neq n} \frac{V_{ml} V_{ln}}{E_n^{(0)} - E_l^{(0)}}, \quad (4.33)$$

e

$$E_n^{(1)} \langle m^{(0)} | n^{(1)} \rangle = V_{nn} \langle m^{(0)} | \left(\sum_{l \neq n} |l^{(0)}\rangle \frac{V_{ln}}{E_n^{(0)} - E_l^{(0)}} \right) = \frac{V_{nn} V_{mn}}{E_n^{(0)} - E_m^{(0)}}, \quad (4.34)$$

da cui

$$\langle m^{(0)} | n^{(2)} \rangle = \sum_{l \neq n} \frac{V_{ml} V_{ln}}{(E_n^{(0)} - E_l^{(0)})(E_n^{(0)} - E_m^{(0)})} - \frac{V_{nn} V_{mn}}{(E_n^{(0)} - E_m^{(0)})^2}. \quad (4.35)$$

Infine possiamo scrivere

$$|n^{(2)}\rangle = \sum_{m \neq n} |m^{(0)}\rangle \langle m^{(0)} | n^{(2)} \rangle, \quad (4.36)$$

ovvero

$$\boxed{\begin{aligned} |n^{(2)}\rangle &= \sum_{m \neq n} \sum_{l \neq n} |m^{(0)}\rangle \frac{V_{ml} V_{ln}}{(E_n^{(0)} - E_l^{(0)})(E_n^{(0)} - E_m^{(0)})} \\ &- \sum_{m \neq n} |m^{(0)}\rangle \frac{V_{nn} V_{mn}}{(E_n^{(0)} - E_m^{(0)})^2}. \end{aligned}}$$

E così abbiamo ottenuto esplicitamente sia l'energia che gli autostati al secondo ordine.

La procedura può essere estesa agli ordini superiori, ma di solito bastano le correzioni al primo e al secondo ordine per ottenere risultati buoni, sia in termini di accuratezza che di comprensione della fisica dei sistemi esaminati, sempre che questi ammettano sviluppi perturbativi.

Prima di passare agli spettri degeneri, vale la pena sottolineare alcune conseguenze di quanto appena visto. La prima è che, come si vede nella (4.24), già al primo ordine lo stato imperturbato $|n^{(0)}\rangle$ viene corretto aggiungendo pezzi che contengono gli stati $|m^{(0)}\rangle$ con $m \neq n$; dunque la perturbazione “mescola” i livelli, almeno quelli per i quali l'elemento di matrice V_{nm} è diverso da zero. La seconda è che, quando la perturbazione mescola due soli livelli, la correzione all'energia prodotta dagli elementi di matrice $V_{12} = V_{21} \neq 0$ è sempre tale da aumentare il *gap* tra di loro, indipendentemente dal segno di V_{12} , come si vede nella (4.30): il livello di energia più bassa scende, quello più in alto sale. La terza considerazione è che la correzione al secondo ordine all'energia dello stato fondamentale, cioè quello che ha l'energia più bassa di tutti, è sempre negativa perché in quel caso tutti i termini nella (4.30) sono negativi.

4.2.2 Esempio: perturbazione all'oscillatore armonico

Facciamo qualche esempio di applicazioni della teoria perturbativa, cominciando da una perturbazione lineare all'oscillatore armonico. Come vedremo, in questo caso il problema è risolubile anche in modo esatto, ma proprio per questo può essere un buon punto di partenza. Dunque, sia $H_0 = p^2/(2m) + (1/2)m\omega^2 x^2$ l'hamiltoniano dell'oscillatore imperturbato, il cui spettro $E_n^{(0)} = \hbar\omega(n + 1/2)$ è già stato calcolato nel capitolo 2. Supponiamo che la particella sia soggetta anche ad un potenziale lineare nella posizione, $\lambda V(x) = \lambda \hbar\omega x/a_{\text{ho}}$, con λ piccolo² e con $a_{\text{ho}} = \sqrt{\hbar/m\omega}$. Per vedere l'effetto di questa perturbazione, almeno fino al secondo ordine in λ , basta prendere il risultato generale (4.30) e inserire gli elementi

²La scelta convenzionale di inserire $\hbar\omega/a_{\text{ho}}$ nella definizione di V fa sì che λ sia adimensionale. Nell'oscillatore armonico, le quantità $\hbar\omega$ e a_{ho} sono le scale naturali di energia e lunghezza.

di matrice di V . Questi si calcolano facilmente se si esprime x in termini degli operatori a scala a e $a^\dagger$, ricordandoci che $x = (a_{\text{ho}}/\sqrt{2})(a + a^\dagger)$. Troviamo

$$V_{nn} = \langle n^{(0)} | V | n^{(0)} \rangle = (\hbar\omega/\sqrt{2}) \langle n^{(0)} | a + a^\dagger | n^{(0)} \rangle = 0, \quad (4.37)$$

dato che gli operatori a scala, agendo sull' n -esimo stato producono lo stato con $n \pm 1$ che è ortogonale a quello di partenza. Invece gli elementi di matrice V_{mn} possono essere diversi da zero, ma solo per i due valori $m = n \pm 1$, dove valgono

$$\langle (n-1)^{(0)} | V | n^{(0)} \rangle = (\hbar\omega/\sqrt{2}) \langle (n-1)^{(0)} | a + a^\dagger | n^{(0)} \rangle = \sqrt{n/2} \hbar\omega \quad (4.38)$$

e

$$\langle (n+1)^{(0)} | V | n^{(0)} \rangle = (\hbar\omega/\sqrt{2}) \langle (n+1)^{(0)} | a + a^\dagger | n^{(0)} \rangle = \sqrt{(n+1)/2} \hbar\omega. \quad (4.39)$$

Ne segue che la correzione al primo ordine dell'energia dei livelli è nulla, mentre al secondo ordine si ottiene

$$\begin{aligned} E_n &= E_n^{(0)} + \lambda^2 \left[\frac{|V_{n-1,n}|^2}{E_n^{(0)} - E_{n-1}^{(0)}} + \frac{|V_{n+1,n}|^2}{E_n^{(0)} - E_{n+1}^{(0)}} \right] \\ &= E_n^{(0)} + \lambda^2 (\hbar\omega)^2 \left[\frac{n}{2\hbar\omega} - \frac{n+1}{2\hbar\omega} \right] \\ &= E_n^{(0)} - \lambda^2 \frac{\hbar\omega}{2}. \end{aligned} \quad (4.40)$$

Gli autostati invece cambiano già al primo ordine. Per vederlo prendiamo la (4.24) e usiamo i due elementi di matrice di prima:

$$\begin{aligned} |n^{(1)}\rangle &= |(n-1)^{(0)}\rangle \frac{V_{n-1,n}}{E_n^{(0)} - E_{n-1}^{(0)}} + |(n+1)^{(0)}\rangle \frac{V_{n+1,n}}{E_n^{(0)} - E_{n+1}^{(0)}} \\ &= \sqrt{\frac{n}{2}} |(n-1)^{(0)}\rangle - \sqrt{\frac{n+1}{2}} |(n+1)^{(0)}\rangle. \end{aligned} \quad (4.41)$$

Come dicevamo, di questo problema esiste anche la soluzione analitica esatta. Infatti basta riscrivere l'hamiltoniano in questo modo

$$\begin{aligned} H &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 + \lambda\hbar\omega \frac{x}{a_{\text{ho}}} \\ &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2 (x + \lambda a_{\text{ho}})^2 - \lambda^2 \frac{\hbar\omega}{2} \end{aligned} \quad (4.42)$$

e già si vede da qui che si tratta di un oscillatore armonico identico a prima, eccetto per uno spostamento verso il basso della scala di energia, di una quantità costante e proporzionale a λ^2 , e per una traslazione spaziale di una distanza λa_{ho} . L'abbassamento dell'energia coincide con il risultato perturbativo al secondo ordine, che era indipendente dall'indice n dei livelli. Invece la traslazione spaziale implica che il nuovo minimo del potenziale armonico si trova in $x = -\lambda a_{\text{ho}}$ e dunque i nuovi autovettori saranno uguali a prima, ma traslati a sinistra della stessa distanza. Per trovarli possiamo applicare l'operatore di traslazione spaziale $\tau(\Delta x) = \exp[(-i/\hbar)p\Delta x]$, definito nel capitolo 1, con $\Delta x = -\lambda a_{\text{ho}}$. Ricordiamo anche, invertendo le definizioni degli operatori a scala, si ha $p = [\hbar/(i a_{\text{ho}} \sqrt{2})](a - a^\dagger)$. Dunque abbiamo

$$\begin{aligned} |n\rangle &= \exp\left(\frac{i}{\hbar}p\lambda a_{\text{ho}}\right) |n^{(0)}\rangle \\ &= \exp\left(\frac{1}{\sqrt{2}}\lambda(a - a^\dagger)\right) |n^{(0)}\rangle \end{aligned} \quad (4.43)$$

e questa espressione contiene la correzione a tutti gli ordini. Se vogliamo solo la correzione al primo ordine, basta fermarsi al primo termine dello sviluppo dell'esponenziale:

$$|n\rangle = |n^{(0)}\rangle + \frac{1}{\sqrt{2}}\lambda(a - a^\dagger)|n^{(0)}\rangle \quad (4.44)$$

e la correzione allo stato imperturbato diventa la stessa calcolata più sopra tramite lo sviluppo perturbativo.

Un test analogo lo potremmo fare per una perturbazione quadratica del tipo $\lambda V = \lambda\hbar\omega(x/a_{\text{ho}})^2$, ma sarebbe davvero poco significativo, dato che in questo caso la soluzione esatta è estremamente diretta. Infatti l'hamiltoniano può essere riscritto così

$$\begin{aligned} H &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 + \lambda\hbar\omega \left(\frac{x}{a_{\text{ho}}}\right)^2 \\ &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2(1 + 2\lambda)x^2 \end{aligned} \quad (4.45)$$

e questo è un nuovo oscillatore armonico con frequenza $\sqrt{1 + 2\lambda}\omega$. Tutti i livelli energetici cambiano dello stesso fattore numerico. Se λ è piccolo si può sviluppare in serie la radice per ottenere le correzioni perturbative a tutti gli ordini. Per esercizio si potrebbe verificare che ciascun termine coincide con quello che si trova applicando le espressioni della teoria delle perturbazioni. Qui non lo facciamo. Piuttosto passiamo al caso della perturbazione cubica, che è più interessante.

Prendiamo il solito oscillatore e stavolta aggiungiamo una perturbazione nella forma $\lambda V = \lambda\hbar\omega(x/a_{\text{ho}})^3$. Con un termine del genere, a differenza del caso lineare e quadratico, il moto non è più armonico. Anche dal punto di vista classico la fisica cambia in modo significativo. Per il calcolo perturbativo conviene al solito scrivere V in termini degli operatori a scala

$$\begin{aligned} V &= \hbar\omega \left(\frac{x}{a_{\text{ho}}}\right)^3 = \frac{\hbar\omega}{\sqrt{2^3}}(a + a^\dagger)(a + a^\dagger)(a + a^\dagger) \\ &= \frac{\hbar\omega}{\sqrt{8}}(a + a^\dagger)(aa + aa^\dagger + a^\dagger a + a^\dagger a^\dagger) \\ &= \frac{\hbar\omega}{\sqrt{8}}(aaa + aaa^\dagger + aa^\dagger a + aa^\dagger a^\dagger + a^\dagger aa + a^\dagger aa^\dagger + a^\dagger a^\dagger a + a^\dagger a^\dagger a^\dagger) \end{aligned}$$

così si vede che la perturbazione ha elementi di matrice non nulli solo tra stati con $m = n \pm 1$ e $m = n \pm 3$. Abbiamo infatti

$$\begin{aligned} \langle(n-3)^{(0)}|V|n^{(0)}\rangle &= \frac{\hbar\omega}{\sqrt{8}}\langle(n-3)^{(0)}|aaa|n^{(0)}\rangle \\ &= \frac{\hbar\omega}{\sqrt{8}}\sqrt{n(n-1)(n-2)} \end{aligned} \quad (4.46)$$

$$\begin{aligned} \langle(n+3)^{(0)}|V|n^{(0)}\rangle &= \frac{\hbar\omega}{\sqrt{8}}\langle(n+3)^{(0)}|a^\dagger a^\dagger a^\dagger|n^{(0)}\rangle \\ &= \frac{\hbar\omega}{\sqrt{8}}\sqrt{(n+1)(n+2)(n+3)} \end{aligned} \quad (4.47)$$

$$\begin{aligned} \langle(n-1)^{(0)}|V|n^{(0)}\rangle &= \frac{\hbar\omega}{\sqrt{8}}\langle(n-1)^{(0)}|(aaa^\dagger + aa^\dagger a + a^\dagger aa)|n^{(0)}\rangle \\ &= \frac{\hbar\omega}{\sqrt{8}}3n\sqrt{n} \end{aligned} \quad (4.48)$$

$$\begin{aligned}
\langle (n+1)^{(0)} | V | n^{(0)} \rangle &= \frac{\hbar\omega}{\sqrt{8}} \langle (n+1)^{(0)} | (aa^\dagger a^\dagger + a^\dagger aa^\dagger + a^\dagger a^\dagger a) | n^{(0)} \rangle \\
&= \frac{\hbar\omega}{\sqrt{8}} 3(n+1)\sqrt{n+1}.
\end{aligned} \tag{4.49}$$

Dato che gli elementi di matrice diagonali sono nulli, la correzione all'energia al primo ordine è nulla. Invece al secondo ordine si trova

$$\begin{aligned}
E_n^{(2)} &= \sum_{m \neq n} \frac{|V_{nm}|^2}{E_n^{(0)} - E_m^{(0)}} \\
&= \frac{\hbar\omega}{8} \left[\frac{1}{3} n(n-1)(n-2) + 9n^3 - 9(n+1)^3 - \frac{1}{3} (n+1)(n+2)(n+3) \right] \\
&= -\frac{\hbar\omega}{8} (30n^2 + 30n + 11),
\end{aligned} \tag{4.50}$$

ovvero

$$E_n = E_n^{(0)} - \lambda^2 \frac{\hbar\omega}{8} (30n^2 + 30n + 11). \tag{4.51}$$

La correzione a tutti i livelli, compreso quello fondamentale ($n = 0$), è negativa e i livelli non sono più equispaziati. La separazione tra livelli vicini diventa

$$E_n - E_{n-1} = \hbar\omega - \lambda^2 \frac{\hbar\omega}{8} \{30[n^2 - (n-1)^2] + 30[n - (n-1)]\} \tag{4.52}$$

ovvero

$$E_n - E_{n-1} = \hbar\omega \left[1 - \frac{15}{2} n \lambda^2 \right]. \tag{4.53}$$

A differenza delle energie, gli autovettori vengono modificati già al primo ordine, come si può vedere inserendo gli elementi di matrice V_{mn} nell'espressione (4.24). Lasciamo il calcolo come esercizio.

4.2.3 Spettro degenere

Nelle espressioni che abbiamo ricavato finora per gli sviluppi perturbativi delle energie e degli autostati compaiono termini del tipo $\lambda V_{mn}/(E_n^{(0)} - E_m^{(0)})$ oppure $\lambda^2 |V_{mn}|^2/(E_n^{(0)} - E_m^{(0)})$, o simili. Per questo è naturale aspettarsi che la serie perturbativa converga rapidamente se gli elementi di matrice di V sono piccoli rispetto alla spaziatura tra i livelli energetici. Per la stessa ragione appare chiaro che lo sviluppo non può funzionare se gli elementi di matrice V_{mn} sono diversi da zero per stati che hanno la stessa energia, $E_n^{(0)} = E_m^{(0)}$. Fin qui avevamo supposto che lo spettro fosse non-degenere e quindi il problema non si poneva. Ma se fosse degenere?

Supponiamo che un generico autovalore $E_D^{(0)}$ dell'hamiltoniano H_0 sia g -volte degenere, cioè esistano g autostati ortonormali $|n_1^{(0)}\rangle, |n_2^{(0)}\rangle, \dots, |n_g^{(0)}\rangle$ per cui vale:

$$H_0 |n_i^{(0)}\rangle = E_D^{(0)} |n_i^{(0)}\rangle, \quad i = 1, 2, \dots, g. \tag{4.54}$$

In questa situazione, gli autovalori di H_0 non bastano per etichettare i vettori della base e si deve ricorrere ad altre osservabili compatibili con H_0 , in modo da usare i loro autovalori come etichette e i loro autovettori come base, anche per i sottospazi degeneri di H_0 . Il simbolo $n_i^{(0)}$ nell'espressione precedente esprime sinteticamente l'insieme degli autovalori delle osservabili compatibili con H_0 sufficienti a individuare univocamente tutti gli stati della base.

Preso uno di questi vettori di base, ne vogliamo calcolare le correzioni perturbative, quando l'hamiltoniano del sistema ha la forma $H = H_0 + \lambda V$ con λ piccolo. Supponiamo di calcolare le correzioni ad uno degli stati nel sottospazio degenerare $\mathcal{D}$. Conviene dapprima scrivere la matrice che rappresenta V in $\mathcal{D}$. Non c'è ragione, a priori, che la matrice sia diagonale. La scelta dell'insieme massimale di osservabili compatibili per definire la base imperturbata, infatti, non è necessariamente condizionata dal tipo di perturbazione che poi si vuole studiare. Ma se la matrice che rappresenta V non è diagonale allora lo sviluppo perturbativo non funziona, dato che gli elementi fuori diagonale non nulli farebbero comparire dei termini divergenti nel calcolo delle correzioni, dato che gli stati connessi da V all'interno del sottospazio degenerare $\mathcal{D}$ hanno tutti la stessa energia. Nulla ci vieta, tuttavia, di cercare una nuova base nello stesso sottospazio tale che V sia diagonale. Una volta trovata, i vettori della nuova base si comporteranno bene nello sviluppo perturbativo. È un semplice problema di algebra lineare. Si tratta di calcolare nuovi vettori $|\nu_i^{(0)}\rangle$ nella forma

$$|\nu_i^{(0)}\rangle = \sum_{j=1}^g |n_j^{(0)}\rangle \langle n_j^{(0)} | \nu_i^{(0)} \rangle \quad (4.55)$$

tali che

$$\langle \nu_i^{(0)} | V | \nu_j^{(0)} \rangle = v_i \delta_{ij}, \quad (4.56)$$

con v_i autovalori reali. Supponiamo di averli trovati e vediamo cosa succede all'equazione di Schrödinger

$$(H_0 + \lambda V) |\nu_i\rangle = E_i |\nu_i\rangle, \quad (4.57)$$

se usiamo gli sviluppi

$$|\nu_i\rangle = |\nu_i^{(0)}\rangle + \lambda |\nu_i^{(1)}\rangle + \lambda^2 |\nu_i^{(2)}\rangle + \dots, \quad (4.58)$$

$$E_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \dots \quad (4.59)$$

Inserendoli nell'equazione di Schrödinger e tenendo solo i termini al primo ordine, troviamo

$$H_0 |\nu_i^{(1)}\rangle + V |\nu_i^{(0)}\rangle = E_D^{(0)} |\nu_i^{(1)}\rangle + E_i^{(1)} |\nu_i^{(0)}\rangle. \quad (4.60)$$

Moltiplichiamo tutti i termini per $\langle \nu_j^{(0)} |$

$$\langle \nu_j^{(0)} | H_0 | \nu_i^{(1)} \rangle + \langle \nu_j^{(0)} | V | \nu_i^{(0)} \rangle = E_D^{(0)} \langle \nu_j^{(0)} | \nu_i^{(1)} \rangle + E_i^{(1)} \langle \nu_j^{(0)} | \nu_i^{(0)} \rangle. \quad (4.61)$$

Nel primo termine facciamo agire l'hamiltoniano sul *bra*, che è un suo autostato. Nel secondo termine usiamo il fatto che V è diagonale in questa base. Nell'ultimo termine usiamo l'ortonormalità. Dunque

$$E_D^{(0)} \langle \nu_j^{(0)} | \nu_i^{(1)} \rangle + v_i \delta_{ij} = E_D^{(0)} \langle \nu_j^{(0)} | \nu_i^{(1)} \rangle + E_i^{(1)} \delta_{ij}, \quad (4.62)$$

da cui

$$\boxed{E_i^{(1)} = v_i = \langle \nu_i^{(0)} | V | \nu_i^{(0)} \rangle} \quad (4.63)$$

In conclusione, la correzione al primo ordine all'energia dei livelli degeneri è fissata dagli autovalori dell'operatore V ! Inoltre, se i v_i sono diversi, la perturbazione rimuove la degenerazione (nel gergo dei fisici, produce uno *splitting* dei livelli).

Gli autostati di H_0 che stanno fuori dal sottospazio $\mathcal{D}$ entrano in gioco nel calcolo della correzione al secondo ordine, ma non al primo. Calcoliamo prima la correzione agli autostati

usando la stessa espressione (4.24) che avevamo ricavato per gli spettri non degeneri. In particolare applicando quell'espressione per calcolare la correzione al primo ordine di ciascuno degli stati $|\nu_i\rangle$ del sottospazio degenere, possiamo scrivere

$$|\nu_i^{(1)}\rangle = \sum_{m \neq \nu_i} |m^{(0)}\rangle \frac{\langle m^{(0)} | V | \nu_i^{(0)} \rangle}{E_D^{(0)} - E_m^{(0)}}. \quad (4.64)$$

Ma gli stati $|\nu_i^{(0)}\rangle$ sono stati scelti proprio in modo tale che gli elementi non diagonali della matrice di V nel sottospazio $\mathcal{D}$ siano nulli. Dunque, la condizione $m \neq \nu_i$ equivale a restringere la somma a tutti gli stati al di fuori di quel sottospazio:

$$|\nu_i^{(1)}\rangle = \sum_{m \notin \mathcal{D}} |m^{(0)}\rangle \frac{\langle m^{(0)} | V | \nu_i^{(0)} \rangle}{E_D^{(0)} - E_m^{(0)}}. \quad (4.65)$$

In questo modo il denominatore non è mai nullo e l'espressione può essere usata anche se lo spettro è degenere. Ora basta ricordare che la correzione al secondo ordine dell'energia era $E_n^{(2)} = \langle n^{(0)} | V | n^{(1)} \rangle$ da cui si otteneva la (4.30). Se lo stato di cui si vuole la correzione è uno degli autostati in $\mathcal{D}$, la stessa espressione diventa

$$E_i = E_i^{(0)} + \lambda v_i + \lambda^2 \sum_{m \notin \mathcal{D}} \frac{|\langle m^{(0)} | V | \nu_i^{(0)} \rangle|^2}{E_D^{(0)} - E_m^{(0)}}. \quad (4.66)$$

In conclusione, per applicare la teoria delle perturbazioni al caso di un hamiltoniano H_0 che ha spettro degenere occorre in primo luogo identificare gli autovalori degeneri e i relativi sottospazi. In ciascuno di questi sottospazi si deve scrivere la matrice della perturbazione V e diagonalizzarla. Gli autostati di V così ottenuti hanno la stessa energia nel limite $\lambda \rightarrow 0$ ma energie diverse per λ finito. Al primo ordine nella perturbazione, le correzioni alle loro energie sono λ volte gli autovalori di V corrispondenti. Nel gergo della spettroscopia questi livelli che erano degeneri, ma che la perturbazione rende non degeneri, formano un *multipletto*. Per le correzioni dal secondo ordine in su all'energia e dal primo ordine in su per gli stati si possono usare le stesse espressioni del caso dello spettro non degenere, ma escludendo dalle sommatorie gli stati del sottospazio a cui appartiene lo stato di cui si cerca la correzione.

La diagonalizzazione di V serve per cercare tra gli stati del sottospazio degenere di H_0 gli stati “buoni” a cui applicare l'ordinaria teoria delle perturbazioni. In molti casi la ricerca di questi stati “buoni” non richiede nemmeno il calcolo della matrice V e la sua diagonalizzazione, perché quegli stati possono essere trovati anche sulla base di argomenti di simmetria. Infatti, se l'insieme massimale di osservabili compatibili che commutano con H_0 viene scelto in modo tale che gli stessi operatori commutino anche con la perturbazione V , allora la matrice che rappresenta V è automaticamente diagonale. Dalla forma operatoriale di V e di H_0 spesso è possibile dedurre quali sono gli operatori che commutano con entrambi; i loro autovettori sono gli stati buoni per lo sviluppo perturbativo.

Un bell'esempio di applicazione della teoria delle perturbazioni è il calcolo della struttura fine e iperfine dello spettro dell'atomo di idrogeno. In effetti lo spettro che avevamo calcolato nella sezione 3.4, pur essendo la soluzione esatta del problema dell'elettrone nel campo coulombiano $-e^2/r$, non è del tutto realistico, dato che trascura alcuni effetti che, benché piccoli, sono misurabili con grande accuratezza con le tecniche di spettroscopia correnti (alcuni lo erano anche con la tecnologia degli anni 20 del secolo scorso). Una correzione che è facile da includere in modo esatto è l'effetto della massa finita del protone: basta usare la massa ridotta del sistema elettrone-protone al posto della massa m_e nell'equazione di

Schrödinger, e il gioco è fatto. Le due correzioni successive sono la cosiddetta *correzione relativistica* e l'interazione di *spin-orbita*. La prima viene dal fatto che, se prendiamo l'espressione dell'energia cinetica valida nella relatività ristretta, $E_K = \sqrt{p^2 c^2 + m^2 c^4} - mc^2$, dove m è la massa a riposo di una particella, p è il suo momento e c la velocità della luce, e la confrontiamo con l'espressione non relativistica $E_K = p^2/(2m)$, la differenza tra le due è $p^4/(8m^3 c^2)$ più termini di ordine superiore nel parametro piccolo $p/(mc)$. Lo spettro che avevamo calcolato era non relativistico, cioè assumeva $p \ll mc$ e teneva per buona l'espressione $p^2/(2m)$. In effetti, le energie cinetiche tipiche dell'elettrone nell'atomo d'idrogeno sono dell'ordine di 10 eV, mentre l'energia a riposo dell'elettrone, mc^2 , è dell'ordine di 0.5 MeV e questo giustifica l'approssimazione. Tuttavia, il fatto che $p/(mc)$ non sia veramente nullo produce un effetto che è calcolabile con la teoria delle perturbazioni. Invece, l'interazione di *spin-orbita* viene dal fatto che l'elettrone possiede spin $1/2$ e a questo spin è associato un momento magnetico. L'elettrone si muove attorno al protone e il protone, visto dall'elettrone, si comporta come una carica in movimento che induce un campo magnetico; dunque l'energia dell'elettrone conterrà un termine del tipo $-\boldsymbol{\mu} \cdot \mathbf{B}$ dove $\boldsymbol{\mu}$ sarà fissato dallo spin $\mathbf{S}$ e $\mathbf{B}$ dal momento angolare orbitale $\mathbf{L}$. Facendo i calcoli per bene si trova un termine proporzionale a $\mathbf{S} \cdot \mathbf{L}$ da aggiungere all'hamiltoniano che, essendo piccolo, può essere trattato perturbativamente. Risulta poi che le due correzioni, quella relativistica e quella di spin-orbita, sono dello stesso ordine e danno luogo alla cosiddetta *struttura fine* dello spettro.

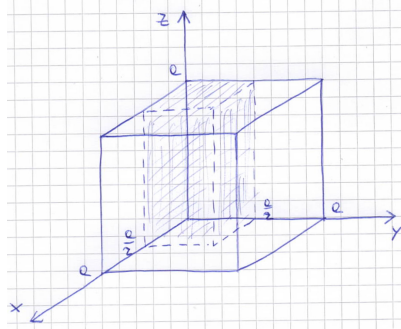
Per i calcoli espliciti gli studenti possono leggersi i capitoli 6.3, 6.4 e 6.5 del testo di D.J. Griffiths, dove si trovano anche le correzioni perturbative causate da campi magnetici esterni (effetto Zeeman) e la separazione *iperfine* dovuta all'accoppiamento con il momento magnetico del protone. Anche una semplice lettura superficiale di questi capitoli può rendere l'idea di quanto sia efficace la teoria delle perturbazioni ai fini della predizione di effetti importanti e misurabili. Si tenga presente che la precisione delle attuali tecniche spettroscopiche è tale da permettere la misura della separazione tra livelli energetici dell'atomo d'idrogeno con un'accuratezza che può arrivare fino alla 15-esima cifra decimale e l'accordo con le predizioni della teoria, esatta o perturbativa, è tale da poter stimare la costante di Rydberg con 12 o più cifre significative, al punto da poter perfino misurare l'effetto prodotto sui livelli elettronici dalla dimensione finita del protone, fornendo una stima del raggio del protone con un errore di poco superiore all'un per cento. Se finora pensavate che la meccanica quantistica, tutto sommato, non fosse altro che un apparato concettuale astratto che fornisce predizioni indeterminate, spero che questo accenno ai risultati della spettroscopia degli atomi vi faccia ricredere.

4.2.4 Esempio: perturbazione alla buca quadra

Come esempio semplice di perturbazioni di uno spettro degenere proponiamo qui l'esercizio 6.2 del Griffiths, tale e quale, a parte qualche variazione di notazione. L'esercizio riguarda una particella confinata in una buca cubica di lato a . Le pareti sono rappresentate da un potenziale $U(x, y, z)$ che vale 0 dentro la buca e infinito fuori. All'interno della buca, ma solo in una porzione del cubo, c'è anche un altro potenziale a scalino che si comporta come una perturbazione.

Scriviamo l'hamiltoniano complessivo in questo modo

$$H = H_0 + \lambda V(x, y) = \frac{p^2}{2m} + U(x, y, z) + \lambda V(x, y) \quad (4.67)$$



con $U(x, y, z) = 0$ se x, y e z sono compresi nell'intervallo $(0, a)$ e infinito altrove, e con

$$V(x, y) = \begin{cases} 1 & \text{per } 0 < x < a/2, 0 < y < a/2 \\ 0 & \text{altrove.} \end{cases} \quad (4.68)$$

Gli stati stazionari dell'hamiltoniano imperturbato sono le soluzioni dell'equazione di Schrödinger $-\hbar^2/(2m)\nabla^2\psi = E\psi$ soggette alla condizione di annullarsi sulle superfici del cubo. Senza perdere di generalità possiamo prenderle reali. Le loro funzioni d'onda sono

$$\psi_{n_x n_y n_z}^{(0)}(x, y, z) = \left(\frac{2}{a}\right)^{3/2} \sin\left(\frac{n_x \pi x}{a}\right) \sin\left(\frac{n_y \pi y}{a}\right) \sin\left(\frac{n_z \pi z}{a}\right) \quad (4.69)$$

con n_x, n_y e n_z interi e positivi. Queste funzioni d'onda rappresentano i ket $|n_x n_y n_z\rangle$ nello spazio delle coordinate. Le corrispondenti energie sono

$$E_{n_x n_y n_z}^{(0)} = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2) = \frac{\pi^2 \hbar^2}{2ma^2}(n_x^2 + n_y^2 + n_z^2). \quad (4.70)$$

Lo stato fondamentale è $|111\rangle$ e la sua energia vale $E_{111}^{(0)} = 3\pi^2 \hbar^2/(2ma^2)$. Questo autovalore corrisponde ad un solo autostato. Il livello energetico successivo è quello di energia

$$E_{211}^{(0)} = E_{121}^{(0)} = E_{112}^{(0)} = \frac{3\pi^2 \hbar^2}{ma^2} \quad (4.71)$$

che è lo stesso per tre autostati diversi. Si tratta quindi di un sottospazio di stati degeneri $\mathcal{D}$ di dimensione 3.

Ora vediamo cosa fa la perturbazione a questi stati. La correzione all'energia al primo ordine per lo stato fondamentale è $\lambda\langle 111|V|111\rangle$. Il valor medio lo possiamo calcolare usando la funzione d'onda in questo modo

$$\begin{aligned} \langle 111|V|111\rangle &= \iiint dxdydz |\psi_{111}^{(0)}(x, y, z)|^2 V(x, y) \\ &= \int_0^{a/2} dx \int_0^{a/2} dy \int_0^a dz |\psi_{111}^{(0)}(x, y, z)|^2 \\ &= \frac{8}{a^3} \int_0^{a/2} dx \sin^2\left(\frac{\pi x}{a}\right) \int_0^{a/2} dy \sin^2\left(\frac{\pi y}{a}\right) \int_0^a dz \sin^2\left(\frac{\pi z}{a}\right) \\ &= \frac{1}{4}. \end{aligned} \quad (4.72)$$

Dunque

$$E_{111} = E_{111}^{(0)} + \frac{\lambda}{4} \quad (4.73)$$

che è quanto ci potevamo aspettare, dato che la perturbazione agisce producendo uno scalino di potenziale di altezza λ in un quarto del volume del cubo (attenzione che qui λ ha le dimensioni di un'energia).

Per calcolare la correzione al primo ordine all'energia degli stati del primo tripletto si deve scrivere la matrice di V nella base degli autostati in $\mathcal{D}$. Gli elementi di matrice diagonali hanno la stessa forma di $\langle 111|V|111\rangle$, eccetto che per un dimezzamento del periodo di uno dei seni. Il risultato non cambia:

$$\langle 211|V|211\rangle = \langle 121|V|121\rangle = \langle 112|V|112\rangle = 1/4. \quad (4.74)$$

Per gli elementi fuori diagonale abbiamo

$$\begin{aligned} \langle 211|V|121\rangle &= \frac{8}{a^3} \int_0^{a/2} dx \sin\left(\frac{2\pi x}{a}\right) \sin\left(\frac{\pi x}{a}\right) \\ &\times \int_0^{a/2} dy \sin\left(\frac{\pi y}{a}\right) \sin\left(\frac{2\pi y}{a}\right) \int_0^a dz \sin^2\left(\frac{\pi z}{a}\right) \\ &= \frac{16}{9\pi^2}, \end{aligned} \quad (4.75)$$

dove i due integrali in dx e dy si risolvono usando la formula $\sin 2\alpha = 2 \sin \alpha \cos \alpha$ e un paio di sostituzioni di variabile. Invece

$$\begin{aligned} \langle 211|V|112\rangle &= \frac{8}{a^3} \int_0^{a/2} dx \sin\left(\frac{2\pi x}{a}\right) \sin\left(\frac{\pi x}{a}\right) \\ &\times \int_0^{a/2} dy \sin^2\left(\frac{\pi y}{a}\right) \int_0^a dz \sin\left(\frac{\pi z}{a}\right) \sin\left(\frac{2\pi z}{a}\right) \\ &= 0, \end{aligned} \quad (4.76)$$

dato che l'integrale in dz è nullo. Per lo stesso motivo è nullo anche l'elemento di matrice $\langle 121|V|112\rangle$. Se prendiamo i tre stati $|211\rangle$, $|121\rangle$ e $|112\rangle$, in questo ordine, come base con cui costruire la matrice simmetrica che rappresenta la perturbazione V nel sottospazio $\mathcal{D}$, allora la matrice diventa

$$V = \begin{pmatrix} 1/4 & \kappa/4 & 0 \\ \kappa/4 & 1/4 & 0 \\ 0 & 0 & 1/4 \end{pmatrix} \quad (4.77)$$

con $\kappa = 64/(9\pi^2)$. A questo punto dobbiamo diagonalizzare la matrice risolvendo l'equazione caratteristica $\det(V - v\mathbb{1}) = 0$ che corrisponde a

$$\left(\frac{1}{4} - v\right) \left[\left(\frac{1}{4} - v\right)^2 - \left(\frac{\kappa}{4}\right)^2 \right] = 0. \quad (4.78)$$

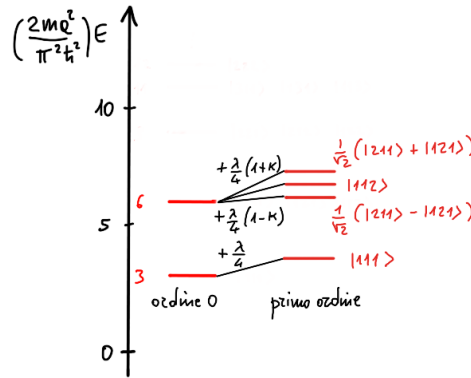
che ha come soluzioni

$$v_1 = \frac{1}{4}(1 - \kappa) \quad , \quad v_2 = \frac{1}{4} \quad , \quad v_3 = \frac{1}{4}(1 + \kappa) \quad (4.79)$$

Dunque le energie del tripletto, al primo ordine perturbativo, diventano

$$\begin{aligned} E_1 &= \frac{3\pi^2\hbar^2}{ma^2} + \frac{\lambda}{4} \left(1 - \frac{64}{9\pi^2}\right) \\ E_2 &= \frac{3\pi^2\hbar^2}{ma^2} + \frac{\lambda}{4} \\ E_3 &= \frac{3\pi^2\hbar^2}{ma^2} + \frac{\lambda}{4} \left(1 + \frac{64}{9\pi^2}\right) \end{aligned} \quad (4.80)$$

e così si vede che la perturbazione rimuove la degenerazione e produce uno *splitting* proporzionale a λ . Naturalmente, affinché la perturbazione si comporti come tale, lo spostamento dei livelli deve essere piccolo rispetto alla loro spaziatura nello spettro imperturbato, ovvero deve essere $\lambda \ll \hbar^2/ma^2$.



Per trovare gli autostati, basta cercare combinazioni lineari del tipo

$$|\nu_i^{(0)}\rangle = c_1|211\rangle + c_1|121\rangle + c_3|112\rangle \quad (4.81)$$

dove la terna di coefficienti $\mathbf{c} = (c_1, c_2, c_3)$ deve soddisfare l'equazione agli autovalori $V\mathbf{c} = v_i\mathbf{c}$ per ciascuno degli autovalori v_i appena trovati. Il risultato è il seguente

$$\begin{aligned} |\nu_1^{(0)}\rangle &= (1/\sqrt{2})(|211\rangle - |121\rangle) \\ |\nu_2^{(0)}\rangle &= |112\rangle \\ |\nu_3^{(0)}\rangle &= (1/\sqrt{2})(|211\rangle + |121\rangle) \end{aligned} \quad (4.82)$$

o, in termini delle funzioni d'onda,

$$\begin{aligned} \psi_1 &= (1/\sqrt{2})(\psi_{211} - \psi_{121}) \\ \psi_2 &= \psi_{112} \\ \psi_3 &= (1/\sqrt{2})(\psi_{211} + \psi_{121}). \end{aligned} \quad (4.83)$$

Le perturbazioni agli ordini superiori si calcolerebbero sommando i contributi degli stati fuori dal tripletto, come per il caso non degenere. Ma ci fermiamo qui.

L'approccio che abbiamo usato in questo esercizio è stato quello di diagonalizzare la matrice di V espressa nella base degli stati $|n_x n_y n_z\rangle$. Avremmo potuto ragionare in termini geometrici. Sia H_0 che V hanno in comune una simmetria per inversione delle coordinate x

e y , che corrisponde alla riflessione del cubo e del parallelepipedo rispetto al piano $x = y$. Chiamiamo P_{xy} l'operatore che, agendo sugli stati del sistema, opera tale inversione. Questo operatore commuta sia con H_0 che con V . Abbiamo inoltre

$$P_{xy}|211\rangle = |121\rangle, \quad P_{xy}|121\rangle = |211\rangle, \quad P_{xy}|112\rangle = |112\rangle \quad (4.84)$$

e dunque $|112\rangle$ è autostato di P_{xy} con autovalore 1 (è uno stato simmetrico), mentre gli altri due non sono autostati. Scambiando x con y i due stati cambiano e le loro funzioni d'onda pure. Per ottenere gli autostati di P_{xy} basta scrivere le combinazioni lineari

$$(1/\sqrt{2})(|211\rangle + |121\rangle), \quad (1/\sqrt{2})(|211\rangle - |121\rangle) \quad (4.85)$$

la prima simmetrica (autovalore 1) e la seconda antisimmetrica (autovalore -1). Lo stato antisimmetrico è uno stato “buono” per applicare la normale teoria delle perturbazioni. I due stati simmetrici sono ancora degeneri nello spettro di P_{xy} . Serve quindi un altro operatore che commuti con H_0 e V . Notiamo che il cubo e il parallelepipedo hanno un'altra simmetria: rimangono uguali a se stessi per riflessione rispetto al piano $z = a/2$. Sia Q l'operatore che porta z in $a - z$. Lo stato $|112\rangle$ è antisimmetrico sotto l'azione di Q (autovalore -1) mentre gli stati con $n_z = 1$ sono simmetrici (autovalore 1), come si vede facilmente dalla forma delle funzioni d'onda. Dunque, i due spettri simultanei di P_{xy} e Q offrono le etichette sufficienti per stabilire univocamente la base degli stati $|\nu_i^{(0)}\rangle$ a cui applicare gli sviluppi perturbativi. Questi coincidono con gli stati (4.83) che avevamo già calcolato in precedenza.

4.3 Perturbazioni dipendenti dal tempo

4.3.1 Probabilità di transizione

Supponiamo che l'hamiltoniano dipenda dal tempo. L'evoluzione temporale può essere calcolata risolvendo l'equazione di Schrödinger dipendente dal tempo, nella versione con i *ket* o in quella con le funzioni d'onda. Soluzioni analitiche ne esistono in alcuni (pochi) casi. In altri casi si può ricorrere a soluzioni numeriche, qualora siano praticabili con le risorse di calcolo a disposizione. La teoria perturbativa, anche in questo caso, offre una valida alternativa. La si può usare quando la dipendenza temporale entra in H in un termine piccolo. Supponiamo che

$$H = H_0 + V(t) \quad (4.86)$$

dove H_0 è un hamiltoniano indipendente dal tempo, di cui si conoscono autovalori e autovettori

$$H_0|n\rangle = E_n|n\rangle \quad (4.87)$$

e $V(t)$ è una perturbazione dipendente dal tempo, ma piccola (in un senso che preciseremo più avanti), che viene accesa a $t = 0$.

Per $t \leq 0$ il comportamento del sistema è interamente descritto da H_0 . Lo stato al tempo $t = 0$, che indichiamo con $|\alpha, 0\rangle$ può essere decomposto nella base degli autostati di H_0 :

$$|\alpha, 0\rangle = \sum_n |n\rangle \langle n|\alpha, 0\rangle = \sum_n |n\rangle c_n(0). \quad (4.88)$$

Se la perturbazione $V(t)$ non venisse accesa, l'evoluzione temporale sarebbe quella già descritta al capitolo 2 nel caso di un hamiltoniano indipendente dal tempo. In particolare varrebbe

$$|\alpha, 0; t\rangle = U(t, 0)|\alpha, 0\rangle = \exp\left(-\frac{iH_0 t}{\hbar}\right) |\alpha, 0\rangle, \quad (4.89)$$

da cui

$$\begin{aligned} c_n(t) &= \langle n|\alpha, 0; t \rangle = \langle n|\exp\left(-\frac{iH_0 t}{\hbar}\right)|\alpha, 0\rangle \\ &= \exp\left(-\frac{iE_n t}{\hbar}\right) \langle n|\alpha, 0 \rangle = \exp\left(-\frac{iE_n t}{\hbar}\right) c_n(0), \end{aligned} \quad (4.90)$$

e tutta l'evoluzione si riduce ad una rotazione della fase di ciascun coefficiente c , come già avevamo visto a suo tempo.

Invece, in presenza della perturbazione $V(t)$ ci aspettiamo che anche i moduli dei coefficienti possano cambiare. L'effetto di $V(t)$ si aggiunge alla rotazione della fase che avverrebbe anche in sua assenza. Per distinguere i due effetti conviene introdurre una nuova rappresentazione della dinamica in cui si isoli l'effetto dominante di H_0 , noto, da quello della perturbazione, incognito. Basta introdurre il nuovo vettore di stato $|\alpha(t)\rangle_I$ definito tramite la relazione

$$|\alpha, 0; t \rangle = \exp\left(-\frac{iH_0 t}{\hbar}\right) |\alpha(t)\rangle_I. \quad (4.91)$$

Questa si chiama *rappresentazione d'interazione*. Coincide con la rappresentazione di Schrödinger in assenza di una perturbazione dipendente dal tempo. Per vedere dove porta nel caso di perturbazione accesa, scriviamo l'equazione di Schrödinger con l'hamiltoniano completo

$$i\hbar \frac{\partial}{\partial t} |\alpha, 0; t \rangle = (H_0 + V(t)) |\alpha, 0; t \rangle \quad (4.92)$$

e inseriamo la definizione di $|\alpha(t)\rangle_I$ appena introdotta

$$i\hbar \frac{\partial}{\partial t} \left[\exp\left(-\frac{iH_0 t}{\hbar}\right) |\alpha(t)\rangle_I \right] = (H_0 + V(t)) \exp\left(-\frac{iH_0 t}{\hbar}\right) |\alpha(t)\rangle_I. \quad (4.93)$$

A sinistra, derivando l'esponenziale si trova un termine uguale al termine con H_0 di destra. Rimane quindi solo il termine in cui si deriva il vettore di stato e l'equazione diventa

$$i\hbar \exp\left(-\frac{iH_0 t}{\hbar}\right) \frac{\partial}{\partial t} |\alpha(t)\rangle_I = V(t) \exp\left(-\frac{iH_0 t}{\hbar}\right) |\alpha(t)\rangle_I. \quad (4.94)$$

Possiamo moltiplicare entrambi i membri da sinistra per $\exp(iH_0 t/\hbar)$. Otteniamo

$$i\hbar \frac{\partial}{\partial t} |\alpha(t)\rangle_I = \exp\left(\frac{iH_0 t}{\hbar}\right) V(t) \exp\left(-\frac{iH_0 t}{\hbar}\right) |\alpha(t)\rangle_I. \quad (4.95)$$

Ora definiamo l'operatore $V_I(t)$ nella rappresentazione d'interazione tramite la relazione

$$V_I(t) = \exp\left(\frac{iH_0 t}{\hbar}\right) V(t) \exp\left(-\frac{iH_0 t}{\hbar}\right). \quad (4.96)$$

In questo modo l'equazione di Schrödinger diventa

$$i\hbar \frac{\partial}{\partial t} |\alpha(t)\rangle_I = V_I(t) |\alpha(t)\rangle_I. \quad (4.97)$$

Quindi siamo riusciti a sintetizzare l'effetto della perturbazione in un'equazione compatta. Ora cerchiamo di capire come utilizzarla per ottenere sviluppi in serie di potenze di V . Per prima cosa trasformiamo l'equazione differenziale nella sua versione integrale

$$|\alpha(t)\rangle_I = |\alpha(0)\rangle + \frac{1}{i\hbar} \int_0^t dt' V_I(t') |\alpha(t')\rangle_I, \quad (4.98)$$

dove abbiamo usato il fatto che lo stato al tempo $t = 0$ nella rappresentazione di interazione coincide con quello nella rappresentazione di Schrödinger, $|\alpha(0)\rangle_I = |\alpha, 0\rangle$. Per semplicità lo indichiamo con $|\alpha(0)\rangle$ omettendo l'indice. Ora, se V è piccolo, la correzione allo stato imperturbato rappresentata dall'integrale sarà piccola. Di conseguenza, sarà piccolo l'errore che commettiamo se, nell'integrale, sostituiamo il vettore incognito $|\alpha(t')\rangle_I$ con il corrispondente stato imperturbato $|\alpha(0)\rangle$ che, essendo indipendente dal tempo, esce dall'integrale:

$$|\alpha(t)\rangle_I \simeq |\alpha(0)\rangle + \frac{1}{i\hbar} \left[\int_0^t dt' V_I(t') \right] |\alpha(0)\rangle. \quad (4.99)$$

In questo modo basta integrare V per ottenere una prima approssimazione di $|\alpha(t)\rangle_I$. Possiamo migliorare la stima inserendo questo risultato di nuovo nell'integrale:

$$\begin{aligned} |\alpha(t)\rangle_I &\simeq |\alpha(0)\rangle \\ &+ \frac{1}{i\hbar} \int_0^t dt' V_I(t') \left[|\alpha(0)\rangle + \frac{1}{i\hbar} \int_0^{t'} dt'' V_I(t'') |\alpha(0)\rangle \right] \\ &= |\alpha(0)\rangle + \frac{1}{i\hbar} \left[\int_0^t dt' V_I(t') \right] |\alpha(0)\rangle \\ &+ \left[\left(\frac{1}{i\hbar} \right)^2 \int_0^t dt' \int_0^{t'} dt'' V_I(t') V_I(t'') \right] |\alpha(0)\rangle \end{aligned} \quad (4.100)$$

e agli ordini successivi funziona allo stesso modo, così la correzione all'ordine k sarà

$$\dots + \left[\left(\frac{1}{i\hbar} \right)^k \int_0^t dt' \int_0^{t'} dt'' \dots \int_0^{t^{(k-1)}} dt^{(k)} V_I(t') V_I(t'') \dots V_I(t^{(k)}) \right] |\alpha(0)\rangle \quad (4.101)$$

dove dobbiamo fare attenzione all'ordine dei prodotti, perché V è un operatore che varia nel tempo e potrebbe non commutare a tempi diversi; dunque deve essere

$$t > t' > t'' > \dots > t^{(k)} \quad (4.102)$$

e il prodotto nell'integrale multiplo è detto *prodotto t-ordinato*. Lo stesso risultato può essere riscritto in questa forma

$$|\alpha(t)\rangle_I = U_I(t, 0) |\alpha(0)\rangle, \quad (4.103)$$

con

$$\begin{aligned} U_I(t, 0) &= \mathbb{1} + \frac{1}{i\hbar} \int_0^t dt' V_I(t') + \left(\frac{1}{i\hbar} \right)^2 \int_0^t dt' \int_0^{t'} dt'' V_I(t') V_I(t'') \\ &+ \dots + \left(\frac{1}{i\hbar} \right)^k \int_0^t dt' \int_0^{t'} dt'' \dots \int_0^{t^{(k-1)}} dt^{(k)} V_I(t') V_I(t'') \dots V_I(t^{(k)}) \\ &+ \dots \end{aligned} \quad (4.104)$$

Questa è nota come *serie di Dyson*.

Per capire come funziona la serie perturbativa, esaminiamo cosa succede al primo ordine e supponiamo che il sistema venga preparato al tempo iniziale $t = 0$ in uno degli autostati dell'hamiltoniano imperturbato

$$|\alpha(0)\rangle = |i\rangle, \quad \text{con } H_0|i\rangle = E_i|i\rangle. \quad (4.105)$$

Poi ci chiediamo quale sia la probabilità che, al tempo $t > 0$, il sistema si trovi in un altro autostato $|n\rangle$, con $n \neq i$, per effetto della perturbazione. Tale probabilità sarà data da

$$\mathcal{P}_{i \rightarrow n} = |\langle n|\alpha, 0; t \rangle|^2 = |c_n(t)|^2 \quad (4.106)$$

e la chiamiamo *probabilità di transizione*. Usiamo la (4.91) per scrivere

$$c_n(t) = \langle n|\alpha, 0; t \rangle = \langle n|\exp\left(-\frac{iH_0 t}{\hbar}\right)|\alpha(t)\rangle_I \quad (4.107)$$

e, facendo agire l'hamiltoniano a sinistra, otteniamo

$$c_n(t) = \exp\left(-\frac{iE_n t}{\hbar}\right) \langle n|\alpha(t)\rangle_I. \quad (4.108)$$

Ora inseriamo l'espressione (4.99) che avevamo già ricavato per $|\alpha(t)\rangle_I$ al primo ordine. Otteniamo

$$c_n(t) = \exp\left(-\frac{iE_n t}{\hbar}\right) \langle n|\left(|i\rangle + \frac{1}{i\hbar} \left[\int_0^t dt' V_I(t')\right]|i\rangle\right). \quad (4.109)$$

e, sfruttando l'ortogonalità della base imperturbata, troviamo

$$c_n(t) = \exp\left(-\frac{iE_n t}{\hbar}\right) \frac{1}{i\hbar} \int_0^t dt' \langle n|V_I(t')|i\rangle. \quad (4.110)$$

L'elemento di matrice lo possiamo calcolare tornando alla rappresentazione di Schrödinger tramite la definizione (4.96):

$$\begin{aligned} \langle n|V_I(t')|i\rangle &= \langle n|\exp\left(\frac{iH_0 t'}{\hbar}\right)V(t')\exp\left(-\frac{iH_0 t'}{\hbar}\right)|i\rangle \\ &= \exp\left(\frac{iE_n t'}{\hbar}\right)\exp\left(-\frac{iE_i t'}{\hbar}\right)\langle n|V(t')|i\rangle, \end{aligned} \quad (4.111)$$

da cui

$$c_n(t) = \exp\left(-\frac{iE_n t}{\hbar}\right) \frac{1}{i\hbar} \int_0^t dt' \exp\left(\frac{i(E_n - E_i)t'}{\hbar}\right) \langle n|V(t')|i\rangle \quad (4.112)$$

e la probabilità di transizione diventa

$$\mathcal{P}_{i \rightarrow n} = \left| \frac{1}{\hbar} \int_0^t dt' \exp\left(\frac{i(E_n - E_i)t'}{\hbar}\right) \langle n|V(t')|i\rangle \right|^2, \quad (4.113)$$

o in forma più compatta

$$\mathcal{P}_{i \rightarrow n} = \left| \frac{1}{\hbar} \int_0^t dt' e^{i\omega_{ni}t'} V_{ni}(t') \right|^2, \quad (4.114)$$

con

$$\omega_{ni} = \frac{E_n - E_i}{\hbar}, \quad V_{ni}(t) = \langle n|V(t)|i\rangle. \quad (4.115)$$

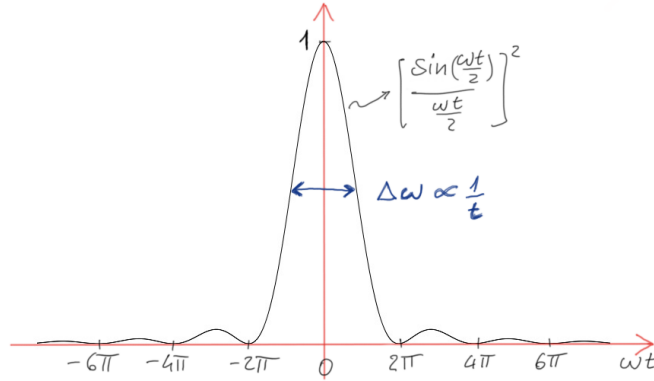
Abbiamo quindi un'espressione semplice per una quantità importante. La probabilità di transizione è misurabile: basta preparare il sistema molte volte allo stesso modo, accendendo la stessa perturbazione al tempo $t = 0$, e poi misurare l'energia dopo un tempo assegnato t . La distribuzione statistica degli esiti ci darà $\mathcal{P}$. La perturbazione è generica e può dipendere anche da altre variabili oltre al tempo.

4.3.2 Perturbazione costante accesa a $t = 0$

Consideriamo il caso semplice di una funzione $V(t)$ a gradino, che passa improvvisamente dal valore 0 al valore costante V al tempo $t = 0$ (naturalmente V può dipendere da altre variabili, ma qui si intende costante nel tempo). In tal caso l'elemento di matrice nell'espressione della probabilità di transizione non dipende dal tempo e possiamo scrivere

$$\begin{aligned}
 \mathcal{P}_{i \rightarrow n} &= \frac{|V_{ni}|^2}{\hbar^2} \left| \int_0^t dt' e^{i\omega_{ni}t'} \right|^2 \\
 &= \frac{|V_{ni}|^2}{\hbar^2} \left| \frac{e^{i\omega_{ni}t} - 1}{i\omega_{ni}} \right|^2 \\
 &= \frac{|V_{ni}|^2}{\hbar^2} \left| \frac{2 \sin(\omega_{ni}t/2)}{\omega_{ni}} \right|^2 \\
 &= \frac{|V_{ni}|^2}{\hbar^2} \left(\frac{\sin(\omega_{ni}t/2)}{\omega_{ni}t/2} \right)^2 t^2.
 \end{aligned} \tag{4.116}$$

Affinché questa probabilità sia non nulla devono verificarsi una paio di condizioni importanti. La prima è che l'elemento di matrice dell'operatore V tra gli stati iniziale e finale deve essere diverso da zero. Se per alcuni n l'elemento di matrice è nullo, allora la transizione tra lo stato i e gli stati n non avverrà mai. Questa condizione è chiamata *regola di selezione*. Le regole di selezione sono generalmente associate a particolari simmetrie del sistema. Ad esempio, se il nostro sistema è un oscillatore armonico unidimensionale con la particella inizialmente nello stato fondamentale e la perturbazione è del tipo λx^2 , con λ piccolo, accesa a $t = 0$, allora la particella a $t > 0$ potrà essere trovata nello stato con $n = 2$ con una certa probabilità, ma mai nello stato con $n = 1$, dato che l'operatore x^2 connette solo stati che differiscono di due quanti $\hbar\omega$ dell'oscillatore. La simmetria alla base di questa regola è la *parità* degli stati, cioè il fatto che essi rimangano uguali o cambino di segno quando si inverte il segno della coordinata ($x \rightarrow -x$). Se la perturbazione è pari, allora permette transizioni a stati che hanno la stessa parità di quello iniziale, che in questo caso è pari.



La seconda condizione riguarda il fatto che, per un tempo t assegnato a partire dall'accensione della perturbazione, gli stati a cui il sistema può effettivamente transire sono limitati dal fattore

$$\left(\frac{\sin(\omega t/2)}{\omega t/2} \right)^2. \tag{4.117}$$

che decresce rapidamente al crescere della frequenza ω . Infatti, questa funzione ha valore massimo pari a 1 in $\omega = 0$, si annulla in $\omega = \pm 2\pi/t$ e i massimi a sinistra e a destra di

quello centrale diventano rapidamente piccoli, con un andamento in $1/\omega^2$. Questo significa che uno stato n viene occupato dopo un tempo t con probabilità significativamente diversa da zero solo se dista dallo stato iniziale meno di una frequenza ω_{ni} dell'ordine di $1/t$ (a meno di un fattore numerico che non ci interessa se ci limitiamo a stimare l'ordine di grandezza). Questo equivale ad una differenza di energia ΔE dell'ordine di $\hbar/t$. È interessante scrivere questa condizione nella forma

$$\Delta E \Delta t \sim \hbar. \quad (4.118)$$

dove $\Delta t = t$ è la durata della perturbazione. Questa assomiglia alla relazione di indeterminazione momento-posizione (principio di indeterminazione di Heisenberg) e per questo possiamo anche chiamarla *relazione d'indeterminazione energia-tempo*. Ma dobbiamo ricordarci che nel caso della posizione e del momento si trattava di una relazione tra le dispersioni dei valori misurati di due osservabili, mentre qui abbiamo a che fare con una quantità, il tempo, che non è un'osservabile (è misurabile con gli orologi, ma entra nella teoria come un parametro indipendente dagli stati fisici dei sistemi). Il significato della relazione che abbiamo trovato è questo: se la perturbazione ha breve durata (piccolo Δt), il sistema può finire in uno stato con energia anche molto diversa da quella iniziale (grande ΔE), anche se la perturbazione è piccola, violando così la conservazione dell'energia. La violazione è concessa solo per tempi piccoli, compatibili appunto con la relazione d'indeterminazione scritta qui sopra. Per tempi lunghi, il sistema mantiene la sua energia, come in fisica classica.

4.3.3 Perturbazione armonica accesa a $t = 0$

Un altro esempio interessante è quello di una perturbazione periodica del tipo

$$V(t) = \begin{cases} 0 & \text{for } t \leq 0 \\ \mathcal{V}e^{i\omega t} + \mathcal{V}^\dagger e^{-i\omega t} & \text{for } t > 0. \end{cases} \quad (4.119)$$

dove $\mathcal{V}$ è un operatore che non dipende dal tempo e ω è una frequenza positiva. Scritta in questo modo, la perturbazione è rappresentata da un operatore hermitiano, come dev'essere. Il calcolo della probabilità di transizione è simile a prima:

$$\begin{aligned} \mathcal{P}_{i \rightarrow n} &= \frac{1}{\hbar^2} \left| \int_0^t dt' e^{i\omega_{ni}t'} (\mathcal{V}_{ni}e^{i\omega t'} + \mathcal{V}_{ni}^\dagger e^{-i\omega t'}) \right|^2 \\ &= \frac{1}{\hbar^2} \left| \int_0^t dt' \left[\mathcal{V}_{ni}e^{i(\omega_{ni}+\omega)t'} + \mathcal{V}_{ni}^\dagger e^{i(\omega_{ni}-\omega)t'} \right] \right|^2 \\ &= \frac{1}{\hbar^2} \left| \left[\mathcal{V}_{ni} \frac{e^{i(\omega_{ni}+\omega)t} - 1}{\omega_{ni} + \omega} + \mathcal{V}_{ni}^\dagger \frac{e^{i(\omega_{ni}-\omega)t} - 1}{\omega_{ni} - \omega} \right] \right|^2. \end{aligned} \quad (4.120)$$

Per lo stesso ragionamento che avevamo fatto in precedenza riguardo alla relazione d'indeterminazione energia-tempo, anche qui si vede che, per tempi sufficientemente lunghi, i due termini in parentesi assumono valori diversi da zero in intervalli di frequenze disgiunti. Il primo termine è apprezzabilmente diverso da zero solo in un intervallo di frequenze ω_{ni} vicine a $-\omega$, con un'ampiezza dell'ordine di $1/t$, mentre il secondo lo è in un intervallo di uguale ampiezza attorno a $+\omega$. In altre parole, il primo addendo nella parentesi descrive i processi in cui si passa da uno stato di energia E_i ad uno stato di energia più bassa E_n con

$$E_n \simeq E_i - \hbar\omega, \quad (4.121)$$

mentre il secondo descrive il processo

$$E_n \simeq E_i + \hbar\omega. \quad (4.122)$$

Nel primo caso si parla di *emissione stimolata* e nel secondo di *assorbimento*. L'emissione stimolata è possibile solo se il sistema si trova già all'inizio in uno stato di energia maggiore di quella dello stato fondamentale, cioè in uno *stato eccitato*. La perturbazione rappresenta l'accoppiamento tra il sistema e il resto dell'universo e si comporta, a seconda dei casi, come un pozzo o come una sorgente di energia. Uno dei vantaggi di questo formalismo è che si adatta perfettamente al caso di sistemi composti da particelle cariche sottoposte all'effetto di radiazione elettromagnetica. Il campo elettromagnetico si comporta proprio come una perturbazione armonica e la condizione $\hbar\omega = \pm(E_n - E_i)$ fornisce l'interpretazione corretta dell'ipotesi di Bohr sui salti tra gli stati stazionari di un elettrone in un atomo.

Dunque, approfittiamo del fatto che i due addendi nella (4.120) possono essere trattati separatamente per tempi t sufficientemente lunghi, ovvero per $1/t \ll \omega$. Consideriamo quindi uno solo dei due, sapendo che nello stesso intervallo di ω l'altro è trascurabile, e scegliamo l'assorbimento. La probabilità di transizione diventa

$$\mathcal{P}_{i \rightarrow n} = \frac{|\mathcal{V}_{ni}^\dagger|^2}{\hbar^2} \left(\frac{\sin[(\omega_{ni} - \omega)t/2]}{(\omega_{ni} - \omega)t/2} \right)^2 t^2, \quad (4.123)$$

che ha lo stesso aspetto di quella con la perturbazione costante, salvo il fatto che la probabilità qui è massima quando $\hbar\omega = E_n - E_i$, ovvero quando la perturbazione è *in risonanza* con i livelli energetici propri del sistema imperturbato. Vale la pena sottolineare che la condizione di risonanza per la frequenza della perturbazione è una condizione necessaria ma non sufficiente per realizzare la transizione tra gli stati. Infatti, per ottenere una probabilità di transizione non nulla occorre anche che l'elemento di matrice della perturbazione sia diverso da zero; in altri termini, la transizione deve essere compatibile con le regole di selezione.

Vale anche qui il discorso che avevamo fatto prima sulla violazione della conservazione dell'energia. Infatti, a partire da uno stato $|i\rangle$ possiamo popolare con probabilità non nulla lo stato $|n\rangle$ anche se l'energia $\hbar\omega$ associata alla perturbazione non è esattamente uguale alla differenza di energia tra i due livelli. Se la transizione avviene, la conservazione dell'energia è violata. Ciò è possibile a condizione che la perturbazione agisca in un tempo finito, di durata compatibile con la relazione d'indeterminazione energia-tempo.

Risultati analoghi valgono per il processo di emissione stimolata. Nulla cambia a parte un segno nelle frequenze e l'elemento di matrice da inserire, che è $\mathcal{V}_{ni}$.

4.3.4 Regola aurea di Fermi

In molti casi pratici, data una perturbazione armonica di frequenza assegnata ω , lo stato di arrivo nel processo di assorbimento appartiene ad un gruppo denso di stati aventi energie vicine tra loro, poco diverse da $E_n = E_i + \hbar\omega$. In effetti, per una durata t della perturbazione, questa può far transire il sistema in tutti gli stati compresi in un intervallo dell'ordine di $\hbar/t$ e, se questi sono molti, non ha molto senso concentrarsi su uno soltanto di loro. Questo può essere il caso di una transizione da uno stato legato ad uno stato libero, quest'ultimo essendo immerso in uno spettro continuo, oppure per transizioni a stati legati in uno spettro discreto ma con spaziatura δE molto minore di $\hbar/t$. In questi casi ha più senso calcolare la probabilità di transire entro il gruppo di stati vicini, opportunamente definiti e compatibili con le regole di selezione. Chiamiamo $\mathcal{D}$ l'insieme di questi stati. La probabilità che cerchiamo è la somma delle probabilità di transire in ciascun degli stati dell'insieme

$$\sum_{\nu \in \mathcal{D}} \mathcal{P}_{i \rightarrow \nu}, \quad (4.124)$$

ma se gli stati, ordinati in energia, sono densi possiamo sostituire la somma con l'integrale

$$\int_{\mathcal{D}} dE_\nu \rho(E_\nu) \mathcal{P}_{i \rightarrow \nu}, \quad (4.125)$$

dove la quantità $\rho(E)$ ha il significato di *densità degli stati* nello spazio dell'energia e il prodotto $\rho(E_\nu)dE_\nu$ è il numero di stati in un intorno infinitesimo di E_ν . Inserendo il risultato (4.123), lo stesso integrale diventa

$$\int_{\mathcal{D}} dE_\nu \rho(E_\nu) \frac{|\mathcal{V}_{\nu i}^\dagger|^2}{\hbar^2} \left(\frac{\sin[(E_\nu - E_i - \hbar\omega)t/2\hbar]}{(E_\nu - E_i - \hbar\omega)t/2\hbar} \right)^2 t^2. \quad (4.126)$$

Ora notiamo che l'espressione nella parentesi tonda, per valori fissati di ω e t , è una funzione di E_ν avente un picco in $E_\nu = E_n = E_i + \hbar\omega$. Per tempi t sufficientemente lunghi il picco è tutto compreso in $\mathcal{D}$. Inoltre, se $|\mathcal{V}_{\nu i}^\dagger|^2$ e $\rho(E)$ variano lentamente in $\mathcal{D}$, allora la stessa funzione si comporta come una delta di Dirac. Basta ricordarci che

$$\lim_{a \rightarrow \infty} \frac{1}{\pi a} \frac{\sin^2 ax}{x^2} = \delta(x), \quad (4.127)$$

e sostituire a e x rispettivamente con $t/2\hbar$ e $E_\nu - E_n$, in modo che

$$\lim_{t \rightarrow \infty} \left(\frac{\sin[(E_\nu - E_n)t/2\hbar]}{(E_\nu - E_n)t/2\hbar} \right)^2 = \frac{2\hbar\pi}{t} \delta(E_\nu - E_n). \quad (4.128)$$

In questo limite l'integrale si riduce a

$$\frac{2\pi}{\hbar} \rho(E_n) |\mathcal{V}_{ni}^\dagger|^2 t \quad (4.129)$$

e questa è la quantità cercata, cioè la probabilità di transire negli stati nell'intorno di $E_n = E_i + \hbar\omega$. La possiamo indicare così: $\mathcal{P}_{i \rightarrow [n]}$. In questa espressione l'elemento di matrice $|\mathcal{V}_{ni}^\dagger|^2$ va inteso come media della stessa quantità nell'intorno $[n]$, piuttosto che come il valore per un singolo elemento di matrice. Notiamo anche che la probabilità che abbiamo trovato è proporzionale a t . Possiamo allora introdurre un *rate* (o velocità) di transizione Γ tale che $\mathcal{P}_{i \rightarrow [n]} = \Gamma_{i \rightarrow [n]} t$. Allora il nostro risultato finale può essere riscritto nella forma

$$\boxed{\Gamma_{i \rightarrow [n]} = \frac{2\pi}{\hbar} \rho(E_n) |\mathcal{V}_{ni}^\dagger|^2}. \quad (4.130)$$

Questa è nota come regola aurea di Fermi (*Fermi golden rule* nei testi inglesi), a causa dell'importanza che riveste in molti casi pratici.

Se ripetessimo gli stessi ragionamenti per l'emissione stimolata, invece dell'assorbimento, troveremmo lo stesso risultato, eccetto sostituire $|\mathcal{V}_{ni}^\dagger|^2$ con $|\mathcal{V}_{ni}|^2$ e definire $[n]$ come l'intorno di $E_n = E_i - \hbar\omega$. Il resto non cambia. Una cosa interessante da osservare è che esiste una ben precisa relazione tra le probabilità di transire tra gli stessi due stati per assorbimento o per emissione stimolata. Per dimostrarlo prendiamo due stati generici $|n_1\rangle$ e $|n_2\rangle$ di uno spettro; poi andiamo a calcolare il *rate* di assorbimento da n_1 a n_2 e lo confrontiamo con il *rate* di emissione da n_2 a n_1 . Notiamo che

$$\mathcal{V}_{21}^\dagger = \langle n_2 | \mathcal{V}^\dagger | n_1 \rangle = \langle n_1 | \mathcal{V} | n_2 \rangle^* = \mathcal{V}_{12}^* \quad (4.131)$$

e dunque i due elementi di matrice hanno lo stesso modulo quadro. Per l'assorbimento possiamo dunque scrivere

$$\Gamma_{n_1 \rightarrow [n_2]} = \frac{2\pi}{\hbar} \rho(E_{n_2}) |\mathcal{V}_{12}|^2 \quad (4.132)$$

e per l'emissione

$$\Gamma_{n_2 \rightarrow [n_1]} = \frac{2\pi}{\hbar} \rho(E_{n_1}) |\mathcal{V}_{12}|^2, \quad (4.133)$$

ma queste implicano che

$$\frac{\Gamma_{n_1 \rightarrow [n_2]}}{\rho(E_{n_2})} = \frac{\Gamma_{n_2 \rightarrow [n_1]}}{\rho(E_{n_1})}. \quad (4.134)$$

Questa relazione è nota come legge del bilancio dettagliato (*detailed balance*).

Le relazioni che abbiamo ricavato per la perturbazione armonica al primo ordine stanno alla base di tutta la *spettroscopia*. Nel caso dei campi elettromagnetici ciò è la conseguenza del fatto che l'azione di un campo classico di radiazione è riconducibile alla forma $\mathcal{V}e^{i\omega t} - \mathcal{V}^\dagger e^{-i\omega t}$ (rimandiamo gli studenti interessati al capitolo 5.8 del Sakurai per approfondimenti). Per questo, come già detto, le transizioni tra stati di particelle cariche sottoposte a radiazione elettromagnetica rientrano pienamente nel formalismo che abbiamo sviluppato. Tuttavia conviene sottolineare un punto cruciale: qui stiamo trattando un sistema quantistico (ad esempio un elettrone in un atomo) sottoposto ad un campo esterno classico dipendente dal tempo. La sua dipendenza dal tempo consente le transizioni tra gli stati del sistema “in presenza” del campo esterno. Per questo si parla di assorbimento e di emissione stimolata. Il fatto di aver incluso nella stessa descrizione un sistema quantistico e un campo classico rende la teoria *semiclassica*. Questo ha un prezzo; ad esempio, non permette di descrivere un processo che invece sappiamo avvenire in natura: un elettrone che si trova in uno stato eccitato può decadere spontaneamente ad uno stato di energia più bassa “in assenza” di un campo esterno. Questo tipo di emissione spontanea non è incluso nella formulazione precedente. Per risolvere il problema dobbiamo trattare *sia* il sistema *sia* il campo elettromagnetico in modo quantistico, sullo stesso piano, in modo che una variazione dello stato del sistema (dell'elettrone) sia accompagnata da una variazione dello stato, quantistico, del campo elettromagnetico. Non è un ostacolo insormontabile. La trattazione quantistica della radiazione è nota come *elettrodinamica quantistica* e i quanti della radiazione sono i *fotoni*. Le perturbazioni al primo ordine saranno quelle che si limitano a processi che coinvolgono un solo fotone alla volta e la teoria includerà anche l'emissione spontanea. Gli studenti portino pazienza se non possiamo trattare proprio tutto in questo corso.

4.4 Metodi variazionali

La teoria delle perturbazioni non sempre funziona. A volte non ha proprio significato. Nel caso di stati stazionari, un metodo alternativo per trovare soluzioni approssimate è quello *variazionale*.

Supponiamo che un sistema sia descritto da un hamiltoniano H . I suoi autostati incogniti siano rappresentati dai vettori $|n\rangle$ e abbiano energia E_n . Al solito possiamo scrivere

$$H = \sum_n E_n |n\rangle \langle n|. \quad (4.135)$$

Poi prendiamo un vettore di stato qualsiasi, non necessariamente normalizzato a 1, e scriviamo

$$\langle \alpha | H | \alpha \rangle = \sum_n E_n \langle \alpha | n \rangle \langle n | \alpha \rangle. \quad (4.136)$$

Ora, supponiamo di aver ordinato gli autostati in base agli autovalori in modo che E_0 sia l'energia dello stato fondamentale. Dato che lo stato fondamentale è quello di energia inferiore, vale $E_n \geq E_0$ per ogni n . Dunque, se nella relazione precedente sostituiamo tutti gli E_n con E_0 otteniamo la disuguaglianza

$$\langle \alpha | H | \alpha \rangle = \sum_n E_n \langle \alpha | n \rangle \langle n | \alpha \rangle \geq \sum_n E_0 \langle \alpha | n \rangle \langle n | \alpha \rangle = E_0 \langle \alpha | \alpha \rangle \quad (4.137)$$

e dunque

$$E_0 \leq \frac{\langle \alpha | H | \alpha \rangle}{\langle \alpha | \alpha \rangle}. \quad (4.138)$$

A questo punto, se la soluzione esatta per lo stato fondamentale è difficile da ottenere, l'alternativa sta nel cercare una forma opportuna di $|\alpha\rangle$ che ne rappresenti una buona approssimazione. La disuguaglianza che abbiamo trovato suggerisce di individuare questa soluzione approssimata andando a cercare il minimo della quantità $\langle \alpha | H | \alpha \rangle / \langle \alpha | \alpha \rangle$ entro uno spazio sufficientemente ampio di vettori. Se questo spazio include la soluzione esatta, la ricerca del minimo porterà proprio a quella. Se non la include, si troverà comunque un limite superiore all'energia esatta (un *upper bound*).

In pratica, si individua una classe di stati “di prova” definiti tramite l'uso di opportuni parametri continui; si calcola $\langle \alpha | H | \alpha \rangle / \langle \alpha | \alpha \rangle$ in funzione dei parametri; si applica una procedura di ricerca del minimo (analitica o numerica); tanto più la classe era buona, tanto più il minimo sarà vicino alla soluzione esatta.

Il metodo è applicabile anche agli stati eccitati sopra quello fondamentale, ma con qualche accortezza: nel cercare il minimo dobbiamo restringere la ricerca ad una classe di stati che sia ortogonale a quelli che stanno sotto a quello che si vuole approssimare. Ad esempio, un atomo ha come stato fondamentale un orbitale s (momento angolare 0), ma se noi vogliamo invece trovare lo stato p (momento angolare 1) con il metodo variazionale, basta imporre che la classe di stati di prova abbia solo e soltanto quel momento angolare. Dunque basta parametrizzare solo la parte radiale della funzione d'onda e fissare invece la parte angolare con le armoniche sferiche Y_{1m} . Così gli stati che esploriamo nella ricerca del minimo sono tutti automaticamente ortogonali allo stato fondamentale e il metodo variazionale funziona correttamente. In generale, si tratta di scegliere opportunamente la classe di stati in modo che abbia la simmetria giusta.

Sembra un metodo grezzo, ma funziona benissimo in molti casi.

Capitolo 5

Più particelle e matrice densità

5.1 Due o più particelle

Finora abbiamo trattato sistemi composti da una sola particella. Ora cominciamo ad aggiungerne altre. Partiamo da due. Consideriamo le osservabili che caratterizzano lo stato di una delle due particelle e quelle che caratterizzano lo stato dell'altra. Possiamo individuare gli insiemi massimali di osservabili compatibili per l'una e per l'altra e costruire, con questi, le basi dei due spazi vettoriali corrispondenti. Il sistema composto da entrambe le particelle sarà descritto da vettori che appartengono allo spazio prodotto diretto, o spazio prodotto tensoriale, dei due spazi associati a ciascuna di esse. Abbiamo accennato al prodotto diretto già nel capitolo sul momento angolare, quando abbiamo incluso i gradi di libertà di spin nella descrizione degli stati di una particella. Qui ci serve dare qualche dettaglio in più.

Prendiamo due sistemi fisici generici. Per ciascuno dei due, scegliamo un insieme massimale di osservabili compatibili tale da definire una base del rispettivo spazio di Hilbert. Siano $\mathcal{H}_1$ e $\mathcal{H}_2$ i due spazi e $\{|m_1\rangle, |m_2\rangle, |m_3\rangle, \dots\}$ e $\{|n_1\rangle, |n_2\rangle, |n_3\rangle, \dots\}$ le rispettive basi ortonormali. Le etichette m e n sintetizzano l'insieme dei numeri quantici che definiscono il singolo autostato di ciascuna base e la dimensione dei due spazi non è necessariamente la stessa. Possiamo definire il prodotto diretto dei due spazi di Hilbert come lo spazio

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \quad (5.1)$$

avente come base gli stati

$$|m_i, n_j\rangle = |m_i\rangle \otimes |n_j\rangle \quad (5.2)$$

tali che

$$\langle m_{i'}, n_{j'} | m_i, n_j \rangle = (\langle m_{i'} | \otimes \langle n_{j'} |) (|m_i\rangle \otimes |n_j\rangle) = \langle m_{i'} | m_i \rangle \langle n_{j'} | n_j \rangle = \delta_{ii'} \delta_{jj'} \quad (5.3)$$

e

$$\sum_{i,j} |m_i, n_j\rangle \langle m_i, n_j| = \mathbb{1} . \quad (5.4)$$

Per definizione, un qualsiasi vettore generico $|\alpha\rangle$ del nuovo spazio di Hilbert potrà essere scritto come la combinazione lineare

$$|\alpha\rangle = \sum_{i,j} c_{ij} |m_i, n_j\rangle = \sum_{i,j} c_{ij} |m_i\rangle \otimes |n_j\rangle . \quad (5.5)$$

dove c_{ij} sono numeri complessi. In questo modo, tutte le proprietà dei prodotti interni tra vettori nei singoli spazi $\mathcal{H}_1$ e $\mathcal{H}_2$ si applicano anche ai prodotti interni tra vettori nello spazio prodotto diretto $\mathcal{H}$.

Un caso particolare è quello in cui i coefficienti nella sommatoria possono essere scritti come prodotto di due numeri complessi che dipendono separatamente dall'indice i e dall'indice j , del tipo $c_{ij} = a_i b_j$. In tal caso le due somme possono essere separate e si può scrivere

$$|\alpha\rangle = \left(\sum_i a_i |m_i\rangle \right) \otimes \left(\sum_j b_j |n_j\rangle \right), \quad (5.6)$$

ottenendo così una fattorizzazione del vettore in $\mathcal{H}$ in due vettori distinti in $\mathcal{H}_1$ e $\mathcal{H}_2$. Se lo stato del sistema nello spazio $\mathcal{H}$ è di questo tipo, allora una qualsiasi misura eseguita in uno dei due sottospazi $\mathcal{H}_1$ o $\mathcal{H}_2$ non modifica lo stato nell'altro sottospazio. Tuttavia, non tutti gli stati di un sistema godono di questa proprietà; la teoria ammette anche vettori di stato i cui coefficienti c_{ij} non sono separabili in prodotti del tipo $a_i b_j$. Questi stati sono detti *entangled*¹. Se il sistema si trova in uno stato *entangled*, allora la misura di un'osservabile in uno dei due sottospazi modifica lo stato anche nell'altro.

Sul concetto di *entanglement* torneremo in seguito. Ora piuttosto vediamo come il formalismo dei prodotti diretti si applica a sistemi con più di una particella.

Prendiamo due particelle e limitiamoci ai gradi di libertà associati al movimento nello spazio (niente spin per il momento). Lo spazio di Hilbert da considerare per ciascuna particella ha dimensioni infinite e possiamo usare come base gli autostati della posizione. Dunque l'operatore che definisce la base è lo stesso, salvo che in uno dei due sottospazi si riferisce alla posizione della particella 1 e nell'altro alla posizione della particella 2, dove 1 e 2 sono etichette convenzionali associate semplicemente all'ordine che utilizziamo nello scrivere la base degli autostati dello spazio prodotto diretto:

$$|\mathbf{r}_1, \mathbf{r}_2\rangle = |\mathbf{r}_1\rangle \otimes |\mathbf{r}_2\rangle. \quad (5.7)$$

Con questa base possiamo rappresentare qualsiasi vettore. Dato che lo spettro è continuo, rispetto alle espressioni precedenti qui dobbiamo usare integrali anziché sommatorie. Dunque, si avrà

$$\langle \mathbf{r}'_1, \mathbf{r}'_2 | \mathbf{r}_1, \mathbf{r}_2 \rangle = \delta(\mathbf{r}'_1 - \mathbf{r}_1) \delta(\mathbf{r}'_2 - \mathbf{r}_2) \quad (5.8)$$

e

$$\iint d^3 r_1 d^3 r_2 |\mathbf{r}_1, \mathbf{r}_2\rangle \langle \mathbf{r}_1, \mathbf{r}_2| = \mathbb{1} \quad (5.9)$$

e un qualsiasi vettore $|\alpha\rangle$ sarà esprimibile come

$$|\alpha\rangle = \iint d^3 r_1 d^3 r_2 |\mathbf{r}_1, \mathbf{r}_2\rangle \langle \mathbf{r}_1, \mathbf{r}_2 | \alpha \rangle. \quad (5.10)$$

Si può dunque definire la funzione d'onda

$$\psi_\alpha(\mathbf{r}_1, \mathbf{r}_2) = \langle \mathbf{r}_1, \mathbf{r}_2 | \alpha \rangle \quad (5.11)$$

il cui modulo quadro è la densità di probabilità di ottenere $\mathbf{r}_1$ e $\mathbf{r}_2$ nella misura delle posizioni delle particelle 1 e 2 quando il sistema è preparato nello stato $|\alpha\rangle$. Ripetiamo qui che l'ordine delle particelle in questa espressione è convenzionale; si tratta essenzialmente di una questione di etichettatura dei vettori. Vedremo poi in quali situazioni l'ordine è davvero rilevante dal punto di vista fisico. Notiamo anche che la ψ_α è una funzione di 6 variabili (più il tempo, quando serve) e la sua normalizzazione richiede una doppia integrazione spaziale:

$$1 = \iint d^3 r_1 d^3 r_2 |\psi_\alpha(\mathbf{r}_1, \mathbf{r}_2)|^2. \quad (5.12)$$

¹In italiano diremmo intrecciati, o ingarbugliati, ma nei testi di fisica si usa sempre il termine inglese.

Tra l'altro, questo fa sì che il modulo quadro della funzione d'onda non possa essere rappresentato da una “nuvoletta” nello spazio euclideo, dato che la densità di probabilità qui non è più una funzione di tre variabili. La raffigurazione pittorica degli orbitali, quella che di solito si usa per disegnare gli stati elettronici nell'atomo d'idrogeno, funziona solo per una particella.

5.1.1 Due particelle isolate

Per prima cosa mostriamo che, se due particelle di massa m_1 e m_2 interagiscono solo tra loro, il problema è riducibile alla soluzione di un problema a un corpo, come in meccanica classica. Supponiamo dunque che l'hamiltoniano sia

$$H = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + V(|\mathbf{r}_1 - \mathbf{r}_2|). \quad (5.13)$$

Allora l'equazione di Schrödinger è

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}_1, \mathbf{r}_2, t) = \left[-\frac{\hbar^2 \nabla_1^2}{2m_1} - \frac{\hbar^2 \nabla_2^2}{2m_2} + V(|\mathbf{r}_1 - \mathbf{r}_2|) \right] \psi(\mathbf{r}_1, \mathbf{r}_2, t). \quad (5.14)$$

Si tratta di un'equazione differenziale in 6 variabili spaziali più il tempo. Per semplificarla conviene introdurre il centro di massa, usando le solite definizioni di posizione del CM

$$\mathbf{r}_{\text{CM}} = \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2}{m_1 + m_2}, \quad (5.15)$$

di momento totale

$$\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2, \quad (5.16)$$

di posizione relativa

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \quad (5.17)$$

e di momento relativo

$$\mathbf{p} = \frac{m_2 \mathbf{p}_1 - m_1 \mathbf{p}_2}{m_1 + m_2}. \quad (5.18)$$

Segue che

$$\frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} = \frac{p^2}{2m} + \frac{P^2}{2M} \quad (5.19)$$

dove $M = m_1 + m_2$ è la massa totale e $m = m_1 m_2 / M$ è la massa ridotta. Lo stesso vale per i corrispondenti operatori differenziali

$$\frac{\nabla_1^2}{2m_1} + \frac{\nabla_2^2}{2m_2} = \frac{\nabla_r^2}{2m} + \frac{\nabla_{r_{\text{CM}}}^2}{2M} \quad (5.20)$$

e possiamo riscrivere l'equazione di Schrödinger così

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, \mathbf{r}_{\text{CM}}, t) = \left[-\frac{\hbar^2 \nabla_r^2}{2m} - \frac{\hbar^2 \nabla_{r_{\text{CM}}}^2}{2M} + V(r) \right] \psi(\mathbf{r}, \mathbf{r}_{\text{CM}}, t). \quad (5.21)$$

A differenza di prima, ora l'hamiltoniano è separabile in due parti, l'una che agisce solo sulla posizione relativa e l'altra solo sulla posizione del CM. Possiamo quindi fattorizzare la funzione d'onda in questo modo

$$\psi(\mathbf{r}, \mathbf{r}_{\text{CM}}, t) = f(\mathbf{r}_{\text{CM}}, t) \varphi(\mathbf{r}, t). \quad (5.22)$$

Inserendo questa funzione nell'equazione di Schrödinger si ottengono le due equazioni disaccoppiate

$$\begin{cases} i\hbar \frac{\partial}{\partial t} f(\mathbf{r}_{\text{CM}}, t) = -\frac{\hbar^2 \nabla_{\mathbf{r}_{\text{CM}}}^2}{2M} f(\mathbf{r}_{\text{CM}}, t) \\ i\hbar \frac{\partial}{\partial t} \varphi(\mathbf{r}, t) = \left[-\frac{\hbar^2 \nabla_{\mathbf{r}}^2}{2m} + V(r) \right] \varphi(\mathbf{r}, t). \end{cases} \quad (5.23)$$

La prima dice che il CM si comporta come una particella libera, la cui funzione d'onda è un'onda piana

$$f(\mathbf{r}_{\text{CM}}, t) = c \exp[i(\mathbf{k}_{\text{CM}} \cdot \mathbf{r}_{\text{CM}} - \omega t)] \quad (5.24)$$

dove c è una costante, $\hbar \mathbf{k}_{\text{CM}} = \mathbf{P}$ e $\hbar \omega = \hbar^2 \mathbf{k}_{\text{CM}}^2 / 2M$. La seconda equazione invece è l'equazione di Schrödinger di una particella di massa ridotta soggetta al potenziale $V(\mathbf{r})$. Le soluzioni stazionarie hanno la forma

$$\varphi(\mathbf{r}, t) = \varphi(\mathbf{r}) \exp\left(-\frac{i}{\hbar} E t\right) \quad (5.25)$$

dove $\varphi(\mathbf{r})$ è la soluzione dell'equazione di Schrödinger stazionaria

$$\left[-\frac{\hbar^2 \nabla_{\mathbf{r}}^2}{2m} + V(r) \right] \varphi(\mathbf{r}) = E \varphi(\mathbf{r}). \quad (5.26)$$

Una volta risolta questa equazione, la soluzione complessiva per il sistema a due particelle è

$$\psi(\mathbf{r}, \mathbf{r}_{\text{CM}}, t) = c \exp(i \mathbf{k}_{\text{CM}} \cdot \mathbf{r}_{\text{CM}}) \exp\left[-\frac{i}{\hbar} \left(E + \frac{\hbar^2 \mathbf{k}_{\text{CM}}^2}{2M}\right) t\right] \varphi(\mathbf{r}). \quad (5.27)$$

Dunque, come in fisica classica, l'interazione entra solo nel moto relativo, a cui corrisponde un'equazione di Schrödinger stazionaria per una sola particella. Ovviamente la cosa non funziona se le due particelle sono soggette anche ad un campo esterno che dipende dalla loro posizione, nel qual caso il problema a due corpi rimane tale.

5.1.2 Particelle identiche

In quanto detto finora, quale particella chiamiamo “prima” particella, quale “seconda”, ecc., non è rilevante. Sicuramente non lo è se le particelle sono diverse come lo sono particelle con massa o carica elettrica diversa, oppure atomi di specie diversa, o differenti isotopi dello stesso atomo. Particelle distinguibili possono sentire potenziali esterni diversi e potranno avere potenziali di interazione reciproca diversi. La situazione cambia se le particelle del sistema sono identiche, come nel caso di un sistema di elettroni, o di atomi della stessa specie tutti nello stato fondamentale. Queste particelle sono indistinguibili; è un fatto sperimentale: non esistono misure che possano distinguere un elettrone da un altro, se si trovano nello stesso stato, né misure che distinguono un atomo di ossigeno da un altro. Nella dinamica classica questa indistinguibilità viene risolta per il fatto che, data la configurazione di un sistema di particelle identiche in un certo istante, è sempre possibile (in linea di principio, non necessariamente in pratica) etichettare le particelle sulla base di dove si trovano in quel momento; tale etichettatura permane valida anche in tutti gli istanti successivi, perché ogni particella seguirà una traiettoria univocamente determinata dalle leggi del moto. In meccanica quantistica ciò non è possibile, a meno di non metterci nella situazione in cui la delocalizzazione spaziale delle particelle (la regione in cui la funzione d'onda della singola particella è non nulla) sia limitata a scale spaziali molto minori di quelle rilevanti ai fini della descrizione del sistema. Ma in generale le funzioni d'onda di singola particella possono sovrapporsi e il metodo classico di etichettatura non funziona. Ne deduciamo che nessuna

misura effettuata su un sistema di particelle identiche può distinguere stati che differiscono solo per una permutazione degli indici. La natura impone che una qualsiasi osservabile dia lo stesso esito se le particelle vengono scambiate.

Prendiamo un sistema di due particelle. In meccanica quantistica diciamo che le particelle sono identiche se lo spazio di Hilbert che descrive il sistema è

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \quad (5.28)$$

dove $\mathcal{H}_1$ e $\mathcal{H}_2$ sono due copie di uno stesso spazio, con le medesime dimensioni, le stesse osservabili compatibili, eccetera (per essere precisi, sono due spazi isomorfi), e inoltre lo scambio delle due particelle lascia invariato il risultato di qualsiasi misura si esegua sul sistema.

Indichiamo con

$$|m, n\rangle = |m\rangle \otimes |n\rangle \quad (5.29)$$

un generico elemento della base dello spazio $\mathcal{H}$. È utile definire un operatore di scambio P_{12} tale che

$$P_{12}(|m\rangle \otimes |n\rangle) = |n\rangle \otimes |m\rangle, \quad (5.30)$$

ovvero

$$P_{12}|m, n\rangle = |n, m\rangle. \quad (5.31)$$

Insistiamo sul fatto che la prima posizione in questi ket corrisponde alla prima particella, per convenzione, e la seconda posizione alla seconda particella. L'azione dell'operatore di scambio quindi va intesa come l'attribuzione delle coordinate della seconda particella alla prima e viceversa, intendendo con coordinate l'insieme degli autovalori (numeri quantici) che caratterizzano univocamente gli stati delle particelle.

Notiamo che applicando due volte lo stesso scambio si torna allo stato iniziale:

$$P_{12}^2|m, n\rangle = P_{12}P_{12}|m, n\rangle = P_{12}|n, m\rangle = |m, n\rangle \quad (5.32)$$

e dunque

$$P_{12}^2 = P_{12}P_{12} = \mathbb{1}, \quad (5.33)$$

che implica anche

$$P_{12}^{-1} = P_{12}. \quad (5.34)$$

Inoltre l'operatore P_{12} è hermitiano. Per dimostrarlo calcoliamo gli elementi della matrice che rappresenta P_{12} nella base assegnata,

$$\langle m', n' | P_{12} | m, n \rangle = \langle m', n' | n, m \rangle = \delta_{m'n} \delta_{n'm}; \quad (5.35)$$

poi calcoliamo gli elementi di matrice dell'hermitiano coniugato $P_{12}^\dagger$,

$$\begin{aligned} \langle m', n' | P_{12}^\dagger | m, n \rangle &= (\langle m, n | P_{12} | m', n' \rangle)^* \\ &= (\langle m, n | n', m' \rangle)^* \\ &= \langle n', m' | m, n \rangle = \delta_{n'm} \delta_{m'n}. \end{aligned} \quad (5.36)$$

Gli elementi di matrice P_{12} e di $P_{12}^\dagger$ sono tutti uguali e dunque P_{12} è hermitiano. Combinando la $P_{12} = P_{12}^{-1}$ con la $P_{12} = P_{12}^\dagger$, si vede che

$$P_{12}^\dagger P_{12} = \mathbb{1}, \quad (5.37)$$

e quindi P_{12} è unitario e conserva la norma dei vettori.

Ora consideriamo una generica osservabile A definita sul sistema di due particelle, e un suo generico autostato $|a\rangle$, in modo che

$$A|a\rangle = a|a\rangle . \quad (5.38)$$

Il vettore $|a\rangle$ rappresenta uno stato del sistema. Se le particelle sono identiche, l'applicazione dell'operatore di scambio a questo stato non deve modificare l'esito della misura di A e, quindi, per lo stato $P_{12}|a\rangle$ deve valere

$$AP_{12}|a\rangle = aP_{12}|a\rangle . \quad (5.39)$$

Se applichiamo l'operatore di scambio ad entrambi i membri e ricordiamo che $P_{12}P_{12} = \mathbb{1}$, otteniamo

$$P_{12}AP_{12}|a\rangle = aP_{12}P_{12}|a\rangle = a|a\rangle . \quad (5.40)$$

Confrontando questa con la precedente, $A|a\rangle = a|a\rangle$, vediamo che

$$P_{12}AP_{12} = A , \quad (5.41)$$

o anche

$$[A, P_{12}] = 0 . \quad (5.42)$$

Questo vale per qualsiasi osservabile e qualsiasi autostato. Concludiamo che, se le particelle sono identiche, le osservabili sono tutte invarianti per scambio di particelle e tutti gli autostati di qualsiasi osservabile sono anche autostati dell'operatore di scambio.

Ma quali sono gli autovalori di P_{12} ? Supponiamo che λ sia autovalore di P_{12} e $|\lambda\rangle$ sia il corrispondente autovettore. Allora

$$P_{12}|\lambda\rangle = \lambda|\lambda\rangle , \quad (5.43)$$

ma, essendo $P_{12}^2 = \mathbb{1}$, è anche vero che

$$|\lambda\rangle = P_{12}^2|\lambda\rangle = P_{12}P_{12}|\lambda\rangle = \lambda P_{12}|\lambda\rangle = \lambda^2|\lambda\rangle , \quad (5.44)$$

da cui segue

$$\lambda^2 = 1 . \quad (5.45)$$

Dunque gli autovalori di P_{12} sono solo 1 e -1 .

Stati in cui le due particelle hanno le stesse coordinate, come gli stati $|m, m\rangle$ e $|n, n\rangle$ sono palesemente autostati di P_{12} con autovalore 1, dato che

$$P_{12}|m, m\rangle = |m, m\rangle . \quad (5.46)$$

Invece uno stato come $|m, n\rangle$, con $m \neq n$, non è autostato di P_{12} , dato che lo scambio modifica lo stato in $|n, m\rangle$. Tuttavia, a partire da questo tipo di stati si possono facilmente trovare autostati di P_{12} tramite le combinazioni lineari

$$\frac{1}{\sqrt{2}}(|m, n\rangle + |n, m\rangle) \quad (5.47)$$

e

$$\frac{1}{\sqrt{2}}(|m, n\rangle - |n, m\rangle) . \quad (5.48)$$

Il primo è simmetrico per scambio e, dunque, corrisponde a $\lambda = 1$, mentre il secondo è antisimmetrico e corrisponde a $\lambda = -1$.

Il tutto è generalizzabile a sistemi di N particelle, con N qualsiasi. In generale possiamo scrivere

$$|1, 2, 3, \dots, N\rangle = |1\rangle \otimes |2\rangle \otimes |3\rangle \dots \otimes |N\rangle \quad (5.49)$$

dove qui, sinteticamente, ogni numero indica l'insieme delle etichette (posizione, spin, altro) che caratterizzano lo stato di quella particolare particella. Non sarebbe necessario usare numeri; potremmo scrivere $|m, n, p, q, \dots\rangle$, quello che conta è che il simbolo che sta al primo posto rappresenta gli autovalori delle osservabili compatibili che individuano lo stato della prima particella, e così via per le altre. Se l'osservabile in questione è soltanto la posizione, allora possiamo scrivere

$$|\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N\rangle = |\mathbf{r}_1\rangle \otimes |\mathbf{r}_2\rangle \otimes |\mathbf{r}_3\rangle \dots \otimes |\mathbf{r}_N\rangle \quad (5.50)$$

e la funzione d'onda associata ad un generico stato $|\alpha\rangle$ sarà

$$\psi_\alpha(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N) = \langle \mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N | \alpha \rangle . \quad (5.51)$$

Questa è una funzione di $3N$ variabili scalari (più il tempo, che qui omettiamo) che in generale obbedisce all'equazione di Schrödinger a molti corpi

$$i\hbar \frac{\partial}{\partial t} \psi_\alpha(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N) = H \psi_\alpha(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N) , \quad (5.52)$$

dove H è l'hamiltoniano del sistema.

Per particelle identiche ogni osservabile deve essere invariante per qualsiasi permutazione delle coordinate. Le permutazioni possibili di N particelle sono $N!$ e possono essere tutte rappresentate da prodotti di operatori P_{ij} che scambiano singole coppie di coordinate in questo modo:

$$P_{ij}|1, 2, \dots, i, \dots, j, \dots, N\rangle = |1, 2, \dots, j, \dots, i, \dots, N\rangle \quad (5.53)$$

ovvero, per le funzioni d'onda,

$$P_{ij}\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots, \mathbf{r}_N) = \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_j, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N) . \quad (5.54)$$

Questa scrittura va intesa in questo modo: l'operatore P_{ij} assegna alla particella i -esima le coordinate che prima aveva la particella j -esima, e viceversa. Questi operatori soddisfano le stesse proprietà che avevamo derivato nel caso di due particelle, in particolare

$$P_{ij}^{-1} = P_{ij}^\dagger = P_{ij} . \quad (5.55)$$

Invece, scambi di coppie diverse non commutano. Per esempio, supponiamo che il sistema sia composto da tre particelle e che gli stati di tali particelle siano etichettabili tramite gli autovalori di un'osservabile A . Uno stato del sistema potrebbe essere $|a, a', a''\rangle$ che corrisponde allo stato in cui una misura di A sulla prima particella darebbe esito a , sulla seconda a' e sulla terza a'' . L'operatore P_{12} agisce in questo modo

$$P_{12}|a, a', a''\rangle = |a', a, a''\rangle \quad (5.56)$$

e analogamente P_{13} e P_{23} . Sono possibili anche permutazioni multiple come questa

$$P_{12}P_{13}|a, a', a''\rangle = P_{12}|a'', a', a\rangle = |a', a'', a\rangle , \quad (5.57)$$

che corrisponde ad una permutazione ciclica dei tre indici, o altre analoghe, tutte esprimibili come prodotti di operatori di permutazione di coppia. Osserviamo poi che le permutazioni generalmente non commutano. Infatti

$$P_{13}P_{12}|a, a', a''\rangle = P_{13}|a', a, a''\rangle = |a'', a, a'\rangle \neq P_{12}P_{13}|a, a', a''\rangle . \quad (5.58)$$

Un generico hamiltoniano per N particelle ha la forma

$$H = \sum_{i=1}^N \frac{p_i^2}{2m_i} + \sum_{i=1}^N U_i(\mathbf{r}_i) + \frac{1}{2} \sum_{i \neq j} V_{ij}(|\mathbf{r}_i - \mathbf{r}_j|), \quad (5.59)$$

dove U e V rappresentano, rispettivamente, l'azione delle forze esterne ed interne al sistema. Se le particelle sono identiche, devono ovviamente avere la stessa massa e, inoltre, la forma dei potenziali U e V deve essere la stessa per tutte. Quindi l'hamiltoniano sarà

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i=1}^N U(\mathbf{r}_i) + \frac{1}{2} \sum_{i \neq j} V(|\mathbf{r}_i - \mathbf{r}_j|). \quad (5.60)$$

È facile convincersi che questo operatore commuta con qualsiasi operatore di scambio. Infatti, nelle prime due somme, lo scambio di due particelle non fa altro che scambiare due addendi, mentre nella terza somma l'argomento della funzione V è simmetrico per scambio delle coordinate. Dunque

$$P_{ij}H = HP_{ij}. \quad (5.61)$$

Se in un sistema di N particelle lo stato $|\alpha\rangle$ è un autostato dell'hamiltoniano H con autovalore E , siccome qualsiasi scambio di coppia commuta con H , si ha

$$HP_{ij}|\alpha\rangle = P_{ij}H|\alpha\rangle = P_{ij}E|\alpha\rangle = EP_{ij}|\alpha\rangle. \quad (5.62)$$

e quindi anche lo stato $P_{ij}|\alpha\rangle$ è autostato di H con lo stesso autovalore E . Questo dev'essere vero non solo per l'hamiltoniano, ma anche per qualsiasi altra osservabile. Il caso dell'hamiltoniano è particolare perché il suo spettro definisce gli stati stazionari del sistema e quello che stiamo dicendo qui è che uno stato stazionario di energia E è anche autostato di P_{ij} e deve quindi essere simmetrico o antisimmetrico per scambio della coppia ij .

A questo punto sorge però un'apparente contraddizione. Come abbiamo già visto, operatori che scambiano coppie diverse da ij non commutano con l'operatore P_{ij} e, dunque, non possono avere autostati in comune. Se $|\alpha\rangle$ è autostato comune di H e P_{ij} , non può essere anche autostato dell'operatore che scambia una coppia diversa, pur essendo ancora autostato di H . L'unico modo per risolvere la questione è ammettere che lo spettro di H sia degenere, in modo tale che ad ogni autovalore E corrispondano tanti autostati diversi quante sono le possibili permutazioni delle N particelle. Questa è nota come *degenerazione di scambio* e si applica allo spettro di qualsiasi osservabile di un sistema di particelle identiche.

In conclusione, ad un autostato di qualsiasi osservabile corrisponde un sottospazio degenere dello spazio di Hilbert costituito dagli autostati degli operatori che scambiano ciascuna coppia di particelle. È anche possibile individuare nello stesso spazio di Hilbert opportuni stati che sono simmetrici o antisimmetrici per scambio di qualsiasi coppia. Uno stato completamente simmetrico o completamente antisimmetrico è rappresentato da un vettore $|\alpha\rangle$ tale che

$$P_{ij}|\alpha\rangle = \pm|\alpha\rangle \quad (5.63)$$

per qualsiasi coppia di indici ij . L'insieme degli stati completamente simmetrici (autovalore $+1$) e quello degli stati completamente antisimmetrici (autovalore -1) formano due sottospazi dello spazio prodotto diretto nei quali tutti gli operatori di scambio di coppia commutano tra loro. Infatti, se ne prendiamo due qualsiasi e li applichiamo in sequenza su stati simmetrici,

$$P_{i'j'}P_{ij}|\alpha\rangle = P_{i'j'}|\alpha\rangle = |\alpha\rangle, \quad (5.64)$$

si vede che l'ordine degli operatori non cambia il risultato, e lo stesso per gli stati antisimmetrici,

$$P_{i'j'}P_{ij}|\alpha\rangle = -P_{i'j'}|\alpha\rangle = |\alpha\rangle. \quad (5.65)$$

Dunque in questi sottospazi tutti gli operatori P_{ij} commutano tra loro e ammettono autostati comuni, rimuovendo così la degenerazione di scambio. Per due particelle abbiamo già visto come possiamo costruire stati simmetrici e antisimmetrici combinando stati che non sono autostati di P_{12} in questo modo:

$$|\alpha\rangle_s = \frac{1}{\sqrt{2}}(|1, 2\rangle + |2, 1\rangle) \quad (5.66)$$

e

$$|\alpha\rangle_a = \frac{1}{\sqrt{2}}(|1, 2\rangle - |2, 1\rangle). \quad (5.67)$$

Qualcosa di analogo, anche se un po' più complicato, si potrà fare per sistemi con più di due particelle.

A questo punto siamo pronti ad introdurre un fatto molto importante:

in natura un qualsiasi sistema di particelle identiche si trova *sempre* in un autostato dell'operatore di scambio. L'autovalore è determinato *solo* dal tipo di particella.
Alcune particelle si trovano *sempre* in stati simmetrici, mentre altre *sempre* in stati antisimmetrici.

La simmetria per scambio ha a che fare con la separabilità delle funzioni d'onda; ad esempio, gli stati antisimmetrici, come nell'espressione (5.67), non sono mai rappresentabili da funzioni d'onda che siano il prodotto di funzioni dipendenti separatamente dalle coordinate delle singole particelle. Inoltre, la simmetria o antisimmetria ha conseguenze nel calcolo dei valori medi delle osservabili che entrano nella descrizione statistica dei sistemi a molte particelle; le particelle che si trovano sempre in stati simmetrici obbediscono alla statistica di Bose-Einstein e vengono chiamate *bosoni*, mentre quelle che si trovano sempre in stati antisimmetrici obbediscono alla statistica di Fermi-Dirac e vengono chiamate *fermioni*. La natura impone un ulteriore vincolo:

tutte le particelle con spin 0 o intero sono bosoni; tutte le particelle con spin semintero sono fermioni.

A questo fatto si dà il nome di *connessione spin-statistica*. Ad esempio, le particelle elementari come gli elettroni, i protoni, o i neutroni hanno spin $1/2$ e sono fermioni, mentre i quanti della radiazione elettromagnetica (fotoni) hanno spin 1 e sono bosoni. Sottolineiamo che non si tratta di una conseguenza della commutatività di P_{ij} con H ma, nell'ambito della meccanica quantistica non-relativistica, è un dato empirico che viene assunto come postulato della teoria. Diventa una predizione della teoria se invece si passa alla versione relativistica della meccanica quantistica, dove lo spin delle particelle emerge come una proprietà dei campi.

Una conseguenza notevole della connessione spin-statistica è il *principio di esclusione di Pauli*². Se prendiamo N fermioni e li mettiamo in uno stato individuato dal vettore $|1, 2, 3, \dots, N\rangle$, questo deve essere antisimmetrico, cioè deve valere

$$|1, 2, 3, \dots, N\rangle = -|2, 1, 3, \dots, N\rangle \quad (5.68)$$

²Va da sé che, trattandosi di una conseguenza, non può essere un principio, ma i nomi che si danno ai concetti riflettono spesso l'origine storica delle idee più che la loro collocazione all'interno della teoria compiuta, formulata a posteriori.

e lo stesso per gli scambi delle altre coppie. Ricordiamo che ogni numero i in questo *ket* rappresenta in modo sintetico l'insieme dei valori che si ottengono nella misura delle osservabili (l'insieme massimale di osservabili compatibili) nello spazio degli stati della singola particella i -esima. Possiamo chiamarle coordinate, intese in senso generale. Possono essere semplicemente le coordinate spaziali, ma anche lo spin, o il momento angolare orbitale, o combinazioni di queste, o altro. Ora supponiamo che due particelle, ad esempio la prima e la seconda, condividano lo stesso stato di singola particella, cioè abbiano gli stessi valori delle coordinate. In tal caso possiamo scrivere lo stato nella forma $|1, 1, 3, \dots, N\rangle$, dove il secondo 1 va inteso nel senso che la seconda particella ha le stesse coordinate della prima. Ma questo stato è palesemente simmetrico per scambio delle due particelle, il che non è possibile dovendo essere antisimmetrico. L'unica possibilità che abbiamo per evitare la contraddizione $|1, 1, 3, \dots, N\rangle = -|1, 1, 3, \dots, N\rangle$ è che si tratti di un vettore nullo:

$$|1, 1, 3, \dots, N\rangle = 0 \quad (5.69)$$

e lo stesso per tutte le coppie. Ne deduciamo una regola generale: *due fermioni non possono mai occupare lo stesso stato di singola particella*, cioè non possono avere entrambi le stesse coordinate. Questo principio di esclusione venne introdotto da Pauli nel 1925 su basi empiriche per spiegare la tabella periodica degli elementi: le proprietà chimiche degli elementi, infatti, si interpretano facilmente ammettendo che gli elettroni aggiunti mano a mano nelle orbite di Bohr si sistemino uno alla volta nelle orbite libere. Alla luce della meccanica quantistica e dell'equazione di Schrödinger, la stessa idea diventa ancora più efficace, trovando un'interpretazione più profonda nelle proprietà di simmetria degli stati fisici e delle funzioni d'onda. Con la versione relativistica della meccanica quantistica, poi, anche la connessione spin-statistica troverà una giustificazione teorica, chiudendo il cerchio.

Un esempio semplice può illustrare bene la differenza notevole che c'è tra bosoni e fermioni. Supponiamo di avere un sistema composto da due particelle in uno spazio limitato a due cassette, il cassetto a e il cassetto b . La misura che facciamo è la misura di posizione per l'una e per l'altra, ristretta ai due soli valori a e b , assumendo che non vi siano altri gradi di libertà. Se si tratta di palline classiche, allora gli stati possibili sono quattro:

$$a, a ; a, b ; b, a ; b, b \quad (5.70)$$

dove in ogni coppia il primo valore dice in quale cassetto si trova la prima particella, e il secondo la seconda. Come si vede, le particelle stanno nello stesso cassetto in 2 casi su 4, cioè il 50% dei casi. Se invece si tratta di due bosoni in regime quantistico, allora gli stati ammessi sono quelli rappresentati dai *ket* simmetrici

$$|a, a\rangle ; |b, b\rangle ; \frac{1}{\sqrt{2}}(|a, b\rangle + |b, a\rangle). \quad (5.71)$$

In questo caso, le particelle stanno nello stesso "cassetto" 2 volte su 3, cioè nel 67% dei casi circa, più che nel caso classico. Ma se sono fermioni, allora l'unico stato ammesso è quello antisimmetrico

$$\frac{1}{\sqrt{2}}(|a, b\rangle - |b, a\rangle) \quad (5.72)$$

e questi non staranno mai assieme nello stesso cassetto. Questa riluttanza dei fermioni giustifica la tabella di Mendeleev e rende conto essenzialmente di tutta la varietà della materia di cui è fatto il mondo che ci circonda. Il fatto che, al contrario, i bosoni amino stare assieme ha conseguenze meno evidenti, ma non meno interessanti, come la superconduttività e la superfluidità.

È importante sottolineare che il requisito di simmetria e antisimmetria per bosoni e fermioni si applica ai loro stati fisici completi, che includono tutti i gradi di libertà rilevanti.

Nell'esempio precedente avevamo assunto che l'unico grado di libertà fosse l'essere in un cassetto o in un altro (l'osservabile posizione ristretta a due valori discreti); ma se le particelle hanno anche spin, i loro stati completi saranno il prodotto diretto dei vettori che ne descrivono la posizione e dei vettori che ne descrivono lo spin. Così, ad esempio, potremmo ammettere lo stato

$$\frac{1}{\sqrt{2}}|a, a\rangle \otimes (|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle) \quad (5.73)$$

che è globalmente antisimmetrico, come ci si aspetta per due fermioni di spin $1/2$, ma descrive due fermioni che stanno nello stesso cassetto. In altri termini, per due particelle di spin $1/2$ che si muovono nello spazio, la funzione d'onda può essere simmetrica per scambio delle coordinate spaziali se il loro stato di spin è antisimmetrico, e viceversa. Per questo motivo, due elettroni con spin opposto possono occupare, ad esempio, lo stesso stato $|n, l, m\rangle$ di un atomo e la stessa regola spiega il grosso delle proprietà degli elementi della tavola periodica. Spiega anche il motivo per cui i legami covalenti nelle molecole si formano solo con gli elettroni nello stato di singoletto, antisimmetrico, di spin zero. Infatti, questo permette agli stessi elettroni di avere una funzione d'onda spaziale simmetrica per scambio dei due, come fossero bosoni; di conseguenza la probabilità di trovare i due elettroni contemporaneamente nella regione centrale della molecola, tra i due nuclei positivi, è maggiore rispetto al caso di una funzione d'onda antisimmetrica. Questa maggior probabilità di trovarli vicini è equivalente, come effetto, ad un eccesso di carica negativa tra i due nuclei; in questo modo i nuclei sentono una forza attrattiva che dà luogo al legame molecolare.

Il fatto che due fermioni in uno stato di tripletto di spin (simmetrico) stiano mediamente più lontani tra loro di quanto lo sarebbero gli stessi fermioni in uno stato di singoletto, oppure di quanto lo sarebbero due bosoni, lo possiamo immaginare come l'effetto dell'azione di una sorta di interazione repulsiva o attrattiva. Questa è detta *forza di scambio*. Però non va confusa con una vera forza, dato che si tratta di un puro effetto geometrico dovuto al requisito di simmetrizzazione dei vettori di stato; è una conseguenza della statistica fermionica o bosonica propria delle particelle, cioè qualcosa di intrinsecamente quantistico, senza alcuna analogia classica.

La connessione spin-statistica vale anche per particelle composite, formate a loro volta da altre particelle. Ad esempio vale per gli atomi, che sono composti di elettroni, protoni e neutroni. Questi ultimi, presi singolarmente, sono tutti fermioni; ma gli atomi, in quanto tali, sono rappresentati da stati simmetrici o antisimmetrici? Prendiamo ad esempio due atomi di idrogeno, che hanno soltanto un elettrone e un protone ciascuno. Lo stato complessivo del sistema è esprimibile tramite le coordinate di ciascuna particella elementare, nella forma di combinazioni lineari di stati del tipo $|1_p, 1_e, 2_p, 2_e\rangle$, dove il numero specifica quale atomo e gli indici indicano gli elettroni e i protoni. Ora, le combinazioni lineari devono essere tali che lo scambio di due fermioni cambi il segno del vettore. Ma scambiare le coordinate dei due atomi corrisponde ad effettuare due scambi di coppie di fermioni; quando scambiamo le coordinate dei due protoni, il vettore di stato cambia segno, e quando scambiamo anche quelle degli elettroni, cambia segno di nuovo. Così troviamo che scambiando i due atomi il vettore di stato deve rimanere uguale; dunque si tratta di un vettore simmetrico e ciascun atomo si comporta come un bosone. D'altra parte, l'atomo è composto da particelle con $s_e = s_p = 1/2$ e la somma degli spin dà $s = 0$ o $s = 1$. Dunque l'atomo d'idrogeno ha spin intero. Se aggiungiamo un neutrone, l'atomo diventa quello del deuterio. Lo scambio di coordinate tra due atomi di deuterio comporta tre scambi di fermioni e lo stato complessivo sarà antisimmetrico, mentre lo spin totale sarà semintero. In generale, se un atomo contiene un numero dispari di fermioni è esso stesso un fermione, altrimenti è un bosone. Nel limite classico, quando gli atomi sono mediamente molto lontani tra loro e si comportano come particelle indistinguibili, il fatto che siano bosoni o fermioni è irrilevante. Diventa invece cruciale quando le loro distanze sono comparabili con la lunghezza d'onda di de Broglie; in

tal caso, il comportamento dei sistemi dipende fortemente dalla statistica. Ad esempio, ^3He e ^4He sono atomi chimicamente identici, che differiscono solo per la massa. Al di sotto di una certa temperatura manifestano effetti quantistici e si comportano in modo totalmente diverso: il liquido composto da ^4He (bosone) diventa superfluido, cioè capace di scorrere senza alcuna viscosità, intorno ai 2 K, mentre l' ^3He (fermione) rimane un liquido normale. Occorre scendere alla scala dei mK per ottenere la superfluidità dell' ^3He , ma il tipo di superfluidità è completamente diverso da quello dell' ^4He .

Per capire come entrano in gioco le distanze, riprendiamo la definizione (5.11) della funzione d'onda di un sistema di due particelle, $\psi_\alpha(\mathbf{r}_1, \mathbf{r}_2) = \langle \mathbf{r}_1, \mathbf{r}_2 | \alpha \rangle$. Supponiamo di mettere un rivelatore nel punto $\mathbf{r}$, che scatti quando una particella si trova in un volumetto d^3r centrato in quel punto. Se le particelle sono distinguibili, il rivelatore può essere congegnato opportunamente da identificare l'una o l'altra. Potremo così avere, separatamente, il conteggio degli eventi in cui la particella 1 si trova in $\mathbf{r}$, indipendentemente da dove si trova la particella 2, e degli eventi in cui la particella 2 si trova in $\mathbf{r}$, indipendentemente da dove si trova la particella 1. In tante repliche dell'esperimento, avendo preparato ogni volta il sistema nello stato $|\alpha\rangle$, il numero di volte in cui il rivelatore scatta per la particella 1 sarà proporzionale a

$$d^3r \int d^3r_2 |\psi(\mathbf{r}, \mathbf{r}_2)|^2 \quad (5.74)$$

e il numero di volte in cui scatta per la particella 2 sarà proporzionale a

$$d^3r \int d^3r_1 |\psi(\mathbf{r}_1, \mathbf{r})|^2. \quad (5.75)$$

Ma se le due particelle sono identiche, nessun rivelatore saprà indicarci se scatta per l'una o per l'altra. Potremo solo contare quante volte una particella viene rivelata indipendentemente da dove si trova l'altra, ma senza distinguere quale delle due ha fatto realmente scattare il rivelatore. Il numero di conteggi sarà dunque proporzionale all'unione dei due casi:

$$d^3r \int d^3r_2 |\psi(\mathbf{r}, \mathbf{r}_2)|^2 + d^3r \int d^3r_1 |\psi(\mathbf{r}_1, \mathbf{r})|^2, \quad (5.76)$$

che, rinominando la variabile d'integrazione, può essere scritto così:

$$d^3r \int d^3r' (|\psi(\mathbf{r}, \mathbf{r}')|^2 + |\psi(\mathbf{r}', \mathbf{r})|^2). \quad (5.77)$$

La cosa diventa interessante nel caso di particelle soggette a due buche di potenziale identiche e sufficientemente distanti. Si tratta del problema discusso precedente dei due cassetti a e b , ma affrontato in modo più quantitativo nello spazio delle coordinate. Supponiamo che le particelle siano identiche e non interagiscano tra loro; in tal caso, la funzione d'onda complessiva è fattorizzabile nella forma di un prodotto di due funzioni d'onda di singola particella, soluzioni dell'equazione di Schrödinger per una particella in un potenziale a doppia buca. Consideriamo gli stati legati di energia più bassa, le cui funzioni d'onda di singola particella siano non nulle solo nell'intorno di ciascuna buca. Sia dunque $\varphi_a(\mathbf{r})$ la funzione d'onda localizzata nell'intorno della buca a e $\varphi_b(\mathbf{r})$ quella localizzata nell'intorno della buca b , entrambe di norma unitaria. Con queste assunzioni, la funzione d'onda $\psi_{aa}(\mathbf{r}_1, \mathbf{r}_2) = \langle \mathbf{r}_1, \mathbf{r}_2 | aa \rangle$, in cui entrambe le particelle stanno nella buca a , assume la forma $\psi_{aa}(\mathbf{r}_1, \mathbf{r}_2) = \varphi_a(\mathbf{r}_1)\varphi_a(\mathbf{r}_2)$, e analogamente $\psi_{bb}(\mathbf{r}_1, \mathbf{r}_2) = \varphi_b(\mathbf{r}_1)\varphi_b(\mathbf{r}_2)$. Invece per i due stati $(1/\sqrt{2})(|a, b\rangle \pm |b, a\rangle)$, l'uno simmetrico e l'altro antisimmetrico, la funzione d'onda sarà

$$\psi_{S/A}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}[\varphi_a(\mathbf{r}_1)\varphi_b(\mathbf{r}_2) \pm \varphi_b(\mathbf{r}_1)\varphi_a(\mathbf{r}_2)]. \quad (5.78)$$

Ora applichiamo il ragionamento di prima sul numero di conteggi dei rivelatori e poniamoci il problema di sapere in quali condizioni i rivelatori sono in grado distinguere gli stati simmetrici da quelli antisimmetrici. Gli stati $|aa\rangle$ e $|bb\rangle$ non sono utili allo scopo, essendo entrambi simmetrici. Prepariamo quindi le particelle in uno stato $|\alpha\rangle$ rappresentato da una o l'altra delle funzione d'onda $\psi_{S/A}(\mathbf{r}_1, \mathbf{r}_2)$. Per quanto detto sopra, la probabilità $\mathcal{P}$ che un rivelatore in $\mathbf{r}$ scatti è proporzionale a

$$\mathcal{P} \propto d^3r \int d^3r' (|\psi_{S/A}(\mathbf{r}, \mathbf{r}')|^2 + |\psi_{S/A}(\mathbf{r}', \mathbf{r})|^2) \quad (5.79)$$

ovvero

$$\mathcal{P} \propto \frac{d^3r}{2} \int d^3r' [|\varphi_a(\mathbf{r})\varphi_b(\mathbf{r}') \pm \varphi_b(\mathbf{r})\varphi_a(\mathbf{r}')|^2 + |\varphi_a(\mathbf{r}')\varphi_b(\mathbf{r}) \pm \varphi_b(\mathbf{r}')\varphi_a(\mathbf{r})|^2] ; \quad (5.80)$$

sviluppando i calcoli dei moduli quadri e tenendo conto della normalizzazione delle funzioni φ , si ottiene

$$\begin{aligned} \mathcal{P} &\propto d^3r |\varphi_a(\mathbf{r})|^2 + d^3r |\varphi_b(\mathbf{r})|^2 \\ &\pm 2 d^3r \operatorname{Re} \left(\varphi_a^*(\mathbf{r})\varphi_b(\mathbf{r}) \int d^3r' \varphi_a(\mathbf{r}')\varphi_b^*(\mathbf{r}') \right). \end{aligned} \quad (5.81)$$

In questa espressione, l'effetto della simmetrizzazione o antisimmetrizzazione degli stati fisici delle due particelle identiche sta tutto nel termine che contiene l'integrale

$$\int d^3r' \varphi_a(\mathbf{r}')\varphi_b^*(\mathbf{r}'). \quad (5.82)$$

Se questo integrale è nullo o trascurabile, allora la scelta tra il segno $+$ e il segno $-$ nell'espressione (5.78) non cambia il risultato e le particelle si comportano come fossero distinguibili. Questo succede se le buche sono abbastanza distanti tra loro da rendere trascurabile la sovrapposizione tra le funzioni d'onda φ_a e φ_b . In altri termini, l'indistinguibilità, e con essa la necessità di simmetrizzare o antisimmetrizzare gli stati, entra in gioco soltanto quando le funzioni d'onda hanno una sovrapposizione spaziale non trascurabile, ad esempio quando due elettroni convivono nella stessa molecola, ma non quando stanno in atomi diversi, lontani tra loro. Per lo stesso motivo non dobbiamo preoccuparci di dover scrivere un'unica enorme combinazione antisimmetrica per tutti gli elettroni dell'universo, anche se sono tutti, in linea di principio, fermioni indistinguibili.

5.1.3 Entanglement

L'ultimo argomento ci tranquillizza: se le particelle (gli elettroni, gli atomi, le molecole, ecc.) sono abbastanza lontane tra loro, possiamo trattarle come particelle classiche almeno dal punto di vista della loro distinguibilità. Ma la natura è più subdola (o più interessante, a seconda dei punti di vista) di quanto ci aspettiamo e riesce a proporci situazioni meno tranquillizzanti. Nel nostro discorso di prima abbiamo ragionato sulla sovrapposizione delle funzioni d'onda, ma se includiamo anche i gradi di libertà di spin, dovremmo tener conto anche di quale sia, appunto, lo stato di spin delle particelle. Ora, si potrebbe dire (ed è vero nella gran parte dei casi) che lo stato di spin di una particella che si trova in una buca di potenziale sia indipendente da quello di una particella che si trova in un'altra buca spazialmente lontana, ma non possiamo darlo per scontato.

Prendiamo ad esempio due fermioni di spin $1/2$ preparati al tempo iniziale $t = 0$ in una stessa buca e in uno stato rappresentato da una funzione d'onda simmetrica per scambio del coordinate spaziali. Essendo fermioni, dovranno trovarsi in uno stato di singoletto di spin,

come in (5.73), per garantire l'antisimmetria globale. Poi supponiamo che i due fermioni vengano allontanati l'uno dall'altro tramite un potenziale esterno che dipende da $\mathbf{r}$ e da t ; il potenziale non agisce sullo spin delle particelle, semplicemente le separa spazialmente³. Dopo un certo tempo le due particelle saranno abbastanza lontane da far sì che le loro funzioni d'onda spaziali non si sovrappongono, tuttavia il vettore d'onda che le descrive contiene ancora traccia del loro stato iniziale, dovendo mantenere la forma antisimmetrica

$$|\alpha\rangle = \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle) = \frac{1}{\sqrt{2}}(|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle) \quad (5.83)$$

per la parte di spin. A questo punto, immaginiamo di avere a disposizione un apparato di Stern-Gerlach, o un dispositivo equivalente, con cui possiamo misurare lo spin lungo z della particella che si trova in una delle due buche e solo in quella. Supponiamo di trovare il valore $\hbar/2$. In tal caso lo stato del sistema collasserà dal vettore scritto sopra al vettore $|\uparrow, \downarrow\rangle$. Ma il collasso riguarda tutto il sistema, non solo la particella di cui abbiamo misurato lo spin. Anche l'altra, dopo la misura si troverà in uno stato diverso da quello iniziale. Se misurassimo il suo spin lungo z troveremmo con certezza $-\hbar/2$. Dunque la misura (locale) eseguita su una delle due particelle modifica anche lo stato dell'altra: questa è una conseguenza del fatto che il vettore che rappresenta il sistema non è scrivibile come un prodotto diretto della forma

$$|\alpha\rangle \neq |m_1\rangle \otimes |m_2\rangle, \quad (5.84)$$

dove m_1 e m_2 sono gli autovalori dello spin lungo z di ciascuna particella, ma è invece una combinazione lineare di prodotti diretti

$$|\alpha\rangle = \sum_{m_1, m_2} c_{m_1, m_2} |m_1\rangle \otimes |m_2\rangle, \quad (5.85)$$

dove le costanti numeriche c_{m_1, m_2} non possono essere scritte come prodotto di una costante che dipende solo da m_1 e un'altra che dipende solo da m_2 . Se le costanti potessero essere fattorizzate come $c_{m_1} c_{m_2}$ allora la somma precedente potrebbe essere separata in questo modo

$$|\alpha\rangle = \left(\sum_{m_1} c_{m_1} |m_1\rangle \right) \otimes \left(\sum_{m_2} c_{m_2} |m_2\rangle \right), \quad (5.86)$$

e lo stato complessivo sarebbe il prodotto diretto di due vettori in spazi vettoriali diversi. Di questo avevamo già parlato all'inizio del capitolo, dove avevamo detto che stati rappresentati da vettori non separabili, come il vettore in (5.85), sono detti *entangled*. Qui vediamo che il singoletto di spin è uno stato puro nello spazio prodotto diretto dei due sottospazi dello spin di ciascuna particella, $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, ma non è separabile (fattorizzabile) in stati puri di ciascun sottospazio $\mathcal{H}_1$ e $\mathcal{H}_2$. Si tratta dunque di uno stato *entangled*.

L'*entanglement* è una caratteristica peculiare della meccanica quantistica ed è un concetto che può essere applicato a qualsiasi sistema composto da sottosistemi. Una delle sue conseguenze è quella che Einstein chiamava *spooky action at a distance*: la misura locale delle proprietà di una particella produce un effetto istantaneo sullo stato di un'altra particella, lontana quanto si vuole. La discussione sulle implicazioni di tale idea si è protratta per decenni ed è ancora viva. Probabilmente avrete già sentito parlare della disuguaglianza di Bell, delle variabili nascoste e delle diatribe relative alla validità del *realismo locale*, quanto meno perché questi concetti emergono spesso anche a livello di comunicazione e divulgazione

³Potenziali del genere esistono; ad esempio, per spostare atomi si possono usare le cosiddette pinzette ottiche, o *optical tweezer*. Ma i ragionamenti che facciamo qui si applicano anche quando i meccanismi che separano spazialmente le particelle sono altri, come ad esempio nel caso di particelle generate a coppie in processi di decadimento e che si allontanano spontaneamente in direzioni diverse.

scientifica. Una trattazione rigorosa di questi temi richiederebbe più tempo di quanto ne sia rimasto per queste lezioni. Avrete modo di apprezzare il senso e l'utilità dell'*entanglement* nei corsi più avanzati.

5.2 Matrice densità

Nel primo capitolo, introducendo la definizione di misura in meccanica quantistica e facendo l'esempio dell'esperimento di Stern-Gerlach, avevamo distinto gli stati puri dagli stati misti. Da quel momento in poi tutta la nostra trattazione ha riguardato soltanto gli stati puri. Ora riprendiamo gli stati misti e cerchiamo di inserirli nella teoria.

Per cominciare ricordiamo che uno stato puro è uno stato rappresentato da un vettore del tipo

$$|\alpha\rangle = \sum_j c_j |a_j\rangle, \quad (5.87)$$

dove i coefficienti c_j sono numeri complessi e i vettori $\{|a_1\rangle, |a_2\rangle, |a_3\rangle, \dots\}$ sono gli autovettori di un'osservabile A che costituiscono una base ortonormale completa per gli stati che descrivono il sistema. La quantità $|c_j|^2$ è la probabilità di ottenere il valore a_j nella misura di A in una sequenza di misure effettuate su repliche dello stesso sistema preparato sempre nello stato $|\alpha\rangle$. Il valore di aspettazione di una qualsiasi osservabile Q sarà dunque

$$\langle Q \rangle = \langle \alpha | Q | \alpha \rangle = \sum_{i,j} c_i^* c_j \langle a_i | Q | a_j \rangle, \quad (5.88)$$

che nel caso di $Q = A$ si riduce a

$$\langle A \rangle = \langle \alpha | A | \alpha \rangle = \sum_j |c_j|^2 a_j, \quad (5.89)$$

con la condizione

$$\sum_j |c_j|^2 = 1. \quad (5.90)$$

Come sappiamo, il tutto può essere generalizzato per includere gli spettri continui, ma qui limitiamoci a quelli discreti per semplicità.

Ora supponiamo che il meccanismo che genera repliche dello stesso sistema non produca sempre lo stesso stato puro ma ne produca diversi, $|\alpha_1\rangle, |\alpha_2\rangle, |\alpha_3\rangle, \dots$, in modo casuale, con una distribuzione statistica tale per cui la probabilità di ottenere lo stato $|\alpha_1\rangle$ sia w_1 , quella di ottenere $|\alpha_2\rangle$ sia w_2 , ecc., dove i pesi statistici w_p sono numeri reali compresi tra 0 e 1 e soggetti alla condizione $\sum_p w_p = 1$. L'insieme (o *ensemble*) delle repliche corrisponde ad una *miscela* di stati puri. Possiamo chiamarlo *stato misto* o *ensemble misto*. Se vogliamo ottenere il valor medio dell'osservabile generica Q su questo *ensemble*, dobbiamo sommare sui valori di aspettazione che otteniamo per ciascuno stato puro, pesandoli con le probabilità w_i :

$$\langle Q \rangle_m = \sum_p w_p \langle \alpha_p | Q | \alpha_p \rangle = \sum_p w_p \sum_{i,j} c_i^{*(p)} c_j^{(p)} \langle a_i | Q | a_j \rangle, \quad (5.91)$$

dove i coefficienti $c_j^{(p)} = \langle a_j | \alpha_p \rangle$ sono quelli della decomposizione dello stato $|\alpha_p\rangle$ sulla solita base degli autovettori di A . Possiamo anche scrivere

$$\begin{aligned} \langle Q \rangle_m &= \sum_p w_p \sum_{i,j} \langle \alpha_p | a_i \rangle \langle a_j | \alpha_p \rangle \langle a_i | Q | a_j \rangle \\ &= \sum_{i,j} \left(\sum_p w_p \langle a_j | \alpha_p \rangle \langle \alpha_p | a_i \rangle \right) \langle a_i | Q | a_j \rangle, \end{aligned} \quad (5.92)$$

dove abbiamo riordinato le somme in modo da far comparire l'espressione in parentesi tonda, che non dipende dalla scelta dell'osservabile Q , ma solo da com'è stato preparato l'*ensemble* e da quale base si è scelta per rappresentare i vettori di stato. A questo punto viene naturale definire un nuovo operatore, detto *operatore densità*, fatto così

$$\rho = \sum_p w_p |\alpha_p\rangle\langle\alpha_p| \quad (5.93)$$

in modo che la parentesi di prima corrisponde al suo elemento di matrice

$$\rho_{ji} = \langle a_j | \rho | a_i \rangle = \sum_p w_p \langle a_j | \alpha_p \rangle \langle \alpha_p | a_i \rangle \quad (5.94)$$

e la media dell'osservabile Q nell'*ensemble* diventa

$$\begin{aligned} \langle Q \rangle_m &= \sum_{i,j} \langle a_j | \rho | a_i \rangle \langle a_i | Q | a_j \rangle \\ &= \sum_j \langle a_j | \rho \left(\sum_i |a_i\rangle\langle a_i| \right) Q | a_j \rangle = \sum_j \langle a_j | \rho Q | a_j \rangle, \end{aligned} \quad (5.95)$$

ma questa è la somma degli elementi diagonali della matrice che rappresenta ρQ , e quindi

$$\langle Q \rangle_m = \text{tr}(\rho Q). \quad (5.96)$$

La matrice che ha come suoi elementi le espressioni (5.94) si chiama *matrice densità* e contiene tutta l'informazione fisica che possiamo avere sull'*ensemble misto*. Inoltre, dato che la traccia di una matrice non dipende dalla base che si sceglie per scriverla, anche il valor medio di una qualsiasi osservabile Q non dipende dalla scelta della base.

Essendo i w_p numeri reali, l'operatore densità definito in (5.93) è hermitiano: $\rho^\dagger = \rho$. Inoltre per la sua traccia vale

$$\text{tr}(\rho) = \sum_i \langle a_i | \rho | a_i \rangle = \sum_i \sum_p w_p \langle a_i | \alpha_p \rangle \langle \alpha_p | a_i \rangle \quad (5.97)$$

ovvero, riordinando i termini e usando la normalizzazione degli stati puri,

$$\begin{aligned} \text{tr}(\rho) &= \sum_p w_p \langle \alpha_p | \left(\sum_i |a_i\rangle\langle a_i| \right) | \alpha_p \rangle \\ &= \sum_p w_p \langle \alpha_p | \alpha_p \rangle = \sum_p w_p = 1. \end{aligned} \quad (5.98)$$

Dunque, la matrice densità ha sempre traccia pari a 1.

Possiamo mostrare che $\langle \beta | \rho | \beta \rangle \geq 0$ per qualsiasi stato generico $|\beta\rangle$. Infatti,

$$\langle \beta | \rho | \beta \rangle = \sum_p w_p \langle \beta | \alpha_p \rangle \langle \alpha_p | \beta \rangle = \sum_p w_p |\langle \beta | \alpha_p \rangle|^2 \geq 0, \quad (5.99)$$

dato che la somma contiene solo addendi positivi o al più nulli. Questa disuguaglianza rimane la stessa indipendentemente dalla base che si usa per rappresentare i vettori di stato e gli operatori. Infatti, se U è un operatore unitario ($UU^\dagger = \mathbb{1}$) che cambia la base, possiamo usare l'unitarietà per scrivere la disuguaglianza precedente così:

$$0 \leq \langle \beta | \rho | \beta \rangle = \langle \beta | UU^\dagger \rho UU^\dagger | \beta \rangle = (\langle \beta | U) (U^\dagger \rho U) (U^\dagger | \beta) \quad (5.100)$$

e definire il nuovo operatore $\rho' = U^\dagger \rho U$ e il nuovo vettore $|\beta'\rangle = U^\dagger |\beta\rangle$, per cui vale

$$0 \leq \langle \beta' | \rho' | \beta' \rangle. \quad (5.101)$$

Ora, supponiamo di aver scelto la trasformazione di base in modo tale che la matrice densità diventi diagonale e, quindi, che si possa scrivere

$$\rho' = \sum_j \lambda_j |j\rangle \langle j|, \quad (5.102)$$

dove λ_j sono gli autovalori e $|j\rangle$ sono gli autovettori. Inserendo questa forma nella disuguaglianza precedente otteniamo

$$0 \leq \langle \beta' | \rho' | \beta' \rangle = \sum_j \lambda_j \langle \beta' | j \rangle \langle j | \beta' \rangle = \sum_j \lambda_j |\langle j | \beta' \rangle|^2. \quad (5.103)$$

La quantità a destra è una somma dei λ_j moltiplicati per dei numeri positivi che dipendono dallo stato $|\beta'\rangle$; ma la scelta di questo stato è del tutto arbitraria e, dunque, l'unico modo per garantire che la somma sia globalmente non negativa per qualsiasi $|\beta'\rangle$ consiste nell'imporre che tutti i λ_j siano non negativi: $\lambda_j \geq 0$. Altrimenti, se anche uno solo dei λ_j fosse negativo, si potrebbe ad esempio identificare $|\beta'\rangle$ con $|j\rangle$ per violare la disuguaglianza. Inoltre, la loro somma è la traccia della matrice densità, che vale 1 indipendentemente dalla base. Questo implica che ciascun λ_j è un numero reale compreso tra 0 e 1:

$$0 \leq \lambda_j \leq 1. \quad (5.104)$$

Una conseguenza interessante è che la traccia dell'operatore ρ^2 non può essere maggiore di 1. Infatti,

$$\text{tr}(\rho^2) = \text{tr}(\rho'^2) = \sum_j \lambda_j^2 \leq 1. \quad (5.105)$$

Se le probabilità w_p sono tutte nulle tranne quella per un certo p' per cui vale $w_{p'} = 1$, allora l'*ensemble* corrisponde ad un insieme di repliche di un sistema nello stesso stato puro $|\alpha_{p'}\rangle$. In tal caso l'operatore densità si riduce al semplice prodotto esterno $|\alpha_{p'}\rangle \langle \alpha_{p'}|$. Inoltre vale

$$\rho^2 = |\alpha_{p'}\rangle \langle \alpha_{p'}| \alpha_{p'}\rangle \langle \alpha_{p'}| = |\alpha_{p'}\rangle \langle \alpha_{p'}| = \rho, \quad (5.106)$$

ovvero

$$\rho(\rho - 1) = 0. \quad (5.107)$$

e anche

$$\text{tr}(\rho^2) = \text{tr}(\rho) = 1. \quad (5.108)$$

Questo insieme di repliche di uno stato puro lo possiamo anche chiamare *ensemble puro*. Per rappresentarlo è possibile scegliere una base tale che ρ sia una matrice diagonale, con un solo valore pari a 1 e tutti gli altri 0. Una matrice siffatta soddisfa le condizioni appena scritte. Inoltre, si verifica anche la condizione $\text{tr}(\rho^2) = 1$, che rappresenta il valore massimo concesso dalla disuguaglianza (5.105), valida per qualsiasi *ensemble*. In effetti, la traccia di ρ^2 può essere usata come stima del grado di purezza di un *ensemble*: quanto più si avvicina a 1, tanto più l'*ensemble* approssima uno stato puro.

È anche possibile studiare la dinamica dell'operatore densità. Supponiamo che al tempo $t = 0$ venga preparato un *ensemble misto* di stati puri $|\alpha_p, 0\rangle$ con probabilità assegnate w_p . Supponiamo anche che le condizioni che determinano le probabilità w_p non dipendano dal tempo, ovvero, in altri termini, il meccanismo che genera gli stati in modo casuale è

stazionario, indisturbato. Allora l'evoluzione temporale riguarda solo gli stati puri, che obbediscono all'equazione di Schrödinger

$$i\hbar \frac{\partial}{\partial t} |\alpha_p, 0; t\rangle = H |\alpha_p, 0; t\rangle. \quad (5.109)$$

Possiamo quindi calcolare la derivata temporale dell'operatore densità in questo modo:

$$\begin{aligned} i\hbar \frac{\partial \rho}{\partial t} &= i\hbar \frac{\partial}{\partial t} \sum_p w_p |\alpha_p, 0; t\rangle \langle \alpha_p, 0; t| \\ &= \sum_p w_p (H |\alpha_p, 0; t\rangle \langle \alpha_p, 0; t| - |\alpha_p, 0; t\rangle \langle \alpha_p, 0; t| H) \\ &= H\rho - \rho H, \end{aligned} \quad (5.110)$$

e dunque

$$\boxed{\frac{\partial \rho}{\partial t} = -\frac{1}{i\hbar} [\rho, H]}. \quad (5.111)$$

Questa assomiglia ad una equazione del moto di Heisenberg per l'operatore ρ , ma non lo è. Per esserlo, avrebbe pure il segno sbagliato. In realtà qui stiamo lavorando nella rappresentazione di Schrödinger per i vettori $|\alpha\rangle$ con cui ρ è definito. Piuttosto, è molto utile l'analogia con un'altra equazione, l'*equazione di Liouville* valida in meccanica statistica classica

$$\frac{\partial \rho_{\text{cl}}}{\partial t} = -[\rho_{\text{cl}}, H]_{\text{Poisson}}, \quad (5.112)$$

dove H è la funzione di Hamilton e ρ_{cl} è la densità dei punti che rappresentano il sistema nello spazio delle fasi. In effetti, il nome stesso di operatore densità in meccanica quantistica viene dalla somiglianza con la densità classica ρ_{cl} . Entrambe si prestano ad essere usate, con un ruolo simile, nella descrizione statistica dei sistemi a molte particelle. Come gli studenti avranno modo di vedere in altri corsi (o leggendo il Sakurai), l'operatore densità ρ può essere usato per definire l'*entropia* di un sistema quantistico, ad esempio tramite la relazione $S = -k_B \text{tr}(\rho \ln \rho)$, detta *entropia di von Neumann*, dove k_B è la solita costante di Boltzmann.

Da queste considerazioni si capisce l'importanza dell'operatore densità: se noto, dà l'informazione completa su un sistema fisico, sia in termini di distribuzione statistica classica (le probabilità w_p) che di probabilità quantistiche, permettendo di calcolare il valor medio di qualsiasi osservabile; inoltre se ne conosce l'evoluzione temporale e lo si può usare come base per la meccanica statistica e la termodinamica.

Queste sono definizioni e proprietà generali. Per dare concretezza e per tornare al nostro solito esempio dello Stern-Gerlach, proviamo a scrivere la matrice densità per un sistema di particelle identiche, neutre, di spin $1/2$, non interagenti tra loro, come gli atomi di argento che escono dal forno e passano attraverso le sequenze di magneti descritte nel primo capitolo. Gli stati di spin appartengono ad uno spazio vettoriale bidimensionale. Il caso più semplice è quello degli atomi all'uscita di uno Stern-Gerlach lungo z in cui si lascia passare solo la componente nello stato $|\uparrow\rangle$. Si tratta di uno stato puro. L'operatore densità è semplicemente $\rho = |\uparrow\rangle \langle \uparrow|$ e la matrice densità, espressa nella base degli autostati di S_z è

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (5.113)$$

Analogamente, all'uscita di uno Stern-Gerlach diretto lungo x possiamo selezionare un fascio di atomi (un *ensemble puro*) tutti nello stato $|S_x \uparrow\rangle$, ad esempio. Dato che $|S_x \uparrow\rangle = (1/\sqrt{2})(|\uparrow\rangle$

$\rangle + |\downarrow\rangle\rangle$, l'operatore densità è

$$\rho = \frac{1}{2} (|\uparrow\rangle + |\downarrow\rangle)(\langle\uparrow| + \langle\downarrow|) \quad (5.114)$$

e la matrice densità, nella solita base degli autostati di S_z , diventa

$$\rho = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (5.115)$$

ed è facile vedere che questa matrice diventa identica alla precedente se si cambia la base da quella degli autostati di S_z a quella degli autostati di S_x , com'è ragionevole aspettarsi. La traccia della matrice resta comunque uguale a 1 come dev'essere. Infine, moltiplicando la matrice per se stessa, si vede che $\rho^2 = \rho$, come previsto per gli *ensemble puri*.

Fin qui non c'è nulla di speciale. Gli stati puri li sapevamo già trattare e possiamo usarli per calcolare valori di aspettazione di osservabili senza ricorrere alla matrice densità. Ora però supponiamo di preparare una miscela di due stati puri. Ad esempio prendiamo due apparati di Stern-Gerlach, non in sequenza, ma in parallelo. Dal primo selezioniamo gli atomi in $|\uparrow\rangle$ e gli altri li blocchiamo; dal secondo selezioniamo quelli in $|\downarrow\rangle$. Facciamo in modo che le intensità dei due fasci siano uguali e, inoltre, disponiamo gli apparati in modo che i due fasci si uniscano in un unico fascio. Quest'ultimo sarà dunque rappresentato da una miscela, un *ensemble misto*, di due stati puri, con peso statistico $w = 1/2$ uguale per entrambi. L'operatore densità è

$$\rho = \frac{1}{2} |\uparrow\rangle\langle\uparrow| + \frac{1}{2} |\downarrow\rangle\langle\downarrow| \quad (5.116)$$

e la matrice densità è

$$\rho = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2} \mathbb{1}. \quad (5.117)$$

Metà degli atomi si trovano in $|\uparrow\rangle$ e metà in $|\downarrow\rangle$. Notiamo, tra l'altro, che la matrice densità $\rho = (1/2)\mathbb{1}$ si ottiene anche con una miscela bilanciata di stati $|S_x \uparrow\rangle$ e $|S_x \downarrow\rangle$, oppure $|S_y \uparrow\rangle$ e $|S_y \downarrow\rangle$. In generale, essa descrive qualsiasi miscela di stati per cui la probabilità di ottenere i valori $\pm\hbar/2$ è la stessa in qualsiasi direzione si misuri lo spin. Rappresenta quindi un sistema di particelle *non polarizzate*, in cui l'orientazione dello spin ha distribuzione completamente casuale e isotropa, come nel caso degli atomi di argento che escono dal forno nell'esperimento di Stern-Gerlach. Tra l'altro, se $\rho = (1/2)\mathbb{1}$, allora il valor medio di un qualsiasi operatore che agisce nello spazio dello spin vale $\langle Q \rangle_m = \text{tr}(\rho Q) = (1/2)\text{tr}(Q)$, che applicato alle osservabili S_x , S_y e S_z , aventi traccia nulla, implica l'annullarsi dei loro valori medi, come ci si aspetta per una distribuzione isotropa.

Occorre insistere (e qui si gioca la nostra comprensione della meccanica quantistica) sul fatto che una miscela di stati puri *non* è equivalente ad una combinazione lineare degli stessi stati puri. Un conto è mescolare stati in senso statistico, un altro è farne combinazioni lineari con coefficienti fissati. I coefficienti di una combinazione lineare, oltre ad avere dei moduli, hanno anche una fase relativa fissata (lo stato complessivo ha una sua *coerenza di fase*); la miscela non contiene alcuna fase relativa (i pesi statistici sono numeri reali e gli stati sono mescolati in modo *incoerente*). Alcune osservabili possono essere insensibili alla differenza. Ad esempio, nel caso della miscela descritta dall'operatore densità (5.116), un terzo apparato di Stern-Gerlach che misuri il valore di S_z sull'*ensemble misto* troverebbe due fasci in uscita di uguale intensità, corrispondenti agli autovalori $\pm\hbar/2$, esattamente come farebbe lo stesso apparato se analizzasse l'*ensemble puro* di atomi in $|S_x \uparrow\rangle$, dato che i moduli

quadri dei coefficienti della decomposizione di $|S_x \uparrow\rangle$ nella base di S_z valgono $1/2$ come le probabilità w nella miscela. Inoltre, il valor medio di S_z sulla miscela è nullo, dato che

$$\langle S_z \rangle_m = \text{tr}(\rho S_z) = \frac{1}{2} \text{tr}(\mathbb{1} S_z) = \frac{\hbar}{4} \text{tr} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 0, \quad (5.118)$$

esattamente come lo è il valore di aspettazione di S_z nello stato puro $|S_x \uparrow\rangle$:

$$\langle S_x \uparrow | S_z | S_x \uparrow \rangle = \frac{\hbar}{4} (1, 1) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0. \quad (5.119)$$

Dunque misurando S_z non distinguiamo i due casi e questo avviene perché la miscela è composta solo da autostati di S_z . Le cose cambiano però se misuriamo lo spin lungo una direzione diversa, ad esempio S_x . In tal caso, se lo stato iniziale è $|S_x \uparrow\rangle = (1/\sqrt{2})(|\uparrow\rangle + |\downarrow\rangle)$ e la matrice densità è quella in (5.115), il valor medio dell'osservabile è

$$\langle S_x \rangle_m = \text{tr}(\rho S_x) = \frac{\hbar}{4} \text{tr} \left[\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right] = \frac{\hbar}{2}, \quad (5.120)$$

consistentemente con il fatto che lo stesso valor medio può essere calcolato come valore di aspettazione sullo stato puro:

$$\langle S_x \uparrow | S_x | S_x \uparrow \rangle = \frac{\hbar}{2}. \quad (5.121)$$

Invece, se il sistema si trova nell'ensemble misto dei due stati $|\uparrow\rangle$ e $|\downarrow\rangle$ con uguale probabilità, e la matrice densità è $(1/2)\mathbb{1}$, allora

$$\langle S_x \rangle_m = \text{tr}(\rho S_x) = \frac{\hbar}{4} \text{tr} \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right] = 0. \quad (5.122)$$

In termini di macchie prodotte in un apparato di Stern-Gerlach, nel primo caso lo stato puro è esso stesso un autostato di S_x e la misura produce una macchia sola in corrispondenza del valore $\hbar/2$, mentre nel secondo caso si ottengono due macchie uguali, corrispondenti ai valori $\pm\hbar/2$, perché gli stati puri $|\uparrow\rangle$ e $|\downarrow\rangle$ che costituiscono la miscela contribuiscono allo stesso modo a ciascuna macchia. Questo è un esempio di misura che distingue le combinazioni lineari *coerenti* di stati puri dalle miscele *incoerenti* degli stessi stati. La fase conta e non può essere ignorata. In termini di matrice densità la differenza sta nel fatto che nel caso dello stato puro ci sono elementi non-diagonali diversi da zero, mentre nell'ensemble misto non ci sono. Sono proprio i termini fuori diagonale nella matrice ρ a produrre gli effetti di coerenza e di interferenza propri della meccanica quantistica.

In termini qualitativi, la situazione è analoga a quella di un sacchetto contenente un gran numero di palline. Estraiamo una pallina alla volta infilando la mano nel sacchetto, ma senza guardarci dentro. Una pallina sarà rossa, una blu, un'altra ancora blu, e così via. Alla fine contiamo le palline in base al colore e supponiamo di averne trovate metà rosse e metà blu. Da questo risultato è corretto dedurre che nel sacchetto c'era una miscela di palline rosse e blu in uguale quantità. In effetti le palline sono oggetti macroscopici a cui possiamo applicare la fisica classica senza alcun problema. Ma se invece di palline il sacchetto contiene particelle che obbediscono alla meccanica quantistica tali da esibire uno spettro a due stati, che possiamo ancora chiamare “blu” e “rosso”, allora le cose cambiano assai. Infatti, anche nel caso in cui si eseguisse una estrazione, una per volta, senza guardare dentro al sacchetto, e si trovasse in metà dei casi “blu” e in metà “rosso”, non potremmo dedurre che nel sacchetto, prima dell'estrazione, ci fosse una miscela bilanciata di stati “blu” e “rosso”. Nel sacchetto poteva esserci un insieme di particelle né blu né rosse, ma ciascuna in una sovrapposizione del tipo $(1/\sqrt{2})(\text{blu} + \text{rosso})$ e il risultato sarebbe stato lo stesso. Ogni

particella contribuirebbe al conteggio per uno dei due colori solo al momento dell'estrazione e nulla ci è concesso affermare sul suo essere blu o rosso prima dell'estrazione. Si potrebbe dire che la questione non è rilevante, perché una teoria, in fondo, deve fare predizioni su ciò che viene misurato e, quindi, siamo autorizzati ad ignorare il problema di cosa esista o non esista prima della misura. Ma questa risposta riflette una visione troppo semplicistica, che non tiene conto del fatto che possiamo avere a disposizione osservabili diverse per caratterizzare lo stesso sistema. L'esempio precedente dello spin ci mostra che, se invece di classificare le particelle in base al colore blu e rosso osservato nell'atto dell'estrazione, le classifichiamo in base ad una caratteristica diversa (in quel caso lo spin lungo x invece dello spin lungo z), allora è possibile distinguere i due casi, quello in cui nel sacchetto prima dell'estrazione c'era una miscela incoerente di particelle blu e di particelle rosse da quello in cui c'era un insieme di particelle tutte nella stessa combinazione lineare di blu e rosso.

Se questa idea, o meglio, questa evidenza osservativa, vi sembra incomprensibile malgrado siate arrivati alla fine di questi appunti, allora è possibile che questi appunti non siano stati scritti in modo sufficientemente chiaro. Ma non c'è troppo da preoccuparsi: la sensazione di non capire la meccanica quantistica è assai diffusa e tende a permanere nel tempo. In ogni caso, prendersela con la Natura non vale, che quella se ne infischia dei nostri pregiudizi culturali.

Capitolo 6

Conclusioni

Bene, siamo arrivati in fondo. Eravamo partiti da Stern-Gerlach e lì siamo ritornati. Se aveste saltato la parte in mezzo potreste avere l'impressione che tutto lo sforzo fatto per introdurre la nuova teoria sia mirato a spiegare un solo esperimento. Ovviamente non è così. In quasi cento anni dalla sua formulazione la meccanica quantistica ha permesso di comprendere una vastità di fenomeni naturali con uno straordinario accordo tra predizioni e osservazioni, dimostrandosi estremamente robusta ai tentativi di invalidarla. In questo senso, le lezioni presentate in questi appunti rappresentano l'inizio di un percorso. Avrete probabilmente modo di comprendere appieno il significato e le implicazioni della teoria quando ne studierete le applicazioni ai vari ambiti della fisica moderna.

La fine di un corso è un punto di sosta. Si può ripensare alle cose viste e prefigurare quelle che si vedranno nel percorso che rimane. Tra le cose viste in questi capitoli ci sono i postulati della teoria. Qui proviamo a riassumerli, ma non senza una premessa essenziale: la scelta di quali siano i postulati in una teoria fisica non è rigida. Quello che conta è il quadro completo che la teoria offre con l'insieme dei suoi postulati, o principi, e delle sue leggi. La teoria, nel suo complesso, fornisce uno schema razionale per la comprensione della natura, o almeno per una parte dei fenomeni naturali. Entro tale schema sono concesse formulazioni che siano diverse anche nell'ordine gerarchico delle leggi e dei principi, mantenendo invariate le predizioni per le grandezze fisiche misurabili. In queste lezioni abbiamo scelto di strutturare la teoria attorno ad alcune definizioni e alcuni principi così sintetizzabili:

- *Stati fisici e principio di sovrapposizione.* Dato un sistema fisico, inteso come la porzione di universo che si vuole studiare, tutta l'informazione che possiamo avere su di esso è contenuta nello "stato del sistema", rappresentato da un elemento di uno spazio vettoriale complesso appropriatamente definito in cui vale il principio di sovrapposizione. Vettori che differiscono per la norma rappresentano lo stesso stato. Le grandezze fisiche, definite tramite le procedure che ne permettono la misura, sono rappresentate da operatori lineari hermitiani che agiscono nello spazio dei vettori di stato, aventi autovalori reali e i cui autovettori costituiscono una base ortonormale dello stesso spazio. Possedere l'informazione completa sullo stato di un sistema equivale quindi a conoscere i coefficienti della decomposizione del vettore che rappresenta quello stato nella base scelta.
- *Postulato della misura.* Una misura è una procedura per mezzo della quale, agendo con un opportuno apparato (strumento di misura) su un sistema fisico, se ne ricava un numero reale. Il numero reale che si ottiene è uno degli autovalori dell'operatore hermitiano che rappresenta la grandezza misurata (osservabile). Se prima della misura il sistema si trovava già in uno degli autostati di quell'osservabile, il processo di

misura non cambia lo stato del sistema e dà come esito l'autovalore corrispondente. Se invece il sistema si trovava in uno stato diverso, combinazione lineare degli autostati dell'osservabile, allora la misura determina una modifica dello stato: dopo la misura il sistema si troverà nello stato rappresentato dall'autovettore dell'osservabile corrispondente all'autovalore ottenuto come esito della misura ("collasso della funzione d'onda", oppure "riduzione dello stato"). La probabilità di ottenere quel particolare valore è pari al modulo quadro del corrispondente coefficiente della decomposizione dello stato iniziale sulla base degli autostati dell'osservabile. Nel caso di spettri degeneri, la misura agisce come una proiezione in un sottospazio dei vettori di stato, in modo che esiti di misure distinti producono proiezioni in sottospazi tra loro ortogonali; ma l'idea di fondo non cambia. Questo postulato è cruciale per garantire che la teoria abbia contenuto empirico, ovvero che sia ancorata alla realtà tramite il confronto tra le sue predizioni e gli esiti degli esperimenti. Data la natura intrinsecamente probabilistica delle predizioni, il confronto va effettuato su repliche di uno stesso sistema preparato nello stesso stato.

- *Relazione d'indeterminazione.* Operatori associati a osservabili diverse definite per lo stesso sistema fisico possono commutare o non commutare. Se commutano, allora ammettono una base ortonormale completa nello stesso spazio vettoriale e la misura di una delle due non perturba l'esito della misura dell'altra (sono osservabili compatibili). Se non commutano, allora l'ordine con cui si misurano è cruciale (sono osservabili incompatibili) e per loro vale una relazione di indeterminazione: se ΔA e ΔB sono gli scarti quadratici medi delle misure ottenute per due osservabili A e B effettuate su molte repliche di un sistema preparato sempre nello stesso stato generico, allora il prodotto $\Delta A \Delta B$ ha come limite inferiore $1/2$ volte il modulo del valor medio del commutatore $[A, B]$ calcolato nello stesso stato. Una conseguenza diretta è che possedere tutta l'informazione sullo stato di un sistema non implica conoscere con esattezza il valore di tutte le osservabili che possiamo definire per quel sistema, ma solo di quelle che sono tra loro compatibili. L'algebra degli operatori associati alle osservabili dunque pone limiti intrinseci all'informazione che possiamo estrarre da un sistema. In realtà questo non è un postulato a sé, ma piuttosto una conseguenza dei due precedenti e della definizione di osservabili incompatibili. L'ho messo comunque in evidenza per l'importanza che la relazione d'indeterminazione ha nella comprensione della teoria.
- *Posizione e momento.* La posizione spaziale di una particella è un'osservabile con spettro di autovalori continuo. Il momento di una particella è anch'esso un osservabile con spettro continuo; l'operatore ad esso associato è il generatore delle traslazioni spaziali. La non commutatività della misura di posizione in una generica direzione e della traslazione spaziale nella stessa direzione implica che i due operatori associati alla posizione e al momento non commutano e le due osservabili sono incompatibili. In particolare, la posizione lungo una generica direzione x e la componente del momento nella stessa direzione, p_x , obbediscono alla regola di commutazione canonica $[x, p_x] = i\hbar$. Ne segue la disuguaglianza $\Delta x \Delta p_x \geq \hbar/2$, nota come relazione d'indeterminazione (o d'incertezza) di Heisenberg. Le regole di commutazione canonica introducono nella teoria la costante di Planck.
- *Evoluzione temporale.* L'evoluzione nel tempo di un vettore di stato generico $|\alpha\rangle$ è data dall'equazione di Schrödinger $i\hbar(\partial/\partial t)|\alpha\rangle = H|\alpha\rangle$, dove H è l'operatore di Hamilton, o hamiltoniano. Quest'ultimo può essere visto come il generatore delle traslazioni temporali. Gli autostati dell'hamiltoniano sono gli stati stazionari del sistema; questi dipendono dal tempo solo tramite un fattore di fase $\exp[-(i/\hbar)Et]$, dove E è il corrispondente autovalore di H , che ha il significato di energia del sistema in quello stato.

La trattazione quantistica di un sistema che ammette un limite classico si ottiene sostituendo le coordinate spaziali e i momenti coniugati che compaiono nella funzione di Hamilton classica con i corrispettivi operatori quantistici, tenendo conto delle regole di commutazione canonica.

- *Ensemble puri e misti.* Gli stati definiti al primo punto di questo elenco sono detti stati puri. Gli esiti di misure di osservabili eseguite su repliche di un sistema nello stesso stato puro seguono distribuzioni probabilistiche fissate dai coefficienti della decomposizione dello stato puro sulla base scelta per lo spazio degli autovettori. In generale, tuttavia, un *ensemble* può essere costituito da miscele incoerenti di diversi stati puri, ciascuno con la sua probabilità di essere trovato nella miscela; in tal caso si parla di *ensemble* misto e la distribuzione statistica degli esiti delle misure di un'osservabile deve tener conto sia delle probabilità quantistiche proprie degli stati puri che delle probabilità classiche della miscela. A tale scopo si introduce la matrice densità ρ , opportunamente definita in modo tale che il valor medio di un'osservabile generica Q eseguito sull'*ensemble* misto è dato dalla traccia del prodotto di ρ per la matrice che rappresenta l'osservabile Q .
- *Connessione spin-statistica.* Se le particelle di un sistema sono identiche, l'esito della misura di qualsiasi osservabile non deve dipendere dall'ordine delle particelle nella definizione dei vettori di stato. In particolare, l'energia non può cambiare per scambio di coordinate tra coppie di particelle. L'operatore che scambia due particelle qualsiasi commuta con l'hamiltoniano H e dunque i due operatori ammettono una base comune. L'operatore che scambia due particelle ammette solo gli autovalori $+1$, per stati simmetrici, e -1 , per stati antisimmetrici. In natura si osserva che gli stati di particelle con spin intero (bosoni) sono sempre simmetrici, mentre quelli di particelle con spin semintero (fermioni) sono sempre antisimmetrici. Una conseguenza è che due fermioni non possono condividere le stesse coordinate, cioè gli stessi identici valori per l'insieme massimale di osservabili compatibili; questo è noto come principio di esclusione di Pauli.

Tutto ciò sta alla base della meccanica quantistica non-relativistica. Introducendo la dinamica quantistica, all'inizio del secondo capitolo, avevamo detto che ci saremmo dedicati al caso di particelle con massa a riposo non nulla e che viaggiano a velocità piccole rispetto alla velocità della luce. Abbiamo così formulato una teoria che, laddove esista il corrispettivo classico, in quel limite si riduce alla dinamica newtoniana. Se vogliamo invece formularla anche per particelle relativistiche (nell'accezione della relatività ristretta), allora occorre esprimere le equazioni in modo che siano invarianti per trasformazioni di Lorentz. Tempo e spazio, essendo le componenti di uno stesso quadrivettore, devono entrare allo stesso ordine nelle equazioni, mentre così non è nel caso dell'equazione di Schrödinger. Non è un problema insormontabile. Dirac introdusse la teoria per elettroni relativistici già nel 1928, ad appena due anni dalla prima formulazione della meccanica quantistica non-relativistica. Da lì è nata la teoria quantistica dei campi (*quantum field theory*) che avrete modo di imparare se continuerete gli studi in fisica nella laurea magistrale. Fin d'ora tenete presente che quando si parla di meccanica quantistica in generale si intende sia la versione non-relativistica sia la teoria quantistica dei campi, trattandosi di due versioni di uno stesso quadro concettuale.

Una cosa che sarebbe interessante fare, e che non richiede troppe complicazioni, è la trattazione quantistica dei campi elettromagnetici. Ne avevamo accennato nel capitolo dedicato alle perturbazioni dipendenti dal tempo e la regola aurea di Fermi. Non abbiamo il tempo di affrontarla in questo corso, ma è uno dei classici argomenti da primo semestre della laurea magistrale. Sarà l'occasione per dare un senso compiuto all'idea di "pacchetto di energia", $h\nu$, per la radiazione elettromagnetica nella cavità del corpo nero, da cui tutto

era partito con Planck. Inserire i *fotoni* nella descrizione fisica del mondo è un bel passo in avanti, anche in considerazione del fatto che un sacco di fisica molto interessante oggi si fa proprio con i fotoni, ad esempio nel campo della fisica atomica, dell'ottica quantistica, della fotonica, e dell'informazione quantistica.

Una cosa che invece non si può fare, o almeno nessuno è in grado di fare fino ad oggi, è mettere d'accordo la meccanica quantistica con la relatività generale. Si pensi semplicemente a questo fatto: la meccanica quantistica ammette situazioni nelle quali un elettrone si trova in una combinazione lineare di possibili esiti della misura della sua posizione. In altri termini, finché non avviene la misura, l'elettrone è descritto da una funzione d'onda che sta un po' qua e un po' là. Ma la relatività generale si basa sull'idea di uno spazio-tempo che viene deformato dalla presenza di masse. L'elettrone ha massa, dunque deforma, seppur di poco, lo spazio-tempo attorno a sé. Bene, ma lo deforma qua o là? La risposta non c'è. Forse il problema è mal posto. Se trovate la soluzione ditemelo. Nel frattempo, se il nostro scopo è quello di indagare il mondo alle scale di distanze ed energie che conosciamo, allora non dobbiamo preoccuparcene troppo. Gli effetti dell'incompatibilità tra le due teorie, infatti, si vedrebbero solo in situazioni assai estreme, per il momento fuori dalla portata delle nostre osservazioni empiriche, anche le più avanzate.

Al netto di questo problema, peraltro interessante e che merita gli sforzi delle generazioni di cervelli che lavorano (e si logorano) nel tentativo di risolverlo, la meccanica quantistica è una teoria che funziona, funziona a meraviglia. Possiamo certamente immaginare che un giorno venga superata da una teoria ancora più ampia, perché questo è il destino delle teorie scientifiche. Ma intanto cerchiamo di capirne il significato e la portata.

Si è scritto e detto molto nei decenni passati sull'insoddisfazione che la meccanica quantistica genera in termini di "interpretazione". Anche questo è un problema che ha logorato fior fior di cervelli, ma su un fronte diverso, diciamo, più filosofico. Intanto ci sarebbe da discutere su cosa significhi esattamente interpretare una teoria fisica. Una teoria è quel che è, con il suo schema di definizioni, principi, leggi, eccetera. Interpretarla significa probabilmente tentare di connetterla a nozioni e teorie precedenti, più familiari, sperando di comprenderla alla luce di un linguaggio comunemente accettato in una certa epoca e in un certo contesto culturale. Se è così, allora insistere su questioni di interpretazione è una strategia perdente. Per sua natura una nuova teoria ha in sé elementi di rottura con le precedenti, nella sua formulazione, nel suo linguaggio; altrimenti sarebbe solo un ritocco dell'esistente. Nel caso della meccanica quantistica la rottura con la fisica classica c'è già alle prime battute e questo crea un evidente disturbo in chi è affezionato alla buona e vecchia fisica che c'era prima. Alla prima esecuzione della Sagra della Primavera di Stravinsky il pubblico rumoreggiava, fischiava e tirava le sedie. Era il 1913, lo stesso anno in cui Bohr pubblicava la sua trilogia di articoli sul suo modello atomico. Ecco, più o meno è lo stesso meccanismo. Ora non vorrei sminuire troppo la questione (è un pezzo della storia della scienza del '900), ma come fisici siamo tenuti ad essere fedeli al nostro metodo di lavoro: una teoria è buona se permette un accordo tra predizioni e osservazioni e se offre gli strumenti per comprendere quello che osserviamo. Se il prezzo da pagare è quello di introdurre concetti che appaiono inusuali a prima vista, lo paghiamo. Poi, se la teoria è buona, quegli stessi concetti ci diventeranno familiari, esattamente come ci si è abituati alla musica di Stravinsky, all'idea che le onde possono propagarsi anche nel vuoto, o che la terra è una minuscola sferetta che si muove in un universo brulicante di stelle e pianeti. L'unica differenza, semmai, è che Stravinsky venne portato in trionfo nella stessa città meno di un anno dopo lo scandalo della prima, mentre invece la meccanica quantistica deve tuttora difendersi da ampie schiere di detrattori¹.

¹Un commento molto lucido su questa linea è quello di P.A.M. Dirac nel primo capitolo del suo libro "The Principles of Quantum Mechanics": [...] *it may be remarked that the main object of physical science is not the provision of pictures, but is the formulation of laws governing phenomena and the application of these laws to the discovery of new phenomena. If a picture exists, so much the better. But whether a*

Uno dei detrattori, come si sa, era Einstein, che riteneva la teoria incompleta in quanto incapace (la teoria, non lui) di rendere conto di tutti gli “elementi di realtà”, definibili localmente, che una teoria completa dovrebbe saper descrivere. A differenza di altri critici meno competenti, Einstein, assieme ai suoi collaboratori Podolsky e Rosen, riuscì a formulare il problema in modo tale da farne nascere un filone di ricerche ricco di idee e di esperimenti brillanti. Se la teoria è incompleta, allora vuol dire che ignora alcune variabili, proprio quelle che, essendo ignorate, impediscono alla meccanica quantistica di soddisfare il requisito di realismo locale e le permettono, al contrario, di predire la *spooky action at a distance* che Einstein aborrisce. Si tratta quindi di cercare una nuova teoria che dia gli stessi risultati della meccanica quantistica, ma che includa anche queste “variabili nascoste”. John Bell pose la questione in termini di correlazioni tra esiti di misure effettivamente eseguibili in sistemi concreti e dimostrò che le predizioni di qualsiasi teoria che abbia i requisiti richiesti da Einstein, cioè includa variabili nascoste compatibilmente con il realismo locale, dovevano soddisfare una certa disuguaglianza, mentre quelle della meccanica quantistica no. Gli esperimenti effettuati finora danno ragione alla meccanica quantistica: sembra proprio che dobbiamo rinunciare all’idea confortante di realtà che Einstein aveva in mente. Questa bella storia di ingegno teorico e sperimentale è meno semplice di come ve l’ho descritta qui, ma il messaggio che emerge, al di là delle sottigliezze e complicazioni del caso, sta nel fatto che le dispute sulle teorie fisiche devono risolversi confrontando le predizioni delle teorie (numeri) con gli esiti delle misure (numeri anch’essi); il resto è dibattito filosofico, per quanto interessante possa essere.

Un punto apparentemente debole della teoria, che non possiamo ignorare, è il fatto che essa prevede due meccanismi diversi per l’evoluzione degli stati fisici di un sistema: la misura, regolata da una dinamica non unitaria e non reversibile, che fa precipitare lo stato del sistema in uno degli autostati dell’osservabile misurata, o in un sottospazio di autostati, e l’evoluzione temporale in assenza di misure, unitaria e reversibile, che obbedisce all’equazione di Schrödinger. Peggio ancora, la misura per essere definita richiede l’intervento di strumenti che si comportano necessariamente, ad un qualche livello, in modo classico. Gli apparati di misura, in senso lato, sono essi stessi sistemi fisici composti da particelle che possono manifestare un comportamento quantistico ma, agendo come strumenti di misura, obbediscono complessivamente alle leggi classiche. Dunque, per definire le osservabili e per eseguire concretamente le misure, abbiamo bisogno della fisica classica. Detto in altri termini: per dare contenuto empirico alla meccanica quantistica dobbiamo ricorrere a sistemi che obbediscono alla fisica classica. L’opposto non vale.

Questa asimmetria può turbare, ma prima di farci prendere dal panico ragioniamo un attimo su cos’è una teoria a partire da questo punto di vista: noi che, in quanto essere pensanti e curiosi, vogliamo formulare ipotesi su come vanno le cose nel mondo dei fenomeni naturali, per prima cosa raccogliamo da questi fenomeni delle informazioni; poi memorizziamo queste informazioni in qualche forma e le rielaboriamo secondo schemi mentali appresi, per esperienza, per studio, per intuito. Da questa elaborazione ne ricaviamo correlazioni e regolarità che traduciamo in leggi; con queste leggi, opportunamente strutturate, diamo forma alle teorie. Per fare tutto questo, in linea di principio, non serve nemmeno un essere cosciente, nel senso normalmente attribuito al termine coscienza. Infatti, potremmo pensare che la stessa procedura possa essere eseguita da un apparato dotato di sensori, in grado di raccogliere le informazioni che arrivano dai sensori in forma di numeri, o di *bit*, e poi

picture exists or not is a matter of only secondary importance. In the case of atomic phenomena no picture can be expected to exist in the usual sense of the word ‘picture’, by which is meant a model functioning essentially on classical lines. One may, however, extend the meaning of the word ‘picture’ to include any way of looking at the fundamental laws which makes their self-consistency obvious. With this extension, one may gradually acquire a picture of atomic phenomena by becoming familiar with the laws of the quantum theory.

di memorizzarle in appositi registri e di analizzarle con algoritmi finalizzati alla ricerca di correlazioni tra i dati. E se questo apparato, o rete di apparati, fosse anche dotato di sistemi di intelligenza artificiale tali da fargli apprendere autonomamente le strategie migliori per correlare i dati tra di loro (si badi bene che sistemi del genere esistono già), allora il passo per arrivare alla formulazione di teorie fisiche potrebbe essere breve. Magari non sarebbero teorie espresse in forma di leggi matematiche eleganti ma, come pare dicesse Boltzmann, “l’eleganza è di pertinenza dei sarti e dei calzolari” e una teoria poco elegante che funziona è sempre preferibile ad una elegantissima teoria che non funziona². Comunque, che siano esseri umani o apparati meccanici/elettronici, i sistemi fisici che si fanno carico di formulare teorie sono necessariamente sistemi classici, quanto meno per il fatto che la raccolta, archiviazione, elaborazione di molti *bit* richiede sistemi fatti di molti sottosistemi interagenti tra loro, segnali tipicamente elettrici che viaggiano tra una parte e l’altra (che siano neuroni o registri di memoria digitale poco importa), sistemi di controllo e raffreddamento (biologici o meccanici), e così via. In questo senso, non c’è nulla di strano nel fatto che per formulare la meccanica quantistica si debba passare attraverso la conoscenza della fisica classica e non viceversa: è implicito nella natura stessa di chi le teorie le formula. Un elettrone in un atomo è un sistema quantistico, ma non è in grado di elaborare alcuna teoria.

Vedere una teoria come l’esito di una raccolta ed elaborazione di informazioni fornisce anche una chiave di lettura diversa dei postulati della meccanica quantistica. In fondo, riletti sotto questa luce, essi forniscono le regole per estrarre dalla natura le informazioni utili e necessarie ad ampliare la nostra conoscenza. Gli stati fisici di un sistema altro non sono che l’insieme massimo di informazioni che possiamo raccogliere su di esso. La relazione di indeterminazione pone limiti al tipo di informazioni che possiamo raccogliere. L’equazione di Schrödinger dice qual è il legame tra le informazioni che possiamo ottenere interrogando il sistema in tempi diversi. Possiamo pensare che dietro tutto ciò ci sia la Realtà oggettiva, indipendente dalle informazioni che possono essere raccolte ed elaborate da esseri umani o computer dotati di sensori, ma la nostra conoscenza della Realtà non può prescindere dalle regole con cui le informazioni su di essa possono essere ottenute in pratica. La meccanica quantistica indica tali regole meglio di altre teorie. Lo fa senza troppo assecondare il nostro istinto di aggrapparci ad un realismo ingenuo, antropocentrico, fatto di palline che viaggiano seguendo ciascuna la propria traiettoria. Ci impone, ad esempio, di codificare le informazioni usando anche i numeri immaginari, che già il loro nome (scelta sfortunata, col senno di poi) evoca un’astrazione dalla realtà. Ma la meccanica quantistica non ci allontana affatto dalla realtà, semplicemente ne parla nei termini consentiti dal metodo scientifico. Peraltro, il fatto che i sistemi quantistici si comportino in un certo modo, secondo quanto risulta dalle informazioni che ne estraiano, ci porta anche all’idea di utilizzare quegli stessi sistemi come parte dei processi di elaborazione e trasmissione di informazioni, sfruttando le peculiarità degli stati fisici (il principio di sovrapposizione, l’entanglement, ecc.) come risorsa per realizzare operazioni altrimenti impossibili con dispositivi classici. Da questa idea nascono gli sviluppi odierni dell’informazione quantistica, uno dei settori di ricerca più attivi della scienza moderna.

Tolti di mezzo i possibili fraintendimenti sul significato di teoria fisica in generale e sul ruolo centrale dell’informazione, rimane comunque una questione di fondo: dove sta il confine tra quantistico e classico? L’apparato di misura è esso stesso un sistema fisico che interagisce con il sistema su cui eseguiamo la misura; non potremmo forse inglobarlo in una descrizione puramente quantistica? Perché non possiamo applicare un’evoluzione temporale unitaria, che preservi la coerenza degli stati puri quantistici, anche al sistema complessivo, trattando l’apparato di misura come una parte del tutto? Se questo non è possibile, allora vuol dire

²Mi spiace per Dirac, che considerava la bellezza come metodo, ma su questo tenderei a parteggiare per Boltzmann.

che esiste una soglia, un livello di complessità oltre al quale un sistema non può più essere descritto da una sovrapposizione coerente di stati puri. Ma i postulati della teoria non dicono nulla sull'esistenza di tale soglia. Questo, più nel passato che in tempi recenti, ha generato discussioni interessanti, quasi tutte incentrate sulla palese non esistenza in natura di un gatto simultaneamente vivo e morto, o varianti di quello. Qualcuno, anche tra gli scienziati più autorevoli, ha ipotizzato che il passaggio da uno stato coerente (sovrapposizione del gatto vivo e morto) allo stato classico (vivo o morto) richiedesse l'atto di misura da parte di un individuo cosciente. Visto il fascino che emana dall'accostamento tra scienza e coscienza, tale ipotesi, infelice direi, ha avuto una risonanza tale da radicare nella società l'idea che la meccanica quantistica rimettesse l'uomo, con la sua coscienza, al centro dell'universo. Tuttavia, purtroppo, c'è chi ne parla in questi termini.

Un approccio epistemologicamente corretto, e quello più seguito in questi anni, consiste nell'andare in laboratorio e tentare di riprodurre, se non proprio un gatto, almeno qualcosa di equivalente ad un minuscolo gattino, ma che stia effettivamente in una sovrapposizione tra due stati puri per un tempo abbastanza lungo. Si scopre così che la maggiore difficoltà sta nel proteggere gli stati coerenti dall'interazione con l'ambiente, inteso come sorgente di perturbazioni indesiderate, esterne al sistema, il cui effetto è quello di trasformare rapidamente uno stato puro, coerente, in uno stato misto, incoerente. Questi processi di *decoerenza* sono difficili da eliminare, e tanto più lo sono quanto più il sistema diventa grande, macroscopico. Con le tecniche attuali, ad esempio, si riesce a mettere in uno stato *entangled* due oscillatori meccanici identici, fatti ciascuno di circa 10^{10} atomi, distanti tra loro una ventina di centimetri all'interno di un criostato; ma lo stato di sovrapposizione coerente dura solo per qualche micro-secondo, perché bastano le fluttuazioni termiche dell'ambiente (per quanto sia tenuto a poche decine di mK di temperatura) per indurre decoerenza. È da tentativi come questo, anche con altri dispositivi, altri tipi di "gatti", che possiamo esplorare la zona di confine tra meccanica quantistica e fisica classica e sperare di capire meglio l'origine della cesura esistente, nella formulazione della meccanica quantistica, tra l'evoluzione unitaria degli stati in assenza di misura e la riduzione degli stati nell'atto della misura. Quello che si imparerà seguendo questa strada è che non esiste un confine netto tra il mondo microscopico dominato dalla meccanica quantistica e il mondo macroscopico governato dalla meccanica classica, né tra l'evoluzione unitaria degli stati e i loro collasso nella misura; si tratta piuttosto di transizioni graduali, linee d'ombra che vanno esplorate con il metodo giusto (idee ed esperimenti), senza farsi fuorviare da questioni estetiche o filosofiche, con tutto il rispetto per l'estetica e la filosofia.

Il disagio che la meccanica quantistica produce in chi l'affronta, tutto sommato, è un buon segno e anche una risorsa. Vuol dire che la teoria costringe a tenere la mente aperta e lo spirito critico vigile; ci parla dei limiti della nostra conoscenza ma, allo stesso tempo, ci permette di esplorare mondi che stanno ben oltre la nostra immaginazione. Non sarà la miglior teoria, non sarà senz'altro l'ultima, ma intanto studiamola con la passione che merita.

Nota: le immagini in copertina rappresentano l'esito di osservazioni sperimentali effettuate su gas di atomi, in particolare circa 10^7 atomi di sodio, raffreddati al di sotto della temperatura critica per ottenere un condensato di Bose-Einstein, a temperature dell'ordine di 10^{-7} K. In questo regime lo stato degli atomi è rappresentato da un'unica funzione d'onda macroscopica. Le immagini di colore rosso mostrano la differenza tra la densità misurata e quella predetta dalla teoria per lo stato fondamentale del sistema. Il rumore che si vede è principalmente dovuto al rumore tipico della tecnica di *imaging* utilizzata. Il segnale più evidente è costituito da linee scure che attraversano il condensato. Ciascuna linea corrisponde ad un filamento vorticale attorno al quale gli atomi ruotano in modo tale che il momento angolare di ogni atomo valga esattamente $\hbar$. Questa quantizzazione della vorticità, che si manifesta su scala macroscopica è una manifestazione tipica della superfluidità. Le immagini in blu e verde rappresentano invece una misura della fase della funzione d'onda macroscopica, ottenuta con una procedura di interferenza tra il condensato e una sua copia. In queste immagini la presenza di un vortice è segnalata dal comparire di dislocazioni nel *pattern* di interferenza. Tali dislocazioni corrispondono ad una variazione di fase di 2π in una rotazione attorno al filamento vorticale. Gli esperimenti sono stati condotti nel laboratorio di atomi ultrafreddi del BEC Center di Trento (ringrazio Gabriele Ferrari e Giacomo Lamporesi per averle messe a disposizione).