

APPUNTI DI
MECCANICA QUANTISTICA

autore: VALERIANO BARASSI
(docente: *Francesco Nicodemi*)

Anno Accademico 1990-91

Quest'opera è soggetta alla Creative Commons Public License versione 3.0 o posteriore. L'enunciato integrale della Licenza in versione 3.0 è reperibile all'indirizzo internet:

<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.it>

- Si è liberi di riprodurre, distribuire, comunicare al pubblico, esporre, in pubblico, rappresentare, eseguire e recitare quest'opera alle seguenti condizioni:

Attribuzione. Bisogna attribuire la paternità dell'opera nei modi indicati dall'autore o da colui al quale è stata data quest'opera in licenza; in questo caso si tratta dell'autore,

Non commerciale. Non si può usare quest'opera per fini commerciali,

Non opere derivate. Non si può alterare o trasformare quest'opera, né usarla per crearne un'altra.

- Ogni volta che si usa o si distribuisce quest'opera, lo si deve fare secondo i termini di questa licenza, che va comunicata con chiarezza.
- In ogni caso si possono concordare con il titolare dei diritti d'autore (in questo caso l'autore) utilizzi di quest'opera non consentiti da questa licenza.

Quest'opera si avvale del diritto di citazione a scopo accademico e di critica previsto dall'Articolo 10 della Convenzione di Berna sul diritto d'autore.

“È sbagliato pensare che lo scopo della fisica sia di trovare com'è la natura. La fisica riguarda ciò che possiamo dire riguardo alla natura.”

– Niels Bohr

Indice

NOTE	vii
CONVENZIONI TIPOGRAFICHE	viii
1 La crisi della fisica classica	1
1.1 La radiazione di corpo nero	1
1.2 L'effetto fotoelettrico	5
1.3 Il problema dei calori specifici	7
1.4 Conclusioni	8
2 I primi modelli atomici e la vecchia teoria dei quanti	11
2.1 Il modello di Thomson	11
2.2 Il modello di Rutherford	12
2.3 Gli spettri atomici	14
2.4 Il modello di Bohr-Sommerfeld	14
2.5 L'ipotesi di de Broglie. Il dualismo onda-particella	24
3 La meccanica ondulatoria	29
3.1 Il pacchetto d'onda	29
3.2 L'equazione di Schrödinger	30
3.3 La funzione d'onda e sua interpretazione	33
3.4 Calcolo della funzione d'onda in alcuni casi di rilevanza pratica	38
3.4.1 La barriera di potenziale	38
3.4.1.1 L'effetto tunnel	41
3.4.2 La buca di potenziale	44
3.4.3 La doppia buca di potenziale	48
4 La struttura formale della meccanica quantistica	51
4.1 La trasformata di Fourier. La funzione d'onda per l'impulso	51
4.2 La funzione <i>delta di Dirac</i> $\delta(x)$	52
4.3 Valore atteso di una grandezza. La rappresentazione operatoriale	55
4.4 Il principio di indeterminazione di Heisemberg	57
4.5 Operatori Hermitiani e osservabili. L'apparato formale della teoria	58
4.5.1 Proprietà dei commutatori	63
4.5.2 Densità e corrente di probabilità	65
4.5.3 Degenerazione e numeri quantici	66
4.6 Note riassuntive sulla struttura formale della teoria	68
4.7 Una dimostrazione generale della relazione di indeterminazione	69
4.8 Evoluzione classica dei valori medi	70
4.9 Costanti del moto e evoluzione temporale	71

5	L'oscillatore armonico	75
5.1	Metodo analitico	75
5.2	Metodo operatoriale. Operatori di <i>creazione</i> e <i>distruzione</i>	79
5.3	Valori medi e stati coerenti	82
6	Il momento angolare	85
6.1	Il momento angolare in meccanica quantistica	85
6.2	Autofunzioni di l^2 e l_z : le armoniche sferiche	87
6.3	Gli <i>operatori gradino</i> L_+ e L_-	93
6.4	Lo <i>spin</i>	97
6.5	Il momento angolare come generatore infinitesimo delle rotazioni	100
6.6	Addizione di momenti angolari	102
7	L'atomo di idrogeno	107
7.1	Equazione di Schrödinger per l'atomo di idrogeno	107
7.2	Soluzione dell'equazione radiale	110
7.3	Autofunzioni dell'atomo di idrogeno	113
7.3.1	Sullo spin dell'elettrone	116
7.3.2	Sulla stabilità degli autostati	116
8	Teoria delle perturbazioni	119
8.1	Perturbazioni indipendenti dal tempo	119
8.1.1	Un esempio concreto: La correzione relativistica all'atomo di idrogeno	126
8.2	Perturbazioni dipendenti dal tempo	129
8.3	Metodo variazionale	133
A	L'interpretazione di Feynman e i <i>Path Integral</i>	137
A.1	Impostazione concettuale	137
A.2	Formulazione matematica	140
A.3	Le relazioni di de Broglie e di Planck	144
A.4	L'equazione di Schrödinger	145

NOTE

Il presente lavoro è costituito dagli appunti raccolti durante il corso di Istituzioni di Fisica Teorica tenuto dal prof. FRANCESCO NICODEMI nell'Anno Accademico 1990-1991 presso la Facoltà di Fisica dell'Università di Napoli "Federico II".

Queste note non possono certamente essere considerate come una trattazione esaustiva dell'argomento, vista anche la grande quantità di testi sull'argomento. L'autore ha voluto però soffermarsi su alcuni aspetti concettuali e su diversi aspetti qualitativi utili agli studenti alle prime armi con la teoria. È richiesta una conoscenza di base della meccanica Hamiltoniana e dell'algebra lineare, in particolare l'algebra delle matrici, gli spazi vettoriali e i concetti di autovettore e autovalore.

Il materiale presentato è basato sulle lezioni tenute dal docente, in alcune parti ampliate e dettagliate dall'autore. Ampie parti che descrivono l'interpretazione della funzione d'onda sono basate sul testo "La fisica di Berkeley, vol. 4 Meccanica Quantistica", E. H. Wichmann, Ed. Zanichelli, 1973. Le parti riguardanti l'addizione dei momenti angolari e la teoria delle perturbazioni sono state ampliate con gli appunti di Fabio Grazioso. Altri testi di riferimento per le parti descrittive sono il libro "Filosofia della fisica", G. Boniolo, Bruno Mondadori Editore, 1997 e alcuni estratti di Wikipedia. L'appendice A sui path integral è basata sul testo "Quantum Mechanics and Path Integral", R. P. Feynman and A. R. Hibbs, McGraw-Hill, 1965. Questo lavoro è distribuito liberamente nella speranza che possa essere utile nella preparazione degli esami, con la sola condizione che si citi il nome dell'autore e la fonte.

CONVENZIONI TIPOGRAFICHE

Nella tabella seguente sono riportate le convenzioni tipografiche utilizzate nel testo. Si è cercato fin dove possibile di evitare ambiguità laddove tradizionalmente si utilizzano gli stessi simboli per diverse quantità fisiche (ad esempio la “phi” utilizzata sia per le funzioni d’onda che per il potenziale che per la variabile angolare).

<i>Simbolo utilizzato</i>	<i>Quantità fisica</i>
$\psi, \varphi, \chi, \Psi, \Phi$	Funzione d’onda
ϕ	Fattore di fase (funzioni d’onda)
θ, ϕ	Variabili Angolari
Ω	Angolo Solido
\mathcal{V}	Volume
V	Potenziale generico
E	Autovalore dell’Energia
E, ε	Energia
\mathcal{E}	Campo elettrico
ϵ	Quantità infinitesima
H	Hamiltoniana
\mathcal{H}	Spazio di Hilbert
M	Momento angolare
J	Momento angolare (per sistemi composti)
L, l_x, l_y, l_z	Momento angolare e sue componenti
s	Spin
ϕ	Potenziale Elettrostatico
ρ	Densità
μ	Momento magnetico
\hat{O}	Operatore(*)
\vec{v}	Vettore(**)
$\int_{\mathcal{O}}$	Integrale esteso ad una sfera

(*)Il simbolo di operatore $\hat{}$ è generalmente omissso per non appesantire inutilmente la notazione, è tuttavia utilizzato per rimuovere ambiguità (generalmente fra l’operatore e il suo autovalore) o quando essenziale ai fini della comprensione.

(**)Il simbolo $\vec{}$ è sempre utilizzato dove occorre.

La crisi della fisica classica

Alla fine del XIX° secolo, la cosiddetta “concezione classica del mondo” aveva raggiunto così grandi successi che la scienza fisica si considerava fondamentalmente conclusa e tutti i fenomeni fondamentalmente noti e compresi.

Questa “concezione classica del mondo” si basava su due grossi pilastri elaborati lungo il corso dell’evoluzione scientifica, a partire da Galileo per giungere fino al 1800. Questi due pilastri sono la meccanica newtoniana e l’elettromagnetismo di Maxwell.

Il primo di questi, la meccanica newtoniana, trae origine dal lavoro di Newton e fu portata a massimo compimento dai lavori successivi di Lagrange e Hamilton. Essa si occupa del moto dei corpi materiali e delle forze che agiscono su di essi. Tramite la meccanica newtoniana si era riuscito anche a trovare spiegazione delle leggi della termodinamica, grazie ai lavori di Boltzmann, Gibbs e Maxwell.

Il secondo di questi pilastri è l’elettromagnetismo di Maxwell. Esso ha permesso di unificare due fenomeni apparentemente diversi fra loro, l’elettricità ed il magnetismo, grazie ai lavori di Faraday e Maxwell, il quale introdusse anche il concetto di *campo elettromagnetico*. Compare così nella fisica il concetto di *campo*, un’entità che ha una sua esistenza materiale al pari delle particelle e che può trasportare energia. Esso si poteva manifestare in diversi modi, come onde radio, come luce o come raggi X, permettendo di interpretare tutti questi fenomeni con un’unica entità. La caratteristica di tutti questi fenomeni, riassunti dalle equazioni di Maxwell, è il loro comportamento ondulatorio.

Questo era il quadro generale della fisica alla fine dell’800: particelle materiali e campo elettromagnetico avevano entrambi un’esistenza reale e rispondevano a delle precise leggi fisiche nei propri domini di applicazione (le leggi della meccanica le prime, le leggi di Maxwell il secondo). In linea di principio, tutti i fenomeni fisici potevano essere descritti tramite queste leggi, al punto che Michelson poté dire: “*The most important fundamental laws and facts of physical science have all been discovered, and these are now so firmly established that the possibility of their ever being supplemented by new discoveries is exceedingly remote.*”

Restavano, certamente, alcuni *piccoli* problemi in sospeso, ma era sicuro che si trattasse di problemi di calcolo che sarebbero stati risolti in breve tempo. Questi problemi erano tre: trovare l’effettiva forma della radiazione di corpo nero, dare una spiegazione dell’effetto fotoelettrico (rivelato per la prima volta da Hertz nel 1887) e spiegare perché il calore specifico dei corpi va a zero con la temperatura.

Di lì a poco, questi tre problemi avrebbero rivoluzionato la “concezione classica del mondo”.

1.1 La radiazione di corpo nero

Il fenomeno per cui molti corpi (ad esempio i metalli) quando sono scaldati cominciano prima ad emettere calore per poi cambiare colore dal rosso via via verso il giallo ed il bianco era al centro dell’interesse dei fisici. Il grande interesse derivava dal fatto che per via termodinamica era stato possibile dimostrare che il rapporto fra la potenza emessa e la potenza assorbita è indipendente dalla

natura del corpo e dipende esclusivamente dalla frequenza e dalla temperatura. Questo implica che la legge di emissione deve essere una funzione universale, e gli scienziati sono sempre interessati a fenomeni e leggi di natura universale perché spesso permettono di comprendere nuovi fenomeni.

Per un corpo che assorba tutte le radiazioni che incidono su di esso la funzione universale di cui sopra coincide con il potere emissivo, per di più un corpo di questo tipo (detto *corpo nero*) è facilmente realizzabile in laboratorio.¹ Anche se i processi in gioco sono complessi, le interazioni fra le pareti e la radiazione seguono leggi fisiche ben note, pertanto il calcolo della funzione deve essere possibile senza particolari difficoltà concettuali.

Lo spettro di un corpo nero (intensità o densità della radiazione emessa in funzione della frequenza) è quindi uno spettro con una forma caratteristica dipendente unicamente dalla sua temperatura T .

Dimensionalmente, la densità di energia – che sarà indicata con $u(\nu, T)$ – ha le dimensioni di un'energia per unità di volume, cioè:

$$[u(\nu, T)] = [E][l]^{-3}[t]$$

Siccome è una funzione universale, le uniche variabili per descrivere questo sistema sono $k_B T$, che rappresenta un'energia "universale", la lunghezza d'onda λ e la frequenza ν (e la velocità della luce $\lambda\nu = c$). Per motivi dimensionali, si vede che l'unica combinazione possibile è:

$$[u] = \frac{[k_B T]}{[\nu][\lambda]^3} \quad \Rightarrow \quad u(\nu, T) \propto \frac{\nu^2 k_B T}{c^3}$$

Il corpo nero può essere schematizzato come un insieme di oscillatori armonici, dei quali si deve considerare la potenza media irradiata. Questa potenza dipende dall'accelerazione a , dalla carica q , dalla velocità della luce c e dalla frequenza di oscillazione ν , in particolare $P = f(q^2, a^2, c, \nu)$. La potenza totale irradiata non può però dipendere dalla frequenza, per cui questa è espressa infine dalla:

$$P_{\text{irr}} = \frac{q^2 a^2}{c^3} \quad \text{Formula di Larmor}$$

Nel caso la carica sia la carica dell'elettrone $e^{(2)}$ e l'accelerazione a quella dell'oscillatore armonico, la potenza media irradiata è:

$$\bar{P}_{\text{irr}} = \frac{8\pi^2 e^2 \nu^2}{3 m c^3} \langle \varepsilon \rangle_T$$

dove $\langle \varepsilon \rangle_T$ è l'energia media dell'oscillatore armonico in equilibrio alla temperatura T , intesa come media termodinamica. La potenza assorbita è data invece da:

$$P_{\text{ass}} = \frac{\pi e^2}{3 m} u(\nu, T)$$

All'equilibrio queste due quantità devono evidentemente essere uguali $\bar{P}_{\text{irr}} = P_{\text{ass}}$ e questo permette di ricavare la forma della densità di energia:

$$u(\nu, T) = \frac{8\pi}{c^3} \nu^2 \langle \varepsilon \rangle_T \quad (1.1)$$

classicamente, l'energia media alla temperatura T è data dal teorema di equipartizione dell'energia. Per un oscillatore armonico questa è uguale a $k_B T$, per cui si ricava la:

$u(\nu, T) = \frac{8\pi}{c^3} \nu^2 k_B T$	Legge di Rayleigh-Jeans
--	--------------------------------

¹La realizzazione tipica di un *corpo nero* è una cavità con un piccolissimo foro di entrata. Ogni raggio che entra nella cavità ha una probabilità minima di essere riflesso esattamente verso il foro di ingresso, per cui esso prima o poi sarà assorbito dalle pareti della cavità. La radiazione all'interno sarà quindi in equilibrio con la cavità, alla temperatura della cavità stessa.

²Qui e spesso nel seguito con e^2 si intende la carica elettrica in unità gaussiane $e^2 \equiv e^2/4\pi\epsilon_0$

Questa relazione è una conseguenza inevitabile della fisica classica, ma mentre descrive correttamente il comportamento alle basse frequenze, è clamorosamente errata per le alte. Le conseguenze più gravi di questa particolare forma della densità di energia sono che non si riesce a riprodurre il massimo fenomenologico (legge di Stephan-Boltzmann) e – soprattutto – che l’irraggiamento totale alla temperatura T è infinito:

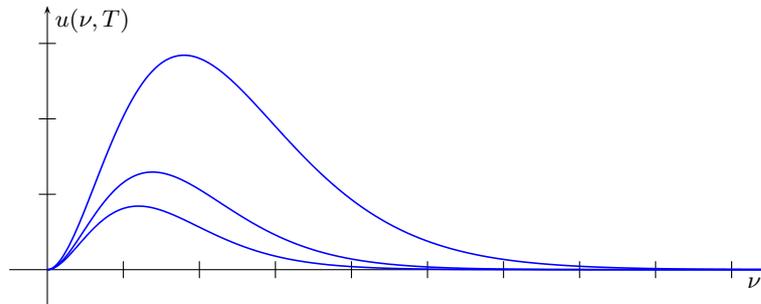
$$u(T) = \int_0^{+\infty} u(\nu, T) d\nu = \frac{8\pi}{c^3} k_B T \int_0^{+\infty} \nu^2 d\nu = \infty$$

Questo comportamento divergente alle alte frequenze è chiamato *catastrofe ultravioletta*. In un certo senso, il risultato classico è corretto: siccome esso rappresenta l’energia per unità di volume e di frequenza e l’interno della cavità del corpo nero è assimilabile a infiniti oscillatori armonici in equilibrio, il risultato infinito ha una sua logica interna.

Dall’altro lato, sperimentalmente si vedeva che l’andamento della densità di energia era totalmente diverso. Erano state stabilite diverse leggi empiriche, fra cui la prima di quelle che portano il nome di Wien, valida alle alte frequenze:

$$u(\nu, T) = \nu^3 f(\nu/T) \quad \text{I}^a \text{ Legge di Wien}$$

Dalla forma sperimentale della funzione:



Wien fu in grado di ricavare anche un’altra legge fenomenologica, ovvero che la funzione $f(\nu/T)$ ha un massimo tale che $\nu_{\max}/T = \text{costante}$:³

$$f(\nu/T): \lambda_{\max} T = \text{costante} = 2898 \mu\text{m} \cdot \text{K} \quad \text{II}^a \text{ Legge di Wien}$$

Dalla prima legge di Wien, si ha che:

$$u(T) = \int_0^{+\infty} \nu^3 f(\nu/T) d\nu$$

e posto $x \equiv \nu/T$:

$$u(T) = T^4 \int_0^{+\infty} x^3 f(x) dx$$

tale integrale, se finito, è quindi indipendente dalla temperatura. La prima legge di Wien conduce quindi a sua volta in maniera naturale all’altra legge empirica:

$$u(T) = \sigma \cdot T^4 \quad \sigma = 5.67 \times 10^{-8} \text{Wm}^{-2} \text{K}^{-4} \quad \text{Legge di Stephan-Boltzmann}$$

Questa relazione fornisce la legge con cui si raffredda un corpo nero rimanendo istante per istante in equilibrio.

In analogia alle curve dei gas, Wien propose una relazione che tenesse conto della prima legge di Wien e che presentasse un massimo, del tipo seguente:

$$u(\nu, T) = b\nu^3 e^{-a\nu/T}$$

³Si noti per inciso che l’esistenza di un massimo rende ragione della terminologia di “fondo cosmico a 3°K”, in quanto la radiazione di fondo cosmica è assimilabile ad una radiazione di corpo nero con massimo proprio a 3°K.

Questa relazione presenta però due grossi inconvenienti. In primo luogo, se si fissano i valori delle costanti a e b in modo da avere accordo alle alte temperature, c'è un grosso disaccordo alle basse temperature, dove la legge di Rayleigh-Jeans è invece corretta. In secondo luogo, occorre una “strana” costante dimensionale a per rendere l'esponente un numero puro.

Planck accettò come base la relazione proposta da Wien e rifiutò il concetto di equipartizione dell'energia classico. Cercò invece il numero di modi in cui l'energia si potesse suddividere fra i diversi oscillatori armonici e nel fare questo fece l'assunzione che l'energia totale potesse essere divisa in porzioni finite di energia ε_0 che lui chiamò “elementi di energia”.

Ora, la probabilità che un sistema si trovi in uno stato di energia n è data dalla statistica di Maxwell-Boltzmann:

$$\mathcal{P}(\varepsilon_n) = \frac{1}{N} e^{-\varepsilon_n/k_B T}$$

Questa distribuzione deve essere normalizzata, ovvero:

$$\sum_{n=0}^{\infty} \mathcal{P}(\varepsilon_n) = \sum_{n=0}^{\infty} \frac{1}{N} e^{-\varepsilon_n/k_B T} = 1 \quad \rightarrow \quad N = \sum_{n=0}^{\infty} e^{-\varepsilon_n/k_B T}$$

Per un oscillatore armonico unidimensionale risulta:

$$\langle \varepsilon \rangle = \sum_{n=0}^{\infty} \varepsilon_n \mathcal{P}(\varepsilon_n) = \frac{\sum_{n=0}^{\infty} \varepsilon_n e^{-\varepsilon_n/k_B T}}{\sum_{n=0}^{\infty} e^{-\varepsilon_n/k_B T}} = \frac{\sum_{n=0}^{\infty} n \varepsilon_0 e^{-n \varepsilon_0/k_B T}}{\sum_{n=0}^{\infty} e^{-n \varepsilon_0/k_B T}}$$

Si noti che in quest'ultimo passaggio è presente la grande novità: nello scrivere cioè l'energia dello stato n come $n\varepsilon_0$, ovvero come multiplo di questo “elemento di energia”.⁴

Posto $\alpha \equiv 1/k_B T$ si ricava:

$$\begin{aligned} \langle \varepsilon \rangle &= \frac{\varepsilon_0 \sum_{n=0}^{\infty} n e^{-n \varepsilon_0 \alpha}}{\sum_{n=0}^{\infty} e^{-n \varepsilon_0 \alpha}} = - \frac{d}{d\alpha} \frac{\sum_{n=0}^{\infty} n e^{-n \varepsilon_0 \alpha}}{\sum_{n=0}^{\infty} e^{-n \varepsilon_0 \alpha}} = - \frac{d}{d\alpha} \left[\ln \sum_{n=0}^{\infty} e^{-n \alpha \varepsilon_0} \right] = \\ &= \frac{d}{d\alpha} \ln (1 - e^{-\varepsilon_0 \alpha}) = \frac{\varepsilon_0}{e^{-\alpha \varepsilon_0} - 1} \end{aligned}$$

la sommatoria è infatti una serie geometrica. Ricordando la sostituzione fatta, la densità di energia (1.1) è data quindi dalla seguente:

$$u(\nu, T) = \frac{8\pi}{c^3} \nu^2 \frac{\varepsilon_0}{e^{\varepsilon_0/k_B T} - 1}$$

Per essere in accordo con la I° legge di Wien, occorre fare la posizione:

$$\boxed{\varepsilon_0 = h\nu} \tag{1.2}$$

da cui la densità di energia per unità di frequenza a temperatura T :

$$\boxed{u(\nu, T) = \frac{8\pi}{c^3} \frac{h\nu^3}{e^{h\nu/k_B T} - 1}}$$

⁴Per correttezza storica, è bene notare che in quel momento questo era per Planck poco più di un esercizio matematico, le implicazioni di questa posizione furono evidenti solo in seguito. A proposito della scoperta di Planck, si veda per esempio l'ottimo articolo “Max Planck: the reluctant revolutionary”, Physics World, physicsworld.com/cws/article/print/373.

Se in questa relazione si manda ν a zero, ovvero quando $h\nu \ll k_B T$, l'esponenziale si può espandere in serie e si trova:

$$u(\nu, T) = \frac{8\pi}{c^3} \frac{h\nu^3}{e^{\frac{h\nu}{k_B T}} - 1} \simeq \frac{8\pi}{c^3} \frac{h\nu^3}{\left[1 + \frac{h\nu}{k_B T} - \frac{1}{2}\left(\frac{h\nu}{k_B T}\right)^2 + \dots\right]} = \frac{8\pi}{c^3} h\nu^3 \frac{k_B T}{h\nu} = \frac{8\pi}{c^3} \nu^2 k_B T$$

ovvero proprio la legge di Rayleigh-Jeans. Lo stesso accade mandando h a zero (se si rendono cioè questi "elementi di energia" piccolissimi), perché in questo caso si può fare lo stesso sviluppo in serie dell'esponenziale.

Per alte frequenza, $h\nu \gg k_B T$, si può invece trascurare l'uno a denominatore e ricavare:

$$u(\nu, T) = \frac{8\pi}{c^3} \frac{h\nu^3}{e^{\frac{h\nu}{k_B T}} - 1} \simeq \frac{8\pi}{c^3} h\nu^3 e^{-\frac{h\nu}{k_B T}}$$

ovvero la prima legge di Wien e la forma della funzione da lui proposta. Considerando invece l'integrale sulla frequenza:

$$u(T) = \int_0^{+\infty} u(\nu, T) d\nu = \frac{8\pi}{c^3} \int_0^{+\infty} \frac{h\nu^3}{e^{\frac{h\nu}{k_B T}} - 1} d\nu$$

e con la posizione $x = h\nu/k_B T$:

$$u(T) = 8\pi \frac{(k_B T)^4}{h^3 c^3} \int_0^{+\infty} \frac{x^3}{e^x - 1} dx = \frac{8}{15} \pi^5 \frac{k_B^4}{h^3 c^3} T^4 \equiv \sigma \cdot T^4$$

ottenendo la legge di Stephan-Boltzmann.

La novità assoluta è qui la comparsa di una nuova costante fondamentale dimensionale, h . Planck osservò che per un valore particolare, ovvero:

$h = 6.62 \times 10^{-27} \text{ erg} \cdot \text{s}$	costante di Planck
---	---------------------------

l'accordo con tutti i dati sperimentali era perfetto. Tale costante è quindi divenuta nota con il nome di *costante di Planck* ed ha le dimensioni di un'azione, per questo è chiamata a volte anche *quanto d'azione*.

1.2 L'effetto fotoelettrico

L'effetto fotoelettrico, ovvero l'emissione di un elettrone da un metallo quando questo viene colpito da radiazione, presentava diverse caratteristiche non spiegabili in base alla teoria classica dell'elettromagnetismo, che vedeva la radiazione elettromagnetica come delle onde.

Le caratteristiche dell'effetto fotoelettrico sono:

- L'effetto si verifica solo se la radiazione ha una frequenza ν superiore ad una certa frequenza di soglia ν_0 , dipendente dal materiale;
- L'energia degli elettroni emessi è indipendente dalla intensità;
- Il numero di elettroni emessi è invece proporzionale alla intensità;
- L'energia cinetica massima dell'elettrone emesso dipende dalla frequenza ν secondo una legge del tipo $T_{\max} = a\nu - b$;
- L'effetto è istantaneo.

Questo comportamento non è spiegabile con la teoria ondulatoria della luce.

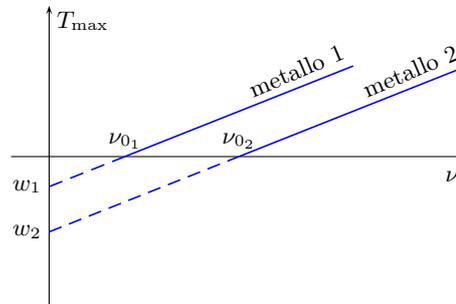
Per afferrare le difficoltà dell'approccio classico, si consideri una sorgente di 1 cd a distanza di 3 metri: in un centimetro quadro vi sarà allora un flusso di energia di circa 1 erg/s. In un materiale tipico l'energia degli elettroni è di circa 10^{-12} erg, mentre l'area che un atomo espone alla radiazione

è circa 10^{-16} cm per cui riceve un flusso di circa 10^{-16} erg/s. Un elettrone dovrebbe quindi essere espulso in 10^4 s, mentre l'effetto è istantaneo. Inoltre, siccome l'energia di un'onda è legata alla sua intensità, la teoria ondulatoria non riesce a spiegare né l'esistenza di una frequenza di soglia, né perché l'energia degli elettroni emessi debba dipendere dalla frequenza e non dall'intensità, mentre invece è il numero di elettroni a dipendere da essa.

Einstein⁵ si rese conto che l'ipotesi di Planck significava – in pratica – che lo scambio di energia con il campo elettromagnetico avveniva per piccole quantità discrete, che battezzò *quanti*. Egli fece notare che se si prendeva sul serio questa ipotesi molti fenomeni, fra cui l'effetto fotoelettrico, avevano una spiegazione semplice. Egli pertanto prese come base il fatto che l'energia sia distribuita nello spazio non in maniera continua, ma in modo discreto, per quanti. Detto con le sue parole, “*Secondo l'ipotesi che voglio qui proporre, quando un raggio di luce si espande partendo da un punto, l'energia non si distribuisce su volumi sempre più grandi, bensì rimane costituita da un numero finito di quanti di energia localizzati nello spazio, che si muovono senza suddividersi e che non possono essere assorbiti o emessi parzialmente*”.⁶

Dallo studio dell'entropia di un gas e dell'entropia del corpo nero, Einstein dedusse che il quanto di energia deve avere proprio la forma data in (1.2), $E = h\nu$.

Una volta fatto questo importante passo, la spiegazione dell'effetto fotoelettrico diviene semplice. Quando un quanto di radiazione elettromagnetica colpisce la superficie di un metallo, esso può essere assorbito e la sua energia ceduta ad un elettrone. Di conseguenza la massima energia (cinetica) con cui un elettrone può essere emesso è data dall'energia del quanto incidente diminuita dell'energia w necessaria per estrarre l'elettrone dal metallo (*lavoro di estrazione*). In altre parole, la massima energia cinetica dell'elettrone è la parte dell'energia del quanto che resta dopo averne spesa una parte nel lavoro di estrazione. L'ipotesi di Einstein prevede quindi in maniera naturale che l'energia massima segua una relazione lineare $T_{\max} = h\nu - w$, con coefficiente angolare dato dalla costante di Planck h . Esiste quindi anche una frequenza di soglia $\nu_0 = w/h$ al di sotto della quale l'elettrone non può essere emesso: questo accade quando l'energia del quanto è inferiore al lavoro di estrazione del metallo.



Questa relazione è stata perfettamente confermata sperimentalmente,⁷ con un valore della costante di Planck identico a quello richiesto dalla distribuzione di corpo nero.

L'intensità della radiazione incidente è data dal numero di quanti per unità di volume e di tempo, e questo giustifica il fatto che il numero di elettroni emessi dipende dall'intensità e non dalla frequenza, perché evidentemente maggiore è il numero di quanti e maggiore è la probabilità di “colpire” un elettrone.

Infine, l'effetto è istantaneo perché il quanto di energia può sia avere energia sufficiente ed essere assorbito (e in questo caso l'elettrone è espulso dal metallo istantaneamente), sia non essere assorbito o non avere energia sufficiente.

Il quanto di luce risulta essere estremamente piccolo, e questo rende ragione del fatto che in condizioni normali esso non è percepibile. Si consideri ad esempio una lampadina da 100 Watt ed

⁵Nel suo famoso articolo del 1905 (detto impropriamente sull'effetto fotoelettrico), “Un punto di vista euristico relativo alla generazione e trasformazione della luce”.

⁶Storicamente, il grande salto concettuale di Einstein fu di considerare l'analogia fra la distribuzione dei gas e quella del corpo nero come una corrispondenza sostanziale. In altri termini egli ipotizzò che come un gas è formato da tante particelle, così il campo elettromagnetico – proprio perché segue sostanzialmente la stessa legge – doveva essere composto da tanti piccoli “quanti”.

⁷Da Millikan nel 1916.

uno schermo con un foro di 1 centimetro quadrato posto ad un metro: in queste condizioni, ogni secondo circa 24×10^{15} quanti di luce attraversano il foro.

1.3 Il problema dei calori specifici

All'inizio dell'ottocento, Dulong e Petit osservarono che il calore specifico molare dei solidi semplici è praticamente costante e vale $C_m \simeq 3R = 5.96 \text{ cal/mol}^\circ\text{C}$. Per esempio:

	Pb	Pt	Ag	Cu	Bi	Zn	Al
C_m	6.43	6.21	6.03	5.92	6.22	6.14	5.83

Questa legge empirica poté trovare giustificazione nell'ambito della fisica classica grazie all'interpretazione statistica di Boltzmann.

Il reticolo cristallino dei solidi può essere schematizzato come un insieme di oscillatori armonici che vibrano per agitazione termica intorno alla posizione di equilibrio. Il numero di modi di oscillazione per mole è dato da $3\mathcal{N}$, dove \mathcal{N} è il numero di Avogadro, per cui l'energia totale è data da:

$$U(T) = 3\mathcal{N}\langle E \rangle$$

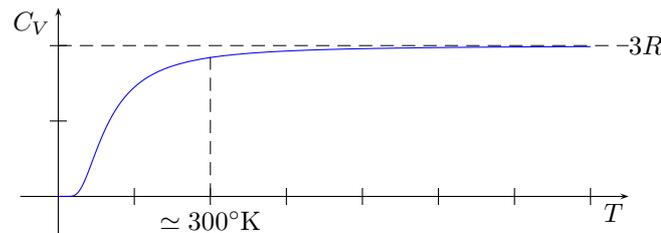
Per il teorema di equipartizione dell'energia, ad ogni oscillatore armonico compete un'energia media $\langle E \rangle = k_B T$, per cui l'energia totale è:

$$U(T) = 3\mathcal{N}k_B T$$

Il calore specifico a volume costante molare è definito come la derivata dell'energia totale rispetto alla temperatura, per cui:

$$C_V = \frac{dU}{dT} = 3\mathcal{N}k_B \equiv 3R = 5.96 \text{ cal/mole}^\circ\text{C}$$

Questo risultato costituisce un successo del modello classico, però non riesce a spiegare perché gli elementi leggeri hanno un calore specifico molto più basso,⁸ né perché esistano comunque delle discrepanze rispetto al valore teorico. Inoltre, esso prevede per il calore specifico un valore costante rispetto alla temperatura, che è in disaccordo evidente con l'esperienza, in quanto questo va invece a zero con un andamento tipo T^3 :



La legge di Dulong e Petit dunque curiosamente *funziona in generale, ma non abbastanza*.

Einstein si rese conto che il problema era lo stesso di quello che aveva affrontato Planck per il corpo nero, ovvero che il teorema di equipartizione di energia non sembra essere del tutto corretto. In particolare, Einstein fece l'assunzione che anche gli scambi di energia fra gli atomi avvenissero per quanti⁹ ed utilizzò l'espressione dell'energia media trovata da Planck:

$$U(T) = 3\mathcal{N}\langle E \rangle = 3\mathcal{N} \frac{h\nu}{e^{\frac{h\nu}{k_B T}} - 1}$$

⁸Ecco per esempio i valori per il Boro, il Carbonio in forma di grafite e diamante e per il Silicio:

	B	C (gr.)	C (diam.)	Si
C_m	2.84	2.60	1.44	5.00

⁹Oltre al campo elettromagnetico postulato da Planck e lo scambio fra atomi e campo elettromagnetico postulato dallo stesso Einstein per l'effetto fotoelettrico.

Si noti che questa forma dell'energia media introduce una dipendenza dalla frequenza, il che significa che in linea di principio ogni atomo può avere una frequenza diversa. Einstein ipotizzò quindi una frequenza media $\bar{\nu}$ comune a tutti gli atomi, ottenendo:

$$U(T) = 3\mathcal{N} \frac{h\bar{\nu}}{e^{\frac{h\bar{\nu}}{k_B T}} - 1}$$

il che permette di ricavare per il calore specifico:

$$C_V = \frac{dU}{dT} = \frac{d}{dT} 3\mathcal{N} \frac{h\bar{\nu}}{e^{\frac{h\bar{\nu}}{k_B T}} - 1} = 3\mathcal{N} k_B \left(\frac{h\bar{\nu}}{k_B T} \right)^2 \frac{e^{\frac{h\bar{\nu}}{k_B T}}}{\left[e^{\frac{h\bar{\nu}}{k_B T}} - 1 \right]^2}$$

il limite ad alte temperature ($h\bar{\nu} \ll k_B T$) dell'espressione precedente si ottiene sviluppando in serie l'esponenziale:

$$\begin{aligned} C_V &\simeq 3\mathcal{N} k_B \left(\frac{h\bar{\nu}}{k_B T} \right)^2 \frac{1 + \frac{h\bar{\nu}}{k_B T}}{\left[\cancel{\lambda} + \frac{h\bar{\nu}}{k_B T} - \cancel{\lambda} \right]^2} = 3\mathcal{N} k_B \left(\frac{h\bar{\nu}}{k_B T} \right)^2 \left(\frac{k_B T}{h\bar{\nu}} \right)^2 \left(1 + \frac{h\bar{\nu}}{k_B T} \right) \\ &= 3\mathcal{N} k_B \left(1 + \frac{h\bar{\nu}}{k_B T} \right) \end{aligned}$$

dunque:

$$\lim_{T \rightarrow +\infty} C_V = 3\mathcal{N} k_B \equiv 3R$$

Viceversa, a basse temperature (ovvero $h\bar{\nu} \gg k_B T$), l'uno a denominatore è trascurabile e si ottiene per il calore specifico:

$$\lim_{T \rightarrow 0} C_V = 3\mathcal{N} k_B \left(\frac{h\bar{\nu}}{k_B T} \right)^2 e^{-\frac{h\bar{\nu}}{k_B T}}$$

Il modello di Einstein dei calori specifici rende ragione del perché questo diminuisce con la temperatura. Infatti, siccome l'energia può essere scambiata solo per quanti discreti, quando questi hanno energia media inferiore a quella classica i gradi di libertà relativi restano per così dire "congelati" e non contribuiscono al calore specifico. La temperatura alla quale questo diventa evidente è $T_c = \frac{h\bar{\nu}}{k_B}$. Questo modello permette anche di spiegare perché alcuni materiali, come il boro, il silicio ed il carbonio, hanno un calore specifico anomalo: in effetti, basta supporre che la frequenza media sia più elevata degli altri materiali, per cui la T_c cada in un range più alto di temperatura. Questa ipotesi è confermata dal fatto sperimentale che aumentando la temperatura anche il calore specifico di questi elementi anomali tende al valore $3R$.

L'accordo del modello è ottimale, ma non perfetto e questo è dovuto all'approssimazione sulle frequenze. Infatti l'uso di una frequenza media significa in pratica tagliare le basse frequenze, che sono quelle che danno contributo alle basse temperature.¹⁰ In effetti un modello più completo dovuto a Debye che tiene conto delle diverse frequenze di oscillazione – e che esula completamente dallo scopo di questo corso – risulta in ottimo accordo con i dati sperimentali.

1.4 Conclusioni

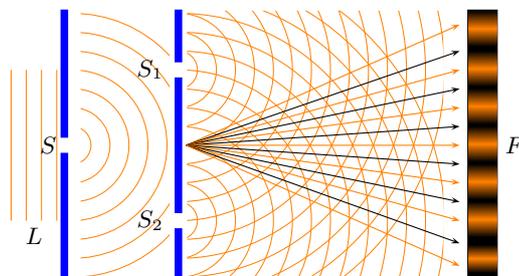
È il momento di trarre le dovute conclusioni di quanto esposto in questo capitolo.

Si è visto che alcuni fenomeni inspiegabili dalla "concezione classica del mondo" divengono chiari e semplici se si ipotizza un comportamento corpuscolare della radiazione elettromagnetica. Secondo questa visione, l'energia non si distribuisce con continuità nello spazio, ma è localizzata in una serie di pacchetti indivisibili ciascuno di energia $E = h\nu$, denominati **fotoni**.

Non si può d'altra parte ignorare il fatto che i campi elettromagnetici si comportano effettivamente come delle onde, come dimostrò in particolare l'esperimento che Young condusse nel 1801.

¹⁰Per rendersene conto, si consideri la forma dell'esponente $\frac{h\bar{\nu}}{k_B T}$.

Convinto che la luce fosse composta da onde e non da corpuscoli come aveva sostenuto Newton, Young pensò di far interferire due raggi di luce per vedere cosa accadesse e dimostrare così la sua teoria. Questo esperimento, insieme allo sviluppo teorico successivo dovuto a Maxwell, sembrava aver chiuso definitivamente la questione (allora aperta) sulla natura della luce e delle onde elettromagnetiche. Ecco schematicamente l'esperimento di Young, noto anche come esperimento delle due fenditure:



L è un raggio di luce rappresentato da onde piane¹¹ che incide su una fenditura S , su cui fa diffrazione. Il fronte d'onda che si crea in seguito al principio di Huygens–Fresnel viene fatto passare attraverso due fenditure praticate in uno schermo, S_1 e S_2 , per produrre due fronti d'onda in fase, che vengono poi raccolti su uno schermo F . Si osserva che la distribuzione della luce su F quando entrambe le fenditure sono aperte è completamente diversa dalla semplice somma delle distribuzioni ottenute con le fenditure aperte una alla volta, originando un fenomeno di interferenza. Questo significa che nel generico punto p dello schermo non sono le intensità $I_1(p)$ e $I_2(p)$ dei fasci passanti per S_1 e S_2 che si sommano, ma le ampiezze $A(p) = A_1(p) + A_2(p)$, e questo fenomeno è caratteristico e peculiare della propagazione per onde. Nei punti in cui si ha l'interferenza costruttiva dei fronti d'onda appare una zona luminosa, dove si ha interferenza distruttiva si ha una banda buia.

Ma se si mettono dei rivelatori sufficientemente sensibili in S_1 o S_2 , si osserva che la luce si comporta in maniera corpuscolare e passa sempre o attraverso l'una o attraverso l'altra fenditura, cioè ciascuno di essi rivela un fotone intero (di energia pari ad $h\nu$) oppure nulla. I rivelatori in S_1 e S_2 non scattano mai insieme, il fotone non si divide mai in due: o segue un percorso o segue l'altro. Ma nel momento in cui si posizionano i rivelatori per individuare il percorso del fotone, l'interferenza scompare: anche se i fotoni non vengono assorbiti dai rivelatori, l'intensità della luce in F diventa la somma delle intensità $I(p) = I_1(p) + I_2(p)$. Inoltre è impossibile prevedere attraverso quale foro passerà un fotone.

Ne consegue che le onde elettromagnetiche oltre ai tipici aspetti ondulatori che le caratterizzano, possiedono anche delle proprietà che nella fisica classica sono associate alle particelle, ovvero – detto in termini classici – *i campi possiedono una doppia natura, sia di onda che di particella*. È questo il primo indizio di quella che viene chiamata *dualità onda-particella* e che sarà studiata approfonditamente in §2.5.

La teoria quantistica deve tuttavia essere vista come una estensione della teoria classica e non come una sua negazione. In altri termini, nel limite in cui la granularità dell'energia non è percepibile le leggi della fisica classica conservano la loro validità, questo fatto si esprime formalmente richiedendo che al limite per $h \rightarrow 0$ le leggi quantistiche restituiscano le leggi della fisica classica. Questa è l'essenza del cosiddetto *Principio di Corrispondenza di Bohr*.

¹¹Attualmente si usa un LASER come sorgente di luce monocromatica, Young nell'esperimento originale utilizzò un raggio di luce solare, isolato tramite la fenditura S .

Capitolo 2

I primi modelli atomici e la vecchia teoria dei quanti

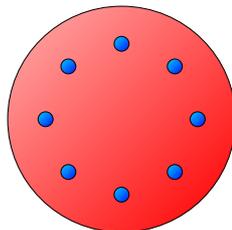
Un'applicazione fondamentale dell'idea di quantizzare l'energia tramite la costante di Planck h è data dalla spiegazione della struttura atomica. Si passeranno in rassegna in questo capitolo i primi modelli di Thomson e Rutherford ed il problema degli spettri atomici, per poi passare al primo modello quantistico di Bohr-Sommerfeld.

Infine, si discuterà l'ipotesi di de Broglie, che ha gettato le basi della moderna meccanica quantistica e si discuterà della dualità onda-particella.

2.1 Il modello di Thomson

Questo modello fu proposto da J. J. Thomson nel 1904 e prevede una carica positiva diffusa in una sfera di dimensioni atomiche, all'interno della quale gli elettroni (chiamati da Thomson "corpuscoli") sono liberi di muoversi. Questo modello è anche detto "a panettone" perché gli elettroni sono sparpagliati nella sfera positiva come l'uvetta in un panettone.

Gli elettroni del modello di Thomson non sono tuttavia disposti a caso e non sono statici, sono invece posizionati in una sorta di anelli orbitali. L'intera struttura è resa stabile dal fatto che quando un elettrone tende ad allontanarsi esso risente di una maggiore carica positiva che lo richiama con maggiore forza.



Secondo il modello di Thomson, le righe spettrali degli atomi sono generate dalle oscillazioni degli elettroni intorno alle posizioni di equilibrio. Infatti, all'interno di una sfera uniformemente carica di raggio R , la forza cui è soggetto un elettrone è data da:

$$F = \frac{4}{3} r^3 \frac{e^2}{r^2} = Ze^2 \frac{r}{R^3}$$

che è una forza di tipo elastico, per cui l'elettrone può vibrare con frequenza:

$$\nu = \frac{1}{2\pi} \sqrt{\frac{k}{m_e}} \approx \frac{\sqrt{Z}}{2\pi} \cdot 10^{16} \text{ Hz}$$

che risulta dell'ordine delle frequenze ottiche.

Il modello atomico di Thomson presentava anche altri vantaggi, in particolare il fatto che gli atomi avessero una dimensione fissa e che gli elettroni si dispongono in strati, rendendo ragione delle proprietà periodiche degli elementi.

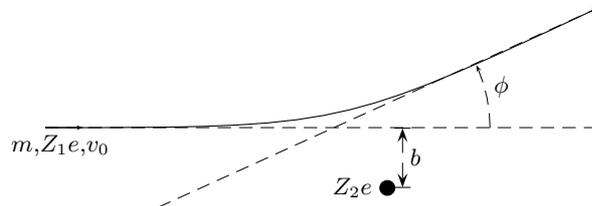
Questo modello fu però escluso dall'esperienza di Rutherford del 1909.

2.2 Il modello di Rutherford

L'esperimento di Rutherford¹ fu concepito per sondare la struttura dell'atomo e provare la validità del modello atomico di Thomson. Questo esperimento diede però dei risultati in completo disaccordo con questo modello e portò alla concezione del modello atomico planetario di Rutherford.

Rutherford utilizzò un fascio di particelle α incidenti su una sottile lamina di oro di circa 4×10^{-5} cm (per ridurre le riflessioni multiple) circondato da un foglio circolare di ZnS utilizzato come rivelatore. Quest'ultimo infatti scintilla quando colpito da particelle α .

Si consideri una particella di massa m e carica Z_1e , incidente con velocità v_0 e con parametro d'urto b , deflessa da una carica Z_2e . Essa sarà deflessa di un angolo ϕ , in genere piccolo.



Se gli angoli di deflessione sono piccoli, si può approssimare la tangente con l'angolo $\tan \phi \simeq \phi$. Ne segue che si può scrivere:

$$\tan \phi \simeq \phi = \frac{p_{\perp}}{p} = \frac{1}{p} \int_{-\infty}^{+\infty} F_{\perp} dt$$

dove p_{\perp} e F_{\perp} sono la componente perpendicolare rispettivamente del momento e della forza. La forza risentita dalla particella è la forza di Coulomb e chiamando b il parametro d'urto (distanza che separa la particella Z_1e dal nucleo Z_2e), si ha:

$$\vec{F} = \frac{e^2 Z_1 Z_2}{r^3} \vec{r} = \frac{e^2 Z_1 Z_2 b}{(\sqrt{x^2 + b^2})^3}$$

da cui:

$$\phi \simeq \frac{e^2 Z_1 Z_2}{p} \int_{-\infty}^{+\infty} \frac{b}{(x^2 + b^2)^{3/2}} dt$$

ponendo ora $dx = v_0 dt$, da cui $dt = \frac{1}{v_0} dx = \frac{m}{p} dx$, e ponendo $y = x/b$ si ricava:

$$\phi \simeq \frac{e^2 Z_1 Z_2 m}{p^2 b} \int_{-\infty}^{+\infty} \frac{y}{(1 + y^2)^{3/2}} dy$$

che risolto fornisce:

$$\boxed{\phi \simeq \frac{2me^2 Z_1 Z_2}{p^2 b}}$$

Per verificare se l'approssimazione $\tan \phi \simeq \phi$ è lecita, si consideri il caso tipico dell'esperimento di Rutherford, ovvero $Z_1 = 2$, $Z_2 = 50$ e $E = p^2/2m \simeq 1$ MeV. In questo caso risulta:

$$\phi \simeq \frac{10^{-13}}{b}$$

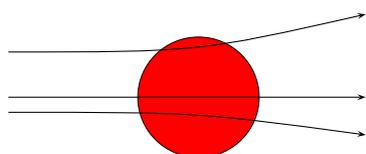
¹Più propriamente, esperimento di Geiger e Marsden. Fu eseguito da Hans Geiger e Ernest Marsden nel 1909, sotto la direzione di Ernest Rutherford presso il laboratorio di fisica dell'Università di Manchester.

con b in metri. Nel modello di Thomson, la massima deflessione si ha quando la particella α passa sul bordo esterno dell'atomo, all'interno dell'atomo di Thomson infatti la forza di attrazione diminuisce proporzionalmente al raggio e all'esterno essa risente di una forza minore a causa della distanza. Ponendo $b \simeq 10^{-10}$ m si ricava $\tan \phi \simeq 10^{-3}$, l'approssimazione è dunque perfettamente lecita e l'angolo di deflessione atteso è molto piccolo. Se invece la particella α passa vicino un elettrone, la deflessione è trascurabile a causa della differenza di massa.

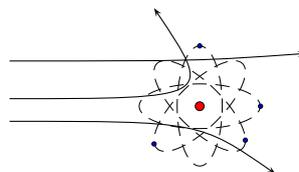
L'esperienza condotta da Rutherford contraddisse clamorosamente questa previsione.

Rutherford si aspettava che tutte le particelle passassero con una deflessione di al più pochi gradi, anche considerando le diffusioni multiple. Invece, egli osservò che alcune particelle (circa 1/8000) erano deflesse ad angoli anche superiori a 90° . Il risultato fu totalmente inaspettato, per usare le parole di Rutherford stesso: *“Fu l'evento più incredibile mai successo in vita mia. Era quasi incredibile quanto lo sarebbe stato sparare un proiettile da 15 pollici su un foglio di carta velina e vederlo tornare indietro e colpirti. Pensandoci, ho capito che questa diffusione all'indietro doveva essere il risultato di una sola collisione e quando feci il calcolo vidi che era impossibile ottenere qualcosa di quell'ordine di grandezza a meno di considerare un sistema nel quale la maggior parte della massa dell'atomo fosse concentrata in un nucleo molto piccolo. Fu allora che ebbi l'idea di un atomo con un piccolissimo centro massiccio e carico.”*²

Egli fu in grado di calcolare il raggio della carica che avrebbe potuto produrre una tale deflessione, e trovò il valore di circa 10^{-14} m, mentre la dimensione di un atomo di oro era di circa 10^{-10} m: ne concluse che la maggior parte della massa doveva essere concentrata in un piccolo nucleo centrale e propose quindi un modello di tipo planetario, con un nucleo centrale carico positivamente ed una nuvola di elettroni su orbite analoghe ai pianeti per compensare la carica del nucleo e rendere l'atomo elettricamente neutro.



Modello di Thomson



Modello di Rutherford

Il modello di Rutherford appare affascinante e logico: la legge di Coulomb prende il posto della legge di gravitazione di Newton e gli elettroni possono quindi orbitare intorno al nucleo come i pianeti intorno al sole. Tuttavia, questo modello non può essere corretto secondo l'elettrodinamica classica, per diversi motivi.

In primo luogo, non si riesce a spiegare l'uguaglianza degli atomi di uno stesso elemento: infatti i dettagli delle orbite dipendono in maniera cruciale dalle condizioni iniziali e quindi atomi diversi dello stesso elemento avrebbero orbite diverse. Tutta l'esperienza chimica dimostra invece la perfetta equivalenza degli atomi di un determinato elemento, indipendentemente dal loro metodo di preparazione.

In secondo luogo, le leggi classiche non riescono a spiegare perché rimuovendo o aggiungendo un elettrone ad un atomo “planetario” le caratteristiche fisiche e chimiche di un atomo possano cambiare in maniera drastica. Si pensi ad esempio al gas Neon, inerte chimicamente, cui basta rimuovere un elettrone per ottenere il Fluoro, uno degli elementi più reattivi. La rimozione o l'aggiunta di un pianeta in un sistema planetario non ha conseguenze altrettanto drammatiche.

Infine, il problema più grave consiste nel fatto che secondo le leggi dell'elettrodinamica classica una carica in moto accelerato, come è un elettrone che orbita intorno ad un nucleo, irraggia perdendo energia. Gli atomi del modello di Rutherford sono quindi altamente instabili, perché la perdita di energia per irraggiamento produce un'orbita a spirale e la caduta dell'elettrone sul nucleo in tempi estremamente brevi.

L'atomo planetario proposto da Rutherford contrasta quindi con l'elettrodinamica classica e pose le basi per la costruzione da parte di Niels Bohr del suo modello atomico quantistico. Prima di

²Citato da David C. Cassidy, Gerald James Holton, Gerald Holton, Floyd James Rutherford, *Understanding Physics* Harvard Project Physics Published by Birkhäuser, 2002.

affrontare questo modello, tuttavia, si accennerà un altro aspetto degli atomi che non era possibile spiegare né con il modello di Rutherford, né con la fisica classica: gli spettri atomici.

2.3 Gli spettri atomici

Diversamente dalla radiazione di corpo nero, gli atomi emettono radiazioni a frequenze ben precise e determinate, costituendo così una serie di *righe spettrali* definite e caratteristiche di ciascun atomo. Ancora una volta, questo fenomeno era inspiegabile con le leggi della fisica classica e con il modello di Rutherford, sebbene (in linea principio) spiegabile con il modello di Thomson.

Già nel 1855 Balmer aveva stabilito una legge empirica che permetteva di calcolare le righe emesse nel visibile dall'idrogeno atomico. Egli aveva notato che esisteva una particolare lunghezza d'onda che era legata a tutte le linee dello spettro secondo la relazione:

$$\lambda = B \left(\frac{n^2}{n^2 - 4} \right) \quad \begin{cases} n > 2 \\ B = 3645.6 \text{ \AA} \end{cases}$$

L'insieme di queste righe è noto come *serie di Balmer*.

Rydberg riuscì a generalizzare questa formula riscrivendola in termini del numero d'onda invece della lunghezza, nella forma:

$$\frac{1}{\lambda} = \frac{4}{B} \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \equiv R_H \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \quad \begin{cases} n_2 > n_1 \\ R_H = 1.09737353 \times 10^7 m^{-1} \end{cases}$$

dove R_H è detta *costante di Rydberg per l'atomo di idrogeno*. Questa formula si riduce evidentemente alla formula di Balmer con $n_1 = 2$ e permise di localizzare le serie con $n_1 = 1$ (serie di Lyman), $n_1 = 3$ (serie di Paschen), $n_1 = 4$ (serie Brackett) e $n_1 = 5$ (serie di Pfund).

Ritz osservò ulteriormente che le frequenze tipiche di tutte le sostanze possono essere espresse da una formula simile a quella di Rydberg:

$$\nu = \frac{R}{c} \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \quad n_2 > n_1$$

ed enunciò quello che va sotto il nome di *principio di combinazione di Ritz*: le linee spettrali di un dato elemento possono essere ricavate sottraendo (o sommando) le frequenze di altre due linee spettrali. L'insieme delle frequenze di base che possono essere utilizzate per determinare tutte le altre sono dette *termini spettroscopici*. Tutte le linee spettrali possono quindi essere espresse come differenza di due di questi termini:

$$\nu_{ij} = F_i - F_j$$

L'aspetto significativo di queste regole empiriche è che le linee spettrali possono essere ricavate come differenza di due numeri interi e che i termini spettroscopici sono in numero molto minore delle linee spettrali realmente osservate.

Un modello atomico che fosse stabile, compatibile con l'esperienza di Rutherford e che potesse spiegare le caratteristiche degli spettri atomici fu proposto da Bohr nel 1913.

2.4 Il modello di Bohr-Sommerfeld

Sulla scia della rivoluzione introdotta da Planck riguardante i campi elettromagnetici, ovvero la quantizzazione dell'energia, Bohr propose un modello che introduceva ugualmente un'ipotesi senza precedenti: la quantizzazione di sistemi materiali. In pratica, Bohr propose di risolvere i problemi del modello di Rutherford *postulando* che non tutte le orbite "planetarie" "classiche" sono possibili, ma solo alcune di esse secondo certe regole precise.

Questo approccio permette di spiegare l'uniformità degli atomi di uno stesso elemento perché se solo determinate orbite sono possibili, queste saranno le stesse per tutti gli atomi; permette anche di spiegare gli spettri a righe degli atomi, perché nel salto di un elettrone da un'orbita possibile

ad un'altra di differente energia,³ viene emesso un fotone di frequenza precisa data dalla relazione di Planck $\Delta E = h\nu$. Se si prende infatti per buona l'ipotesi di Planck, un elettrone in un'orbita con energia E_2 potrà andare solo in un'orbita con energia data da $E_1 = \pm h\nu$ (o multipli di $h\nu$). Si vedranno quindi solo le frequenze relative ai salti fra le orbite permesse, generando così una struttura a righe.

Bohr notò che nel modello di Rutherford è assente una lunghezza caratteristica del sistema, il che poteva rendere conto del fatto che gli atomi di Rutherford non avessero dimensioni prestabilite. Se invece si teneva conto della costante di Planck h , era possibile costruire la quantità:

$$l = \frac{h^2}{m_e e^2} \approx 10^{-10} \text{ m}$$

che è grosso modo dell'ordine di grandezza delle dimensioni atomiche. Bohr immaginò allora che la costante di Planck dovesse avere un ruolo in un modello atomico che fosse corretto, e basò quindi il suo modello su tre postulati.

Primo postulato: Un elettrone può muoversi solo su determinate orbite permesse, sulle quali, per qualche ragione ancora ignota, l'elettrone non irraggia.

Secondo postulato: Un atomo irraggia solo quando per qualche motivo l'elettrone passa da un'orbita all'altra, e nel fare questo emette una frequenza data dalla relazione di Planck:

$$\nu = \frac{E_2 - E_1}{h}$$

dove E_2 e E_1 rappresentano rispettivamente l'energia dell'orbita origine e di destinazione. A meno di fattori costanti, questo postulato coincide concettualmente con la regola di Ritz.

Sulla base dell'ipotesi di Planck, Bohr propose che le orbite permesse fossero solo quelle la cui energia soddisfa la relazione:

$$E_n = -\frac{1}{2}nhf \quad (2.1)$$

dove f è la frequenza associata al sistema atomo di idrogeno (frequenza di rivoluzione) e n è un intero positivo. In un moto circolare, uguagliando la forza coulombiana all'accelerazione centripeta si ha:

$$\frac{Ze^2}{r^2} = m \frac{v^2}{r} \quad \rightarrow \quad mv^2 = \frac{Ze^2}{r}$$

per cui l'energia in una determinata orbita di raggio a può essere espressa da:

$$E = \frac{1}{2}mv^2 - \frac{Ze^2}{a} = -\frac{1}{2} \frac{Ze^2}{a}$$

Segue immediatamente:

$$a = \frac{Ze^2}{2|E|}$$

Considerando ora la frequenza di rivoluzione di un moto circolare $f = \frac{v}{2\pi a}$ e sostituendo in essa il valore di a appena scritto e il valore di $v = \sqrt{Ze^2/ma}$, si ricava:

$$f = \frac{v}{2\pi a} = \frac{2|E|}{2\pi Ze^2} \sqrt{\frac{Ze^2}{ma}} = \frac{|E|}{\pi Ze^2} \sqrt{\frac{Ze^2 2|E|}{m Ze^2}} = \frac{\sqrt{2}}{\pi} \frac{|E|^{3/2}}{Ze^2 \sqrt{m}}$$

la condizione di quantizzazione sull'energia fornisce quindi:

$$\boxed{|E_n| = \frac{2\pi^2 m Z^2 e^4}{h^2 n^2} \quad n \neq 0}$$

³Nel modello di Bohr non c'è *passaggio* da un'orbita all'altra, ma una sorta di *salto* perché gli stati intermedi non sono permessi.

Si noti come la differenza di due energie permette di ritrovare proprio la formula di Balmer e di Rydberg, prevedendo per la costante di Rydberg il valore:

$$R_H = \frac{2\pi^2 m e^4}{h^3 c}$$

in ottimo accordo con il valore sperimentale. Nello stato fondamentale dell'atomo di idrogeno, $Z = 1$ e $n = 1$, si trova:

$$\begin{cases} E_0 = -13.6 \text{ eV} \\ a_0 = \frac{h^2}{4\pi^2 m e^2} \equiv \frac{\hbar^2}{m e^2} \simeq 0.5 \times 10^{-10} \text{ m} \end{cases} \quad \text{raggio di Bohr}$$

che è in eccellente accordo con i dati sperimentali. Si noti che il rapporto $h/2\pi$ è così comune in meccanica quantistica da essere indicato con un segno particolare detto *acca tagliato*, \hbar .

Ricordando che si lavora sempre in unità gaussiane (nelle quali la carica elettrica $e \equiv e/4\pi\epsilon_0$), si noti che i livelli energetici dell'atomo di idrogeno possono essere scritti nella forma:

$$E_n = -\frac{1}{n^2} \frac{1}{2} \frac{m c^2 e^2}{\hbar^2 c^2} \equiv \frac{1}{2n^2} m c^2 \alpha$$

dove si è fatta la posizione:

$$\alpha \equiv \frac{e^2}{\hbar c} \simeq \frac{1}{137} \quad \text{costante di struttura fine}$$

La costante di struttura fine è adimensionale. Analogamente, la costante di Rydberg si può riscrivere nella forma:

$$R_H = \frac{1}{4\pi} \frac{\alpha}{a_0}$$

Sul significato della costante di struttura fine

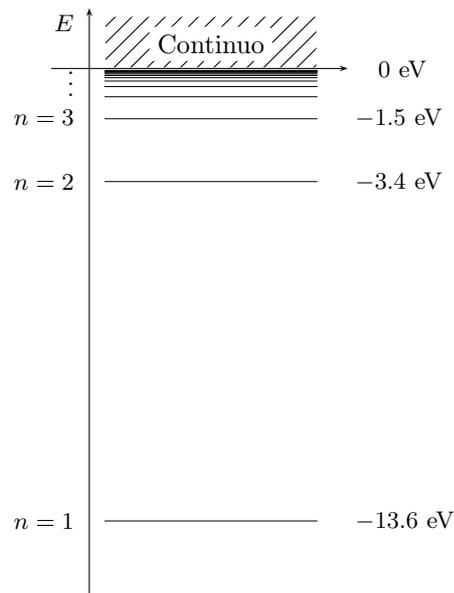
La costante di struttura fine misura l'intensità dell'interazione del campo elettromagnetico. L'energia di Bohr dipende da α^2 perché in questo caso intervengono due interazioni elettromagnetiche: la prima del campo con l'elettrone e la seconda del campo con il nucleo. Nel caso della costante di Rydberg, invece, esiste una sola interazione, ovvero quella del campo con l'elettrone e per questo R_H dipende solo da α .

Dunque, per ogni interazione che coinvolge il campo elettromagnetico interviene un fattore α e la sua piccolezza garantisce che le interazioni multiple siano relativamente rare. Inoltre, dall'espressione dell'energia di Bohr in termini della costante di struttura fine si vede che, grazie alla piccolezza della costante, per valori di Z non troppo elevati il moto dell'elettrone è non relativistico.

Si noti inoltre che la piccolezza della costante di struttura fine permette di trattare le interazioni del campo elettromagnetico tramite sviluppo perturbativo, cosa che non è possibile nel caso della costante $\alpha_F \simeq 1$ che caratterizza le interazioni forti.

Storicamente, questa costante è stata introdotta da Sommerfeld nel 1916 per tenere conto dello splittamento delle linee spettrali dell'atomo di idrogeno in seguito alla correzione relativistica dell'energia cinetica. Questa correzione sarà trattata in dettaglio in §8.1.1.

Se si mettono in grafico i valori dell'energia del modello di Bohr si ottiene la configurazione seguente (raffigurazione in scala corretta):

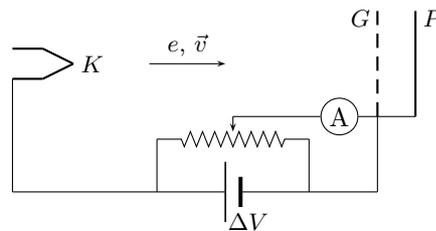


Per energie positive invece, non c'è quantizzazione dell'energia perché le orbite non sono chiuse.

Questa interpretazione della struttura atomica rende ragione facilmente degli spettri atomici, la serie di Balmer si interpreta infatti come la serie di transizioni dai livelli $n > 2$ al livello $n = 2$.

È possibile anche il procedimento inverso. L'idrogeno nello stato naturale si trova nel livello fondamentale $n = 1$ e non può né irradiare né assorbire l'energia corrispondente allo spettro di Balmer (le righe del visibile).⁴ L'idrogeno atomico può quindi emettere nel visibile in seguito ad eccitamento per urto (ad esempio, aumentando la temperatura).

Si noti che in realtà il modello di Bohr non dice nulla riguardo gli atomi a più elettroni. Tuttavia, il fatto che per l'atomo di idrogeno si possa scrivere la relazione $E_2 - E_1 = h\nu$ per i livelli energetici suggerisce che anche gli altri termini spettroscopici potessero essere interpretati con l'esistenza di livelli interni. L'esperimento che mise effettivamente in evidenza l'esistenza questi livelli fu condotto da Franck e Hertz con l'apparato sperimentale illustrato nella figura seguente:⁵



Gli elettroni emessi dal catodo K sono accelerati dalla differenza di potenziale ΔV fino alla griglia G e raccolti dalla placca P , posta ad un potenziale leggermente inferiore (e regolabile) alla griglia. La corrente che arriva sulla placca è misurata da un amperometro. Gli elettroni sono quindi accelerati fino alla griglia e poi sono leggermente decelerati prima di giungere sulla placca. La velocità massima raggiungibile sulla griglia, trascurando la velocità iniziale di emissione dal catodo, è data da:

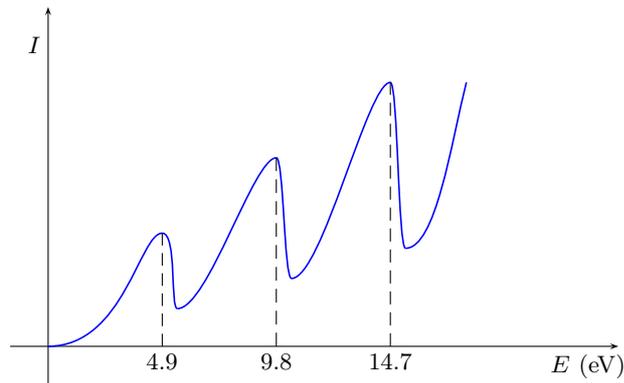
$$T_{\max} = \frac{1}{2} m_e v_{\max}^2 = e\Delta V$$

Regolando la differenza di potenziale fra la griglia e la placca, è possibile regolare l'intensità della corrente che giunge su questa e fare in modo che sia appena sufficiente a far giungere gli elettroni sulla placca. Tutto l'apparato era immerso in un tubo a vuoto contenente vapori di mercurio.

Il risultato dell'esperimento è illustrato nel grafico seguente:

⁴Fra l'altro, la lunghezza d'onda del visibile è di circa 5000\AA mentre l'atomo di idrogeno nello stato fondamentale è approssimativamente 1\AA , pertanto questo non può essere perturbato.

⁵Una simulazione dell'esperimento è disponibile al sito <http://phys.educ.ksu.edu/vqm/free/FranckHertz.html>.



Fino ad un'energia di 4.9 eV c'è un aumento di corrente, come ci si attende. Una volta raggiunti i 4.9 eV si trova una brusca caduta di corrente: questo fu interpretato da Franck e Hertz con il fatto che gli urti degli elettroni diventavano anelastici. Questo è proprio quanto ci si aspetta nel caso gli atomi abbiano dei livelli interni: l'energia cinetica si converte completamente in energia potenziale dell'atomo che può così eccitarsi e aumentare il suo livello di energia. A 4.9 eV l'eccitazione avviene esattamente davanti la griglia, aumentando ancora la differenza di potenziale gli elettroni raggiungeranno questo valore più presto e l'eccitazione avverrà sempre prima. Al valore di 9.8 eV, gli elettroni potranno eccitare due volte gli atomi di mercurio: quando raggiungono i 4.9 eV a metà percorso – perdendo completamente energia cinetica – e quando raggiungeranno di nuovo i 4.9 eV in corrispondenza della griglia.

Per confermare definitivamente che l'effetto è dovuto all'esistenza di un livello atomico, si deve trovare una linea spettrale di emissione corrispondente alla diseccitazione degli atomi di mercurio. Effettivamente, appena la differenza di potenziale supera i 4.9 eV, si riscontra una riga di emissione a 2546\AA , corrispondente esattamente ad un quanto di energia di 4.9 eV.

Gli atomi hanno dunque realmente dei livelli energetici interni, come ipotizzato da Bohr.

La quantizzazione di Bohr può essere espressa anche in termini di momento angolare. Si consideri un'orbita circolare, il cui momento angolare è dato da:

$$\vec{M} = \vec{r} \times \vec{p} = mvr$$

ora, la frequenza di un moto circolare è data da $f = v/2\pi r$, ovvero $r = v/2\pi f$. Il momento angolare si può quindi riscrivere nella forma:

$$M = mvr = \frac{mv^2}{2\pi f} \quad \rightarrow \quad M = \frac{1}{2}mv^2 \frac{nh}{2\pi|E|}$$

dalla relazione (2.1). Ma $\frac{1}{2}mv^2$ è l'energia cinetica e $|E|$ è l'energia totale: $E = T + U$. In un moto circolare vale il teorema del Viriale secondo cui $|E| = T$, per cui $|E|/T = 1$. Segue immediatamente:

$$\boxed{M = n\hbar}$$

ovvero *il momento angolare è quantizzato in unità \hbar* ⁽⁶⁾ e le sole orbite permesse sono quelle il cui momento angolare è multiplo di \hbar . Questo rappresenta il

Terzo postulato: Il momento angolare è quantizzato in unità \hbar e di conseguenza anche la quantità di moto.

Se si tratta di un moto periodico in un potenziale, il momento angolare può essere scritto nella forma:

$$M = \oint P_\phi d\phi$$

con P_ϕ momento coniugato della variabile angolare ϕ .

⁶In effetti, \hbar ha le dimensioni di un momento angolare, oltre che di un'azione.

Richiami di meccanica classica

Un sistema può essere identificato da una funzione detta *lagrangiana*:

$$L = L(q, \dot{q}, t) \quad \text{in } 2n \text{ variabili}$$

Il moto di questo sistema è descritto da $2N$ equazioni alle derivate totali (equazioni di Lagrange):

$$\begin{cases} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 & i = 1, \dots, N \\ \dot{q}_i = \frac{dq_i}{dt} \end{cases}$$

dove le \dot{q}_i sono considerate variabili indipendenti. Nella formulazione hamiltoniana della meccanica compaiono i *momenti coniugati* definiti come $p_i \equiv \partial L / \partial \dot{q}_i$. L'insieme delle variabili q_i e p_i è detto *variabili canoniche*. Prese le p_i come variabili indipendenti, la funzione di Hamilton è definita come:

$$H(q, p, t) = \sum_{i=1}^N p_i \dot{q}_i - L$$

e le equazioni di Hamilton espresse tramite le parentesi di Poisson sono date da:

$$\begin{cases} \dot{q}_i = \frac{\partial H}{\partial p_i} = \{q_i, H\} \\ \dot{p}_i = -\frac{\partial H}{\partial q_i} = \{p_i, H\} \end{cases}$$

dove le parentesi di Poisson sono definite come:

$$\{F, G\} = \sum_{i=1}^N \left\{ \frac{\partial F}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q_i} \right\}$$

in questa particolare notazione il moto di un qualunque sistema descritto da una $F(q, p)$ è dato dall'equazione:

$$\dot{F}(q, p) = \{F, H\}$$

Si consideri ancora un moto in un potenziale centrale:

$$L = T - V = \frac{1}{2} m v^2 - V(q)$$

In coordinate polari c'è una componente radiale ed una angolare, per cui:

$$v^2 = \dot{r}^2 + r^2 \dot{\phi}^2 \quad \Rightarrow \quad L = \frac{1}{2} m \left(\dot{r}^2 + r^2 \dot{\phi}^2 \right) - V(q)$$

Il momento angolare, d'altra parte, è definito come $\vec{M} = \vec{r} \times \vec{p} = m \vec{r} \times \vec{v}$. Essendo un prodotto vettoriale, la componente radiale non contribuisce al momento angolare e si ritrova $\vec{M} = m r^2 \dot{\phi}$. Questo coincide con il momento coniugato della variabile angolare:

$$P_\phi = \frac{\partial L}{\partial \dot{\phi}} = \frac{\partial}{\partial \dot{\phi}} \left[\frac{1}{2} \left(\dot{r}^2 + r^2 \dot{\phi}^2 \right) V(q) \right] = m r^2 \dot{\phi}$$

Essendo:

$$\dot{P}_\phi = -\frac{\partial H}{\partial \phi} = 0 \quad \Rightarrow \quad P_\phi \text{ costante del moto}$$

Sommerfeld scrisse il momento angolare in termini di integrale in ϕ . P_ϕ è una costante e l'integrale è periodico (da cui un integrale ciclico), si ottiene quindi:

$$\oint P_\phi d\phi = 2\pi M$$

infatti P_ϕ è costante. Il terzo postulato di Bohr impone che questa quantità sia quantizzata, per cui Sommerfeld pose questo integrale uguale a nh , in modo da riottenere esattamente la relazione di Bohr. Sommerfeld andò oltre e postulò che per i moti multiperiodici valesse la quantizzazione di tutti gli integrali di azione:

$$\oint p_i dq_i = n_i h$$

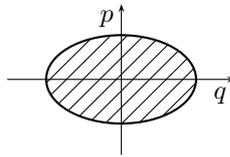
che è in definitiva una quantizzazione sulle traiettorie possibili. Questo modello di Sommerfeld ipotizzava che gli elettroni viaggiassero intorno al nucleo in orbite ellittiche, a differenza del modello di Bohr originale che ipotizzava orbite circolari.

Si vedranno ora due esempi di quantizzazione applicando le regole di Sommerfeld.

Oscillatore armonico. L'oscillatore armonico è caratterizzato dall'hamiltoniana:

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2$$

che nello spazio delle fasi⁷ è rappresentata da un'ellisse:



la cui area, che rappresenta l'energia del sistema, è evidentemente data da:

$$\oint pdq = \text{area nello spazio delle fasi}$$

Posto ora:

$$P = \frac{p}{\sqrt{2m}} \quad ; \quad Q = \sqrt{\frac{k}{2}} q \quad \Rightarrow \quad p = \sqrt{2m} P \quad ; \quad q = \sqrt{\frac{2}{k}} Q$$

segue:

$$\oint pdq = 2\sqrt{\frac{m}{k}} \oint PdQ$$

questa trasformazione è in effetti una riduzione a forma canonica che trasforma l'ellisse nel cerchio rappresentato da $P^2 + Q^2 = E$. L'ultimo integrale scritto è allora l'area di un cerchio di raggio \sqrt{E} :

$$\oint pdq = 2\sqrt{\frac{m}{k}} \pi E$$

Le regole di quantizzazione di Sommerfeld richiedono che questo integrale sia uguale a nh ,⁸ per cui:

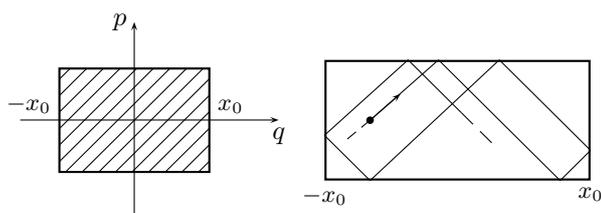
$$2\sqrt{\frac{m}{k}} \pi E = nh \quad \Rightarrow \quad E = hn \cdot \frac{1}{2\pi} \sqrt{\frac{k}{m}} \equiv nh\nu$$

ovvero, esattamente la quantizzazione dell'energia dell'oscillatore armonico ipotizzata da Planck.

Particella libera in una scatola. Si consideri una particella in una scatola soggetta ad urti elastici e quindi con conservazione dell'energia e della quantità di moto. Il moto risulta *pseudoperiodico*.

⁷Si chiama spazio delle fasi di un sistema lo spazio i cui punti rappresentano univocamente tutti e soli i possibili stati del sistema. Nella meccanica classica lo spazio delle fasi di solito rappresenta tutte le possibili posizioni e velocità di ogni punto materiale. In meccanica hamiltoniana lo spazio delle fasi (chiamato anche spazio degli stati, per non creare confusione di termini) è lo spazio rappresentativo del moto del sistema formato dalle $2n$ coordinate canoniche dove p_i sono i momenti coniugati alle coordinate generalizzate q_i .

⁸Per inciso, graficamente questo significa che non tutte le aree dell'ellisse sono ammesse, ma solo alcune. L'ellisse non può allargarsi o restringersi con continuità, ma solo per "scatti" di nh .



La lagrangiana di questo sistema è data da $L = \frac{1}{2}mv^2 + V(q)$ e il momento coniugato di q è dato da $p = \partial L / \partial \dot{q} = \dot{q}$, coincide pertanto con la quantità di moto. Si consideri per esempio l'asse x , da un punto di vista classico, p_x può assumere qualunque valore. Per le regole di quantizzazione di Sommerfeld deve invece risultare:

$$\oint p_x dq_x = 2p_x \cdot 2x_0 = 4p_x x_0 = n_x h$$

da cui:

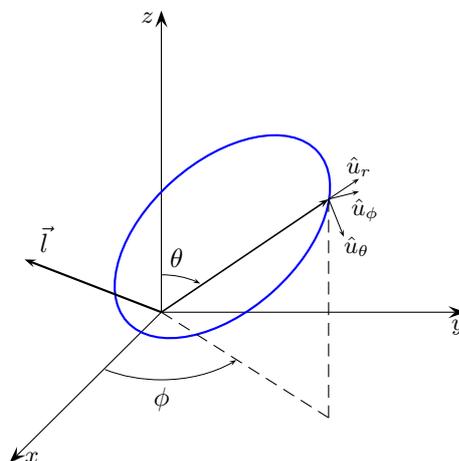
$$p_x = \frac{n_x h}{4x_0}$$

considerando che la medesima regola di quantizzazione si ha anche lungo y e z , si ottiene per l'energia $E = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2)$:

$$E = \frac{h^2}{32m} \left(\frac{n_x^2}{x_0^2} + \frac{n_y^2}{y_0^2} + \frac{n_z^2}{z_0^2} \right) \quad n_x, n_y, n_z \in \mathbb{N}$$

Questo risultato permette di giustificare l'approssimazione di moto unidimensionale o bidimensionale: se, ad esempio, z_0 è molto più piccolo, i suoi contributi dipendono fortemente da n_z quindi la particella tende a disporsi nello stato $n_z = 1$ e a rimanerci. Di conseguenza, in queste condizioni il moto dipenderà solo da x e y .

Si consideri ora il moto in un potenziale coulombiano e una generica orbita.



La lagrangiana è espressa come al solito da :

$$L = \frac{1}{2}mv^2 - V(r)$$

esprimendo \vec{v} in coordinate polari:

$$\vec{v} = \dot{r} \cdot \hat{u}_r + r\dot{\theta} \cdot \hat{u}_\theta + r \sin \theta \dot{\phi} \cdot \hat{u}_\phi$$

la lagrangiana diventa:

$$L = \frac{1}{2}m \left(\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\phi}^2 \right) - V(r)$$

siccome la lagrangiana non dipende da ϕ , ma solo da $\dot{\phi}$, si ricava $P_\phi = \text{costante}$. Per definizione infatti:

$$P_\phi = \frac{\partial L}{\partial \dot{\phi}} = mr^2 \sin^2 \theta \dot{\phi}$$

Dall'equazione di Hamilton risulta $\{P_\phi, H\} = 0$, dunque P_ϕ è una costante del moto. Ma P_ϕ coincide con la componente z del momento angolare, infatti:

$$\vec{l} = \vec{r} \times \vec{p} = m(\vec{r} \times \dot{\vec{r}})$$

questa relazione mostra che la componente radiale l_r è nulla,⁹ per cui:

$$\vec{l} = l_\theta \cdot \hat{u}_\theta + l_\phi \cdot \hat{u}_\phi$$

\hat{u}_ϕ è parallelo al piano \widehat{xy} , dunque l'unico contributo a l_z può venire dalla componente lungo θ , per cui $l_z = l_\theta \cdot (\hat{u}_\theta)_z$ con $l_\theta = -mrv_\phi = mr^2 \sin \theta \dot{\phi}$. Infatti $v_\phi = r \sin \theta \dot{\phi}$.

La componente z di \hat{u}_θ è $\sin \theta$, per cui:

$$l_z = mr^2 \sin^2 \theta \dot{\phi}$$

dunque

$$\boxed{P_\phi = l_z}$$

ne consegue da un lato che in un potenziale centrale il momento angolare è costante, dall'altro che le condizioni di Sommerfeld diventano:

$$\oint P_\phi d\phi = mh \quad \Rightarrow \quad \oint l_z d\phi = mh \quad \Rightarrow \quad 2\pi l_z = mh$$

da cui:

$$\boxed{l_z = m\hbar}$$

che coincide con la condizione di Bohr sulle orbite circolari, ma che qui è stata ottenuta in maniera del tutto generale.

P_θ non è di per sé conservato a causa del fattore $\sin^2 \theta$, tuttavia è esprimibile tramite quantità conservate: la quantizzazione su P_θ porta alla quantizzazione del momento angolare L^2 , mentre quella su P_r porta a quella dell'energia E .

Si noti che curiosamente anche se l_z è quantizzato, esso *dipende dalla scelta del sistema di riferimento*.

La conferma sperimentale della quantizzazione del momento angolare si ottenne con l'esperienza di Stern e Gerlach del 1922,¹⁰ anche se in modo del tutto inaspettato.

L'esperienza di Stern e Gerlach aveva in effetti lo scopo specifico di verificare l'ipotesi di Bohr-Sommerfeld secondo cui la direzione del momento angolare di un atomo di argento è quantizzata.

Essi trovarono il seguente sistema per determinare se le particelle avessero un momento angolare. Siccome un elettrone in moto circolare si comporta come una spira percorsa da corrente, immergendo questo in un campo magnetico si osserva una precessione di Larmor.

L'energia di interazione di un dipolo in un campo magnetico è data da:

$$H = -\vec{\mu} \cdot \vec{B}$$

tenendo presente che il momento magnetico è esprimibile in termini di momento angolare tramite il fattore giromagnetico $\vec{\mu} = g\vec{l}$, si trova:

$$H = -g\vec{l} \cdot \vec{B}$$

e se il campo è diretto lungo z , $H = -gl_z B_z$. Siccome il momento angolare deve conservarsi, il dipolo non può orientarsi, ma deve precedere intorno all'asse z (si tratta appunto della precessione

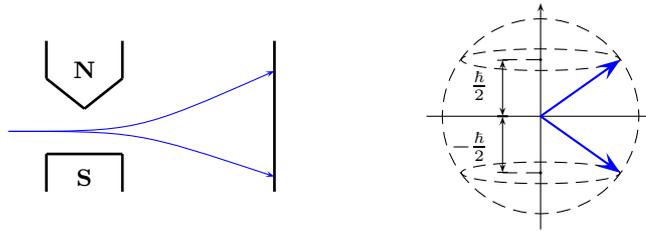
⁹ $\vec{r} \times \dot{\vec{r}} = r \frac{dr}{dt} \hat{u}_r \times \hat{u}_r = 0$.

¹⁰Walther Gerlach e Otto Stern, *Das magnetische Moment des Silberatoms*, Zeitschrift für Physik 9, 353-355 (1922).

di Larmor). Se la particella si muove in un campo magnetico omogeneo, le forze esercitate sulle estremità opposte del dipolo si cancellano a vicenda e la traiettoria della particella non è modificata. Se invece il campo non è omogeneo, questo dipolo risente di una forza lungo z data da:

$$F_z = -gl_z \frac{\partial B_z}{\partial z}$$

dunque la forza dipende dalla componente del momento angolare lungo z . Classicamente, tutti i valori di l_z sono possibili, per cui si deve osservare una banda continua corrispondente a tutti i valori ammessi. Se invece la teoria quantistica è corretta, l_z può assumere solo valori discreti e sullo schermo si deve osservare una serie di punti. L'esperimento, molto schematicizzato, è illustrato nella figura seguente:



Il risultato dell'esperimento diede ragione ancora una volta alle previsioni quantistiche, anche se in una maniera del tutto inaspettata. Un atomo di argento allo stato fondamentale, secondo i calcoli della teoria quantistica, possiede un momento angolare totale $l = 0$, per cui non dovrebbe deviare nel campo magnetico. Nel caso di $l = 1$ si sarebbero invece dovuti osservare tre fasci.¹¹ L'esperimento mostrava invece una divisione in due fasci. Per escludere la possibilità che questo fosse un comportamento tipico dell'atomo di argento, nel 1927 fu ripetuto questo esperimento utilizzando gli atomi di idrogeno, quello conosciuto meglio all'epoca, per i quali vale ancora $l = 0$ nello stato fondamentale. Il risultato sperimentale fu lo stesso: per cui se da un lato era fuori di dubbio che il momento angolare fosse quantizzato come effettivamente previsto dalla teoria di Bohr, dall'altro non si riusciva a spiegare la divisione del fascio in due. Fu grazie agli studenti Goudsmit e Uhlenbeck che si comprese quello che effettivamente accadeva: essi attribuirono il comportamento anomalo all'elettrone stesso, essi ipotizzarono cioè che l'elettrone avesse un momento angolare intrinseco, quantizzato anch'esso: era stato scoperto lo *spin*.¹² Questo momento intrinseco poteva assumere solo due valori $\hbar/2$ e $-\hbar/2$ e cambiare solo di un'unità \hbar .

L'esperimento diede risultati ancora più sorprendenti: se si prende un fascio uscente da un magnete di Stern e Gerlach orientato lungo z e lo si fa passare per un altro magnete ancora orientato lungo z , si ottiene un solo fascio, come è logico che sia. Se però lo si fa passare in un magnete orientato, per esempio, lungo x , il fascio si divide ancora in due, esattamente come se il valore di l_x non fosse assolutamente definito prima di entrare nel magnete. Cosa ancora più sorprendente, se il fascio uscente dal magnete orientato lungo x viene fatto passare di nuovo in un magnete orientato lungo z il fascio si divide di nuovo: ovvero, è come se la misura lungo x avesse di nuovo disturbato (e mischiata) la componente lungo z .

Per concludere questo *excursus* sulla teoria di Bohr-Sommerfeld, è importante sottolineare che Bohr ebbe il coraggio di presentare comunque un modello atomico che, anche se spiegava i fenomeni osservati, era internamente incoerente perché richiedeva di calcolare le orbite secondo le leggi della meccanica classica, alle quali però andavano aggiunte delle "regole esotiche" riassunte dai suoi tre postulati.

Una prima spiegazione delle regole di Bohr venne dagli studi di de Broglie.

¹¹La regola di quantizzazione di Bohr fornisce $2l + 1$ valori per la componente l_z , quindi fra l'altro necessariamente in numero dispari.

¹²Sebbene possa essere immaginato come un equivalente del momento angolare, quindi come una carica in rotazione, in realtà non è semplice darne una descrizione corretta. Questo problema fu risolto infine da Dirac verso il 1930, impostando una trattazione relativistica della meccanica quantistica, in questo modo fu in grado di prevederne i valori ed il comportamento corretto. Non esiste un modo per rappresentare classicamente lo *spin*, e il lavoro di Dirac – che rappresenta la base della moderna fisica delle particelle – è completamente al di là di questo corso. Più avanti, tuttavia, sarà dedicato un intero paragrafo (§6.4) a questo importante argomento.

2.5 L'ipotesi di de Broglie. Il dualismo onda-particella

Nel 1924, Louis de Broglie conseguì il dottorato con una tesi dal titolo *Recherches sur la Théorie des Quanta*.¹³ La tesi conteneva un'ipotesi sensazionale, che sarà alla base della meccanica ondulatoria. Questa ipotesi, integrata nella vecchia teoria quantistica, darà origine alla meccanica quantistica vera e propria.

De Broglie notò che il fenomeno della “quantizzazione” appare spontaneamente nel caso del formalismo ondulatorio, perché in esso intervengono in maniera naturale dei numeri interi. Inoltre, era ormai chiaro che il campo elettromagnetico doveva presentare una doppia natura, ondulatoria, che rendeva ragione degli effetti di interferenza, e corpuscolare, che rendeva conto della radiazione di corpo nero, dell'effetto fotoelettrico e di tutti gli altri casi studiati finora.

De Broglie fece quindi un grande passo, quello di *ipotizzare che anche le particelle presentassero un comportamento ondulatorio*. Questa ipotesi riportava una certa simmetria nella descrizione dei fenomeni fisici, perché in questo modo sia il campo elettromagnetico che le particelle presentavano una doppia natura corpuscolare e ondulatoria.

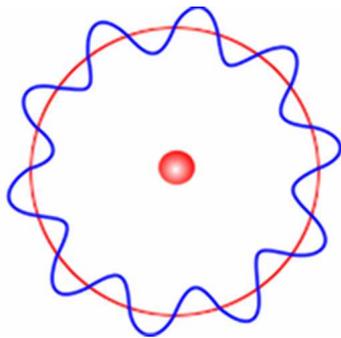
Partendo dalle relazioni note:

$$E = h\nu \quad p = \frac{E}{c} = \frac{h}{\lambda} \quad E^2 = p^2c^2 + m_0^2c^4$$

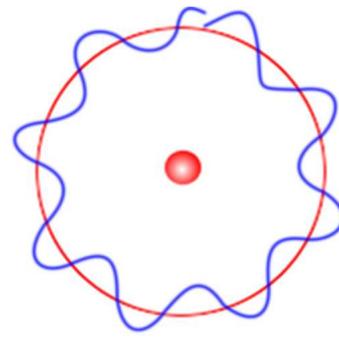
egli propose che le particelle avessero un comportamento ondulatorio con caratteristiche:

$$\boxed{\nu = \frac{E}{h} \quad \lambda = \frac{h}{p} \quad \lambda\nu = v < c} \quad (2.2)$$

ovvero, che ad una particella fosse possibile associare una frequenza pari alla sua energia diviso per la costante di Planck ed una lunghezza d'onda data dalla costante di Planck divisa per l'impulso e che il prodotto della frequenza e della lunghezza d'onda – in analogia alle onde – fornisse la velocità della particella. Questa ipotesi permetteva di identificare le condizioni di Bohr-Sommerfeld come stati di onde (associate all'elettrone) stazionari, ovvero quegli stati per i quali un'onda riesce a raccordarsi perfettamente su un'orbita:



L'onda si raccorda dopo un'orbita



L'onda non si raccorda dopo un'orbita

Un'onda è infatti espressa dalla forma $A(x)e^{i\phi(x)}$. Dopo che x ha terminato un'orbita, si deve ritrovare la medesima condizione affinché l'onda sia stazionaria e interferisca costruttivamente.¹⁴ Deve cioè risultare:

$$\phi(x + l) = \phi(x) + n2\pi$$

¹³“Ricerche sulla Teoria dei Quanti”

¹⁴Si noti che un elettrone che ruota intorno al nucleo è assimilato in questo approccio ad un'onda che curva nel campo elettromagnetico generato dal nucleo stesso. Il fenomeno in sé è comune, basti pensare ai miraggi. Qui il ruolo dell'indice di rifrazione è giocato dal campo elettromagnetico.

la condizione di stazionarietà dell'onda porta quindi in maniera naturale a introdurre dei numeri interi (numeri quantici). L'ampiezza dell'onda è $e^{i\vec{k}\cdot\vec{r}}$ e la sua variazione di fase è espressa da:

$$\Delta\phi = \oint \vec{k} \cdot d\vec{r} = 2n\pi$$

essendo ora $k \equiv 2\pi/\lambda$ e dalla relazione di de Broglie:

$$\oint \vec{k} \cdot d\vec{r} = \oint \frac{2\pi}{\lambda} d\vec{r} = \oint 2\pi \frac{\vec{p}}{h} d\vec{r} = n2\pi \quad \Rightarrow \quad \oint \vec{p} \cdot d\vec{r} = nh$$

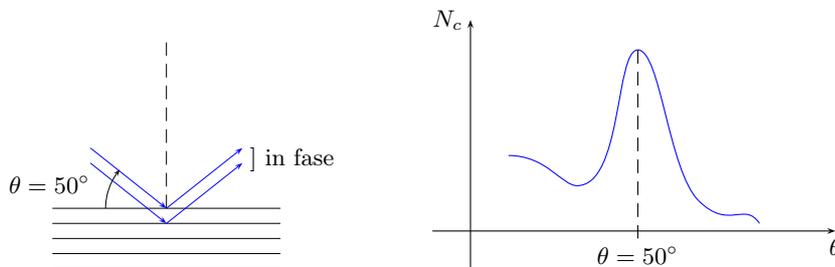
quindi l'ipotesi di de Broglie riesce a giustificare sia le orbite di Bohr che le regole di quantizzazione di Sommerfeld. Le orbite permesse sono dunque quelle per cui la velocità dell'elettrone è tale che un'orbita contiene un numero intero di lunghezze d'onda.

Sorge spontaneo chiedersi perché, se questa ipotesi è corretta, non ci si accorge normalmente del comportamento ondulatorio della materia. Per capirne il motivo, si consideri una particella con massa 1 mg, velocità 10^{-1} cm/s. In questo caso, $\lambda = h/mv \simeq 6 \cdot 10^{-23}$ cm. Occorrerebbe quindi una fenditura di 10^{-23} cm per rivelarne il comportamento ondulatorio, il che è al di là di qualunque esperienza pratica. Maggiore è la massa del corpo, più piccola è la lunghezza d'onda associata ad esso: ma allora per verificare la correttezza dell'ipotesi di de Broglie si può prendere una particella di materia sufficientemente piccola. La più piccola nota all'epoca era l'elettrone e se si svolgono i calcoli si trova che per un elettrone con $m_e \simeq 10^{-27}$ g accelerato da una differenza di potenziale ΔV vale:

$$\lambda = \frac{h}{m_e v} \simeq \frac{1.22 \times 10^{-9}}{\sqrt{\Delta V}} \text{ m}$$

che è dell'ordine della distanza fra i piani dei reticoli cristallini.

Questo è esattamente quello che fecero Davisson e Germer nel 1927: essi spararono un fascio di elettroni lenti su una lamina di nichel cristallino. Misurarono quindi la dipendenza dall'angolo di incidenza dell'elettrone riflesso e si trovò che aveva lo stesso pattern di diffrazione dei raggi X, esattamente come previsto dalla legge di Bragg.¹⁵



Questo esperimento confermò le idee di de Broglie, diede una base sperimentale alla *dualità onda-particella* e gettò le basi di quella che sarà la *meccanica quantistica* di Schrödinger. Vale pertanto la pena spendere qualche parola su questo nuovo aspetto anti-intuitivo della natura.

Sul dualismo onda/particella

L'ipotesi di de Broglie permise finalmente di unificare due aspetti fino ad allora distinti della realtà fisica: le onde e le particelle, ma questo ha richiesto un pesante tributo. Si è visto che i

¹⁵Facendo incidere un'onda elettromagnetica su di un cristallo si osservano fenomeni di interferenza causati dalla riflessione di onde da parte dei diversi piani cristallini paralleli. Questo fenomeno fu interpretato per la prima volta da William Henry Bragg e suo figlio William Lawrence nel 1913 e riassunto dalla cosiddetta legge di Bragg $n\lambda = 2d \sin \theta$, dove θ è l'angolo che il fascio incidente forma col piano cristallino, λ è la lunghezza d'onda della radiazione, d è la distanza tra due piani adiacenti e n indica l'ordine della diffrazione (in genere si riesce ad osservare il primo). Il fenomeno si spiega considerando che le due onde riflesse hanno una differenza di cammino ottico pari a $2d \sin \theta$. Quando questo cammino permette alle onde riflesse di interferire costruttivamente si ha un picco di intensità, che si rivela all'angolo particolare per cui $2d \sin \theta = n\lambda$.

campi si comportano a volte in modo particellare (discreto) e che le particelle si comportano a volte in modo ondulatorio. È impossibile d'altra parte immaginare come un'entità fisica possa comportarsi contemporaneamente sia come onda che come particella.

In realtà, la natura appare ancora più strana. È effettivamente possibile vedere un elettrone come una particella in un esperimento di diffusione e come un'onda in un esperimento alla Davisson e Germer, così come è possibile vedere la luce fare interferenza in un esperimento alla Young e comportarsi come corpuscoli nell'effetto fotoelettrico. È però *impossibile effettuare un esperimento che evidenzi contemporaneamente gli aspetti corpuscolari e ondulatori della stessa entità*.

Ogni volta che si prepara un apparato sperimentale si "sceglie" in qualche modo quale aspetto dell'elettrone o del fotone privilegiare: se la particella o l'onda, e non mai è possibile vedere i due insieme. Se si sottopone un elettrone ad un esperimento alla Young¹⁶ si ottiene una figura di interferenza: questo anche se l'intensità del fascio è così bassa che *un solo elettrone per volta* passa attraverso le fenditure. È sufficiente però introdurre un meccanismo che permetta di individuare attraverso quale fenditura è passato l'elettrone per distruggere la figura di interferenza, e questo anche se viene fatto in modo da non disturbare assolutamente il moto dell'elettrone. Nel momento in cui si rivela l'elettrone sullo schermo esso è una particella, ma il sistema nel suo complesso, *se non osservato*, mostra proprietà ondulatorie. Questa esperienza porta a una conclusione sconcertante: se non si osserva il suo comportamento, è come se l'elettrone non sapesse decidere attraverso quale fenditura passare e *le scegliesse entrambe*. È questa l'unica conclusione che sia in accordo con i risultati sperimentali.¹⁷

In effetti, la questione in questi termini è mal posta. In questo ragionamento si tenta infatti di applicare all'elettrone (come al fotone, al campo elettromagnetico e ad ogni altra entità a livello atomico) le categorie mentali nate per descrivere il mondo come appare ai nostri sensi, che segue la meccanica classica. Quello che tutte queste esperienze mostrano è semplicemente il fatto che *le categorie mentali della fisica classica, onde e particelle, non si applicano ai componenti atomici e subatomici: quello che noi identifichiamo con questi termini è solo quella parte delle loro proprietà che di volta in volta cade sotto i nostri sensi. Detto in altro modo, i termini onda e particella non sono adatti ad esprimere il comportamento dei costituenti atomici, e per evidenziare bene questo fatto è stato addirittura coniato un nuovo termine: wavicle*.¹⁸

Associare un'onda materiale direttamente ad un elettrone porta però a delle difficoltà. Una particella di energia E e momento p è definita da:

$$E = \frac{p^2}{2m} \quad v = \frac{p}{m} = \frac{\partial E}{\partial p}$$

Si consideri ora un'onda piana di de Broglie $e^{i(\vec{k}\cdot\vec{r}-\omega t)}$. Le superfici a fase costante sono dei piani definiti da $\vec{k}\cdot\vec{r} = \text{cost}$. Se \vec{k} è diretto lungo x , l'onda prende la forma $e^{i(kx-\omega t)}$, ovvero $e^{ik(x-\omega/kt)}$. Si supponga ora di mandare $x \rightarrow x + \Delta x$ quando $t \rightarrow t + \Delta t$, allora affinché la fase risulti costante:

$$x + \Delta x - \frac{\omega}{k}(t + \Delta t) = x + \frac{\omega}{k}t \equiv x + v_f t \quad \Rightarrow \quad v_f = \frac{\omega}{k}$$

v_f prende il nome di *velocità di fase*. Dalle relazioni di de Broglie (2.2) nella forma $\omega = 2\pi\nu$ e $k = 2\pi/\lambda$, si trova:

$$\frac{\omega}{k} = \lambda\nu = \frac{E}{p} = \frac{p}{2m} \quad \left(\omega = \frac{E}{\hbar} \quad ; \quad k = \frac{p}{\hbar} \right)$$

ma la velocità della particella è $v = p/m$.

Inoltre, se si considera la relazione relativistica $E^2 = p^2c^2 + m_0^2c^4$, ed applicando ad essa le relazioni di de Broglie, si trova:

$$\hbar\omega = \sqrt{\hbar^2k^2c^2 + m_0^2c^4} = c\sqrt{\hbar^2k^2 + m_0^2c^2} \Rightarrow \omega = c\sqrt{k^2 + \frac{m_0^2c^2}{\hbar^2}} \quad (2.3)$$

¹⁶Un interessante articolo sull'argomento è riportato in <http://digilander.libero.it/roberto20129/esperimenti/interferenzaquantistica.html>.

¹⁷La questione è in realtà molto più sottile e complessa e sembra legata alla *possibilità* di ottenere informazioni sul moto dell'elettrone. Per approfondimenti interessanti si possono studiare gli esperimenti di *delayed choice* di Wheeler, di *quantum eraser* di Scully e Drühl e *delayed choice quantum eraser* sempre di Scully e Drühl.

¹⁸Questo termine deriva dai termini inglesi *wave+particle*. In italiano, suonerebbe pressappoco come *ondicelle*.

la velocità di fase risulta:

$$v_f = \frac{\omega}{k} = c \frac{\sqrt{k^2 + \frac{m_0^2 c^2}{\hbar^2}}}{k} > c$$

dunque la velocità di fase è sempre superiore a c .

Il problema di questo discorso è che l'onda considerata va da $-\infty$ a ∞ ed è sempre uguale a sé stessa: un'onda monocromatica di questo tipo non può trasportare alcun tipo di informazione. La velocità delle "informazioni" trasportate da un'onda è invece quella di gruppo, definita come:

$$v_g \equiv \frac{\partial \omega}{\partial k} = \frac{\partial E}{\partial p} = v$$

per cui la velocità di una particella coincide con quella di gruppo. Si noti che per le onde elettromagnetiche non c'è questo problema: $\omega = 2\pi\nu$, $k = 2\pi/\lambda$, $\lambda\nu = c$, da cui $k = 2\pi\nu/c$, il che permette di ricavare $\omega/k = c = \partial\omega/\partial k$. Il punto è che la relazione $\omega = kc$ è lineare e quindi il rapporto coincide con la derivata: per un'onda elettromagnetica velocità di fase e di gruppo coincidono e valgono c . Per una particella risulta invece:

$$\boxed{\omega = \frac{kE}{p} = \frac{kp}{2m} = \frac{\hbar k^2}{2m}} \quad (2.4)$$

che è la relazione di dispersione per particelle di massa m . La velocità di fase risulta dipendere da k :

$$v_f = \frac{\hbar k}{2m}$$

quindi in un gruppo di onde la forma non è conservata nel tempo e tende a slargarsi. Occorre quindi riferirsi alla velocità del "picco", che corrisponde alla velocità di gruppo. Dalla relazione relativistica (2.3) la velocità di gruppo risulta essere:

$$v_g = \frac{d\omega}{dk} = c \frac{1}{2} \frac{2k}{\sqrt{k^2 + \frac{m_0^2 c^2}{\hbar^2}}} < c$$

quindi la velocità di gruppo è sempre inferiore a c . Questo è coerente con quanto ci si aspetta, cioè che nessun segnale viaggi a $v > c$.

Si consideri ora come esempio un'onda non monocromatica composta da un termine monocromatico più un altro termine della forma:

$$\psi(x, t) = \sin 2\pi(kx - \omega t) + \sin 2\pi[(k + \Delta k)x - (\omega + \Delta\omega)t]$$

queste onde interferiscono costruttivamente e distruttivamente tra loro. Dalla formula di addizione del seno:¹⁹

$$\psi(x, t) = 2 \cos 2\pi \left(\frac{\Delta k}{2} x - \frac{\Delta\omega}{2} t \right) \sin 2\pi \left(\frac{2k + \Delta k}{2} x - \frac{2\omega + \Delta\omega}{2} t \right)$$

rispetto a $2k$ e 2ω si possono trascurare Δk e $\Delta\omega$, che si suppongono piccoli, per cui:

$$\psi(x, t) = 2 \cos 2\pi \left(\frac{\Delta k}{2} x - \frac{\Delta\omega}{2} t \right) \sin 2\pi(kx + \omega t)$$

dunque consiste nell'onda monocromatica di partenza, ma con un'ampiezza non costante modulata da una funzione di x e t (in effetti, si tratta del fenomeno dei battimenti). Il massimo dell'ampiezza si ha quando:²⁰

$$\frac{\Delta k}{2} x - \frac{\Delta\omega}{2} t = 0 \Rightarrow \frac{dx}{dt} = \frac{\Delta\omega}{\Delta k} = \frac{\partial\omega}{\partial k}$$

¹⁹Formula di addizione del seno:

$$\sin \alpha + \sin \beta = 2 \cos \frac{\alpha - \beta}{2} \sin \frac{\alpha + \beta}{2}$$

²⁰Si noti che in realtà i massimi sono dati dalla condizione:

$$\frac{\Delta k}{2} x - \frac{\Delta\omega}{2} t = 0 \quad [\text{mod } 1]$$

ovvero modulo $2n\pi$.

ovvero, la velocità con cui si spostano i massimi dell'intensità coincide proprio con la velocità di gruppo. Se inoltre ω dipende linearmente da k , questa coincide anche con la velocità di fase.

Capitolo 3

La meccanica ondulatoria: l'equazione di Schrödinger e la funzione d'onda

3.1 Il pacchetto d'onda

Si generalizzerà ora il concetto espresso nel paragrafo precedente. Qualunque segnale è scrivibile come somma di segnali sinusoidali:

$$f(x) = \int_{-\infty}^{+\infty} g(k)e^{ikx} dk$$

che rappresenta la cosiddetta *antitrasformata di Fourier* della $g(k)$.¹ Basandosi su questa osservazione, è possibile costruire un *pacchetto d'onda* sommando delle componenti monocromatiche nel seguente modo:

$$\psi(\vec{x}, t) = \int_{-\infty}^{+\infty} \chi(\vec{k})e^{i(\vec{k}\cdot\vec{r}-\omega t)} d\vec{k} \quad (3.1)$$

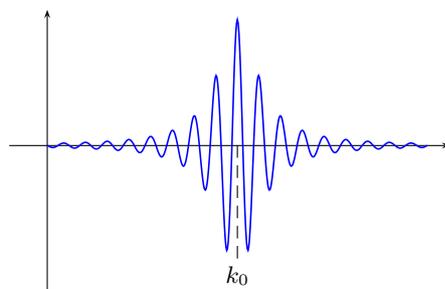
con la funzione $\chi(\vec{k})$ data da:

$$\chi(\vec{k}) = A(\vec{k})e^{i\alpha(\vec{k})}$$

per cui:

$$\psi(\vec{x}, t) = \int_{-\infty}^{+\infty} A(\vec{k})e^{i[\vec{k}\cdot\vec{x}-\omega t+\alpha(\vec{k})]} d\vec{k}$$

le diverse componenti monocromatiche risultano quindi modulate da una funzione $A(\vec{k})$, che dovrà avere una forma piccata intorno un valore k_0 , ottenendo così un risultato apprezzabile solo in un piccolo intervallo di k_0 . Un pacchetto d'onda $\psi(\vec{x}, t)$ si presenta quindi come in figura:

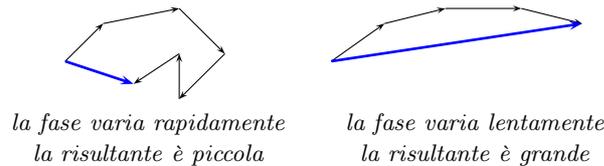


¹Si tratta in sostanza della cosiddetta *analisi di Fourier*. Se la funzione è periodica, la trasformata si riduce alla *serie di Fourier*.

In questo modo è possibile associare ad una particella un'onda con sufficienti proprietà di localizzazione, cosa che un'onda monocromatica alla de Broglie non ha.

Si procederà ora a determinare il massimo del pacchetto e la sua velocità.

Per comprendere la procedura, si consideri il campo complesso. Un numero complesso è rappresentato da un vettore, dotato di un modulo e di una fase che rappresenta l'angolo rispetto al sistema di assi coordinati $\Re - \Im$.² Se la fase varia rapidamente in k , in termini di somma di vettori il risultato è piccolo, mentre se le fasi sono pressappoco identiche la risultante è grande, come mostrato nella figura seguente (la risultante è indicata in blu):



La condizione per cui la risultante sia diversa da zero è quindi che la fase del pacchetto d'onda $\phi(k) = kx - \omega t + \alpha(k)$ vari molto lentamente, al limite che sia stazionaria.³ Supponendo per comodità di trattare un problema unidimensionale, sia x la coordinata:

$$\frac{d\phi}{dk} = x - \frac{d\omega}{dk}t + \frac{d\alpha}{dk} = 0$$

Se questa relazione risulta verificata per un determinato \bar{x} ma per un valore $k \neq k_0$, le fasi si sommano bene, ma la funzione di modulazione $A(\vec{k})$ ne annulla il valore. Ne consegue che il massimo del pacchetto di onda, che è un numero complesso, si ha quando:

$$\boxed{x - \frac{d\omega}{dk}t + \frac{d\alpha}{dk} \Big|_{k=k_0} = 0}$$

per cui la velocità di propagazione del picco del pacchetto d'onda coincide con la velocità di gruppo:

$$v_{\text{pacchetto}} = \frac{d\omega}{dk} \equiv v_g$$

La velocità di fase influisce invece sulla costanza della forma della funzione $\psi(\vec{x}, t)$, infatti quest'ultima – come dice il nome stesso – rappresenta la velocità alla quale si “sfasano” le diverse componenti monocromatiche all'interno del pacchetto.⁴

3.2 L'equazione di Schrödinger

Una volta compreso che le particelle possono comportarsi in maniera ondulatoria oltre che corpuscolare e dopo aver associato un pacchetto d'onda ad esse, sorge il problema di trovare un'equazione che ne descriva il comportamento e l'evoluzione nel tempo. Questa equazione deve quindi avere come soluzione un particolare pacchetto d'onda, detto **funzione d'onda**, che descrive in termini ondulatori il comportamento della particella in esame.

Questa equazione fu trovata da Schrödinger nel 1926 e rappresenta una delle massime conquiste della fisica. Essa è ricavabile in base ad argomenti di plausibilità e costituisce la base della *Meccanica Quantistica*.

²In altri termini: un numero complesso $z = x + iy$ può essere visto come un vettore di lunghezza (modulo) $\sqrt{x^2 + y^2}$ che forma un angolo (fase) $\alpha = \arctan y/x$.

³Questo modo di procedere si chiama appunto *metodo della fase stazionaria*.

⁴Si è visto che questa velocità è sempre superiore a c , tuttavia non è un quantità misurabile fisicamente: il solo effetto è quello di cambiare forma al pacchetto d'onda.

Schrödinger ipotizzò innanzitutto che fosse sufficiente avere una funzione d'onda *scalare* e non vettoriale come per il campo elettromagnetico. Si consideri una particella libera, per la quale la relazione di dispersione è data dalla (2.4):

$$\omega = \frac{\hbar k^2}{2m}$$

Una funzione d'onda che sia scalare e che sia retta dalla relazione di dispersione qui sopra soddisfa una determinata equazione differenziale. Si consideri il pacchetto in (3.1) e si consideri la derivata temporale:

$$\frac{\partial}{\partial t} \psi(\vec{x}, t) = \int_{-\infty}^{+\infty} (-i\omega) \chi(\vec{k}) e^{i(\vec{k} \cdot \vec{x} - \omega t)} d\vec{k}$$

ma dalla relazione di dispersione, questo termine è uguale a:

$$\int_{-\infty}^{+\infty} (-i\omega) \chi(\vec{k}) e^{i(\vec{k} \cdot \vec{x} - \omega t)} d\vec{k} = -\frac{i\hbar}{2m} \int_{-\infty}^{+\infty} \vec{k}^2 \chi(\vec{k}) e^{i(\vec{k} \cdot \vec{x} - \omega t)} d\vec{k} = -\frac{i\hbar}{i^2 2m} \frac{\partial^2}{\partial x^2} \psi(\vec{x}, t)$$

da cui si vede che risulta in linea generale:

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{x}, t) = -\frac{\hbar^2}{2m} \nabla^2 \psi(\vec{x}, t)$$

Dunque, un pacchetto d'onda per una particella libera deve soddisfare questa equazione differenziale. Si noti che questa forma è dettata dalla relazione di dispersione (2.4), che deriva direttamente dall'ipotesi di de Broglie.⁵

Questa è l'equazione che fu assunta come base di partenza da Schrödinger e costituisce *la generalizzazione delle relazioni di de Broglie alle onde non monocromatiche*. Si noti che per essere vera la $\psi(\vec{x}, t)$ **deve** essere complessa, e questo è legato al fatto che compare un fattore immaginario i davanti la derivata prima rispetto al tempo.⁶ Si tratta in effetti di funzioni complesse $\vec{r} \in \mathbb{R}^3$: $\psi(\vec{r}, t) \in \mathbb{C}$. Dunque l'equazione adatta per i sistemi quantistici è:

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{x}, t) = -\frac{\hbar^2}{2m} \nabla^2 \psi(\vec{x}, t)$$

Da notare che *le soluzioni di questa equazione non possono essere interpretate come oscillazioni perché appartengono al campo complesso*.

Il *principio di corrispondenza*, enunciato da Bohr, afferma che i risultati della meccanica quantistica devono ridursi a quelli della meccanica classica nelle situazioni in cui l'interpretazione classica

⁵Si noti anche che nel caso si utilizzi la relazione di dispersione tipica delle onde elettromagnetiche si ricaverebbe la relazione:

$$\frac{1}{c} \frac{\partial}{\partial t} \psi(x, t) = -\frac{\partial}{\partial x} \psi(x, t)$$

Questa equazione non è corretta perché la derivata spaziale $\partial \psi(x, t) / \partial x$ non è invariante per cambiamento di sistema di riferimento. Infatti se $x'_i = R_{ij} x_j$, si ricava:

$$\frac{\partial}{\partial x_i} \psi(\vec{x}, t) = \frac{\partial}{\partial x'_j} \psi(\vec{x}, t) \frac{\partial x'_j}{\partial x_i} = R_{ji} \frac{\partial}{\partial x'_j} \psi(\vec{x}, t) \quad \Rightarrow \quad \frac{\partial}{\partial x_i} \psi(\vec{x}, t) \neq \frac{\partial}{\partial x'_j} \psi(\vec{x}, t)$$

ma un'equazione delle onde non può dipendere dal sistema di riferimento. Le quantità invarianti per rotazioni sono le derivate seconde, ovvero il laplaciano, perché R_{ij} è ortogonale: $R_{ij} R_{jk}^T = \delta_{ik}$, da cui $\nabla^2 = \nabla'^2$. Analogamente, per il principio di relatività devono comparire le derivate seconde rispetto al tempo, per cui in definitiva è il d'alambertiano \square che risulta invariante per cambiamento di sistema di riferimento.

Per dovere di precisione, occorre precisare che un'equazione del tipo:

$$\frac{1}{v} \frac{\partial}{\partial t} \psi(\vec{x}, t) = \nabla^2 \psi(\vec{x}, t)$$

esiste, ed è un'equazione di diffusione (*equazione di diffusione del calore* di Fourier).

⁶Storicamente, si racconta – e sembra vero – che Schrödinger cercasse in realtà una equazione reale ma comprese, dopo 3-4 settimane di tentativi infruttuosi, di dover aggiungere la i alla sua equazione.

può essere considerata valida. Questo principio insieme alla relazione di de Broglie $\vec{p} = \hbar\vec{k}$ permette di identificare l'equazione di sopra con la classica relazione:

$$E = \frac{p^2}{2m}$$

a patto di stabilire una relazione fra la derivata temporale della funzione d'onda e l'energia:

$$\begin{cases} E \rightarrow i\hbar \frac{\partial}{\partial t} \\ \vec{p} = \hbar\vec{k} \rightarrow -i\hbar\nabla \end{cases}$$

Nel caso di particelle soggette a forze, secondo la fisica classica vale la relazione $E = p^2/2m + V(x)$. L'idea di Schrödinger fu che bastasse aggiungere un termine reale⁷ di potenziale all'equazione differenziale:

$$\boxed{\left[-\frac{\hbar^2}{2m} \nabla^2 + V(\vec{x}) \right] \psi(\vec{x}, t) = i\hbar \frac{\partial}{\partial t} \psi(\vec{x}, t) \quad \text{Equazione di Schrödinger}} \quad (3.2)$$

Questa equazione ha tre importanti caratteristiche:

- l'equazione è lineare, per cui la somma di due soluzioni è ancora una soluzione. Questo rende conto del *principio di sovrapposizione*,
- le soluzioni hanno la proprietà di propagarsi, ovvero di evolvere nel tempo,
- dipende solo dalla derivata prima del tempo, per cui fissato un istante iniziale l'evoluzione temporale della funzione d'onda è completamente determinata dal valore iniziale $\psi(\vec{r}, t_0)$.

Se si vuole inoltre che la funzione d'onda sia realmente associabile ad una particella localizzata, essa non può divergere (nel paragrafo successivo sarà data una giustificazione matematica di questa affermazione). Si deve dunque sempre imporre alle soluzioni dell'equazione di Schrödinger la condizione aggiuntiva:

$$\lim_{|\vec{x}| \rightarrow +\infty} |\psi(\vec{x}, t)| = 0$$

Una classe di soluzioni di questa equazione è quella delle funzioni fattorizzabili come $\Psi(x, t) = \psi(x)\varphi(t)$.⁸ In questo caso l'equazione di Schrödinger diventa:

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi(x)\varphi(t) = i\hbar\psi(x) \frac{d}{dt} \varphi(t)$$

ovvero:

$$\frac{1}{\psi(x)} \left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi(x) = i\hbar \frac{1}{\varphi(t)} \frac{d}{dt} \varphi(t)$$

siccome a sinistra compare solo la variabile x e a destra solo la variabile t , ne consegue che la parte temporale deve essere una costante. Quindi:

$$\frac{d}{dt} \varphi(t) = -\frac{i}{\hbar} \text{cost.} \cdot \varphi(t) \quad \Rightarrow \quad \varphi(t) = e^{-i \frac{\text{cost.}}{\hbar} t}$$

che è una funzione oscillante con frequenza $\nu = \text{cost.}/\hbar$. Ma la frequenza associata all'onda di una particella è $\omega = E/\hbar$, per cui la parte temporale è rappresentata da:

$$\varphi(t) = e^{-i \frac{E}{\hbar} t} = e^{-i\omega t}$$

⁷Nel senso: $V(x) \in \mathbb{R}$.

⁸Si tratta del cosiddetto metodo di Fourier per separazione delle variabili.

Si tratta quindi di un sistema stazionario,⁹ in cui il tempo compare solo come un fattore di fase. Inserita nell'equazione originale fornisce ($\hbar\omega = E$):

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{x}) + V(\vec{x})\psi(\vec{x}) = -(i \cdot i)\hbar\omega\psi(\vec{x}) \equiv E\psi(\vec{x})$$

ovvero:

$$\boxed{\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{x})\right]\psi(\vec{x}) = E\psi(\vec{x})} \quad (3.3)$$

Questa è l'equazione di Schrödinger indipendente dal tempo. Gli stati stazionari di un sistema¹⁰ sono quindi retti da questa equazione.

Questa equazione ha la forma di un'equazione agli autovalori e la particolare forma richiede che le soluzioni siano almeno di classe $C^{(1)}$, ovvero continue fino almeno alla derivata prima.

3.3 La funzione d'onda e sua interpretazione

Il fatto che l'equazione di Schrödinger indipendente dal tempo sia un'equazione agli autovalori significa richiedere, dal punto di vista matematico, che la funzione risultante dopo aver applicato l'operatore sia un multiplo di quella di partenza, il fattore di proporzionalità essendo l'autovalore.¹¹

*Si vedrà che quando si impongono le opportune condizioni al contorno, non tutte le soluzioni di questa equazione saranno accettabili fisicamente, questo accadrà invece solo per determinati valori dell'energia E . Sono queste le soluzioni che rappresentano gli stati e le energie possibili del sistema.*¹²

Resta ancora da capire cosa rappresenti effettivamente la funzione d'onda.

De Broglie riteneva che l'onda fosse la particella, ma questo conduce immediatamente a delle difficoltà, come visto in §2.5. In effetti, un pacchetto d'onda retto dalla relazione di dispersione (2.4) tende a slargarsi nel tempo, per cui non può "essere" una particella che invece risulta sempre localizzata. Inoltre, se la relazione di dispersione non è lineare la forma della funzione d'onda non è conservata, per cui quest'ultima non può nemmeno "guidare" la particella.

Un pacchetto d'onda si slarga nel tempo

Si consideri un pacchetto d'onda, la cui funzione modulante sia per semplicità una gaussiana centrata su k_0 di larghezza σ :

$$\psi(x, t) = \int_{-\infty}^{+\infty} f(k)e^{i(kx-\omega t)} dk = \int_{-\infty}^{+\infty} e^{-(k-k_0)^2/2\sigma^2} \cdot e^{i(kx-\omega t)} dk$$

e considerando la relazione di dispersione $\omega = \hbar k^2/2m$:

$$\psi(x, t) = \int_{-\infty}^{+\infty} e^{-(k-k_0)^2/2\sigma^2} \cdot e^{i(kx-\frac{\hbar k^2}{2m}t)} dk$$

ponendo per semplicità $\chi \equiv k - k_0$, si ha $k = \chi + k_0$ e $dk = d\chi$, per cui:

$$\psi(x, t) = \int_{-\infty}^{+\infty} e^{-\chi^2/2\sigma^2} \cdot e^{i(\chi+k_0)x} \cdot e^{-i\frac{\hbar}{2m}t(\chi^2+2k_0\chi+k_0^2)} d\chi$$

si faccia l'ulteriore posizione $\tau \equiv \hbar t/2m$, estraendo dall'integrale i termini costanti e raggruppando opportunamente i termini a esponente:

$$\psi(x, t) = e^{i(k_0x-\omega_0t)} \int_{-\infty}^{+\infty} e^{-\left(\frac{1}{2\sigma^2}+i\tau\right)\chi^2+i(x-2k_0\tau)\chi} d\chi$$

⁹La giustificazione matematica di questa affermazione deve essere rimandata al paragrafo seguente.

¹⁰In particolare, si hanno in mente qui gli stati dell'atomo di idrogeno.

¹¹Non ci si dilungherà tuttavia qui nello spiegare cosa sia un autovalore, un autovettore o un'autofunzione. Queste conoscenze di base di algebra lineare sono date per scontate.

¹²Schrödinger riuscì a risolvere la sua equazione per l'atomo di idrogeno, ritrovando i livelli energetici di Bohr. Questo calcolo sarà naturalmente effettuato in questo corso, ma molto più avanti, quando si avrà ormai una conoscenza sufficiente della meccanica ondulatoria.

questo integrale risulta avere la forma:

$$\int_{-\infty}^{+\infty} e^{-a\chi^2 + b\chi} d\chi \quad \begin{cases} a = \frac{1}{2\sigma^2} + i\tau \\ b = i(x - 2k_0\tau) \end{cases}$$

nell'ottica di giungere ad un integrale gaussiano, si completa l'esponente in modo che risulti un quadrato perfetto:

$$-a\chi^2 + b\chi = -a\chi^2 + b\chi - \frac{b^2}{4a^2} + \frac{b^2}{4a^2} = -\left(a\chi - \frac{b}{2a}\right)^2 + \frac{b^2}{4a^2}$$

il termine $b^2/4a^2$ nell'esponente si può portare fuori il segno di integrale, per cui si ottiene:

$$\psi(x, t) = e^{i(k_0x - \omega_0t)} \cdot e^{-\frac{(x - 2k_0\tau)^2}{4\left(\frac{1}{2\sigma^2} + i\tau\right)^2}} \int_{-\infty}^{+\infty} e^{-(a\chi - \frac{b}{2a})^2} d\chi$$

l'integrale è un integrale gaussiano del tipo $\int e^{-y^2} dy$ che, se $\Re(a) > 0$ e $b \in \mathbb{C}$, vale $\sqrt{\pi/a}$, per cui in definitiva:

$$\psi(x, t) = e^{i(k_0x - \omega_0t)} \cdot e^{-\frac{(x - 2k_0\tau)^2}{4\left(\frac{1}{2\sigma^2} + i\tau\right)^2}} \cdot \frac{\sqrt{\pi}}{\sqrt{\frac{1}{2\sigma^2} + i\tau}}$$

ovvero, ricordando le posizioni fatte:

$$\psi(x, t) = \frac{\sqrt{\pi} e^{i(k_0x - \omega_0t)}}{\sqrt{\frac{1}{2\sigma^2} + i\frac{\hbar t}{2m}}} \cdot e^{-\frac{\left(x - \frac{\hbar k_0}{m}t\right)^2}{4\left(\frac{1}{2\sigma^2} + i\frac{\hbar t}{2m}\right)^2}}$$

Il primo esponenziale rappresenta l'onda portante. Il secondo è un pacchetto d'onda centrato in $x_c = \frac{\hbar k_0}{m}t$ che si muove a velocità pari alla velocità di gruppo $\frac{\hbar k_0}{m} = \frac{\partial \omega}{\partial k} = v_g$, quindi in modo "classico". Ma si vede anche che la larghezza di questo pacchetto, rappresentata da $\sigma_p = \frac{1}{2\sigma^2} + i\frac{\hbar t}{2m}$, aumenta con il tempo, ovvero il pacchetto si allarga fino a diventare un'onda monocromatica.

Nel 1927, Max Born ebbe l'idea risolutiva.¹³ In analogia al fatto che il quadrato dell'ampiezza di un'onda elettromagnetica è legato alla sua intensità, *egli propose che la funzione d'onda fosse legata alla probabilità di trovare la particella in un determinato punto dello spazio ad un determinato istante. In particolare, la sua interpretazione – che è anche quella moderna – fu la seguente: il quadrato della funzione d'onda rappresenta una densità di ampiezza di probabilità, in altri termini la densità di probabilità di trovare la particella in un volume $d\tau$ è data da $|\psi(\vec{r}, t)|^2 d\tau$. L'idea del modulo quadro era suggerita fra l'altro anche dal fatto che la probabilità deve essere ovunque non negativa.*

La probabilità di trovare una particella in una posizione $\vec{r}_1 < \vec{r}_0 < \vec{r}_2$ all'istante t_0 è dunque data da:

$$\mathcal{P}(\vec{r}_0, t_0) = N \int_{\vec{r}_1}^{\vec{r}_2} |\psi(\vec{r}, t_0)|^2 d\vec{r}$$

dove la costante di proporzionalità N è indipendente da \vec{r} . Questa rappresenta semplicemente la cosiddetta *costante di normalizzazione*, ricavata imponendo che la probabilità estesa a tutto lo spazio sia 1:

$$\int_{-\infty}^{+\infty} \mathcal{P}(\vec{r}, t) d\vec{r} = 1 \quad \Rightarrow \quad N \int_{-\infty}^{+\infty} |\psi(\vec{r}, t)|^2 d\vec{r} = 1 \quad \Rightarrow \quad N \equiv \frac{1}{\int_{-\infty}^{+\infty} |\psi(\vec{r}, t)|^2 d\vec{r}}$$

in termini matematici, questo integrale si chiama *norma* e viene indicato con $\|\psi(\vec{r}, t)\|$. Se l'integrale non convergesse, dovrebbe valere $N = 0$ affinché la probabilità su tutto lo spazio sia ancora 1. Ciò

¹³Per questa idea ebbe il premio Nobel nel 1954. Fu lui, tra l'altro, a coniare il termine *meccanica quantistica*.

significa che la probabilità di trovare la particella in un intervallo finito arbitrario sarebbe nulla: questo non ha significato fisico, per cui l'integrale deve essere finito. Questo dà una base matematica al fatto che la funzione d'onda debba andare (abbastanza velocemente) a zero all'infinito. Le funzioni che sono a quadrato sommabile costituiscono uno spazio lineare denominato \mathcal{L}^2 .

Lo spazio \mathcal{L}^2

\mathcal{L}^2 , letto *elle due*, è lo spazio funzionale lineare delle funzioni a quadrato sommabile. Esso è un caso particolare degli spazi \mathcal{L}^p , lo spazio delle funzioni a p -sima potenza sommabile.

Formalmente, sia $p < \infty$ e denotando con S un insieme aperto dello spazio euclideo \mathbb{R}^n e \vec{x} un suo elemento, l'insieme delle funzioni che verifica la proprietà:

$$\|f(\vec{x})\|_p = \left(\int_S |f(\vec{x})|^p d\vec{x} \right)^{1/p} < \infty$$

e che gode della proprietà

$$|f(\vec{x}) + g(\vec{x})|^p \leq |f(\vec{x})|^p + |g(\vec{x})|^p$$

è uno spazio vettoriale denominato \mathcal{L}^p , formato dalle funzioni su S a quadrato p -sommabile. Ne consegue che sono ereditate tutte le proprietà della norma e della disuguaglianza triangolare.

Il caso \mathcal{L}^2 è un caso estremamente particolare, perchè per le funzioni a quadrato sommabile la norma è indotta dal prodotto scalare:

$$\langle f(\vec{x}), g(\vec{x}) \rangle = \int_S f^*(\vec{x})g(\vec{x})d\vec{x} \Rightarrow \|f(\vec{x})\| = \int_S f^*(\vec{x})f(\vec{x})d\vec{x} = \langle f(\vec{x}), f(\vec{x}) \rangle$$

ed è l'unico valore di p per cui questo accade. In particolare, questo rende lo spazio \mathcal{L}^2 uno spazio di Hilbert.¹⁴

Tutte le funzioni d'onda di Schrödinger devono quindi appartenere a \mathcal{L}^2 .

La funzione $\psi_n(\vec{r}, t) = \frac{1}{\sqrt{N}}\psi(\vec{r}, t)$ è detta essere *normalizzata a 1*. Si può sempre formare una funzione d'onda normalizzata partendo da una data funzione d'onda fisica e, volendo, si potrebbe lavorare esclusivamente con funzioni d'onda normalizzate a uno.

L'equazione di Schrödinger ammette soluzione solo se la costante di normalizzazione N è indipendente dal tempo e questo è effettivamente il caso. Infatti (per comodità di notazione si considera una funzione d'onda unidimensionale):

$$\frac{d}{dt} \int_{-\infty}^{+\infty} |\psi(x, t)|^2 dx = \int_{-\infty}^{+\infty} \frac{\partial}{\partial t} |\psi(x, t)|^2 dx$$

ora, risulta:¹⁵

$$\frac{\partial}{\partial t} |\psi(x, t)|^2 dx = \frac{\partial}{\partial t} \psi^*(x, t) \cdot \psi(x, t) + \psi^*(x, t) \cdot \frac{\partial}{\partial t} \psi(x, t)$$

L'equazione di Schrödinger (3.2) fornisce un'espressione per la derivata temporale della $\psi(x, t)$. Analogamente, la sua complessa coniugata permette di ricavare un'espressione per la $\frac{\partial}{\partial t} \psi^*(x, t)$:

$$i\hbar \frac{\partial}{\partial t} \psi^*(x, t) = \left[\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} - V(x) \right] \psi^*(x, t)$$

In questo discorso si deve aggiungere come ipotesi che $V(x)$ sia reale, perchè il potenziale dell'equazione di Schrödinger corrisponde al potenziale fisico. Questa ipotesi è sempre fatta nella teoria di Schrödinger.

Sostituendo l'espressione delle derivate temporali:

$$\begin{aligned} \frac{\partial}{\partial t} |\psi(x, t)|^2 &= \frac{i\hbar}{2m} \left[\psi^*(x, t) \cdot \frac{\partial^2}{\partial x^2} \psi(x, t) - \frac{\partial^2}{\partial x^2} \psi^*(x, t) \cdot \psi(x, t) \right] = \\ &= \frac{\partial}{\partial x} \left[\psi^*(x, t) \cdot \frac{\partial}{\partial x} \psi(x, t) - \frac{\partial}{\partial x} \psi^*(x, t) \cdot \psi(x, t) \right] \end{aligned}$$

¹⁴Questo aspetto sarà approfondito in §4.5.

¹⁵Si ricorda qui che, essendo la $\psi(x, t)$ una funzione complessa, risulta $|\psi(x, t)|^2 = \psi^*(x, t)\psi(x, t)$.

ottenendo:

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{\partial}{\partial t} |\psi(x, t)|^2 dx &= \frac{i\hbar}{2m} \int_{-\infty}^{+\infty} \frac{\partial}{\partial x} \left[\psi^*(x, t) \cdot \frac{\partial}{\partial x} \psi(x, t) - \frac{\partial}{\partial x} \psi^*(x, t) \cdot \psi(x, t) \right] dx = \\ &= \frac{i\hbar}{2m} \left[\psi^*(x, t) \cdot \frac{\partial}{\partial x} \psi(x, t) - \frac{\partial}{\partial x} \psi^*(x, t) \cdot \psi(x, t) \right]_{-\infty}^{+\infty} \end{aligned}$$

ma la funzione d'onda si annulla all'infinito per cui vale effettivamente:

$$\frac{d}{dt} N = 0$$

Occorre fare attenzione al fatto che la meccanica quantistica non è una teoria statistica. In una teoria statistica occorre sommare le probabilità, mentre in meccanica quantistica sono le ampiezze che si sommano:

$$|\psi_1(x, t) + \psi_2(x, t)|^2 = |\psi_1(x, t)|^2 + |\psi_2(x, t)|^2 + \underbrace{\psi_1^*(x, t)\psi_2(x, t) + \psi_1(x, t)\psi_2^*(x, t)}_{\text{termini di interferenza}} \neq |\psi_1(x, t)|^2 + |\psi_2(x, t)|^2$$

La somma delle ampiezze dà origine a fenomeni di interferenza dovuti al termine $\psi_1^*(x, t)\psi_2(x, t) + \psi_1(x, t)\psi_2^*(x, t)$, ed in questo risiede l'aspetto ondulatorio della materia. *I sistemi non sono descritti in termini di onde o corpuscoli, ma in termini di ampiezze di probabilità.*

Si consideri l'equazione di Schrödinger indipendente dal tempo:

$$-\frac{\hbar^2}{2m} \nabla^2 \psi(\vec{x}) = [E - V(\vec{x})] \psi(\vec{x})$$

Questa equazione è valida evidentemente solo per particelle di energia E . In quanto segue si considererà per semplicità il caso unidimensionale, per cui l'equazione diviene:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = [E - V(x)] \psi(x)$$

Questa equazione era stata ricavata nel paragrafo precedente nell'ipotesi che la funzione d'onda fosse fattorizzabile nella forma $\Psi(x, t) = \psi(x)\varphi(t)$. Si può ora comprendere perchè essa rappresenta uno stato stazionario del sistema: in questa ipotesi, la funzione d'onda di Schrödinger assume la forma $\Psi(x, t) = \psi(x)e^{-itE/\hbar}$. Nell'interpretazione probabilistica, quello che conta è il modulo quadro e evidentemente risulta:

$$|\Psi(x, t)| = |\psi(x)||e^{-itE/\hbar}| = |\psi(x)|$$

dunque una componente temporale fattorizzata come una fase non ha alcun ruolo nell'evoluzione del sistema: si tratta quindi di uno stato stazionario. Una volta ricavata la soluzione $\psi(x)$ si trova dunque senza difficoltà la $\Psi(x, t)$.

In sostanza, l'equazione di Schrödinger indipendente dal tempo è una banale equazione differenziale del secondo ordine che dal punto di vista matematico ammette sempre una soluzione se i coefficienti non hanno singolarità. Questa soluzione necessita di due costanti,¹⁶ inoltre essendo omogenea un multiplo di $\psi(x)$ è ancora una soluzione. Oltre alle condizioni iniziali per determinare le costanti si deve imporre che la funzione d'onda appartenga a \mathcal{L}^2 , ovvero che la soluzione vada a zero all'infinito:

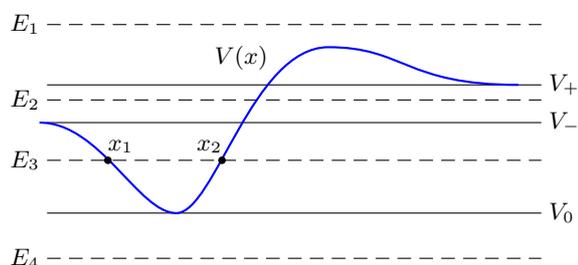
$$\lim_{x \rightarrow \pm\infty} |\psi(x)| = 0$$

inoltre, come si vedrà in seguito, se la funzione d'onda viene ricavata per pezzi, occorre imporre anche delle condizioni di raccordo fra le varie parti. *Tutte queste condizioni sono in genere incompatibili fra di loro: occorre quindi cercare i particolari valori delle costanti per i quali esse risultino*

¹⁶Essendo un'equazione differenziale del secondo ordine, essa ammette due soluzioni indipendenti, dunque la generica soluzione è scritta come una combinazione lineare di queste due, da cui le due costanti arbitrarie da determinare in base alle condizioni iniziali.

verificate contemporaneamente. In altri termini, alle soluzioni puramente matematiche dell'equazione di Schrödinger vanno aggiunte delle richieste di natura fisica che permettono di accettare come possibili solo alcune fra esse.

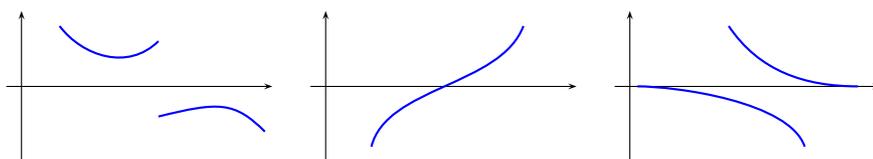
Per comprendere il significato concreto di questo discorso, si consideri un generico potenziale come quello in figura:



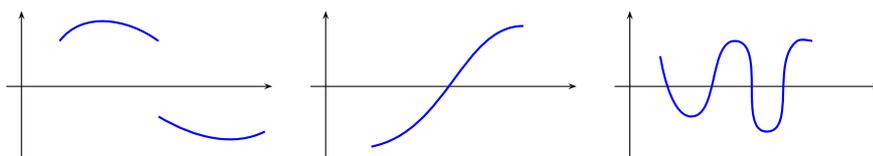
Questa forma di potenziale tende a V_+ per $x \rightarrow +\infty$ e a V_- per $x \rightarrow -\infty$. Inoltre si suppone $V_+ > V_-$. L'equazione di Schrödinger indipendente dal tempo ha la forma:

$$\frac{d^2}{dx^2}\psi(x) = -\frac{2m}{\hbar} [E - V(x)] \psi(x)$$

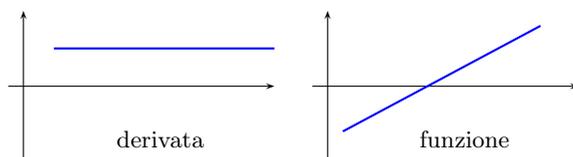
Se $E - V(x) < 0$, la derivata seconda ha lo stesso segno della funzione: pertanto se la funzione non interseca l'asse la convessità è verso di esso, se lo interseca si allontana da esso. Può anche capitare che la soluzione tenda asintoticamente all'asse. In questo caso, il modulo della funzione deve crescere indefinitamente a sinistra o a destra. Pertanto, per $E < V_0$ non esiste alcuna soluzione fisicamente accettabile (caso E_4).



Se invece $E - V(x) > 0$, la derivata seconda ha segno opposto alla funzione e in questo caso accade il contrario. La funzione può intersecare l'asse più volte ed assumere quindi un andamento oscillatorio.



Se accade invece $E - V(x) = 0$, il che può accadere solo se il potenziale è costante, la derivata prima è una costante e la funzione è una retta:

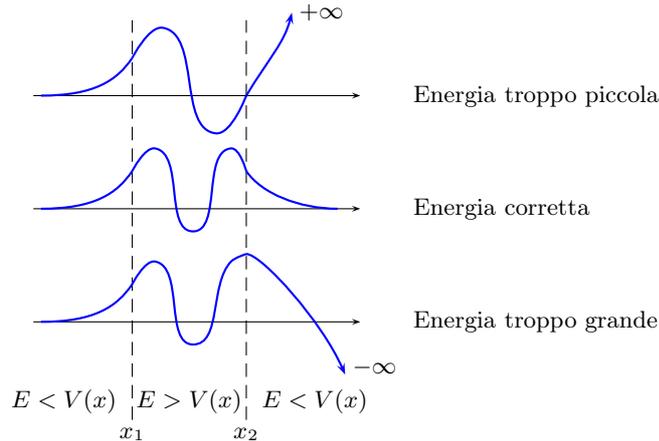


in questo caso la funzione interseca l'asse ma non lo "tocca", perché se la funzione si annulla in un punto e la derivata prima è nulla, allora la funzione è ovunque nulla.

Se $E > V_+$ (caso E_1) le soluzioni sono ovunque oscillatorie, anche a $\pm\infty$. Se invece E è inferiore al massimo, ma $E > V_+$ le soluzioni sono ancora oscillatorie e si ha un problema di barriera di potenziale. In entrambi i casi la funzione rappresenta una particella viaggiante da sinistra a destra e viceversa, con una riflessione parziale dell'onda.

Se risulta $V_- < E < V_+$ (caso E_2) si ha ancora un problema di barriera: l'onda è in parte trasmessa e in parte riflessa.

Nel caso $V_0 < E < V_-$ (caso E_3), c'è una regione centrale in cui $E - V(x) > 0$ e due regioni a sinistra e a destra in cui $E - V(x) < 0$. I punti di separazione, indicati con x_1 e x_2 sono i *punti di inversione classici*. Ai due lati di questa buca la funzione d'onda deve tendere all'asse x , mentre fra i punti x_1 e x_2 le due soluzioni sono fisicamente accettabili e la funzione d'onda ha un comportamento oscillatorio. Il problema in questo caso è *raccordare le due funzioni in modo che la $\psi(x,t)$ totale sia continua con la sua derivata prima*. È precisamente questo che per una energia E arbitraria non è possibile in genere: le soluzioni sono fisicamente accettabili solo in corrispondenza di determinati valori discreti di E , a ciascuno dei quali corrisponde uno stato stazionario legato del sistema. L'insieme dei valori possibili di E prende il nome di **spettro di energia**. Questo aspetto può essere compreso grazie alle figure seguenti:



Si supponga di avere un'energia arbitraria E e che si soddisfino le condizioni a sinistra di x_1 . Nel punto x_1 la soluzione a sinistra va raccordata con la soluzione oscillatoria fra x_1 e x_2 : siccome la funzione e la derivata prima sono continue, si otterrà un'unica soluzione. Similmente accade a destra di x_2 , e si avrà un'unica soluzione. Se il valore di E non è quello corretto, questa soluzione di destra non tenderà asintoticamente all'asse, ma divergerà allontanandosi da esso rendendola fisicamente inaccettabile. Le richieste che la funzione d'onda decresca sia a sinistra che a destra della buca sono incompatibili, tranne che per un insieme discreto (e forse vuoto) di livelli di energia. Questo accade solo se $E < V_-$, se invece $E > V_-$ esiste un continuo di energia, ed ogni soluzione è accettabile. Si comprende qui un aspetto fondamentale della meccanica quantistica: *la quantizzazione dell'energia appare solo per i sistemi legati*.¹⁷

In ciascun caso, il problema di trovare gli stati stazionari di un sistema può essere considerato come il problema di adattare una funzione d'onda oscillatoria fra due punti classici di inversione.

Gli esempi illustrati qui in modo generico saranno risolti in maniera completa nel paragrafo successivo.

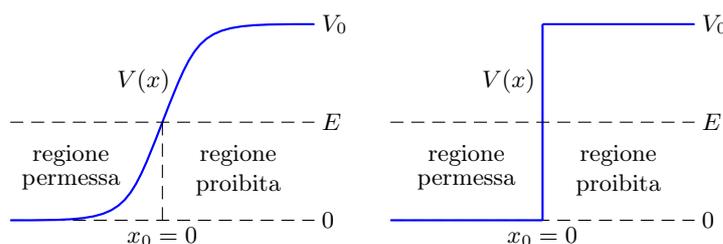
3.4 Calcolo della funzione d'onda in alcuni casi di rilevanza pratica

Per illustrare concretamente il calcolo della funzione d'onda, si analizzeranno ora tre casi di rilevanza pratica.

3.4.1 La barriera di potenziale

Si consideri una barriera di potenziale come in figura seguente:

¹⁷È questo, per esempio, il caso di particelle confinate in una buca di potenziale o vincolate in orbite chiuse dalla forza coulombiana.



Il punto x_0 è il cosiddetto *punto classico di inversione*: una particella classica di energia E si fermerebbe qui e tornerebbe indietro. Per semplificare lo studio di questo caso si considera la barriera a gradino mostrata a destra, la quale risulta definita da:

$$\begin{cases} V(x) = 0 & x < 0 \\ V(x) = V_0 & x > 0 \end{cases}$$

con $V_0 > E$. Naturalmente, questo è un caso limite.

Per ragioni fisiche, la funzione d'onda $\psi(x, t)$ deve essere continua insieme alla sua derivata prima, mentre la derivata seconda in genere può essere discontinua.

Per $x > 0$ l'equazione diventa:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = [E - V(x)] \psi(x)$$

che ha due soluzioni indipendenti:

$$\psi_1(x) = e^{-qx} \quad ; \quad \psi_2(x) = e^{qx} \quad \quad q = \frac{\sqrt{2m(V_0 - E)}}{\hbar}$$

per cui la generica soluzione ha la forma:

$$\psi(x) = Ce^{qx} + De^{-qx}$$

la soluzione e^{qx} cresce esponenzialmente con x , pertanto è da scartare per motivi fisici (non appartiene infatti a \mathcal{L}^2). Dunque (il pedice d indica la soluzione a destra della barriera):

$$\psi_d(x) = De^{-qx} \quad x > 0$$

Per $x < 0$ l'equazione ha la forma:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = E\psi(x)$$

le cui soluzioni indipendenti sono:

$$\psi_1(x) = e^{-ikx} \quad ; \quad \psi_2(x) = e^{ikx} \quad \quad k = \frac{\sqrt{2mE}}{\hbar}$$

queste due soluzioni sono oscillatorie, esse non divergono e sono pertanto *entrambe* accettabili fisicamente. La soluzione a sinistra è data quindi da (il pedice s indica la soluzione a sinistra della barriera):

$$\psi_s(x) = Ae^{ikx} + Be^{-ikx}$$

dove A e B sono due costanti che si ricavano in base alle condizioni fisiche:

$$\begin{cases} \psi_d(0) = \psi_s(0) \\ \psi'_d(0) = \psi'_s(0) \end{cases}$$

queste condizioni traggono origine dal fatto che esse rappresentano la *stessa* funzione d'onda fisica in regioni diverse e pertanto *devono* raccordarsi in $x = 0$. Da queste condizioni si ricava:

$$\begin{cases} A + B = D \\ ik(A - B) = -qD \end{cases}$$

che rappresentano due equazioni in tre incognite, il terzo parametro tuttavia risulta fissato dalla normalizzazione. Risolvendo quindi in funzione di D si ricava:

$$\begin{cases} A = \frac{1}{2} \left(1 - i\frac{q}{k}\right) D \\ B = \frac{1}{2} \left(1 + i\frac{q}{k}\right) D \end{cases}$$

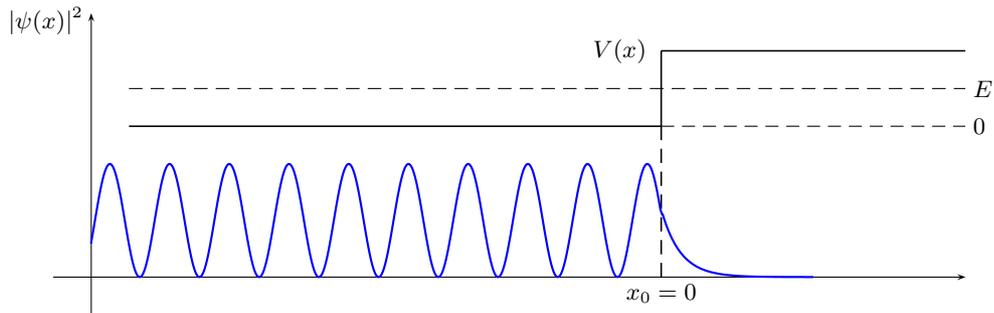
Se ora si moltiplica ovunque la funzione d'onda per $1/A^{(18)}$ si ricava, tenendo presente le espressioni di q e k :

$$\begin{cases} \psi(x) = e^{ikx} + \left[\frac{1 - i\sqrt{V_0/E - 1}}{1 + i\sqrt{V_0/E - 1}} \right] e^{-ikx} & x < 0 \\ \psi(x) = \frac{2e^{-qx}}{1 + i\sqrt{V_0/E - 1}} & x > 0 \end{cases}$$

nella regione $x < 0$ c'è un termine che si propaga verso destra (e^{ikx}) e uno verso sinistra (e^{-ikx}), in altre parole una parte rappresenta un'onda che avanza verso la barriera, l'altra rappresenta un'onda riflessa. Il modulo del rapporto fra il coefficiente B e A si chiama *coefficiente di riflessione* ed è dato da:

$$R = \left| \frac{\text{ampiezza riflessa}}{\text{ampiezza incidente}} \right| = \frac{|B|^2}{|A|^2} = \frac{B^* B}{A^* A} = \frac{\left(1 + i\frac{q}{k}\right) \left(1 - i\frac{q}{k}\right)}{\left(1 - i\frac{q}{k}\right) \left(1 + i\frac{q}{k}\right)} = 1$$

Si vede quindi che l'onda incidente e l'onda riflessa hanno lo stesso modulo, ovvero la particella viene riflessa (interamente) dalla barriera. È notevole (e estremamente importante) però che la probabilità di trovare la particella dentro la barriera non è nulla (la soluzione per $x > 0$ non è nulla), ma finita e diversa da zero, anche se decresce esponenzialmente con la distanza. Nella meccanica classica, la legge di conservazione dell'energia impone che una particella non possa superare un ostacolo (barriera) se non ha l'energia necessaria per farlo, ovvero $E > V_0$, per cui essa viene riflessa dalla barriera. Quantisticamente, anche una particella con energia inferiore alla barriera ha una probabilità finita di penetrare prima di essere riflessa. Questo comportamento è alla base dell'effetto tunnel illustrato più avanti.

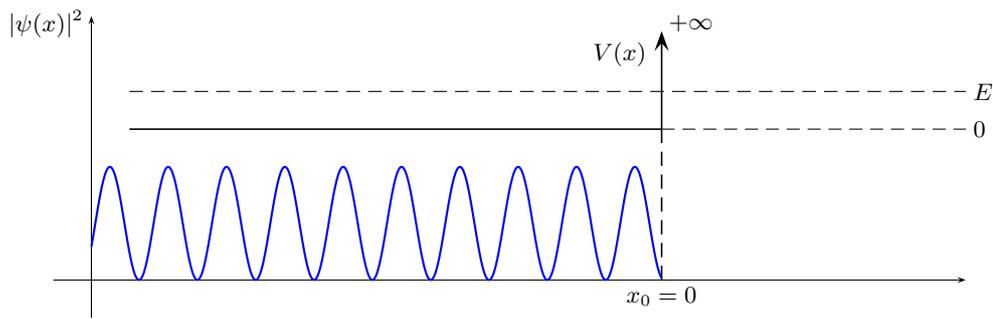


Se l'altezza della barriera V tende a infinito, allora q tende a infinito e la soluzione in questo caso diventa:

$$\begin{cases} \psi_s(x) = e^{ikx} - e^{-ikx} & x < 0 \\ \psi_d(x) = 0 & x > 0 \end{cases}$$

ovvero, in caso di barriera infinita la funzione d'onda si annulla sulla barriera e la probabilità di trovare la particella all'interno della barriera è effettivamente nulla.

¹⁸L'equazione di Schrödinger è lineare.



Se l'energia della particella è superiore alla barriera $E > V_0$, le soluzioni sono oscillatorie da entrambi i lati della barriera. Esse sono pertanto le seguenti:

$$\begin{cases} \psi_s(x) = Ae^{ik_1x} + Be^{-ik_1x} \\ \psi_d(x) = Ce^{ik_2x} + De^{-ik_2x} \end{cases} \quad \begin{cases} k_1 = \frac{\sqrt{2mE}}{\hbar} \\ k_2 = \frac{\sqrt{2m(E - V_0)}}{\hbar} \end{cases}$$

Dopo la barriera non può esserci riflessione, per cui supponendo la particella incidente da sinistra risulta $D = 0$. Le condizioni di raccordo sono quindi:

$$\begin{cases} \psi_s(0) = \psi_d(0) \\ \psi'_s(0) = \psi'_d(0) \end{cases} \Rightarrow \begin{cases} A + B = C \\ ik_1A - ik_1B = ik_2C \end{cases} \Rightarrow \begin{cases} A + B = C \\ A - B = \frac{k_2}{k_1}C \end{cases}$$

come al solito, si tratta di due equazioni per tre costanti, la terza essendo fissata dalla condizione di normalizzazione. Le costanti sono dunque:

$$\begin{cases} A = \frac{1}{2} \left(1 + \frac{k_2}{k_1} \right) C \\ B = \frac{1}{2} \left(1 - \frac{k_2}{k_1} \right) C \end{cases}$$

È utile mettere queste costanti anche nella forma:

$$\begin{cases} B = \frac{k_1 - k_2}{k_1 + k_2} A \\ C = \frac{2k_1}{k_1 + k_2} A \end{cases}$$

in questo modo risulta immediato definire il *coefficiente di riflessione*:

$$R = \left| \frac{\text{ampiezza riflessa}}{\text{ampiezza incidente}} \right| = \frac{|B|^2}{|A|^2} = \frac{B^*B}{A^*A} = \left(\frac{k_1 - k_2}{k_1 + k_2} \right)^2$$

e il *coefficiente di trasmissione*:

$$T = \left| \frac{\text{ampiezza trasmessa}}{\text{ampiezza incidente}} \right| = \frac{|C|^2}{|A|^2} = \frac{C^*C}{A^*A} = \frac{4k_1k_2}{(k_1 + k_2)^2}$$

e si verifica, come deve, che $R + T = 1$.

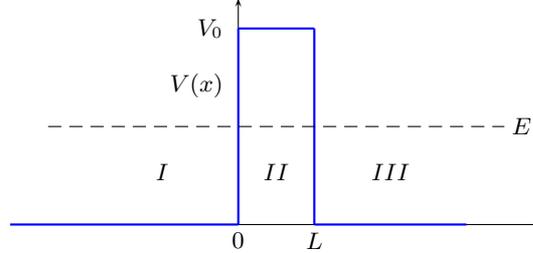
3.4.1.1 L'effetto tunnel

Il fatto che la meccanica quantistica prevede una penetrazione nella barriera anche per una particella con energia non sufficiente per scavalcarla porta a chiedersi: cosa accade se lo spessore della barriera

è tale che la probabilità non si annulli all'interno di essa, ma sia ancora maggiore di zero una volta raggiunto l'altro lato?

La risposta della meccanica quantistica è semplice: in questo caso, *esiste una probabilità non nulla di trovare la particella al di là della barriera*. In altri termini, *se la barriera è abbastanza sottile, la particella può passare dall'altro lato (tunnelling) anche se non ha energia sufficiente per scavalcarla*. Questo è un effetto strettamente quantistico che non ha alcun analogo classico.

Si consideri la barriera mostrata in figura:



Il potenziale è nullo ovunque, tranne in un intervallo $0 \leq x \leq L$, dove vale V_0 :

$$\begin{cases} V(x) = 0 & x \in]-\infty, 0[\times]L, +\infty[\\ V(x) = V_0 & x \in [0, L] \end{cases}$$

L'asse sarà diviso in tre parti e così la funzione d'onda. La situazione che si vuole studiare è costituita da una particella che giunge sulla barriera di potenziale dalla regione di sinistra. L'equazione di Schrödinger per le tre regioni fornisce allora le soluzioni generali:

$$\begin{cases} \psi_I(x) & = Ae^{ikx} + Be^{-ikx} \\ \psi_{II}(x) & = Ce^{qx} + De^{-qx} \\ \psi_{III}(x) & = Fe^{ikx} + Ge^{-ikx} \end{cases} \quad \begin{cases} q = \frac{\sqrt{2m(V_0 - E)}}{\hbar} \\ k = \frac{\sqrt{2mE}}{\hbar} \end{cases}$$

siccome nella situazione in studio non esiste alcuna particelle incidente da destra verso sinistra, nella regione *III* deve risultare, per ragioni fisiche, $G = 0$: si ricordi infatti che il termine e^{-ikx} rappresenta un'onda che si propaga da destra verso sinistra. Restano pertanto cinque coefficienti da determinare in base alle condizioni di raccordo in $x = 0$ (due per la funzione e la sua derivata prima), in $x = L$ (altre due) e la condizione di normalizzazione. Si noti che in questo caso non si deve imporre $D = 0$, perché essendo la larghezza della barriera finita l'esponenziale non può in ogni caso divergere a infinito.

Le condizioni di raccordo in $x = 0$ conducono alle equazioni:

$$\begin{cases} \psi_I(0) = \psi_{II}(0) \\ \psi'_I(0) = \psi'_{II}(0) \end{cases} \Rightarrow \begin{cases} A + B = C + D \\ ikA - ikB = qC - qD \end{cases}$$

le quali risolte per A forniscono:

$$A = \frac{1}{2} \left(1 - i\frac{q}{k}\right) C + \frac{1}{2} \left(1 + i\frac{q}{k}\right) D$$

Le altre due condizioni conducono alle equazioni:

$$\begin{cases} \psi_{II}(L) = \psi_{III}(L) \\ \psi'_{II}(L) = \psi'_{III}(L) \end{cases} \Rightarrow \begin{cases} Ce^{qL} + De^{-qL} = Fe^{ikL} \\ Cqe^{qL} + Dqe^{-qL} = Fike^{ikL} \end{cases}$$

che risolte per C e D forniscono:

$$\begin{cases} C = \frac{1}{2} e^{-qL} F \left(1 + i\frac{k}{q}\right) e^{ikL} \\ D = \frac{1}{2} e^{qL} F \left(1 - i\frac{k}{q}\right) e^{ikL} \end{cases}$$

È possibile quindi esprimere A in termini di F :

$$A = \left[\frac{1}{4} \left(1 - i\frac{q}{k}\right) \left(1 + i\frac{k}{q}\right) e^{-qL} e^{ikL} + \frac{1}{4} \left(1 + i\frac{q}{k}\right) \left(1 - i\frac{k}{q}\right) e^{qL} e^{ikL} \right] F$$

che tenendo conto della definizione del seno e coseno iperbolico e del fatto che $\cosh^2 x - \sinh^2 x = 1$ si può riscrivere nella forma:

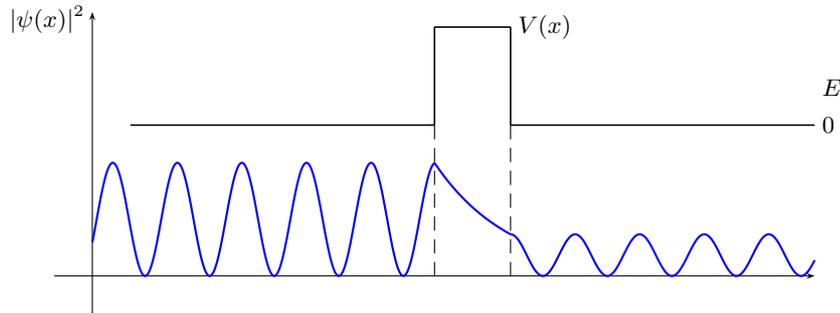
$$A = 2Fe^{ikL} \left[\frac{1}{2} \cosh qL - i\frac{1}{4} \left(\frac{k}{q} - \frac{q}{k} \right) \sinh qL \right]$$

Questa forma è interessante, perché il rapporto $T = |F|^2/|A|^2$ rappresenta la parte di funzione d'onda trasmessa attraverso la barriera, chiamato appunto *coefficiente di trasmissione*:

$$\begin{aligned} T &= \left| \frac{\text{ampiezza trasmessa}}{\text{ampiezza incidente}} \right|^2 = \frac{|F|^2}{|A|^2} = \left| \frac{1}{2e^{ikL} \left[\frac{1}{2} \cosh qL - i\frac{1}{4} \left(\frac{k}{q} - \frac{q}{k} \right) \sinh^2 qL \right]} \right|^2 = \\ &= \frac{1}{1 + \frac{1}{4} \frac{V_0}{E \left(1 - \frac{E}{V_0}\right)} \sinh^2 \frac{\sqrt{2m(V_0 - E)}}{\hbar} L} \end{aligned}$$

Si nota dalla forma del coefficiente di trasmissione che l'effetto è tanto più marcato quanto più piccola è la massa o più grande l'energia.

In pratica, è come se la particella scavasse un *tunnel* sotto la barriera di potenziale ed uscisse dall'altra parte, da cui il nome di *effetto tunnel*. Si noti però che sperimentalmente risulta impossibile misurare la particella *mentre* attraversa la barriera. Ogni volta che si cerca di determinarne la posizione la si trova al di qua o al di là della barriera, ma mai all'interno.



Nel caso risulti $E > V_0$, le soluzioni nelle regioni *I* e *III* rimangono le stesse, mentre nella regione *II* la soluzione diventa:

$$\psi_{II}(x) = Ce^{ik'x} + De^{-ik'x} \quad k' = \frac{\sqrt{2m(E - V_0)}}{\hbar}$$

I calcoli risultano identici al caso visto sopra, ed il coefficiente di trasmissione è dato da:

$$T = \frac{1}{1 + \frac{1}{4} \frac{V_0}{E \left(1 - \frac{E}{V_0}\right)} \sin^2 \frac{\sqrt{2m(E - V_0)}}{\hbar} L}$$

In questo caso T presenta un andamento oscillatorio smorzato, è interessante inoltre notare che $|T| < 1$, per cui anche quando $E < V_0$ si ha una parziale riflessione dell'onda: *anche se l'energia della particella è superiore alla barriera di potenziale, esiste una probabilità finita che essa sia riflessa*.

In particolare, se risulta $\frac{\sqrt{2m(E - V_0)}}{\hbar} L = n\pi$, il coefficiente di trasmissione è uguale a 1.

Anche se non è intuitivo, l'effetto tunnel fornisce la spiegazione di diversi fenomeni.

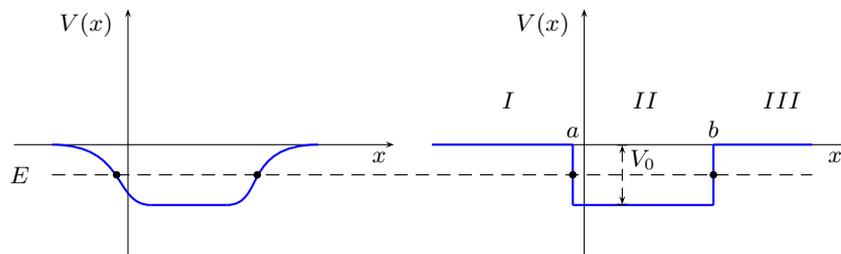
Decadimento Alfa: L'effetto tunnel venne utilizzato per la prima volta nel 1928 dal fisico ucraino George Gamow per spiegare il decadimento alfa, nel quale una particella alfa (un nucleo di elio) viene emessa da un nucleo perché riesce a superarne la barriera di potenziale. Questo fenomeno è in realtà inspiegabile con le leggi della fisica classica, perché le particelle emesse hanno un'energia *inferiore* alla barriera di potenziale rappresentata dal nucleo: la soluzione al problema venne dall'ipotesi che le particelle emesse riuscissero ad evadere dal nucleo grazie all'effetto tunnel. Calcoli precisi mostrano che l'effetto tunnel riesce a spiegare accuratamente la legge ed il rate del decadimento.

Nucleosintesi: Nelle condizioni tipiche degli interni stellari, il gas di protoni liberi è soggetto alle forze repulsive coulombiane a lungo range che ostacolano la sintesi dell'He, mentre la fusione nucleare può realizzarsi solo se le particelle entrano nel raggio d'azione delle forze nucleari. Secondo la fisica classica, la barriera repulsiva coulombiana è abbastanza alta da impedire che i due nuclei di idrogeno si avvicinino abbastanza. L'effetto tunnel è invece quello che permette ad un nucleo di idrogeno di penetrare la barriera coulombiana e dare inizio al processo di fusione nucleare. Ancora una volta, calcoli precisi mostrano che l'effetto tunnel riesce a prevedere accuratamente il rate delle reazioni nucleari.

Diodo a effetto tunnel (o Esaki): In questo particolare tipo di diodo il flusso di elettroni è regolato dall'altezza di una barriera di potenziale ed è fornito proprio grazie all'effetto tunnel. Questo flusso si può interrompere molto velocemente agendo sulla barriera: siccome questa variazione può essere rapidissima (anche inferiore a 5 ps) questo dispositivo è utilizzato quando occorrono risposte estremamente rapide.

3.4.2 La buca di potenziale

Si consideri un elettrone in una buca di potenziale:



la forma a sinistra non è risolvibile analiticamente con nessuna tecnica, per questo si utilizza la forma del potenziale semplificata a destra. In questa schematizzazione l'elettrone risente di una forza solo nei punti di discontinuità, in cui $F = \partial V / \partial x \neq 0$: si tratta di una forza cosiddetta *impulsiva*. Quello che conta è in realtà che l'impulso $\Delta p = \int F dt$ sia finito. Così per l'energia $\Delta T = \int F dl$, si considera una forza fittizia concentrata nei punti di discontinuità. Mentre nel caso reale la particella oscilla fra i due punti di inversione, nel caso schematizzato è piuttosto come se rimbalzasse fra le pareti.

In queste ipotesi, il potenziale presente nell'equazione differenziale è differenziabile infinite volte e di conseguenza le funzioni soluzione devono appartenere anch'esse a $C^{(\infty)}$: $V \in C^{(\infty)} \Rightarrow \psi(x) \in C^{(\infty)}$. Per la precisione, nei punti di discontinuità del potenziale la derivata seconda non è continua per cui la formulazione corretta è che la funzione d'onda appartiene a $C^{(\infty)}$ in ogni singolo tratto, mentre globalmente $\psi(x) \in C^{(1)}$.

Si supponga $E < 0$. Nella zona I risulta $V - E > 0$ per cui le soluzioni sono date da esponenziali reali. Similmente per la soluzione nella zona III. Nella zona II invece risulta $-V_0 - E < 0$ per cui

le soluzioni sono date da esponenziali immaginari e dunque oscillatorie:¹⁹

$$\begin{cases} \psi_I(x) &= A_1 e^{kx} + B_1 e^{-kx} \\ \psi_{II}(x) &= A_2 e^{i\chi x} + B_2 e^{-i\chi x} \\ \psi_{III}(x) &= A_3 e^{kx} + B_3 e^{-kx} \end{cases} \quad \begin{cases} k &= \frac{\sqrt{2mE}}{\hbar} \\ \chi &= \frac{\sqrt{2m(E+V_0)}}{\hbar} \end{cases}$$

occorre ora imporre le condizioni di raccordo nei punti a e b :

$$\begin{cases} \psi_I(a) &= \psi_{II}(a) \\ \psi'_I(a) &= \psi'_{II}(a) \end{cases} \quad \begin{cases} \psi_{II}(b) &= \psi_{III}(b) \\ \psi'_{II}(b) &= \psi'_{III}(b) \end{cases}$$

si tratta di quattro condizioni in sei parametri, tuttavia, essendo un'equazione omogenea, uno dei parametri può essere utilizzato come divisore. Una volta imposto che la soluzione appartenga a \mathcal{L}^2 (ovvero che $\lim_{x \rightarrow \pm\infty} \psi(x) = 0$), si ritrovano sei condizioni in cinque parametri. Ora, le condizioni che le funzioni si annullino all'infinito implicano che $B_1 = A_3 = 0$, per cui si ritrovano quattro coefficienti in quattro equazioni:

$$\begin{cases} e^{ka} A_1 & -e^{i\chi a} A_2 & -e^{-i\chi a} B_2 & & = 0 \\ ke^{ka} A_1 & -i\chi e^{i\chi a} A_2 & -i\chi e^{-i\chi a} B_2 & & = 0 \\ & e^{i\chi b} A_2 & e^{-i\chi b} B_2 & e^{-kb} B_3 & = 0 \\ & i\chi e^{i\chi b} A_2 & -i\chi e^{-i\chi b} B_2 & ke^{-kb} B_3 & = 0 \end{cases}$$

Tuttavia, questo sistema non ha sempre soluzioni. Per avere una soluzione diversa dalla soluzione banale il determinante del sistema deve essere nullo, il che si traduce nella condizione:

$$\frac{k - i\chi}{k + i\chi} e^{i\chi(b-a)} = \frac{k + i\chi}{k - i\chi} e^{-i\chi(b-a)}$$

il secondo membro di questa equazione è il complesso coniugato del primo. Se un numero coincide con il suo complesso coniugato, allora la sua parte immaginaria deve essere nulla ed il numero è reale:

$$\Im \left[\frac{k - i\chi}{k + i\chi} e^{i\chi(b-a)} \right] = 0$$

la parte immaginaria è data da:

$$2k\chi \cos \chi(b-a) = (k^2 - \chi^2) \sin \chi(b-a)$$

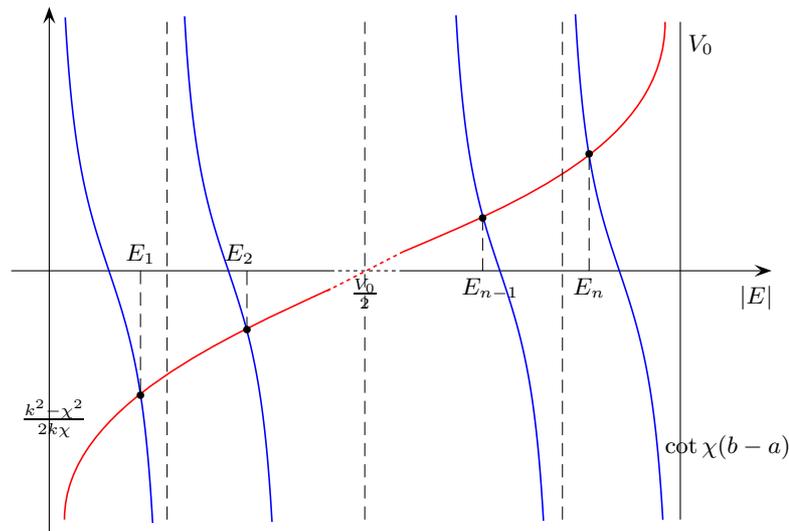
ovvero:

$$\boxed{\frac{k^2 - \chi^2}{2k\chi} = \cot \chi(b-a)}$$

L'unico parametro che può variare in questa equazione è l'energia E . Questa è un'equazione trascendente, quindi la si risolverà per via grafica:

¹⁹Si noti fra l'altro che questa corrisponde all'equazione di un oscillatore armonico:

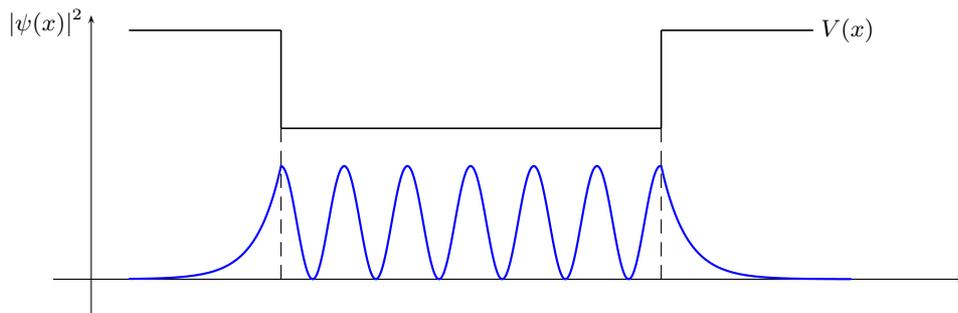
$$\frac{d^2}{dx^2} \psi(x) = \frac{2m}{\hbar^2} (V - E) \psi(x) \quad \Rightarrow \quad \frac{d^2}{dx^2} \psi(x) + k^2 \psi(x) = 0$$



dal grafico si vede che esistono solo alcuni valori possibili per l'energia ($E_1, E_2, \dots, E_{n-1}, E_n$), permette quindi di vedere come si originano i valori quantizzati dell'energia per una particella in una buca di potenziale. Il numero di rami di cotangente presenti nell'intervallo $[0, V_0]$ dipende dal valore del suo argomento e quindi in ultima analisi dal valore dell'energia della particella, oltre che dalla larghezza della buca. Si può notare infatti che all'aumentare dell'energia della particella, il numero di rami di cotangente aumenta sempre più e di conseguenza il numero di autovalori dell'energia E_i possibili aumenta sempre più fino a tendere al continuo. Similmente, per lo stesso ragionamento, se la larghezza della buca aumenta sempre più il numero di autovalori aumenta e al limite tende al continuo. Questo è coerente con il fatto che quando la larghezza tende a infinito, la buca in pratica scompare.

Come detto, queste condizioni rappresentano il fatto che il numero di onde all'interno della buca si deve raccordare con l'andamento esponenziale all'esterno. È importante notare ancora una volta che mentre la particella classica è destinata a muoversi all'interno della buca, una particella quantistica può penetrare in una zona classicamente proibita.

L'andamento tipico *qualitativo* della funzione d'onda in una buca di potenziale è rappresentato nella figura seguente.



Nel caso di $E < V_0$, la sola soluzione del sistema è quella banale, per cui tutti i coefficienti sono zero e la funzione d'onda è nulla ovunque: nessuna particella può esistere.

Se invece risulta $E > 0$, tutte le soluzioni sono di carattere oscillatorio:

$$\begin{cases} \psi_I(x) &= A_1 e^{kx} + B_1 e^{-kx} \\ \psi_{II}(x) &= A_2 e^{k'x} + B_2 e^{-k'x} \\ \psi_{III}(x) &= A_3 e^{kx} + B_3 e^{-kx} \end{cases}$$

In questo caso, la richiesta che la funzione d'onda si annulli ad infinito non ha importanza: le soluzioni nelle tre regioni sono delle onde piane che si propagano verso destra (coefficienti A) e verso sinistra

(coefficienti B). Imponendo le condizioni di continuità su $\psi(x)$ e $\psi'(x)$ si hanno quattro condizioni in sei incognite, si ha perciò sempre soluzione: in questo caso *non si ha quantizzazione*, ma due soluzioni indipendenti che rappresentano le due onde che si propagano nelle due direzioni. Date le condizioni sperimentali, si può porre $B_3 = 0$ (nessuna onda incidente da destra), ma non annullare B_1 e B_2 in quanto è presente un'onda riflessa. Pertanto *le particelle che passano sopra una buca di potenziale sono in parte riflesse e in parte trasmesse*. Il fenomeno è simile alla riflessione e alla trasmissione della luce attraverso una lastra di vetro.

Caso particolare: la buca infinita

Si analizzerà ora in dettaglio il caso importante di una buca di potenziale infinita.

Il potenziale risulta definito in questo caso da:

$$\begin{cases} V(x) = 0 & x \in]-\infty, 0[\times]L, +\infty[\\ V(x) = +\infty & x \in [0, L] \end{cases}$$

Questo implica che nella regione *I* e *III* la funzione d'onda è zero, mentre nella regione *II* la funzione d'onda è espressa da:

$$\psi(x) = Ae^{ikx} + Be^{-ikx} \quad k = \frac{\sqrt{2mE}}{\hbar}$$

per ragioni di simmetria deve risultare $|A| = |B|$,²⁰ per cui le due soluzioni indipendenti assumono la forma:

$$\begin{cases} \psi(x) = A(e^{ikx} + e^{-ikx}) = 2A \cos kx & A = B \\ \psi(x) = A(e^{ikx} - e^{-ikx}) = 2iA \sin kx & A = -B \end{cases}$$

la soluzione generale può quindi essere scritta nella forma di una somma di seni e coseni:

$$\psi(x) = A \cos kx + B \sin kx$$

Le condizioni al contorno da soddisfare sono l'annullamento della funzione d'onda sui bordi della buca:

$$\begin{cases} \psi(0) = 0 & \Rightarrow & A \cos 0 + B \sin 0 = 0 \\ \psi(L) = 0 & \Rightarrow & A \cos kL + B \sin kL = 0 \end{cases}$$

La prima di queste condizioni implica che $A = 0$, perché è l'unico modo per soddisfare questa condizione in $x = 0$. Resta quindi da soddisfare l'altra condizione, che risulta avvertata se:

$$kL = n\pi \quad \frac{\sqrt{2mE}}{\hbar}L = n\pi \quad \frac{2mE}{\hbar^2} = \frac{n^2\pi^2}{L^2}$$

da cui si ricavano le condizioni sull'energia²¹ che devono essere rispettate affinché la funzione d'onda si annulli sulla parete $x = L$ della barriera:

$$E_n = \frac{n^2\pi^2\hbar^2}{2mL^2}$$

con $\psi_n(x) = B \sin k_n x$. Resta ora da calcolare la costante di normalizzazione, che risulta definita da:

$$\int_{-\infty}^{+\infty} B^2 \sin^2 kx dx = \int_0^L B^2 \sin^2 kx dx \equiv 1$$

ovvero:

$$\frac{1}{B^2} = \int_0^L \sin^2 kx dx$$

²⁰Questo risulta più evidente se si cambia sistema di riferimento e si pone il centro della buca in $x = 0$, per cui le due pareti risultano essere a $-L/2$ e $L/2$. Poiché la soluzione rappresenta un'onda viaggiante verso sinistra e una verso destra in una buca a pareti infinite, per la simmetria della situazione i due coefficienti devono essere uguali.

²¹Si tratta dell'unico parametro che può variare.

L'integrale può essere calcolato con la sostituzione $kx \equiv y$ ($x = y/k$, $dx = 1/k dy$, $L \rightarrow kL = n\pi$) e tramite un artificio matematico:

$$\int_0^L \sin^2 kx dx = \frac{1}{k} \int_0^{n\pi} \sin^2 y dy$$

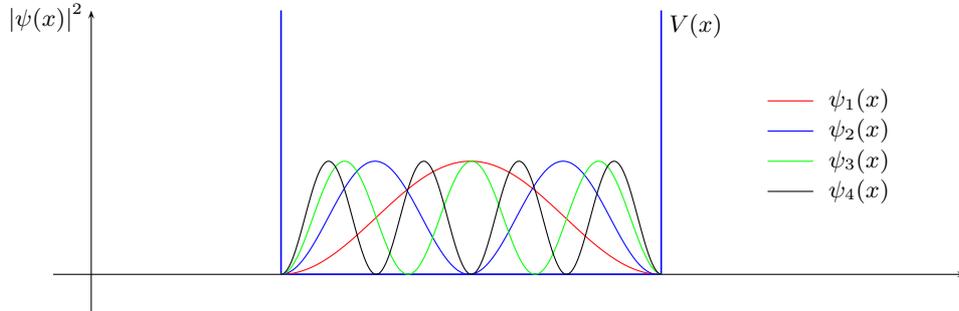
ma l'integrale di $\sin^2 x$ su mezzo periodo è la metà di quello su un periodo e questo è uguale all'integrale di $\cos^2 x$ sullo stesso periodo, per cui:

$$\frac{1}{2k} \int_0^{2n\pi} \sin^2 y dy = \frac{1}{4k} \int_0^{2n\pi} (\sin^2 y + \cos^2 y) dy = \frac{1}{4k} 2n\pi = \frac{1}{2k} kL = \frac{L}{2}$$

ovvero $B = \sqrt{2/L}$, per cui le autofunzioni normalizzate per la buca infinita di potenziale sono:

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi}{L} x = \sqrt{\frac{2}{L}} \sin \left(\frac{\sqrt{2mE_n}}{\hbar} x \right)$$

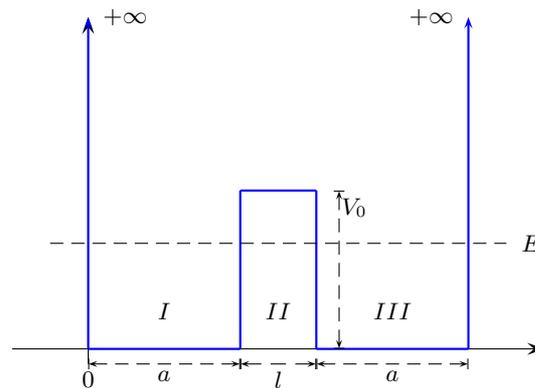
Le prime quattro autofunzioni corrispondenti ai primi quattro livelli di energia sono rappresentate nella figura seguente:



3.4.3 La doppia buca di potenziale

In questa parte si getteranno le linee guida per analizzare la doppia buca di potenziale, che risulta definita dalle relazioni:

$$\begin{cases} V(x) = +\infty & x \in]-\infty, 0[\cup]2a+l, +\infty[\\ V(x) = 0 & x \in [0, a] \cup [a+l, 2a+l] \\ V(x) = V_0 & x \in]a, a+l] \end{cases}$$



con la condizione $a\sqrt{2mV_0} \gg \hbar$. Le funzioni d'onda nelle tre regioni sono date da:

$$\begin{cases} \psi_I(x) &= A_1 e^{ikx} + B_1 e^{-ikx} \\ \psi_{II}(x) &= A_2 e^{\chi x} + B_2 e^{-\chi x} \\ \psi_{III}(x) &= A_3 e^{ikx} + B_3 e^{-ikx} \end{cases} \quad \begin{cases} k \equiv \sqrt{\frac{2mE}{\hbar^2}} \\ \chi \equiv \sqrt{\frac{2m(V_0 - E)}{\hbar^2}} \end{cases}$$

e le sei condizioni da imporre sono:

$$\begin{cases} \psi_I(0) = 0 \\ \psi_I(a) = \psi_{II}(a) \\ \psi'_I(a) = \psi'_{II}(a) \end{cases} \quad \begin{cases} \psi_{II}(a+l) = \psi_{III}(a+l) \\ \psi'_{II}(a+l) = \psi'_{III}(a+l) \\ \psi_{III}(2a+l) = 0 \end{cases}$$

ovvero:

$$\begin{cases} A_1 + B_1 = 0 \\ A_1 e^{ika} + B_1 e^{-ika} = A_2 e^{\chi a} + B_2 e^{-\chi a} \\ ikA_1 e^{ika} - ikB_1 e^{-ika} = \chi A_2 e^{\chi a} - \chi B_2 e^{-\chi a} \\ A_2 e^{\chi(a+l)} + B_2 e^{-\chi(a+l)} = A_3 e^{k(a+l)} + B_3 e^{-k(a+l)} \\ \chi A_2 e^{\chi(a+l)} - \chi B_2 e^{-\chi(a+l)} = kA_3 e^{k(a+l)} - kB_3 e^{-k(a+l)} \\ A_3 e^{k(2a+l)} + B_3 e^{-k(2a+l)} = 0 \end{cases}$$

che può essere messa in una forma più chiara in termini di matrici:

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ e^{ika} & e^{-ika} & -e^{\chi a} & -e^{-\chi a} & 0 & 0 \\ ik e^{ika} & -ik e^{-ika} & -\chi e^{\chi a} & \chi e^{-\chi a} & 0 & 0 \\ 0 & 0 & e^{\chi(a+l)} & e^{-\chi(a+l)} & -e^{k(a+l)} & -e^{-k(a+l)} \\ 0 & 0 & \chi e^{\chi(a+l)} & -\chi e^{-\chi(a+l)} & -k e^{k(a+l)} & k e^{-k(a+l)} \\ 0 & 0 & 0 & 0 & e^{k(2a+l)} & e^{-k(2a+l)} \end{pmatrix} \cdot \begin{pmatrix} A_1 \\ B_1 \\ A_2 \\ B_2 \\ A_3 \\ B_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Come al solito, affinché non si abbia la sola soluzione banale il determinante deve essere nullo. Si tratta di un determinante di ordine 6, applicando quindi la regola di Laplace alla prima riga si ottiene:

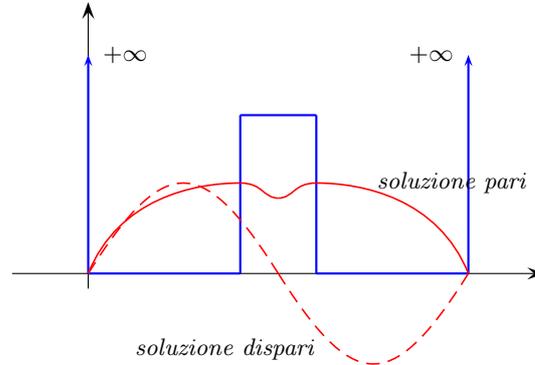
$$\begin{vmatrix} e^{-ika} & -e^{\chi a} & -e^{-\chi a} & 0 & 0 \\ -ik e^{-ika} & -\chi e^{\chi a} & \chi e^{-\chi a} & 0 & 0 \\ 0 & e^{\chi(a+l)} & e^{-\chi(a+l)} & -e^{k(a+l)} & -e^{-k(a+l)} \\ 0 & \chi e^{\chi(a+l)} & -\chi e^{-\chi(a+l)} & -k e^{k(a+l)} & k e^{-k(a+l)} \\ 0 & 0 & 0 & e^{k(2a+l)} & e^{-k(2a+l)} \end{vmatrix} - \begin{vmatrix} e^{ika} & -e^{\chi a} & -e^{-\chi a} & 0 & 0 \\ ik e^{ika} & -\chi e^{\chi a} & \chi e^{-\chi a} & 0 & 0 \\ 0 & e^{\chi(a+l)} & e^{-\chi(a+l)} & -e^{k(a+l)} & -e^{-k(a+l)} \\ 0 & \chi e^{\chi(a+l)} & -\chi e^{-\chi(a+l)} & -k e^{k(a+l)} & k e^{-k(a+l)} \\ 0 & 0 & 0 & e^{k(2a+l)} & e^{-k(2a+l)} \end{vmatrix} = 0$$

Tralasciando il lungo calcolo del determinante, si trova che le condizioni affinché esistano soluzioni sono le seguenti:

$$\begin{cases} A_2 = B_2 & \frac{k}{\chi} \cot ka = -\tanh\left(\chi \frac{l-a}{2}\right) \\ A_2 = -B_2 & \frac{k}{\chi} \cot ka = -\coth\left(\chi \frac{l-a}{2}\right) \end{cases}$$

In un sistema di riferimento con l'origine al centro, le soluzioni $A_2 = B_2$ sono soluzioni pari, mentre quelle che soddisfano la $A_2 = -B_2$ sono soluzioni dispari.

Si ponga ora un sistema di riferimento al centro della doppia buca. Siccome vale $\coth x > \tanh x$ per $x > 0$, allora se esiste una soluzione per la cotangente, ne esiste anche una per la tangente. Queste soluzioni presentano un andamento del tipo rappresentato in figura.



Se ora si manda l o V_0 a infinito, allora $\chi = \frac{1}{\hbar} \sqrt{2m(V_0 - E)}$ va a infinito, la tangente e la cotangente iperbolica tendono a 1 e gli stati sono degeneri. In particolare, se si manda V_0 a infinito, ci si ritrova con due buche di potenziale infinite affiancate, quindi la funzione d'onda in questo caso limite deve ridursi alla combinazione lineare delle funzioni d'onda di singola buca infinita. Denotate queste con $\varphi_0^I(x)$ e $\varphi_0^{II}(x)$ per la prima e seconda buca, si ha per lo stato fondamentale:²²

$$\lim_{V_0 \rightarrow \infty} \psi_0(x) \simeq \varphi_0^I(x) + \varphi_0^{II}(x)$$

e per il primo livello eccitato:

$$\lim_{V_0 \rightarrow \infty} \psi_1(x) \simeq \begin{cases} \varphi_0^I(x) - \varphi_0^{II}(x) \\ -\varphi_0^I(x) + \varphi_0^{II}(x) \end{cases}$$

Se lo stato del sistema è una combinazione lineare dello stato fondamentale e del primo livello eccitato, esplicitando la dipendenza temporale di questa funzione d'onda si ricava:

$$\psi(x, t) = \frac{1}{\sqrt{2}} \left[\psi_0(x) e^{\frac{iE_0 t}{\hbar}} + \psi_1(x) e^{\frac{iE_1 t}{\hbar}} \right] = \frac{e^{\frac{iE_0 t}{\hbar}}}{\sqrt{2}} \left[\psi_0(x) + \psi_1(x) e^{\frac{i(E_1 - E_0)t}{\hbar}} \right]$$

il tempo caratteristico per la transizione verso lo stato fondamentale è quindi dato da $t = \frac{\hbar\pi}{E_1 - E_0}$.

Si noti infine che la probabilità di effetto tunnel fra le due buche decade come e^{-2x} . Per sistemi macroscopici questa probabilità diventa comparabile con $e^{-\mathcal{N}x}$ (\mathcal{N} numero di Avogadro), quindi praticamente mai.

²²Lo stato fondamentale è sempre simmetrico.

Capitolo 4

La struttura formale della meccanica quantistica

4.1 La trasformata di Fourier. La funzione d'onda per l'impulso

Sia $\psi(\vec{x}, t)$ una funzione d'onda. La relazione:

$$\varphi(\vec{k}, t) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} \psi(\vec{x}, t) e^{-i\vec{k}\cdot\vec{x}} d\vec{x}$$

definisce la cosiddetta *trasformata di Fourier della* $\psi(\vec{x}, t)$.¹ Questa relazione può essere invertita:

$$\psi(\vec{x}, t) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} \varphi(\vec{k}, t) e^{i\vec{k}\cdot\vec{x}} d\vec{k}$$

che definisce l'*antitrasformata di Fourier della* $\varphi(\vec{k}, t)$.² Come si vede, le due trasformazioni sono l'una l'inversa dell'altra.

Questa coppia di trasformazioni è vera se una delle due funzioni appartiene a \mathcal{L}^2 , in questo caso si può dimostrare che anche l'altra appartiene a \mathcal{L}^2 . Risulta quindi:

$$\psi(\vec{x}, t): \mathbb{R}^3 \times \mathbb{R} \mapsto \mathbb{C} \qquad \varphi(\vec{k}, t): \mathbb{R}^3 \times \mathbb{R} \mapsto \mathbb{C}$$

La trasformata di Fourier si indica con $\varphi(\vec{k}) = \mathcal{F}[\psi(\vec{x}, t)]$. Una proprietà essenziale delle trasformate di Fourier è il cosiddetto *teorema di Parseval*:³

$$\int_{-\infty}^{+\infty} f^*(\vec{x}) \cdot g(\vec{x}) d\vec{x} = \int_{-\infty}^{+\infty} \mathcal{F}[f^*(\vec{x})] \cdot \mathcal{F}[g(\vec{x})] d\vec{x}$$

Siccome lo spazio \mathcal{L}^2 è dotato di una norma definita da:

$$\|\psi(\vec{x}, t)\| = \int_{-\infty}^{+\infty} \psi^*(\vec{x}, t) \psi(\vec{x}, t) d\vec{x}$$

¹La ragione per cui compare un fattore $1/(2\pi)^{3/2}$ invece del classico $1/\sqrt{2\pi}$ è che qui si è scritta la trasformata tridimensionale e non unidimensionale come si fa in genere.

²Si noti che in pratica questo significa dire ancora una volta che la funzione $\psi(\vec{x}, t)$ è scritta in termini di somma di onde monocromatiche, pesate dalla $\varphi(\vec{k}, t)$.

³In sostanza: la trasformata di una convoluzione è la convoluzione delle trasformate. Per la precisione, questo teorema è più propriamente detto *teorema di Plancherel*, ma in ambito scientifico il caso particolare sulle trasformate di Fourier è comunemente attribuito a Parseval.

se si pone nel teorema di Parseval $f(\vec{x}) = g(\vec{x}) = \psi(\vec{x})$, si ricava che *la norma è conservata dalla trasformata di Fourier*. Questa proprietà è importante per quanto segue.

Nell'interpretazione di de Broglie l'impulso di una particella è dato da $\vec{p} = \hbar\vec{k}$ e ad un numero d'onda definito (come nel caso delle onde monocromatiche) corrisponde un momento definito. In un pacchetto d'onda non si ha invece una sola onda monocromatica, ma una sovrapposizione: risulta pertanto naturale associare il termine $\varphi(\vec{k}, t)$, che pesa il contributo delle onde che costituiscono il pacchetto, alla probabilità di trovare una particella di impulso $\vec{p} = \hbar\vec{k}$. Di conseguenza, nell'interpretazione probabilistica si ha:

$$\mathcal{P}(\vec{p}, t) = \frac{1}{N^2} |\varphi(\vec{k})|^2$$

Il teorema di Parseval garantisce che la norma delle due distribuzioni di probabilità è la stessa e quindi la costante di normalizzazione è unica. Da notare inoltre che la funzione $\varphi(\vec{k})$ è completamente determinata dalla $\psi(\vec{x})$, quindi la funzione d'onda fornisce sia la probabilità della posizione che dell'impulso. La funzione d'onda $\psi(\vec{x})$ permette quindi di determinare completamente lo stato dinamico del sistema, fornendo sia la posizione che l'evoluzione nello spazio delle fasi. Si noti che mentre nel calcolo della probabilità della posizione il fattore di fase è ininfluente, questo non è vero per la probabilità dell'impulso:

$$|\psi(\vec{x})|^2 = |\psi(\vec{x})e^{i\alpha(\vec{x})}|^2 \quad \text{ma} \quad |\mathcal{F}[\psi(\vec{x})]|^2 \neq |\mathcal{F}[\psi(\vec{x})e^{i\alpha(\vec{x})}]|^2$$

dunque si può dire che *il modulo della funzione d'onda è legato alla posizione, la fase all'impulso*.

Si considerino ora le due relazioni sulla trasformata di Fourier:

$$\varphi(\vec{k}) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} \psi(\vec{x}) e^{-i\vec{k}\cdot\vec{x}} d\vec{x} \quad \psi(\vec{x}) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} \varphi(\vec{k}) e^{i\vec{k}\cdot\vec{x}} d\vec{k}$$

questo significa che nella definizione della $\psi(\vec{x})$ si può sostituire la $\varphi(\vec{k})$ con la sua definizione in termini di trasformata di Fourier. Considerando per semplicità di notazione il caso unidimensionale:

$$\psi(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \psi(y) e^{iky} dy e^{-ikx} dk = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \psi(y) \left[\int_{-\infty}^{+\infty} e^{ik(y-x)} dk \right] dy$$

dove lo scambio di integrazione è valido se tutte le funzioni sotto integrale appartengono a \mathcal{L}^2 . Deve quindi risultare:

$$\int_{-\infty}^{+\infty} e^{ik(y-x)} dk = \begin{cases} 0 & y \neq x \\ 1 & y = x \end{cases} \quad (4.1)$$

A rigore di termini una funzione che presenti questo comportamento non esiste, ma la si può definire come limite di una serie di funzioni. Per mostrare come si può fare, anche se in maniera non rigorosa, occorrerà fare una piccola digressione matematica.

4.2 La funzione delta di Dirac $\delta(x)$

Si supponga una successione definita nel seguente modo:

$$\{n_i\} \quad n_i = \sum_j n_j g_{ij} \quad \forall i$$

Questa può essere riscritta nel seguente modo:

$$n_i = n_i g_{ii} + \sum_{i \neq j} n_j g_{ij}$$

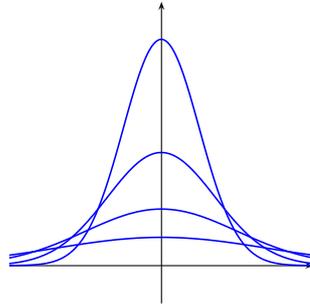
e l'unica soluzione possibile è che valga $g_{ii} = 1$ e $g_{ij} = 0$, deve cioè risultare $g_{ij} = \delta_{ij}$.

Se si passa a considerare degli indici continui, si vede che questa è esattamente la situazione ritrovata per la funzione $\psi(x)$ alla fine del paragrafo precedente. Ne consegue che deve essere verificata la (4.1). Una funzione che si comporti in questo modo è chiamata *delta di Dirac* ed è indicata con $\delta(y-x)$. Quindi, risulta:

$$\int_{-\infty}^{+\infty} e^{ik(y-x)} dk = 2\pi\delta(y-x)$$

Questo integrale definitorio in realtà non esiste. Infatti se deve valere 1 quando $y=x$ e zero altrove, questo significa che la funzione sotto integrando deve valere infinito per $y=x$.

La delta di Dirac deve pertanto essere intesa nel senso di limite di una successione di funzioni che si stringono sempre più intorno a $x=y$ in modo che l'area totale sia sempre 1.



Per avere un esempio concreto, si consideri la funzione gaussiana:

$$f_\alpha(x) = \frac{1}{\alpha} e^{-\frac{x^2}{4\alpha^2}} \quad \forall \alpha \neq 0$$

il cui integrale è dato da (ponendo $y \equiv x/2\alpha$):

$$\int_{-\infty}^{+\infty} f_\alpha(x) dx = \frac{1}{\alpha} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{4\alpha^2}} = 2 \int_{-\infty}^{+\infty} e^{-y^2} dy = 2\sqrt{\pi}$$

Risulta evidente che il valore dell'integrale è indipendente dal parametro α , per cui risulta verificato:

$$\int_{-\infty}^{+\infty} f_\alpha(x) = \frac{1}{2\sqrt{\pi}\alpha} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{4\alpha^2}} dx = 1$$

Risulta anche immediato verificare che:

$$\lim_{\alpha \rightarrow 0} f_\alpha(x) = \begin{cases} 0 & x \neq 0 \\ +\infty & x = 0 \end{cases}$$

e quindi per la definizione data sopra risulta proprio:

$$\lim_{\alpha \rightarrow 0} f_\alpha(x) = \delta(x)$$

Si noti che quella della gaussiana è solo una delle tante rappresentazioni possibili della delta di Dirac. In effetti, data una funzione qualsiasi $F(x)$ tale che $F(0) \neq 0$ e normalizzata tale che $\int_{-\infty}^{+\infty} F(x) dx = 1$, allora il limite:

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} F\left(\frac{x}{\varepsilon}\right) = \delta(x)$$

rappresenta una funzione delta.

Si consideri ora una generica funzione $g(x)$ e l'integrale:

$$\int_{-\infty}^{+\infty} g(x) f_\alpha(x) dx$$

se la funzione $g(x)$ è regolare nell'origine significa che si può espandere in serie di MacLaurin:

$$g(x) = g(0) + g'(0) \cdot x + g''(0) \cdot x^2 + O(3)$$

per cui l'integrale diviene:

$$g(0) \int_{-\infty}^{+\infty} f_{\alpha}(x) dx + g'(0) \int_{-\infty}^{+\infty} \frac{x}{2\sqrt{\pi} \alpha} e^{-\frac{x^2}{4\alpha^2}} dx + g''(0) \int_{-\infty}^{+\infty} \frac{x^2}{2\sqrt{\pi} \alpha} e^{-\frac{x^2}{4\alpha^2}} dx + O(3)$$

ora, i termini in x^n vanno a zero per $\alpha \rightarrow 0$:

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{x^n}{2\alpha} e^{-\frac{x^2}{4\alpha^2}} dx = \int_{-\infty}^{+\infty} (2\alpha y)^n e^{-y^2} dy = \frac{(2\alpha)^n}{\sqrt{\pi}} \int_{-\infty}^{+\infty} y^n e^{-y^2} dy$$

per cui se $g(x)$ è regolare nell'origine risulta effettivamente:

$$\int_{-\infty}^{+\infty} g(x) f_{\alpha}(x) dx = g(0) \int_{-\infty}^{+\infty} f_{\alpha}(x) dx = g(0)$$

ovvero:

$$\lim_{\alpha \rightarrow 0} \int_{-\infty}^{+\infty} g(x) f_{\alpha}(x) dx = \int_{-\infty}^{+\infty} g(x) \lim_{\alpha \rightarrow 0} f_{\alpha}(x) dx = \int_{-\infty}^{+\infty} g(x) \delta(0) dx = g(0)$$

pertanto, per poter "utilizzare" una funzione delta sotto integrale, la funzione a cui è associata deve essere regolare nel punto in cui la delta si annulla.

La delta è una funzione pari e soddisfa la proprietà:

$$\delta(ax) = \frac{1}{a} \delta(x)$$

Si calcoli ora la trasformata di Fourier di una gaussiana.

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\alpha^2 k^2} e^{ikx} dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\alpha^2 k^2 + ikx} dk$$

È possibile completare l'esponente come un quadrato perfetto:

$$\alpha^2 k^2 - ikx = \alpha^2 k^2 - ikx + \frac{x^2}{4\alpha^2} - \frac{x^2}{4\alpha^2} = \left(\alpha k - \frac{ix}{2\alpha} \right)^2 + \frac{x^2}{4\alpha^2}$$

per cui la trasformata di Fourier diventa:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\alpha^2 k^2} e^{ikx} dk = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{4\alpha^2}} \int_{-\infty}^{+\infty} e^{-(\alpha k - ix/2\alpha)^2} dk$$

ponendo ora $y = \alpha k - ix/2\alpha$, da cui $dk = \frac{1}{\alpha} dy$, l'integrale diventa:

$$\frac{1}{\sqrt{2\pi}} \frac{1}{\alpha} e^{-\frac{x^2}{4\alpha^2}} \int_{-\infty}^{+\infty} e^{-y^2} dy = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\pi}}{\alpha} e^{-\frac{x^2}{4\alpha^2}}$$

per cui:

$$\mathcal{F} \left[e^{-\alpha^2 k^2} \right] = \frac{\alpha}{\sqrt{2}} e^{-\frac{x^2}{4\alpha^2}}$$

In conclusione, la trasformata di Fourier di una gaussiana è ancora una gaussiana, ma le loro larghezze sono inversamente proporzionali.

4.3 Valore atteso di una grandezza. La rappresentazione operatoriale

Secondo l'interpretazione probabilistica della meccanica quantistica, se lo stato del sistema è descritto da una funzione d'onda $\psi(\vec{x}, t)$ significa che non si possono attribuire dei valori precisi alle grandezze fisiche, ma solo delle probabilità.

È pertanto importante introdurre il concetto di *valore medio*, che sarà indicato con $\langle x \rangle$. Questo è definito come il rapporto tra il numero di valori ottenuti in un insieme di misure e il numero totale delle misure. Una definizione alternativa, utile quando questo numero è molto grande, è data dalla somma dei valori ottenuti moltiplicati per le rispettive probabilità.

Il valore medio di una grandezza fisica è quindi dato da:⁴

$$\langle x \rangle = \sum_i \mathcal{P}(x_i) x_i$$

cioè il valore atteso è dato dalla media dei valori possibili x_i pesati con la probabilità $\mathcal{P}(x_i)$ di ottenere quel determinato valore. Nel caso di una variabile continua, la probabilità di ottenere un determinato valore è una funzione continua e il valore medio assume la forma:

$$\langle x \rangle = \int_{-\infty}^{+\infty} x \mathcal{P}(x) dx$$

Nel caso di una particella quantistica, la probabilità di trovare la particella nella regione dx è data da $|\psi(x)|^2 dx$, per cui il valore medio della coordinata x_i è dato da:

$$\langle x_i \rangle = \int_{-\infty}^{+\infty} \psi^*(\vec{x}, t) x_i \psi(\vec{x}, t) d\vec{x}$$

Per l'impulso vale naturalmente una relazione analoga, ovvero:

$$\langle p_i \rangle = \int_{-\infty}^{+\infty} \varphi^*(\vec{k}, t) p_i \varphi(\vec{k}, t) d\vec{k} = \hbar \int_{-\infty}^{+\infty} \varphi^*(\vec{k}, t) k_i \varphi(\vec{k}, t) d\vec{k}$$

Si trascuri per ora la dipendenza temporale e si ponga $k_i \varphi(\vec{k}, t) \equiv H(\vec{k})$. Dal teorema di Parseval risulta allora:

$$\int_{-\infty}^{+\infty} \varphi^*(\vec{k}) H(\vec{k}) d\vec{k} = \int_{-\infty}^{+\infty} \mathcal{F}[\varphi^*(\vec{k})] \mathcal{F}[H(\vec{k})] d\vec{x}$$

scrivendo esplicitamente le trasformate (trascurando i fattori costanti che in questo contesto non interessano e indicando con $h(\vec{y})$ la trasformata di $H(\vec{k})$):

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \psi^*(\vec{x}) e^{i\vec{k}\cdot\vec{x}} d\vec{x} \int_{-\infty}^{+\infty} h(\vec{y}) e^{-i\vec{k}\cdot\vec{y}} d\vec{y} d\vec{k} = \int_{-\infty}^{+\infty} \psi^*(\vec{x}) \int_{-\infty}^{+\infty} h(\vec{y}) \int_{-\infty}^{+\infty} e^{i\vec{k}\cdot(\vec{x}-\vec{y})} d\vec{k} d\vec{y} d\vec{x}$$

dove lo scambio di integrazione è lecito se le funzioni appartengono a \mathcal{L}^2 . Come visto:

$$\int_{-\infty}^{+\infty} e^{i\vec{k}\cdot(\vec{x}-\vec{y})} d\vec{k} = \delta(\vec{x} - \vec{y})$$

per cui:

$$\int_{-\infty}^{+\infty} h(\vec{y}) \int_{-\infty}^{+\infty} e^{i\vec{k}\cdot(\vec{x}-\vec{y})} d\vec{k} d\vec{y} = \int_{-\infty}^{+\infty} h(\vec{y}) \delta(\vec{x} - \vec{y}) d\vec{y} = h(\vec{x})$$

ne consegue allora:

$$\int_{-\infty}^{+\infty} \varphi^*(\vec{k}) H(\vec{k}) d\vec{k} = \int_{-\infty}^{+\infty} \psi^*(\vec{x}) h(\vec{x}) d\vec{x}$$

⁴Per la precisione non si parla qui di *valore medio* nel senso di media statistica, ma di *valore di aspettazione*, detto anche *valore atteso*. Spesso però qui si userà la dizione "valore medio".

Ricordando ora la definizione di $H(\vec{k})$:

$$H(\vec{k}) = k_i \varphi(\vec{k}) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} \psi(\vec{x}) k_i e^{-i\vec{k}\cdot\vec{x}} d\vec{x}$$

e notando che risulta:

$$k_i e^{-i\vec{k}\cdot\vec{x}} = i \frac{\partial}{\partial x_i} e^{-i\vec{k}\cdot\vec{x}}$$

si ricava:

$$k_i \varphi(\vec{k}) = \int_{-\infty}^{+\infty} \psi(\vec{x}) i \frac{\partial}{\partial x_i} e^{-i\vec{k}\cdot\vec{x}} d\vec{x}$$

che integrata a sua volta per parti fornisce:

$$k_i \varphi(\vec{k}) = \left[i \psi(\vec{x}) e^{-i\vec{k}\cdot\vec{x}} \right]_{-\infty}^{+\infty} - i \int_{-\infty}^{+\infty} \frac{\partial}{\partial x_i} \psi(\vec{x}) e^{-i\vec{k}\cdot\vec{x}} d\vec{x}$$

il termine è nullo perché $\psi(\vec{x}) \in \mathcal{L}^2$ e quindi si annulla all'infinito. Siccome vale anche:

$$k_i \varphi(\vec{k}) = \int_{-\infty}^{+\infty} h(\vec{x}) e^{-i\vec{k}\cdot\vec{x}} d\vec{x}$$

ne segue:

$$h(\vec{x}) = -i \frac{\partial}{\partial x_i} \psi(\vec{x})$$

moltiplicare la $\varphi(\vec{x})$ per k_i equivale quindi a derivare la $\psi(\vec{x})$ rispetto a x_i :

$$\langle p_i \rangle = \int_{-\infty}^{+\infty} \varphi^*(\vec{k}) k_i \varphi(\vec{k}) d\vec{k} = \int_{-\infty}^{+\infty} \psi^*(\vec{x}) \left(-i \frac{\partial}{\partial x_i} \right) \psi(\vec{x}) d\vec{x}$$

Alla variabile p_i corrisponde l'operatore derivata rispetto a x_i nello spazio delle x . Si preciserà ora questa affermazione.

Si è visto che descrivere il sistema in termini della $\psi(\vec{x})$ o in termini della $\varphi(\vec{k})$ è equivalente, dal momento che dall'una è possibile ricavare l'altra tramite la trasformata di Fourier. Si parla di *rappresentazione delle coordinate o delle \vec{x}* nel caso si utilizzino le $\psi(\vec{x})$ e di *rappresentazione degli impulsi o delle \vec{p}* (o delle \vec{k}) nel caso si utilizzino le $\varphi(\vec{k})$.

Si noti qui un punto importante: *il valore medio della posizione o dell'impulso⁵ è rappresentato matematicamente da un operatore lineare applicato alla funzione d'onda*. Il valore medio dell'impulso è rappresentato dall'operatore moltiplicazione \hat{k}_i nella rappresentazione degli impulsi e dalla derivazione rispetto a x_i nella rappresentazione delle coordinate. Un discorso analogo può essere fatto per il valore medio della posizione, giungendo alla conclusione:

<i>grandezza</i>	<i>operatore</i>	
x_i	\rightarrow	\hat{x}_i
		{
		x_i nella rappresentazione delle \vec{x}
		$i\hbar \frac{\partial}{\partial k_i}$ nella rappresentazione delle \vec{k}
p_i	\rightarrow	\hat{p}_i
		{
		$-i\hbar \frac{\partial}{\partial x_i}$ nella rappresentazione delle \vec{x}
		k_i nella rappresentazione delle \vec{k}

⁵E di qualunque altra variabile, come si vedrà più avanti.

4.4 Il principio di indeterminazione di Heisemberg

Si denotino con x_i una delle n coordinate e con p_j gli impulsi. Si lavori nella rappresentazione delle coordinate.

Siccome gli operatori che rappresentano matematicamente una misura di posizione o di impulso sono lineari, la quantità $x_i p_j$ rappresenta una *misura di impulso lungo la coordinata j seguita da una misura di posizione lungo la coordinata i* . Il valore medio di questa quantità è dato da:

$$\langle x_i p_j \rangle = \int_{-\infty}^{+\infty} \psi^*(x) \hat{x}_i \hat{p}_j \psi(x) d^n x = -i\hbar \int_{-\infty}^{+\infty} \psi^*(x) x_i \frac{\partial}{\partial x_j} \psi(x) d^n x$$

analogamente:

$$\begin{aligned} \langle p_j x_i \rangle &= \int_{-\infty}^{+\infty} \psi^*(x) \hat{p}_j \hat{x}_i \psi(x) d^n x = -i\hbar \int_{-\infty}^{+\infty} \psi^*(x) \frac{\partial}{\partial x_j} [x_i \psi(x)] d^n x = \\ &= -i\hbar \int_{-\infty}^{+\infty} \psi^*(x) \left[\delta_{ij} \psi(x) + x_i \frac{\partial}{\partial x_j} \psi(x) \right] d^n x \end{aligned}$$

si vede quindi che $\langle x_i p_j \rangle \neq \langle p_j x_i \rangle$, in particolare:

$$\langle x_i p_j \rangle - \langle p_j x_i \rangle = i\hbar \delta_{ij} \quad \forall \psi(x) \in \mathcal{L}^2$$

essendo gli operatori lineari $\langle x_i p_j \rangle - \langle p_j x_i \rangle = \langle x_i p_j - p_j x_i \rangle$. Questo implica che gli operatori stessi devono soddisfare la relazione:

$$x_i p_j - p_j x_i = i\hbar \delta_{ij}$$

Introducendo il concetto di *commutatore* di due operatori, definito come $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$, si trova:

$$\boxed{[\hat{q}_i, \hat{p}_j] = i\hbar \delta_{ij}}$$

La posizione e la quantità di moto sono descritti da operatori che non commutano: concretamente, questo significa che non è la stessa cosa misurare prima una grandezza e poi l'altra. L'ordine in cui sono effettuate le misure è importante.

Questa relazione si traduce nel cosiddetto **principio di indeterminazione di Heisemberg**.

Si consideri la quantità:

$$\frac{1}{N} \int_{-\infty}^{+\infty} \left| \left(x - \lambda \hbar \frac{\partial}{\partial x} \right) \psi(x) \right|^2 dx \geq 0 \quad \forall \lambda \in \mathbb{R}$$

esplicitando il modulo si ricava:

$$\begin{aligned} &= \int_{-\infty}^{+\infty} \left(x - \lambda \hbar \frac{\partial}{\partial x} \right) \psi^*(x) \left(x - \lambda \hbar \frac{\partial}{\partial x} \right) \psi(x) dx = \\ &= \int_{-\infty}^{+\infty} \left[x \psi^*(x) - \lambda \hbar \frac{\partial}{\partial x} \psi^*(x) \right] \left[x \psi(x) - \lambda \hbar \frac{\partial}{\partial x} \psi(x) \right] dx = \\ &= \int_{-\infty}^{+\infty} \left\{ \psi^*(x) x^2 \psi(x) - \lambda \hbar \left[\psi^*(x) x \frac{\partial \psi(x)}{\partial x} + \frac{\partial \psi^*(x)}{\partial x} x \psi(x) \right] + \lambda^2 \hbar^2 \frac{\partial}{\partial x} \psi^*(x) \frac{\partial}{\partial x} \psi(x) \right\} dx \end{aligned}$$

il secondo pezzo fra parentesi quadre può essere integrato per parti:

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{\partial \psi^*(x)}{\partial x} x \psi(x) dx &= \cancel{\psi^*(x) x \psi(x)} \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \psi^*(x) \frac{\partial}{\partial x} [x \psi(x)] dx = \\ &= - \int_{-\infty}^{+\infty} \psi^*(x) x \frac{\partial}{\partial x} \psi(x) dx - \int_{-\infty}^{+\infty} \psi^*(x) \psi(x) dx \end{aligned}$$

Allo stesso modo, si integra per parti l'ultimo termine:

$$\int_{-\infty}^{+\infty} \frac{\partial}{\partial x} \psi^*(x) \frac{\partial}{\partial x} \psi(x) dx = \psi^*(x) \frac{\partial}{\partial x} \psi(x) \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \psi^*(x) \frac{\partial^2}{\partial x^2} \psi(x) dx$$

i termini per parti si annullano naturalmente perché le funzioni d'onda appartengono a \mathcal{L}^2 . Si ottiene quindi:

$$\frac{1}{N} \int_{-\infty}^{+\infty} \left[\psi^*(x) x^2 \psi(x) + \lambda \hbar \psi^*(x) \psi(x) - \lambda^2 \hbar^2 \psi^*(x) \frac{\partial^2}{\partial x^2} \psi(x) \right] dx \geq 0$$

Il termine $\psi^*(x) x^2 \psi(x)$ rappresenta il valore medio di x^2 per definizione, $\lambda \hbar \psi^*(x) \psi(x)$ dà $\lambda \hbar$ perché le funzioni d'onda sono normalizzate. Per l'altro termine invece:

$$\begin{aligned} -\lambda^2 \hbar^2 \psi^*(x) \frac{\partial^2}{\partial x^2} \psi(x) &= \lambda^2 \hbar^2 (i)^2 \psi^*(x) \frac{\partial^2}{\partial x^2} \psi(x) = \lambda^2 \left[(i)^2 \hbar^2 \psi^*(x) \frac{\partial^2}{\partial x^2} \psi(x) \right] = \\ &= \lambda^2 \psi^*(x) \left(i \hbar \frac{\partial}{\partial x} \right)^2 \psi(x) \equiv \lambda^2 \hat{p}^2 \end{aligned}$$

dove nell'ultimo passaggio il quadrato deve essere inteso nel senso di "operatore applicato due volte". In definitiva si ha:

$$\langle x^2 \rangle + \lambda \hbar + \lambda^2 \langle p^2 \rangle \geq 0 \quad \forall \lambda \in \mathbb{R}$$

Questa relazione rappresenta una disequazione di secondo grado in λ , ovvero una parabola. Perché la disequazione sia soddisfatta deve accadere che non esistano radici reali, per cui il determinante deve essere minore di zero:

$$\Delta \equiv \hbar^2 - 4 \langle x^2 \rangle \langle p^2 \rangle \leq 0$$

deve quindi risultare:

$$\langle x^2 \rangle \langle p^2 \rangle \geq \frac{\hbar^2}{4}$$

Ora, lo scarto quadratico medio è dato da $x - \langle x \rangle$ e $p - \langle p \rangle$, per cui $(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2$ e $(\Delta p)^2 = \langle p^2 \rangle - \langle p \rangle^2$. Se ci si mette nel sistema di riferimento in cui $\langle x \rangle$ e $\langle p \rangle$ sono zero, e notando che il valore dello scarto quadratico medio non dipende dal sistema di riferimento, si trova infine:

$$\boxed{\Delta x \Delta p \geq \frac{\hbar}{2} \quad \text{Principio di Heisemberg}}$$

Questo principio, che va sotto il nome di **principio di indeterminazione di Heisemberg**, è fondamentale nella meccanica quantistica e stabilisce che *non è possibile attribuire ad uno stato fisico contemporaneamente un valore preciso di posizione e quantità di moto*. Il principio non deve essere interpretato in senso statistico, esso si riferisce infatti allo stesso sistema fisico. Si noti però che nel momento in cui si decide di fare una misura di posizione con precisione arbitraria e subito dopo una misura di momento con precisione arbitraria, si stanno di fatto compiendo due misure su due sistemi differenti, in quanto la prima misura di posizione perturba necessariamente il sistema in modo tale che al momento della seconda misura questo non sia più uguale al primo sistema. *Il principio vieta invece di condurre una misura di posizione e quantità contemporaneamente e con precisione arbitraria*. Il principio di indeterminazione sancisce quindi l'indeterminismo del comportamento degli enti appartenenti al mondo subatomico, che si sottraggono totalmente ai criteri della fisica classica.

Vale la pena di notare esplicitamente che il principio di indeterminazione di Heisemberg discende direttamente dal fatto che la posizione e la quantità di moto sono descritti da operatori che non commutano. Questo punto importante sarà approfondito in seguito.

4.5 Operatori Hermitiani e osservabili. L'apparato formale della teoria

Si consideri l'operatore associato alla quantità di moto nella rappresentazione delle coordinate:

$$p \quad \rightarrow \quad \hat{p} = -i\hbar \frac{\partial}{\partial x}$$

Il valore medio dell'impulso è dato quindi da:

$$\langle p \rangle = \int_{-\infty}^{+\infty} \psi^*(x) [\hat{p}\psi(x)] dx = \int_{-\infty}^{+\infty} \psi^*(x) (-i\hbar) \frac{\partial}{\partial x} \psi(x) dx$$

integrando per parti questa relazione si trova:

$$\begin{aligned} \int_{-\infty}^{+\infty} \psi^*(x) (-i\hbar) \frac{\partial}{\partial x} \psi(x) dx &= \cancel{-i\hbar \psi^*(x) \psi(x) \Big|_{-\infty}^{+\infty}} + i\hbar \int_{-\infty}^{+\infty} \frac{\partial}{\partial x} \psi^*(x) \cdot \psi(x) dx = \\ &= \int_{-\infty}^{+\infty} \left[-i\hbar \frac{\partial}{\partial x} \psi(x) \right]^* \psi(x) dx = \int_{-\infty}^{+\infty} [\hat{p}\psi(x)]^* \psi(x) dx \end{aligned}$$

ovvero:

$$\int_{-\infty}^{+\infty} \psi^*(x) [\hat{p}\psi(x)] dx = \int_{-\infty}^{+\infty} [\hat{p}\psi(x)]^* \psi(x) dx \quad \forall \psi(x) \in \mathcal{L}^2$$

Si consideri un generico operatore \hat{A} , risulta per definizione di *operatore aggiunto*:⁶

$$\int_{-\infty}^{+\infty} \psi^*(x) \hat{A}\psi(x) dx = \int_{-\infty}^{+\infty} [\hat{A}^\dagger \psi^*(x)] \psi(x) dx$$

Un operatore \hat{A} che goda della proprietà di coincidere con il suo aggiunto $\hat{A} = \hat{A}^\dagger$, è definito *autoaggiunto* o *hermitiano*.⁷ L'operatore \hat{p} , e analogamente l'operatore \hat{x} , sono dunque *hermitiani*.

Si considerino due operatori \hat{A} e \hat{B} :

$$\int_{-\infty}^{+\infty} \psi^*(x) \hat{A}\hat{B}\psi(x) dx = \int_{-\infty}^{+\infty} \hat{A}^\dagger \psi^*(x) \hat{B}\psi(x) dx = \int_{-\infty}^{+\infty} \hat{B}^\dagger \hat{A}^\dagger \psi^*(x) \psi(x) dx$$

da cui:

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger \neq \hat{A}^\dagger \hat{B}^\dagger \quad (4.2)$$

ne consegue che la combinazione di due operatori $\hat{A}\hat{B}$ è hermitiana solo se \hat{A} e \hat{B} commutano oltre che essere hermitiani.

Si noti a questo punto che una funzione d'onda $\psi(x)$ determina univocamente uno stato fisico, ma uno stato fisico non determina univocamente una funzione d'onda $\psi(x)$. Infatti, se $\psi(x)$ determina uno stato s , allora anche $\lambda\psi(x)$, $\lambda \neq 0 \in \mathbb{C}$ determina lo stesso stato. Si parla in questo caso di una *corrispondenza proiettiva*.

Nello spazio del sistema è definibile un prodotto scalare:

$$\langle \varphi(x) | \psi(x) \rangle = \int_{-\infty}^{+\infty} \varphi^*(x) \psi(x) dx$$

che permette immediatamente di introdurre una norma definita positiva:

$$\|\psi(x)\| = \langle \psi(x) | \psi(x) \rangle = \int_{-\infty}^{+\infty} \psi^*(x) \psi(x) dx \geq 0$$

In questi termini, il valore medio di un operatore sullo stato $\psi(x)$ è dato da:

$$\langle \hat{A} \rangle_\psi = \frac{\langle \psi(x) | \hat{A}\psi(x) \rangle}{\|\psi(x)\|}$$

L'insieme delle funzioni $\psi(x) \in \mathcal{L}^2$ costituisce dunque uno spazio lineare, denominato *spazio di Hilbert*.⁸

⁶L'aggiunto di un operatore è la generalizzazione del concetto di trasposto coniugato nell'ambito delle matrici e si indica con \hat{A}^\dagger .

⁷E la matrice che li rappresenta è simmetrica e reale.

⁸Questo discorso era stato anticipato nell'inserto in §3.3.

Si ricaveranno ora alcune proprietà estremamente importanti degli operatori associati alle grandezze fisiche.

Se ripetendo più volte una misura si ottiene sempre lo stesso risultato, significa che lo scarto quadratico deve essere nullo: $\langle \hat{A} \rangle$ è certamente $a \Leftrightarrow (\Delta \hat{A})^2 = 0$. Quindi:

$$(\Delta \hat{A})^2 = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2 = \frac{\langle \psi(x) | \hat{A}^2 \psi(x) \rangle}{\|\psi(x)\|} - \frac{\langle \psi(x) | \hat{A} \psi(x) \rangle^2}{\|\psi(x)\|^2} = 0$$

ovvero:

$$\|\psi(x)\| \langle \psi(x) | \hat{A}^2 \psi(x) \rangle = \langle \psi(x) | \hat{A} \psi(x) \rangle^2$$

nell'ipotesi che \hat{A} sia hermitiano si ha $\langle \psi(x) | \hat{A}^2 \psi(x) \rangle = \langle \hat{A} \psi(x) | \hat{A} \psi(x) \rangle = \|\hat{A} \psi(x)\|^2$. Quindi:

$$\|\psi(x)\| \cdot \|\hat{A} \psi(x)\|^2 = \langle \psi(x) | \hat{A} \psi(x) \rangle^2 = |\psi(x)|^2 \cdot |\hat{A} \psi(x)|^2 \cos^2 \theta = \|\psi(x)\| \cdot \|\hat{A} \psi(x)\| \cos^2 \theta$$

il che implica $\cos^2 \theta = 1$,⁹ ovvero $\hat{A} \psi(x)$ deve essere parallelo a $\psi(x)$. Ma per essere parallelo deve risultare $\hat{A} \psi(x) \propto \psi(x)$, ed in particolare:

$$\boxed{\hat{A} \psi(x) = a \psi(x)}$$

cioè se una misura fornisce con certezza un risultato a allora la funzione d'onda $\psi(x)$ deve essere un autostato dell'operatore \hat{A} con autovalore a . Siccome l'autovalore rappresenta il risultato di una misura fisica, esso deve essere un numero reale $a \in \mathbb{R}$ e questo implica che l'operatore sia hermitiano:

$$a^* \langle \psi(x) | \psi(x) \rangle = \langle \hat{A}^\dagger \psi(x) | \psi(x) \rangle = \langle \psi(x) | \hat{A} \psi(x) \rangle = a \langle \psi(x) | \psi(x) \rangle$$

ovvero se deve essere $a^* = a$, allora \hat{A} deve essere hermitiano. Ad una grandezza fisica misurabile è associato un operatore hermitiano detto **osservabile** i cui autovalori (spettro dell'operatore) rappresentano i possibili esiti della misura. Tuttavia, una volta fissato il valore dell'autovalore a , in genere lo stato $\psi(x)$ non è ben determinato, infatti possono esistere diversi stati (vettori nello spazio di Hilbert) linearmente indipendenti relativi allo stesso autovalore:

$$\exists \left\{ \psi_a^{(n)}(x) \right\} \in \mathcal{S}_a: \lambda_1 \psi_a^{(1)}(x) + \dots + \lambda_n \psi_a^{(n)}(x) = 0 \Leftrightarrow (\lambda_1, \dots, \lambda_n) = \vec{0}$$

in questo caso l'operatore \hat{A} è detto essere *degenere*.

Si considerino ora due autostati dell'operatore hermitiano \hat{A} relativi a due autovalori distinti:

$$\hat{A} \psi_1(x) = a_1 \psi_1(x) \quad \hat{A} \psi_2(x) = a_2 \psi_2(x) \quad a_1 \neq a_2$$

moltiplicando scalarmente a sinistra le due relazioni rispettivamente per $\psi_2(x)$ e $\psi_1(x)$:

$$\begin{aligned} \langle \psi_2(x) | \hat{A} \psi_1(x) \rangle &= a_1 \langle \psi_2(x) | \psi_1(x) \rangle \\ \langle \psi_1(x) | \hat{A} \psi_2(x) \rangle &= a_2 \langle \psi_1(x) | \psi_2(x) \rangle \end{aligned}$$

il complesso coniugato della seconda relazione fornisce ($a_2^* = a_2$ e $\hat{A}^\dagger = \hat{A}$):

$$\langle \hat{A} \psi_2(x) | \psi_1(x) \rangle = a_2 \langle \psi_2(x) | \psi_1(x) \rangle$$

ma ancora, grazie all'hermiticità di \hat{A} , $\langle \hat{A} \psi_2(x) | \psi_1(x) \rangle = \langle \psi_2(x) | \hat{A} \psi_1(x) \rangle$ e sottraendo le due equazioni si ritrova:

$$(a_1 - a_2) \langle \psi_2(x) | \psi_1(x) \rangle = 0$$

e siccome per ipotesi gli autovalori sono distinti, deve risultare $\langle \psi_2(x) | \psi_1(x) \rangle = 0$, ovvero *autofunzioni hermitiane relative ad autovalori distinti sono ortogonali*. Se invece l'operatore \hat{A} è degenere,

⁹Generalmente risulta $\langle \varphi | \varphi \rangle \langle \varphi | \psi \rangle \geq |\langle \varphi | \psi \rangle|^2$

le autofunzioni non sono necessariamente ortogonali anche se sono indipendenti. Possono però essere ortogonalizzate, ad esempio con il procedimento di ortonormalizzazione Gram-Schmidt:¹⁰

$$\begin{cases} \tilde{\psi}_1(x) = \psi_1(x) \\ \tilde{\psi}_2(x) = \psi_2(x) - \frac{\langle \psi_2(x) | \psi_1(x) \rangle}{\|\psi_1(x)\|} \psi_1(x) \\ \dots \\ \tilde{\psi}_n(x) = \psi_n(x) - \sum_{i=1}^{n-1} \frac{\langle \psi_n(x) | \psi_i(x) \rangle}{\|\psi_i(x)\|} \psi_i(x) \end{cases}$$

Naturalmente, lo stato di un sistema non è necessariamente un autostato di un osservabile \hat{A} . Tuttavia, scelta una base di autofunzioni relative agli autovalori a_1, \dots, a_n dell'operatore \hat{A} , lo stato del sistema sarà sicuramente esprimibile come combinazione lineare di autostati:¹¹

$$\psi(x) = \sum_{i=1}^n c_i \varphi_i(x)$$

Si supponga ora per semplicità avere una base di due autofunzioni $\varphi_1(x)$ e $\varphi_2(x)$ e che quindi lo stato sia espresso dalla combinazione lineare $\psi(x) = c_1 \varphi_1(x) + c_2 \varphi_2(x)$. Questa combinazione lineare non è naturalmente essa stessa un'autofunzione:¹²

$$\hat{A}\psi(x) = a_1 c_1 \varphi_1(x) + a_2 c_2 \varphi_2(x) \neq \lambda \psi(x)$$

siccome però lo stato deve essere normalizzato, deve risultare:

$$\|\psi(x)\| = 1 \Leftrightarrow |c_1|^2 + |c_2|^2 = 1$$

Il valore medio dell'osservabile \hat{A} in questo stato normalizzato è dato da:

$$\begin{aligned} \langle \hat{A} \rangle_\psi &= \langle \psi(x) | \hat{A} \psi(x) \rangle = \\ &= \langle c_1 \varphi_1(x) + c_2 \varphi_2(x) | \hat{A} (c_1 \varphi_1(x) + c_2 \varphi_2(x)) \rangle = \\ &= \langle c_1 \varphi_1(x) + c_2 \varphi_2(x) | a_1 c_1 \varphi_1(x) + a_2 c_2 \varphi_2(x) \rangle = \\ &= a_1 |c_1|^2 + a_2 |c_2|^2 \end{aligned}$$

tenendo conto dell'ortonormalità degli stati. Ma il valore medio di un operatore è dato dalla media dei possibili valori λ_i pesata per le probabilità di ottenere un determinato valore $\langle \hat{A} \rangle = \sum_i \lambda_i \mathcal{P}(\lambda_i)$, ne consegue:

$$\begin{cases} \lambda_1 \equiv a_1 & \mathcal{P}(a_1) \equiv |c_1|^2 \\ \lambda_2 \equiv a_2 & \mathcal{P}(a_2) \equiv |c_2|^2 \end{cases}$$

ovvero *i possibili risultati di una misura descritta dell'osservabile \hat{A} sono dati dai suoi autovalori, i moduli quadri dei coefficienti dello sviluppo della funzione d'onda $\psi(x)$ in una base ortonormale di autostati $\{\varphi_i(x)\}$ relative all'osservabile \hat{A} rappresentano invece le probabilità $\mathcal{P}(a_i)$ di ottenere l'autovalore a_i come risultato di una misura.*

Se lo stato del sistema non è scrivibile come combinazione lineare di autostati dell'osservabile $\psi(x) = \sum_i c_i \varphi_i(x)$, allora si **postula** che:

$$\{\varphi_i(x)\} \text{ è completo in } \mathcal{L}^2$$

¹⁰Questo procedimento consiste in pratica nel sottrarre da ogni vettore di base la proiezione su tutti gli altri vettori della base.

¹¹Al limite, il numero di autofunzioni nello sviluppo può essere infinito, se la dimensione dello spazio vettoriale è infinita.

¹²Tranne nel caso in cui i due autovalori siano identici, ovvero nella base del sottospazio degenere relativo all'autovalore a .

ed in questo senso l'uguaglianza dello sviluppo non deve essere intesa in modo puntuale, ma nel senso di media quadratica:

$$\forall \psi(x) \in \mathcal{L}^2 \quad \lim_{N \rightarrow +\infty} \left\| \psi(x) - \sum_{n=1}^N c_n \varphi_n(x) \right\| = 0$$

Si supponga assegnato lo stato $\psi(x) = \sum_j c_j \varphi_j(x)$. Risulta allora:

$$\langle \varphi_i(x) | \psi(x) \rangle = \sum_{j=1}^N c_j \langle \varphi_i(x) | \varphi_j(x) \rangle = \sum_{j=1}^N c_j \delta_{ij} = c_i$$

ovvero il prodotto scalare della funzione d'onda rappresentante lo stato e un vettore della base ortonormale fornisce il relativo coefficiente dello sviluppo nella detta base. In sostanza:

$$c_i = \langle \varphi_i(x) | \psi(x) \rangle = \int_{-\infty}^{+\infty} \varphi_i^*(x) \psi(x) dx$$

Se l'insieme è completo, da questa relazione segue immediatamente:¹³

$$\psi(x) = \sum_n c_n \varphi_n(x) = \sum_n \int_{-\infty}^{+\infty} \varphi_n^*(y) \psi(y) dy \cdot \varphi_n(x) = \int_{-\infty}^{+\infty} \left(\sum_n \varphi_n^*(y) \varphi_n(x) \right) \psi(y) dy \quad \forall x, \psi(x)$$

da cui la:

$$\sum_n \varphi_n^*(y) \varphi_n(x) = \delta(x - y) \quad \text{Criterio di completezza}$$

sotto la condizione di ortonormalità:

$$\int_{-\infty}^{+\infty} \varphi_n^*(x) \varphi_m(x) = \delta_{mn}$$

Il postulato fondamentale è quindi che *ad ogni osservabile fisico corrisponde un insieme di autofunzioni completo e ortonormale.*

Si vedrà ora in quali condizioni due osservabili possono avere entrambi un valore preciso su un determinato stato $\psi(x)$. Se due osservabili hanno un valore preciso, significa evidentemente che la funzione d'onda deve essere un autostato simultaneo di entrambi gli osservabili:

$$\begin{cases} \hat{A}\psi(x) = a\psi(x) \\ \hat{B}\psi(x) = b\psi(x) \end{cases}$$

operando rispettivamente con \hat{B} e \hat{A} :

$$\begin{cases} \hat{B}\hat{A}\psi(x) = a\hat{B}\psi(x) = ab\psi(x) \\ \hat{A}\hat{B}\psi(x) = b\hat{A}\psi(x) = ba\psi(x) \end{cases}$$

sottraendo membro a membro e considerando che a e b sono numeri:

$$(\hat{A}\hat{B} - \hat{B}\hat{A})\psi(x) = 0$$

ovvero:

$$[\hat{A}, \hat{B}]\psi(x) = 0$$

¹³Si lascia volontariamente non specificato l'estremo superiore della sommatoria, in quanto questa può essere finita o no.

quindi due osservabili possono avere un valore definito su uno stato se i due operatori commutano. In altre parole, se due osservabili commutano, uno stato può essere autofunzione contemporaneamente dei due operatori. Si noti che questa è una condizione necessaria, ma non strettamente sufficiente.

Si supponga infatti che sia verificato:

$$\hat{A}\psi(x) = a\psi(x) \quad [\hat{A}, \hat{B}]\psi(x) = 0$$

dalla relazione sul commutatore si deduce che:

$$(\hat{A}\hat{B} - \hat{B}\hat{A})\psi(x) = (\hat{A}\hat{B} - \hat{B}a)\psi(x) = 0$$

dunque:

$$\hat{A}[\hat{B}\psi(x)] = a[\hat{B}\psi(x)]$$

che ammette come soluzioni $\hat{B}\psi(x) = 0$ e $\hat{B}\psi(x) \in S_a$, ovvero che $\psi(x)$ è autostato di \hat{B} con autovalore 0 oppure $\hat{B}\psi(x)$ appartiene al sottospazio di \hat{A} dell'autovalore a . Quindi, se \hat{A} e \hat{B} commutano, non si può dire con sicurezza se $\psi(x)$ sia o meno autostato di \hat{B} , ma certamente $\hat{B}\psi(x)$ appartiene al sottospazio relativo a a . Nel caso S_a sia non degenere, allora il sottospazio ha dimensione uno e $\psi(x)$ deve essere anche autostato di \hat{B} .

Sull'operatore parità e l'Hamiltoniana

Ogni autostato dell'Hamiltoniana è anche un autostato dell'operatore parità \hat{P} , poiché risulta:

$$[\hat{P}, H] = 0$$

l'operatore parità è però degenere in quanto $\hat{P}^2 = \mathbb{I}$ e $p \pm 1$, dunque non necessariamente un autostato di \hat{P} è anche un autostato dell'Hamiltoniana.

Anche se i due operatori \hat{A} e \hat{B} commutano, un autostato dell'uno non è necessariamente un autostato dell'altro. Essi hanno però un sistema di autofunzioni in comune. Si supponga:

$$\hat{A}\psi_a(x) = a\psi_a(x) \quad \hat{B}\varphi_{a,b_m}(x) = b_m\varphi_{a,b_m}(x)$$

e si scriva lo sviluppo dell'autofunzione di \hat{A} in termini di autofunzioni di \hat{B} :

$$\psi_a(x) = \sum_m \varphi_{a,b_m}(x)$$

e si consideri la funzione $\chi_m(x) = (\hat{A} - a)\varphi_{a,b_m}(x)$. Queste funzioni sono autofunzioni di \hat{B} relative ad autovalori distinti, infatti $\hat{B}\chi_m(x) = (\hat{A} - a)\hat{B}\varphi_{a,b_m}(x) = b_m(\hat{A} - a)\varphi_{a,b_m}(x)$. Si consideri dunque la somma delle $\chi_m(x)$:

$$\sum_m \chi_m(x) = (\hat{A} - a) \sum_m \varphi_{a,b_m}(x) = (\hat{A} - a)\psi_a(x) = 0$$

ma questo è impossibile perché le $\chi_m(x)$ appartengono ad autovalori distinti, per cui deve essere $\chi_m(x) = 0$, ovvero:

$$\chi_m(x) = 0 \quad \Rightarrow \quad (\hat{A} - a)\varphi_{a,b_m}(x) = 0 \quad \Rightarrow \quad \hat{A}\varphi_{a,b_m}(x) = a\varphi_{a,b_m}(x)$$

ovvero, le $\varphi_{a,b_m}(x)$ sono autofunzioni anche di \hat{A} .

4.5.1 Proprietà dei commutatori

Il concetto di commutatore riveste, come visto, un ruolo fondamentale nella meccanica quantistica. Essi godono delle seguenti proprietà:

1. $[\hat{A}, \hat{A}] = 0$

2. $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$
3. $[\hat{A}, c] = 0 \quad \forall c \in \mathbb{C}$
4. $[\hat{A}, \hat{B} + \hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}]$
5. $[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}]$
6. $[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{B}, [\hat{C}, \hat{A}]] + [\hat{C}, [\hat{A}, \hat{B}]] = 0$

L'insieme dei commutatori costituisce quindi un'algebra sugli spazi di Hilbert. Questa è un'algebra di derivazione¹⁴ che ha come prodotto il commutatore.

In base a quanto detto finora, è chiaro che gli operatori che commutano fra di loro rivestono un ruolo importante in meccanica quantistica. Si consideri ad esempio l'Hamiltoniana \hat{H} e gli operatori impulso \hat{p} e posizione \hat{x} :

$$[\hat{p}^2, \hat{H}] = 0$$

in quanto $\hat{H} = \hat{p}^2/2m$. Per quanto riguarda l'operatore posizione \hat{x} si ha invece:

$$\begin{aligned} [\hat{x}, \hat{H}] &= \frac{1}{2m} [\hat{x}, \hat{p}^2] = \frac{1}{2m} (\hat{x}\hat{p}^2 - \hat{p}^2\hat{x}) = \\ &= \frac{1}{2m} (\hat{x}\hat{p}^2 - \hat{p}^2\hat{x} + \hat{p}\hat{x}\hat{p} - \hat{p}\hat{x}\hat{p}) = \\ &= \frac{1}{2m} ([\hat{x}, \hat{p}]\hat{p} + \hat{p}[\hat{x}, \hat{p}]) = \frac{1}{2m} \cdot 2i\hbar\hat{p} = \\ &= \frac{i\hbar}{m}\hat{p} \end{aligned}$$

ovvero, la posizione e l'energia non possono essere misurate contemporaneamente con precisione arbitraria. Si noti che:

$$[\hat{x}, \hat{p}^2] = 2i\hbar\hat{p} = i\hbar \frac{\partial \hat{p}^2}{\partial \hat{p}}$$

Questa relazione è un caso particolare di una relazione più generale:

$$[\hat{x}, \hat{p}^n] = ni\hbar\hat{p}^{n-1} = i\hbar \frac{\partial}{\partial \hat{p}} \hat{p}^n \quad ; \quad [\hat{p}, \hat{x}^n] = -ni\hbar\hat{x}^{n-1} = -i\hbar \frac{\partial}{\partial \hat{x}} \hat{x}^n$$

e ancora più in generale:

$$\begin{cases} [\hat{x}, F(\hat{x}, \hat{p})] = i\hbar \frac{\partial}{\partial \hat{p}} F(\hat{x}, \hat{p}) \\ [\hat{p}, G(\hat{x}, \hat{p})] = -i\hbar \frac{\partial}{\partial \hat{x}} G(\hat{x}, \hat{p}) \end{cases} \quad (4.3)$$

Si noti che esiste una forte analogia formale con le *parentesi di Poisson* della meccanica classica:

$$\{F(q, p), G(q, p)\} = \sum_{i,j} \left[\frac{\partial F(q, p)}{\partial q_i} \frac{\partial G(q, p)}{\partial p_j} - \frac{\partial F(q, p)}{\partial p_i} \frac{\partial G(q, p)}{\partial q_j} \right]$$

ed in particolare posto $F(q, p) = p_n$ e $F(q, p) = q_h$:

$$\begin{cases} \{p_n, G(q, p)\} = -\frac{\partial G(q, p)}{\partial q_n} \\ \{q_h, G(q, p)\} = -\frac{\partial G(q, p)}{\partial p_n} \\ \{q_i, p_j\} = \delta_{ij} \end{cases}$$

Se il commutatore di due operatori è nullo, allora anche le corrispondenti parentesi di Poisson sono nulle: significa che gli osservabili che commutano corrispondono alle variabili classiche che possono

¹⁴Ha cioè le stesse proprietà delle derivate.

essere scelte come momenti coniugati nella rappresentazione canonica. In effetti, la formulazione della meccanica quantistica è estremamente simile formalmente a quella della meccanica classica. Si confrontino per esempio le equazioni di Hamilton-Jacobi:

$$\begin{cases} \dot{q}_i = \frac{\partial H}{\partial p_i} \\ \dot{p}_i = -\frac{\partial H}{\partial q_i} \end{cases}$$

con le corrispondenti quantistiche (4.5) in §4.9.

4.5.2 Densità e corrente di probabilità

Come è noto, la dinamica nello spazio delle fasi è retta dall'equazione di Schrödinger:

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = \hat{H} \psi(x, t)$$

dove l'operatore hamiltoniano corrisponde all'Hamiltoniana classica se questa è rappresentabile tramite le sole coordinate. Questo non è sempre il caso, ad esempio l'Hamiltoniana completa di un elettrone deve tenere conto anche dei termini di spin. In questi casi, si deve cercare una Hamiltoniana \hat{H} che sia comunque hermitiana.

L'evoluzione nel tempo della probabilità è espressa da:

$$\mathcal{P}_{\text{tot}} = \int_{-\infty}^{+\infty} \psi^*(x, t) \psi(x, t) d^n x \quad \frac{d}{dt} \mathcal{P}_{\text{tot}} = \int_{-\infty}^{+\infty} \left[\frac{\partial \psi^*(x, t)}{\partial t} \psi(x, t) + \psi^*(x, t) \frac{\partial \psi(x, t)}{\partial t} \right] d^n x$$

che utilizzando l'equazione di Schrödinger può essere riscritta come:

$$\mathcal{P}_{\text{tot}} = \frac{1}{i\hbar} \int_{-\infty}^{+\infty} \left\{ \psi^*(x, t) \hat{H} \psi(x, t) - [\hat{H} \psi(x, t)]^* \psi(x, t) \right\} d^n x$$

ma:

$$\int_{-\infty}^{+\infty} [\hat{H} \psi(x, t)]^* \psi(x, t) d^n x = \int_{-\infty}^{+\infty} \psi^*(x, t) \hat{H}^\dagger \psi(x, t) d^n x$$

e siccome l'Hamiltoniana è hermitiana $\hat{H} = \hat{H}^\dagger$, si ricava che:

$$\frac{d}{dt} \mathcal{P}_{\text{tot}} = 0$$

ovvero che la probabilità totale si conserva.

Si supponga ora invece un fascio di particelle che attraversa un materiale: da esso uscirà in genere un flusso più piccolo di quello incidente. Se una parte delle particelle è assorbita, si può descrivere questo fenomeno tramite un potenziale fittizio $V(x)$. In questo caso \hat{H} non è hermitiano e non ha luogo la conservazione della probabilità.

Per una particella in un potenziale risulta:

$$\frac{\partial}{\partial t} \psi(x, t) = \frac{1}{i\hbar} \left[\frac{\hat{p}^2}{2m} + V(x) \right] \psi(x, t)$$

La variazione di probabilità in un volume generico \mathcal{V} è data da:¹⁵

$$\frac{d}{dt} \mathcal{P}_{\text{tot}}(\mathcal{V}) = \int_{\mathcal{V}} \left[\psi^*(x, t) \left(\frac{\hat{p}^2}{2m} + V(x) \right) \psi(x, t) - \left(\frac{\hat{p}^2}{2m} + V^*(x) \right) \psi^*(x, t) \psi(x, t) \right] d^n x$$

¹⁵La notazione qui potrebbe generare confusione, si faccia attenzione a non confondere il potenziale $V(x)$ con il volume di integrazione \mathcal{V} .

Nell'ipotesi che V sia reale, $V^*\psi^*(x, t)\psi(x, t)$ e $\psi^*(x, t)V\psi(x, t)$ si cancellano a vicenda perchè risultano dei semplici prodotti. Si ricava quindi:

$$\frac{d}{dt}\mathcal{P}_{\text{tot}}(\mathcal{V}) = -\frac{1}{2i\hbar m} \int_{\mathcal{V}} \left\{ \psi^*(x, t)\nabla^2\psi(x, t) - [\nabla^2\psi^*(x, t)]\psi(x, t) \right\} d^n x$$

L'espressione sotto integrale si può riscrivere nella forma:

$$\psi^*(x, t)\nabla^2\psi(x, t) - (\nabla^2\psi^*(x, t))\psi(x, t) = \vec{\nabla} \cdot \left\{ \psi^*(x, t)\vec{\nabla}\psi(x, t) - [\vec{\nabla}\psi^*(x, t)]\psi(x, t) \right\}$$

ed è quindi l'integrale di una divergenza. La relazione:

$$\int_{\mathcal{V}} \left\{ \frac{\partial}{\partial t}\psi^*(x, t)\psi(x, t) + \frac{1}{2i\hbar m}\vec{\nabla} \cdot \left\{ \psi^*(x, t)\vec{\nabla}\psi(x, t) - [\vec{\nabla}\psi^*(x, t)]\psi(x, t) \right\} \right\} d^n x = 0$$

deve essere valida per ogni volume di integrazione e pertanto l'integrando deve essere nullo. Definendo:

$$\rho \equiv \psi^*(x, t)\psi(x, t) \quad \text{\textit{Densità di Probabilità}}$$

$$\vec{J} \equiv \frac{1}{2i\hbar m}\vec{\nabla} \cdot \left\{ \psi^*(x, t)\vec{\nabla}\psi(x, t) - [\vec{\nabla}\psi^*(x, t)]\psi(x, t) \right\} \quad \text{\textit{Densità di Corrente di Probabilità}}$$

risulta l'importante *equazione di continuità*:

$$\boxed{\frac{\partial}{\partial t}\rho + \vec{\nabla} \cdot \vec{J} = 0}$$

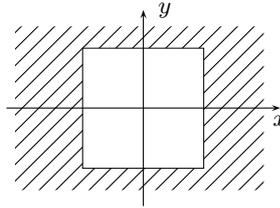
Si noti che la densità di corrente di probabilità si può mettere nella forma $\vec{J} = \frac{1}{\hbar^2}\Re(\psi^*\hat{p}\psi)$.

L'equazione di continuità corrisponde alla legge di conservazione della carica in forma locale. Infatti, se il potenziale è locale, la variazione della probabilità di trovare una carica all'interno di una superficie generica corrisponde al flusso di \vec{J} attraverso questa stessa superficie.

4.5.3 Separazione delle variabili e degenerazione. Numeri quantici.

Per illustrare il concetto di *degenerazione* e di *numero quantico*, si consideri una buca di potenziale infinita in due dimensioni, per semplicità quadrata di lato a e centrata sull'origine:

$$\begin{cases} V(\bar{x}, \bar{y}) = 0 & (\bar{x}, \bar{y}) \in \left[-\frac{a}{2}, \frac{a}{2}\right] \times \left[-\frac{a}{2}, \frac{a}{2}\right] \\ V(\bar{x}, \bar{y}) = +\infty & (\bar{x}, \bar{y}) \notin \left[-\frac{a}{2}, \frac{a}{2}\right] \times \left[-\frac{a}{2}, \frac{a}{2}\right] \end{cases}$$



Detta E l'energia, l'equazione di Schrödinger è:

$$\hat{H}\psi(x, y) = E\psi(x, y) \quad \hat{H} = \frac{\hat{p}_x^2}{2m} + \frac{\hat{p}_y^2}{2m} + V(x, y)$$

dove la funzione d'onda $\psi(x, y)$ si deve annullare ai bordi. L'Hamiltoniana del sistema si può scrivere come la somma di due Hamiltoniane $\hat{H} = \hat{H}_x + \hat{H}_y$ ⁽¹⁶⁾ e risulta:

$$[\hat{H}_x, \hat{H}_y] = 0$$

¹⁶Si noti che questo risulta possibile perché il potenziale stesso è separabile in due componenti indipendenti, una lungo x e l'altra lungo y .

In base al principio di sovrapposizione dell'equazione di Schrödinger, si può allora considerare una soluzione nella forma:

$$\psi(x, y) = \varphi(x)\chi(y)$$

ovvero come il prodotto di funzioni indipendenti che dipendono da una sola variabile. Da questo segue:

$$\hat{H}\psi(x, y) = [\hat{H}_x + \hat{H}_y] \varphi(x)\chi(y) = \hat{H}_x\varphi(x)\chi(y) + \varphi(x)\hat{H}_y\chi(y)$$

Supponendo che $\varphi(x)$ e $\chi(y)$ siano autofunzioni della rispettiva parte di Hamiltoniana:

$$\begin{aligned}\hat{H}_x\varphi_n(x) &= E_x^{(n)}\varphi_n(x) \\ \hat{H}_y\chi_m(y) &= E_y^{(m)}\chi_m(y)\end{aligned}$$

per la posizione fatta sulla $\psi(x, y)$ si ricava:

$$\hat{H}\psi(x, y) = [E_x^{(n)} + E_y^{(m)}]\psi(x, y)$$

ovvero, il prodotto di due autofunzioni è un'autofunzione appartenente all'autovalore somma degli autovalori. Vale il viceversa:

$$\hat{H}\psi(x, y) = [\hat{H}_x + \hat{H}_y] \varphi(x)\chi(y) = E\varphi(x)\chi(y)$$

e dividendo ambedue i membri per $\psi(x, y) = \varphi(x)\chi(y)$:

$$\frac{1}{\varphi(x)}\hat{H}_x + \frac{1}{\chi(y)}\hat{H}_y = E$$

ne segue che i membri a sinistra devono essere separatamente costanti:¹⁷

$$\frac{1}{\varphi(x)}\hat{H}_x\varphi(x) = \lambda_1 \quad ; \quad \frac{1}{\chi(y)}\hat{H}_y\chi(y) = \lambda_2$$

con $\lambda_1 + \lambda_2 = E$. Ma questo significa che:

$$\hat{H}_x\varphi(x) = \lambda_1\varphi(x) \quad ; \quad \hat{H}_y\chi(y) = \lambda_2\chi(y)$$

ovvero, $\varphi(x)$ e $\chi(y)$ sono autofunzioni agli autovalori λ_1 e λ_2 . Questo metodo di ricerca delle soluzioni prende il nome di *metodo di separazione delle variabili*. Questo non significa che ogni autofunzione sia esprimibile come prodotto di due autofunzioni e infatti questo in genere non è possibile. Tuttavia, le funzioni espresse nella forma $\psi_{n,m}(x, y) = \varphi_n(x)\chi_m(y)$ costituiscono una base, cioè:

$$\psi(x, y) = \sum_{n,m} c_{n,m}\psi_{n,m}(x, y) = \sum_{n,m} c_{n,m}\varphi_n(x)\chi_m(y)$$

Gli autovalori della buca bidimensionale sono:

$$E_{n_x} = \frac{n_x^2\hbar^2}{2ma^2} \quad ; \quad E_{n_y} = \frac{n_y^2\hbar^2}{2ma^2}$$

la cui somma è un autovalore di \hat{H} :

$$E = E_{n_x} + E_{n_y} = \frac{\hbar^2}{2ma^2} (n_x^2 + n_y^2) \quad n_i \in \mathbb{N}$$

Si vede in questo caso la cosiddetta *degenerazione*: diverse combinazioni di n_x e n_y possono dare origine allo stesso autovalore dell'energia. In altri termini, *fissato l'autovalore dell'energia, non è possibile da questo risalire ad uno stato quantistico definito*. In questo caso specifico, i numeri i cui quadrati hanno la stessa somma danno origine allo stesso autovalore dell'energia:

¹⁷Il membro a destra è una costante e il membro a sinistra è composto dalla somma di due termini di due variabili diverse e indipendenti.

		E_x	E_y	E
non degeneri	(1,1)	$n_x = 1$	$n_y = 1$	$n = 2$
degeneri	(2,1) ; (1,2)	$n_x = 2$	$n_y = 1$	$n = 5$
		$n_x = 1$	$n_y = 2$	$n = 5$

Con la buca centrata intorno all'origine le autofunzioni dei primi due livelli sono date da:

$$\psi_{1,1}(x, y) = N \cos \frac{\pi x}{a} \cos \frac{\pi y}{a} \quad E = \frac{2\hbar^2}{2ma^2}$$

$$\psi_{2,1}(x, y) = N \sin \frac{2\pi x}{a} \cos \frac{\pi y}{a} \quad E = \frac{5\hbar^2}{2ma^2}$$

$$\psi_{1,2}(x, y) = N \cos \frac{\pi x}{a} \sin \frac{2\pi y}{a} \quad E = \frac{5\hbar^2}{2ma^2}$$

A un \bar{y} fissato risulta $\psi(x, \bar{y}) = \sum_n c_n(\bar{y})\varphi_n(x)$, dove i coefficienti $c_n(\bar{y})$ dipendono da y . Quindi:

$$\psi(x, y) = \sum_{n,m} c_{n,m} \chi_m(y) \varphi_n(x)$$

Una combinazione lineare del tipo:

$$\psi(x, y) = \alpha\psi_{2,1}(x, y) + \beta\psi_{1,2}(x, y)$$

ha ancora lo stesso autovalore, pur non essendo in forma di prodotto di autofunzioni. Lo spettro di \hat{H} è quindi detto *degeneri*. Se si considerano però i due operatori \hat{H} e \hat{H}_x , questi commutano e pertanto ammettono una base di autovettori in comune. Questa base in comune permette di ottenere gli autovalori n e n_x , dai quali si può quindi ottenere n_y : *se si forniscono due osservabili che commutano è possibile rimuovere la degenerazione. Del tutto in generale, però, un insieme di osservabili che commutano riducono solo la degenerazione.*

*Si definisce **insieme di osservabili completo** un insieme che ha un **unico** sistema di autofunzioni in comune. Gli autovalori di questi osservabili sono i cosiddetti **numeri quantici**, che definiscono univocamente il sistema.*

4.6 Note riassuntive sulla struttura formale della teoria

È giunto il momento di tirare le somme e dare un quadro riassuntivo della struttura formale della meccanica quantistica e sul modo matematico con cui rappresenta i sistemi fisici.

- Ad ogni sistema fisico S si associa uno spazio di Hilbert \mathcal{H} .
- Lo stato di un sistema fisico è rappresentato da un vettore normalizzato ψ di questo spazio \mathcal{H} (funzione d'onda), ed è *la descrizione più esauriente possibile permessa da questo sistema formale*.
- Le operazioni di misura sono rappresentate da operatori hermitiani di proiezione \hat{O} (*osservabili*) che agiscono su questi vettori (funzioni d'onda).
- Alcuni osservabili non commutano fra loro: la conseguenza è che le grandezze fisiche rappresentate da questi osservabili non possono essere misurate entrambe con precisione arbitraria contemporaneamente (*principio di indeterminazione*). È bene notare che l'errore implicito nella relazione di indeterminazione non deriva dalle grandezze in sé o da un problema statistico, ma dall'atto di misurarle simultaneamente.
- I soli risultati possibili di una misura sono i valori o_i dello spettro dell'operatore hermitiano \hat{O} associato all'operazione di misura.

- Nel caso lo spettro dell'operatore \hat{O} sia discreto, allora i soli risultati possibili della misura sono dati dagli autovalori o_i dell'operatore.
- Si **postula** che il sistema di autovettori $\varphi_i^{\hat{O}}$ relativo all'operatore \hat{O} sia un sistema ortonormale *completo*: in altri termini, si postula che ogni vettore dello spazio di Hilbert \mathcal{H} associato al sistema sia esprimibile come combinazione lineare di autovettori relativi all'operatore \hat{O} . Nel caso di sistemi a più dimensioni, si assume che il sistema costituito dai prodotti delle singole autofunzioni unidimensionali costituisca una base.
- I coefficienti dello sviluppo $\psi^{\hat{O}} = \sum_i c_i \varphi_i^{\hat{O}}$ rappresentano quindi le probabilità di ottenere il relativo autovalore come risultato di una misura: $\mathcal{P}(o_i) = |c_i|^2$. La probabilità dell'esito di una misura è il massimo di informazione che può fornire la teoria.
- Se una misura fornisce come risultato il valore o_k , allora – nel momento della misura – la funzione d'onda viene istantaneamente proiettata nel sottospazio di \hat{O} relativo all'autovalore o_k , ovvero viene trasformato istantaneamente nell'autofunzione $\varphi_k^{\hat{O}}$ relativa all'autovalore o_k . È questo il cosiddetto *postulato della riduzione del pacchetto d'onda*, uno dei punti più controversi della meccanica quantistica.
- L'evoluzione temporale del sistema è espressa dall'equazione di Schrödinger:

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{x}, t) = \hat{H} \psi(\vec{x}, t)$$

4.7 Una dimostrazione generale della relazione di indeterminazione

La relazione di indeterminazione ricavata in §4.4 era stata introdotta sotto forma di principio da Heisenberg sulla base di motivazioni concrete legate al problema della misura.

È possibile ricavare una forma più generale delle relazioni di indeterminazione a partire dai postulati della meccanica quantistica, e in questo caso le relazioni di indeterminazione assumono la forma di un teorema.

Siano dati due osservabili \hat{A} e \hat{B} e si consideri i valori degli scarti quadratici medi:

$$(\Delta \hat{A})^2 (\Delta \hat{B})^2 = \langle \psi(x) | (\hat{A} - \langle \hat{A} \rangle)^2 \psi(x) \rangle \langle \psi(x) | (\hat{B} - \langle \hat{B} \rangle)^2 \psi(x) \rangle$$

e detto $\tilde{A} \equiv \hat{A} - \langle \hat{A} \rangle$, $\tilde{B} \equiv \hat{B} - \langle \hat{B} \rangle$ dalla hermiticità di \hat{A} e \hat{B} discende l'hermiticità di \tilde{A} e \tilde{B} . Dalla disuguaglianza di Schwarz si ricava:

$$\|\tilde{A}\psi(x)\|^2 \|\tilde{B}\psi(x)\|^2 \geq |\langle \tilde{A}\psi(x) | \tilde{B}\psi(x) \rangle|^2 = |\langle \psi(x) | \tilde{A}\tilde{B}\psi(x) \rangle|^2 \equiv |z|^2 \quad z \in \mathbb{C}$$

ma risulta $|z|^2 \geq \frac{1}{4} |z - z^*|^2$, per cui:

$$(\Delta \hat{A})^2 (\Delta \hat{B})^2 \geq \frac{1}{4} |\langle \psi(x) | \tilde{A}\tilde{B}\psi(x) \rangle - \langle \psi(x) | \tilde{A}\tilde{B}\psi(x) \rangle^*|^2 = \frac{1}{4} |\langle \psi(x) | \tilde{A}\tilde{B}\psi(x) \rangle - \langle \tilde{A}\tilde{B}\psi(x) | \psi(x) \rangle|^2$$

ed essendo hermitiani:

$$(\Delta \hat{A})^2 (\Delta \hat{B})^2 \geq \frac{1}{4} |\langle \psi(x) | \tilde{A}\tilde{B}\psi(x) \rangle - \langle \psi(x) | \tilde{B}\tilde{A}\psi(x) \rangle|^2 = \frac{1}{4} |\langle \psi(x) | [\tilde{A}, \tilde{B}]\psi(x) \rangle|^2$$

ma $[\tilde{A}, \tilde{B}] = [\hat{A}, \hat{B}]$ per cui:

$$\boxed{(\Delta \hat{A})_{\psi}^2 (\Delta \hat{B})_{\psi}^2 \geq \frac{1}{4} |\langle [\hat{A}, \hat{B}] \rangle_{\psi}|^2}$$

che è la formulazione più generale possibile del principio di indeterminazione. Si vede che nel caso i due osservabili non commutino, il prodotto degli scarti quadratici medi non può mai essere inferiore ad un certo valore. È questo il caso della posizione e dell'impulso, nel qual caso si trova:

$$(\Delta \hat{q})_{\psi}^2 (\Delta \hat{p})_{\psi}^2 \geq \frac{1}{4} |[\hat{q}, \hat{p}]|^2 = \frac{\hbar^2}{4}$$

4.8 Evoluzione classica dei valori medi

Si consideri il pacchetto d'onde generico di una particella libera:

$$\psi(x, t) = \int_{-\infty}^{+\infty} \varphi(k) e^{i(kx - \omega t)} dk \quad \omega = \frac{\hbar k^2}{2m}$$

Il valore medio di x è dato da:

$$\langle x(t) \rangle = \frac{1}{N} \int_{-\infty}^{+\infty} \psi^*(x, t) x \psi(x, t) dx$$

Ora, la trasformata di Fourier della $\psi(x, t)$ è data da $\mathcal{F}[\psi(x, t)] = \varphi(k) e^{-i\omega t}$ e per il teorema di Parseval:

$$\langle x(t) \rangle = \frac{1}{N} \int_{-\infty}^{+\infty} \varphi^*(k, t) i\hbar \frac{\partial}{\partial k} \varphi(k, t) dk = \frac{1}{N} \int_{-\infty}^{+\infty} \varphi^*(k) e^{-i\omega t} i\hbar \frac{\partial}{\partial k} [\varphi(k) e^{-i\omega t}] dk$$

la derivata fornisce:

$$\frac{\partial}{\partial k} [\varphi(k) e^{-i\omega t}] = \frac{\partial \varphi(k)}{\partial k} e^{-i\omega t} - it \frac{\partial \omega}{\partial k} \varphi(k) e^{-i\omega t}$$

che sostituita nell'espressione precedente permette di ricavare:

$$\langle x(t) \rangle = \varphi^*(k) i\hbar \frac{\partial}{\partial k} \varphi(k) dk + \frac{1}{N} \int_{-\infty}^{+\infty} \varphi^*(k) \frac{\hbar t}{m} k \varphi(k) dk$$

ma:

$$\frac{1}{N} \int_{-\infty}^{+\infty} \varphi^*(k) i\hbar \frac{\partial}{\partial k} \varphi(k) dk = \left\langle \frac{\partial}{\partial p} \right\rangle_{t=0}$$

è il valore medio all'istante iniziale.

$$\frac{1}{N} \int_{-\infty}^{+\infty} \varphi^*(k) \frac{\hbar t}{m} k \varphi(k) dk = \frac{t}{m} \langle p \rangle$$

inoltre:

$$\langle p \rangle = \hbar \langle k \rangle = \frac{1}{N} \int_{-\infty}^{+\infty} \varphi^*(k) e^{i\omega t} k \varphi(k) e^{-i\omega t} dk$$

quindi $\langle p \rangle$ non dipende dal tempo. Per cui:

$$\langle x(t) \rangle = \langle x_0 \rangle + \frac{t}{m} \langle p \rangle$$

ovvero il valore medio della posizione $\langle x \rangle$ di una particella libera quantistica ha la stessa evoluzione di una particella classica $x = x_0 + tp/m = x_0 + vt$.

Si consideri ora l'evoluzione dello scarto quadratico medio $\langle x^2 \rangle$:

$$\langle x^2 \rangle = \int_{-\infty}^{+\infty} \psi^*(x, t) x^2 \psi(x, t) dx = \frac{1}{N} \int_{-\infty}^{+\infty} \varphi^*(k, t) \left(i \frac{\partial}{\partial k} \right)^2 \varphi(k, t) dk$$

Siccome risulta:

$$\frac{\partial}{\partial k} \varphi(k, t) = \frac{\partial \varphi(k)}{\partial k} e^{-i\omega t} - i \frac{\hbar t}{m} k \varphi(k) e^{i\omega t}$$

e

$$\frac{\partial^2}{\partial k^2} \varphi(k, t) = \frac{\partial^2 \varphi(k)}{\partial k^2} e^{-i\omega t} - i \frac{\hbar t}{m} k \frac{\partial \varphi(k)}{\partial k} e^{-i\omega t} - i \frac{\hbar t}{m} \frac{\partial}{\partial k} [k \varphi(k)] e^{-i\omega t} - \frac{\hbar^2 t^2}{m^2} k^2 \varphi(k) e^{-i\omega t}$$

si ricava:

$$\langle x^2 \rangle = \frac{1}{N} \left[\int_{-\infty}^{+\infty} \varphi^*(k, t) \frac{\partial^2}{\partial k^2} \varphi(k, t) dk + \right. \\ \left. - i \frac{\hbar t}{m} \int_{-\infty}^{+\infty} \varphi^*(k, t) \left\{ k \frac{\partial}{\partial k} \varphi(k, t) + \frac{\partial}{\partial k} [k \varphi(k, t)] \right\} dk + \right. \\ \left. - \frac{\hbar^2 t^2}{m^2} \int_{-\infty}^{+\infty} \varphi^*(k, t) k^2 \varphi(k, t) dk \right]$$

Il primo termine rappresenta il valore medio $\langle \partial^2 / \partial k^2 \rangle$, ovvero il valore medio di x^2 a $t = 0$. Il secondo termine è lineare in t , mentre nel terzo $\hbar^2 k^2 = p^2$ e dunque rappresenta il valore medio $\langle p^2 \rangle$ ed è quadratico nel tempo. Risulta quindi:

$$\langle x^2 \rangle = \langle x^2 \rangle_0 + \beta t + \langle p^2 \rangle \frac{t^2}{m^2}$$

Lo scarto quadratico $(\Delta x)^2(t) = \langle x^2 \rangle(t) - \langle x \rangle^2(t)$ è allora dato da:

$$(\Delta x)^2(t) = \langle x^2 \rangle_0 - \langle x \rangle_0^2 + \gamma t + t^2 \frac{\langle p^2 \rangle - \langle p \rangle^2}{m^2}$$

Lo scarto Δx ha un andamento quadratico che tende a slargare il pacchetto se $\langle p^2 \rangle \neq \langle p \rangle^2$: ma $\langle p^2 \rangle$ non può essere uguale a $\langle p \rangle^2$ perché questo equivale a conoscere p senza incertezze.

Si consideri ora il termine β . Il valore di x è misurabile, dunque tale deve essere il suo scarto quadratico medio. Questo implica che il termine:

$$i \frac{\hbar t}{m} \int_{-\infty}^{+\infty} \varphi^*(k, t) \left\{ k \frac{\partial}{\partial k} \varphi(k, t) + \frac{\partial}{\partial k} [k \varphi(k, t)] \right\} dk$$

deve essere reale. Questo termine rappresenta la quantità:

$$\frac{t}{m} \langle \hat{p} \hat{x} + \hat{x} \hat{p} \rangle$$

e siccome i due operatori non commutano, risulta $\hat{p} \hat{x} \neq \hat{x} \hat{p}$. Infatti, dalla relazione (4.2) si ha $(\hat{x} \hat{p})^\dagger = \hat{p}^\dagger \hat{x}^\dagger = \hat{p} \hat{x} \neq \hat{x} \hat{p}$: ne consegue che $\langle \hat{x} \hat{p} \rangle \in \mathbb{C}$ perché \hat{x} e \hat{p} non commutano. Tuttavia, la combinazione:

$$(\hat{p} \hat{x} + \hat{x} \hat{p})^\dagger = (\hat{x} \hat{p})^\dagger + (\hat{p} \hat{x})^\dagger = \hat{p} \hat{x} + \hat{x} \hat{p}$$

coincide con l'operatore originario. Pertanto, la combinazione $\hat{x} \hat{p} + \hat{p} \hat{x}$ rappresenta un operatore hermitiano e quindi il termine:

$$\frac{t}{m} \langle \hat{p} \hat{x} + \hat{x} \hat{p} \rangle$$

commuta ed ha valore medio reale.

Questo discorso conduce ad un importante concetto: a due grandezze fisiche che non commutano non si può associare solo il prodotto $\hat{x} \hat{p}$ o $\hat{p} \hat{x}$ in quanto i loro valori medi possono risultare complessi. Quando intervengono prodotti di operatori osservabili che non commutano, quindi, occorre usare una combinazione hermitiana, ad esempio simmetrizzando come visto sopra:

$$xp \rightarrow \hat{x} \hat{p} \rightarrow \frac{1}{2} (\hat{x} \hat{p} + \hat{p} \hat{x})$$

4.9 Costanti del moto e evoluzione temporale

Un problema fondamentale della meccanica quantistica è quello delle costanti del moto.

Si consideri un generico osservabile \hat{A} :

$$\frac{d}{dt} \langle \hat{A} \rangle = \frac{\partial}{\partial t} \langle \psi(x, t) | \hat{A} \psi(x, t) \rangle = \langle \frac{\partial}{\partial t} \psi(x, t) | \hat{A} \psi(t) \rangle + \langle \psi(t) | \frac{d\hat{A}}{dt} \psi(x, t) \rangle + \langle \psi(x, t) | \hat{A} \frac{\partial}{\partial t} \psi(x, t) \rangle$$

ma dall'equazione di Schrödinger si ha:

$$\frac{\partial}{\partial t}\psi(x, t) = \frac{1}{i\hbar}\hat{H}\psi(x, t)$$

per cui, se l'hamiltoniana è hermitiana $\hat{H} = \hat{H}^\dagger$:

$$\frac{d}{dt}\langle\hat{A}\rangle = \langle\frac{1}{i\hbar}\hat{H}\psi(x, t) | \hat{A}\psi(x, t)\rangle + \langle\psi(x, t) | \frac{1}{i\hbar}\hat{A}\hat{H}\psi(x, t)\rangle + \langle\frac{d\hat{A}}{dt}\rangle$$

ovvero:

$$\begin{aligned}\frac{d}{dt}\langle\hat{A}\rangle &= \frac{1}{i\hbar}\langle\psi(x, t) | \hat{A}\hat{H}\psi(x, t)\rangle - \frac{1}{i\hbar}\langle\psi(x, t) | \hat{H}^\dagger\hat{A}\psi(x, t)\rangle + \langle\frac{d\hat{A}}{dt}\rangle \\ &= \frac{1}{i\hbar}\langle\psi(x, t) | [\hat{A}, \hat{H}]\psi(x, t)\rangle + \langle\frac{d\hat{A}}{dt}\rangle\end{aligned}$$

e in definitiva l'importante relazione:

$$\boxed{\frac{d}{dt}\langle\hat{A}\rangle = \frac{1}{i\hbar}\langle[\hat{A}, \hat{H}]\rangle + \langle\frac{d\hat{A}}{dt}\rangle}$$

che è una relazione del tutto generale.¹⁸

La grandezza \hat{A} , generalmente, può dipendere dal tempo nel caso di un potenziale variabile nel tempo (come è il caso di una perturbazione). Se si suppone che sia $\partial A/\partial t = 0$, allora si ha:

$$\frac{d}{dt}\langle\hat{A}\rangle = \frac{1}{i\hbar}\langle[\hat{A}, \hat{H}]\rangle \quad (4.4)$$

ne consegue che l'equivalente in meccanica quantistica di un costante del moto è una condizione sui valori medi e sul commutatore:

$$[\hat{A}, \hat{H}] = 0 \quad \Rightarrow \quad \frac{d}{dt}\langle\hat{A}\rangle = 0$$

ovvero *le costanti del moto sono rappresentate dagli osservabili che commutano con l'Hamiltoniana*.

Si consideri il caso particolare in cui $\hat{A} = q_i$. Segue:

$$\frac{d}{dt}\langle q_i \rangle = \frac{1}{i\hbar}\langle [q_i, H] \rangle = \langle \frac{\partial H}{\partial p_i} \rangle$$

e in maniera analoga:

$$\frac{d}{dt}\langle p_i \rangle = -\langle \frac{\partial H}{\partial q_i} \rangle$$

che sono l'analogo quantistico delle equazioni di Hamilton. In realtà, esiste una differenza fondamentale: la presenza di una media fa in modo che nel caso – ad esempio – di un potenziale quartico non ci sia corrispondenza fra le due equazioni a causa della comparsa di termini $\langle q^3 \rangle$ e $\langle q \rangle^3$. In effetti, le equazioni quantistiche coincidono con le equazioni classiche solo fino al termine quadratico. Il sistema di equazioni:

$$\begin{cases} \frac{d}{dt}\langle q_i \rangle = \langle \frac{\partial H}{\partial p_i} \rangle \\ \frac{d}{dt}\langle p_i \rangle = -\langle \frac{\partial H}{\partial q_i} \rangle \end{cases} \quad (4.5)$$

prende il nome di *teorema di Ehrenfest*.

¹⁸Si noti la corrispondenza con il formalismo Hamiltoniano classico:

$$\frac{d}{dt}A(q, p) = \{A, H\} + \frac{\partial A}{\partial t}$$

Se risulta $\hat{H}\psi(x, t) = E\psi(x, t)$, allora dalla (4.4) si ricava $[\hat{A}, \hat{H}] = 0$, da cui:

$$\frac{d}{dt}\langle \hat{A} \rangle = 0$$

ovvero per stati stazionari la situazione non evolve nel tempo, come d'altronde è ragionevole che sia.

Particolare interesse hanno quindi gli autostati dell'Hamiltoniana. Si consideri l'equazione di Schrödinger dinamica:

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = \hat{H}\psi(x, t)$$

e se $\hat{H}\psi(x, t) = E\psi(x, t)$ allora:

$$\frac{\partial}{\partial t} \hat{H} = 0$$

e si può dedurre:

$$\psi(\bar{x}, t) = e^{-iEt/\hbar} \psi(\bar{x}, 0)$$

Per un autostato di \hat{H} vale $\hat{H}\varphi_n(x) = E_n\varphi_n(x)$ e:

$$\psi(x, t) = \sum_n c_n \varphi_n(x) \quad c_n \equiv c_n(t)$$

Siccome vale:

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = i\hbar \sum_n \dot{c}_n(t) \varphi_n(x)$$

l'equazione dinamica assume la forma:

$$i\hbar \sum_n \dot{c}_n(t) \varphi_n(x) = \sum_n E_n c_n(t) \varphi_n(x)$$

ed essendo uguali le basi, devono essere uguali i coefficienti:

$$i\hbar \frac{d}{dt} c_n(t) = E_n c_n(t) \quad \Rightarrow \quad c_n(t) = c_n(0) e^{-iE_n t/\hbar}$$

per cui l'evoluzione temporale di $\psi(x, t)$ è espressa da:

$$\psi(x, t) = \sum_n c_n(0) e^{-iE_n t/\hbar} \varphi_n(x)$$

L'unica cosa che cambia nel tempo sono dunque i fattori di fase relativi. Formalmente, la relazione di sopra può scriversi come segue:

$$\psi(x, t) = \sum_n e^{-i\hat{H}t/\hbar} c_n(0) \varphi_n(x)$$

dove ora il fattore $e^{-i\hat{H}t/\hbar}$ è in comune a tutti termini:

$$\psi(x, t) = e^{-i\hat{H}t/\hbar} \sum_n c_n(0) \varphi_n(x)$$

Da notare che questo fattore è pur sempre un operatore, quindi fornisce un risultato diverso per ogni termine su cui agisce. Formalmente, si ha quindi in questo caso:

$$\boxed{\psi(x, t) = e^{-i\hat{H}t/\hbar} \psi(x, t=0)} \quad (4.6)$$

Questo modo di descrivere l'evoluzione temporale del vettore di stato è chiamato *rappresentazione di Schrödinger* e sarà ripreso più avanti nel contesto della teoria delle perturbazioni dipendenti dal tempo, nel paragrafo §8.2.

Capitolo 5

L'oscillatore armonico

Si studierà ora in dettaglio uno dei sistemi più importanti in fisica: l'*oscillatore armonico*, ovvero il comportamento di un sistema soggetto ad un potenziale quadratico. Il caso preso in esame qui è quello di un oscillatore unidimensionale.

5.1 Metodo analitico

Essendo il potenziale quadratico, l'Hamiltoniana del sistema è data da:

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2$$

l'equazione di Schrödinger per l'oscillatore armonico unidimensionale è quindi data da:

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}kx^2 \right] \psi(x) = E\psi(x)$$

Questa equazione può essere messa in forma adimensionale utilizzando la lunghezza caratteristica del sistema $l = \sqrt{\hbar/m\omega}$. Moltiplicando l'equazione per $\frac{m\omega}{\hbar}$ ed effettuando la sostituzione di variabile $x \rightarrow y \equiv x/l$, e quindi $y^2 = \frac{m\omega}{\hbar}x^2$, si ottiene:

$$\begin{aligned} \left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}kx^2 \right] \psi(x) &= E\psi(x) \rightarrow \\ \left[-\frac{\hbar^2}{2m} \frac{m\omega}{\hbar} \frac{d^2}{dx^2} + \frac{1}{2}k \underbrace{\frac{m\omega}{\hbar}x^2}_{y^2} \right] \psi(x) &= E \frac{m\omega}{\hbar} \psi(x) \rightarrow \\ \left[-\frac{\hbar\omega}{2} \frac{m\omega}{\hbar} \frac{d^2}{d(x/l)^2} + \frac{1}{2}ky^2 \right] \psi(y) &= E \frac{m\omega}{\hbar} \psi(y) \rightarrow \\ \left[-\frac{m\omega^2}{2} \frac{d^2}{k dy^2} + y^2 \right] \psi(y) &= E \frac{m\omega}{\hbar} \frac{2}{k} \psi(y) \end{aligned}$$

ma $\omega = \sqrt{k/m}$ da cui $k = m\omega^2$, per cui:

$$\left[-\cancel{m\omega^2} \frac{1}{\cancel{m\omega^2}} \frac{d^2}{dy^2} + y^2 \right] \psi(y) = E \frac{m\omega}{\hbar} \frac{2}{\cancel{m\omega^2}} \psi(y) = E \underbrace{\frac{2}{\hbar\omega}}_{1/\varepsilon_0} \psi(y) \equiv \varepsilon\psi(y)$$

ovvero in definitiva:

$$\left[-\frac{d^2}{dy^2} + y^2 \right] \psi(y) = \varepsilon\psi(y) \tag{5.1}$$

e l'equazione risulta adimensionale. Si noti che una volta resa adimensionale questa equazione compare un livello di riferimento (un fattore di scala, se si vuole) rispetto al quale l'energia è misurata. Questo valore è dato da $\varepsilon_0 \equiv \frac{\hbar\omega}{2}$.¹

I coefficienti di questa equazione non hanno singolarità al finito, pertanto le soluzioni stesse non hanno singolarità al finito. Per y tendente a infinito, il termine di energia diventa trascurabile e l'equazione diventa:

$$\frac{d^2}{dy^2}\psi(y) = (y^2 - \varepsilon)\psi(y) \quad \xrightarrow{y \rightarrow \infty} \quad \frac{d^2}{dy^2}\psi(y) = y^2\psi(y)$$

che ha per soluzioni:

$$\lim_{y \rightarrow \infty} \psi(y) \sim e^{y^2/2} + e^{-y^2/2}$$

La soluzione con il segno positivo deve essere scartata su base fisica (non appartiene a \mathcal{L}^2), e siccome risulta anche:

$$\frac{d^2}{dy^2}e^{-y^2/2} = (1 - 2y + y^2)e^{-y^2/2}$$

la soluzione all'infinito deve avere la forma di un polinomio moltiplicato l'esponenziale. Poniamo quindi che la soluzione sia nella forma:²

$$\psi(y) = H(y)e^{-y^2/2} \tag{5.2}$$

Sostituendo questa espressione nell'equazione (5.1) si ottiene:

$$\frac{d^2}{dy^2}H(y) - 2yH(y) + (\varepsilon - 1)H(y) = 0$$

La funzione $H(y)$ può certamente essere espressa come una serie di potenze:

$$H(y) = \sum_{n=0}^{\infty} a_n y^n \quad \Rightarrow \quad \begin{cases} \frac{d}{dy}H(y) &= \sum_{n=0}^{\infty} n a_n y^{n-1} \\ \frac{d^2}{dy^2}H(y) &= \sum_{n=0}^{\infty} n(n-1) a_n y^{n-2} \end{cases}$$

che sostituita nell'equazione differenziale fornisce:³

$$\sum_{n=2}^{\infty} n(n-1) a_n y^{n-2} + \sum_{n=0}^{\infty} (-2n + \varepsilon - 1) a_n y^n = 0$$

Mandando $n \rightarrow m = n - 2$, la prima sommatoria si può riscrivere nella forma:

$$\sum_{n=2}^{\infty} n(n-1) a_n y^{n-2} \rightarrow \sum_{m=0}^{\infty} (m+2)(m+1) a_{m+2} y^m$$

che può quindi essere sostituita nella relazione di sopra ottenendo:

$$\sum_{n=0}^{\infty} [(n+2)(n+1) a_{n+2} + (-2n + \varepsilon - 1) a_n] y^n = 0$$

¹Non deve qui sfuggire il fatto che $\hbar\omega/2 \equiv h\nu/2$, ovvero che il valore di riferimento dell'energia è dato in termini di quanti di Planck.

²Nessun passaggio arduo in questa posizione: si sa che all'infinito la soluzione deve comportarsi come un esponenziale decrescente moltiplicato per una potenza, quindi si scrive la forma riportata nel testo e si cerca la forma giusta della funzione $H(y)$ in base alle condizioni che deve rispettare.

³Si noti che derivando due volte, i primi due termini dello sviluppo in serie spariscono e quindi la sommatoria parte da $n = 2$.

Ora, è noto che le potenze sono linearmente indipendenti e quindi l'unico modo per verificare questa relazione è quello in cui tutti i coefficienti sono nulli. Questo permette di ricavare la relazione ricorsiva:

$$a_{n+2} = \frac{2n+1-\varepsilon}{(n+2)(n+1)} a_n$$

che lascia liberi i due parametri a_0 e a_1 . Siccome si tratta una serie di potenze, è ragionevole porre i primi termini a $a_0 = 1$ e $a_1 = y$.

Occorre ora studiare la convergenza della serie. Il rapporto di due termini successivi è dato da:

$$\lim_{n \rightarrow +\infty} \frac{a_{n+2}}{a_n} = \lim_{n \rightarrow +\infty} \frac{2}{n} = 0$$

la serie dunque converge sempre. Tuttavia, se si nota che l'espansione in serie dell'esponenziale è:

$$e^{x^2} = \sum_{n=0}^{\infty} \frac{x^{2n}}{n!} \quad \rightarrow \quad \frac{a_n}{a_{n-1}} = \frac{x^2}{n}$$

si vede che per y tendente a infinito la soluzione tende a e^{y^2} . Ricordando la posizione (5.2) fatta in precedenza, si trova:

$$\lim_{y \rightarrow +\infty} \psi(y) = e^{y^2} e^{-y^2/2} = e^{y^2/2}$$

che non appartiene a \mathcal{L}^2 e pertanto non è normalizzabile. L'unico modo per ottenere una soluzione fisicamente accettabile è quindi che la serie si riduca ad un polinomio, ovvero che esista un intero $n \in \mathbb{N}_0$: $a_n \neq 0 \Rightarrow a_{n+2} = 0$:

$$2n+1-\varepsilon = 0 \quad \Rightarrow \quad \varepsilon \equiv \frac{E}{\varepsilon_0} = 2n+1$$

e ricordando la posizione su ε_0 :

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right)$$

Questa formula esprime gli autovalori dell'energia per l'oscillatore armonico unidimensionale, di cui n rappresenta il numero quantico.

I polinomi $H(y)$ così definiti sono i cosiddetti *polinomi di Hermite*. La funzione generatrice di questi polinomi è:

$$H_n(y) = (-1)^n e^{y^2} \frac{d^n}{dy^n} e^{-y^2}$$

e come si vede hanno la parità corrispondente al loro grado. Essi inoltre risultano ortogonali rispetto alla misura di Lebegue e^{-y^2} :⁽⁴⁾

$$\int_{-\infty}^{+\infty} H_m^*(y) H_n(y) e^{-y^2} dy = N \delta_{n,m}$$

I primi polinomi di Hermite sono:

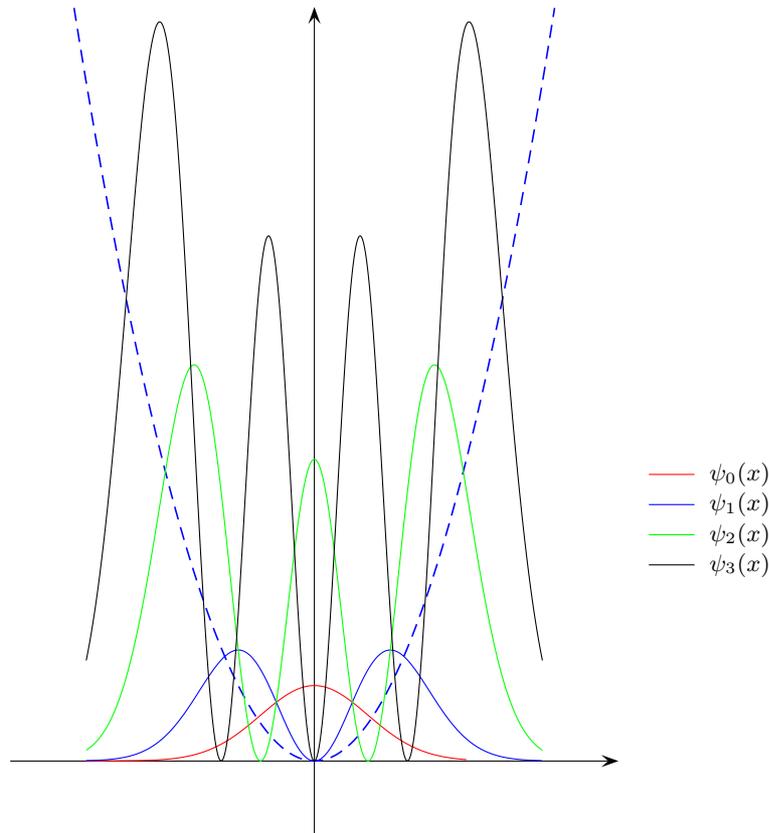
$$\begin{aligned} H_0(x) &= 1 \\ H_1(x) &= 2x \\ H_2(x) &= 4x^2 - 2 \\ H_3(x) &= 8x^3 - 12x \\ H_4(x) &= 16x^4 - 48x^2 + 12 \\ H_5(x) &= 32x^5 - 160x^3 + 120x \end{aligned}$$

⁴Essi costituiscono quindi l'ortogonalizzazione dei polinomi secondo la misura e^{-y^2} .

Le autofunzioni dell'oscillatore armonico sono quindi date da:

$$\psi_n(x) = H_n(x/l)e^{-(x/l)^2/2}$$

a meno di un fattore di normalizzazione. Le densità di probabilità $|\psi(x)|^2$ sono rappresentate in figura seguente per i primi quattro livelli.



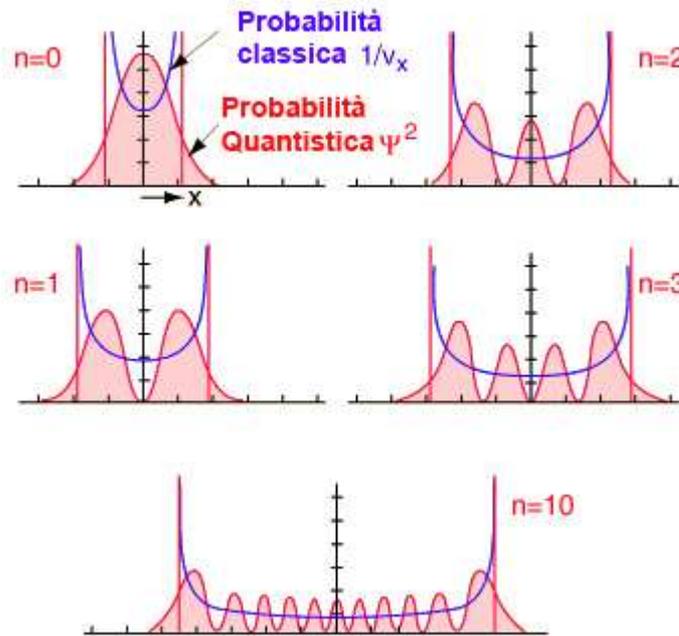
Confronto fra le probabilità quantistiche e classiche dell'oscillatore armonico

È noto che le leggi della meccanica quantistica devono ridursi a quelle della meccanica classica quando i numeri quantici del sistema tendono a infinito. Si vedrà come questo accada concretamente nel caso dell'oscillatore armonico.

Nel caso classico, il massimo di probabilità di trovare la particella oscillante si ha agli estremi del moto perché in quelle zone il moto dell'oscillatore è più lento e si arresta agli estremi. La probabilità relativa di trovare l'oscillatore in un intervallo δx è proprio l'inverso della velocità in questo intervallo. Nel caso quantistico, la probabilità è data dal quadrato della funzione d'onda $|\psi(x)|^2$ e per bassi valori dell'energia è completamente diversa. In particolare, per lo stato fondamentale E_0 , il massimo della probabilità quantistica è al centro, nella posizione di riposo.

Per i primi livelli, non c'è molta somiglianza fra la probabilità classica e quella quantistica. Man mano però che il numero quantico aumenta (e di conseguenza, l'energia), la probabilità si sposta sempre più verso gli estremi e le oscillazioni si fanno sempre più fitte. Al limite per alti numeri quantici, le oscillazioni diventano così fitte che il principio di indeterminazione impedisce di distinguerle, per cui la distribuzione di probabilità appare continua ed il picco di probabilità è intorno ai punti classici di inversione.

Questo è bene illustrato nella figura seguente.



Si noti che Planck postulò che l'energia degli oscillatore armonici fosse $E = n\hbar\omega$, mentre ora si è trovato:

$$E = n\hbar\omega + \frac{1}{2}\hbar\omega$$

Esiste dunque un termine aggiuntivo costante che in realtà non influisce nel momento in cui si calcola il valore medio.⁵ Questo termine è un effetto prettamente quantistico. In effetti, se non fosse presente per $n = 0$ sarebbe possibile il valore $E_0 = 0$ che implica:

$$\langle H \rangle = \frac{\langle p^2 \rangle}{2m} + \frac{1}{2}k\langle x^2 \rangle = 0 \quad E_0 = 0$$

Essendo \hat{p} hermitiano, $\langle p^2 \rangle \geq 0$:

$$\langle p^2 \rangle = \langle \psi(x) | \hat{p}^2 \psi(x) \rangle = \langle \hat{p}\psi(x) | \hat{p}\psi(x) \rangle = \|\hat{p}\psi(x)\| \geq 0$$

e analogamente per $\langle \hat{x}^2 \rangle$. Ne consegue che nello stato $n = 0$ entrambi dovrebbero essere zero:

$$\hat{p}\psi(x) = 0 \quad \hat{x}\psi(x) = 0$$

il che è vietato dal principio di indeterminazione. Dunque, *a causa del principio di indeterminazione, l'energia dello stato fondamentale dell'oscillatore armonico non può essere 0. Questa energia residua in E_0 è detta **energia di punto zero** e vale $E_0 = \frac{1}{2}\hbar\omega$.*

Si noti che una volta opportunamente normalizzata la $\psi(x)$ si estende al di là del potenziale armonico, esiste quindi una probabilità finita di trovare la particella nella zona classicamente proibita.

5.2 Metodo operatoriale. Operatori di *creazione e distruzione*

Questo metodo è stato messo a punto da P.A.M. Dirac.

⁵Per questo Plank non se ne era accorto e non aveva causato problemi nel calcolo della distribuzione di corpo nero.

Richiamo sull'oscillatore armonico classico

È noto che $\omega^2 = k/m$ e che l'Hamiltoniana dell'oscillatore armonico si può scrivere nella forma:

$$H = \frac{1}{2}k \left(x^2 + \frac{p^2}{m^2\omega^2} \right)$$

Le equazioni dell'oscillatore sono:

$$\dot{q} = \frac{\partial H}{\partial p} \propto p \quad \dot{p} = -\frac{\partial H}{\partial q} \propto -q$$

che si risolvono derivando la prima e sostituendola nella seconda:

$$\ddot{q} + q = 0$$

che ha per soluzioni $q(t) = Ae^{it} + Be^{-it}$. È possibile anche passare per un'equazione del primo ordine:

$$\dot{q} + ip = p - iq = -i(q + ip) \quad \Rightarrow \quad q + ip = Ae^{-it}$$

e analogamente per $q - ip$. Dalla conoscenza delle due si ricava la soluzione totale come un problema del primo ordine.

È possibile seguire lo stesso ragionamento in meccanica quantistica, ma si deve tenere presente che non è possibile riscrivere l'Hamiltoniana nella forma:

$$\hat{H} = \left(\hat{q} + i\frac{\hat{p}}{m\omega} \right) \left(\hat{q} - i\frac{\hat{p}}{m\omega} \right)$$

perché i due prodotti incrociati non commutano e quindi non si annullano. Infatti:

$$\begin{aligned} \left(\hat{q} + i\frac{\hat{p}}{m\omega} \right) \left(\hat{q} - i\frac{\hat{p}}{m\omega} \right) &= \hat{q}^2 + \frac{\hat{p}^2}{m\omega} + \frac{i}{m\omega}(\hat{p}\hat{q} - \hat{q}\hat{p}) = \\ &= \hat{q}^2 + \frac{\hat{p}^2}{m^2\omega^2} + \frac{i}{m\omega}[\hat{p}, \hat{q}] = 2\frac{\hat{H}}{k} + \frac{\hbar}{m\omega} \end{aligned}$$

Si introducano gli operatori:

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{q} + i\frac{\hat{p}}{m\omega} \right) \quad \hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{q} - i\frac{\hat{p}}{m\omega} \right) \quad (5.3)$$

con gli opportuni fattori di normalizzazione che li rendono adimensionali. Siccome vale:

$$\begin{aligned} a^\dagger a &= \frac{m\omega}{2\hbar} \left(\hat{q} - i\frac{\hat{p}}{m\omega} \right) \left(\hat{q} + i\frac{\hat{p}}{m\omega} \right) = \frac{m\omega}{2\hbar} \left\{ \hat{q}^2 + \frac{\hat{p}^2}{m\omega} + \frac{i}{m\omega}(\hat{q}\hat{p} - \hat{p}\hat{q}) \right\} = \\ &= \underbrace{\frac{m\omega}{2\hbar} \left(\hat{q}^2 + \frac{\hat{p}^2}{m\omega} \right)}_{\hat{H}/\hbar\omega} + \underbrace{\frac{i}{2\hbar}[\hat{q}, \hat{p}]}_{-1/2} \end{aligned}$$

l'Hamiltoniana si può riscrivere come:

$$\boxed{H = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right)} \quad (5.4)$$

Gli operatori \hat{a} e \hat{a}^\dagger non commutano:

$$[a, a^\dagger] = \frac{m\omega}{2\hbar} \frac{-i}{m\omega} ([\hat{q}, \hat{p}] - [\hat{p}, \hat{q}]) = 1 \quad (5.5)$$

Si consideri ora l'operatore $\hat{N} = \hat{a}^\dagger \hat{a}$.

$$\langle \psi(x) | \hat{N}\psi(x) \rangle = \langle \psi(x) | a^\dagger a\psi(x) \rangle = \langle a\psi(x) | a\psi(x) \rangle \geq 0$$

ed è zero se e solo se $a\psi(x) = 0$. Detta ora $\psi_\lambda(x)$ un'autofunzione di \hat{N} :

$$\hat{N}\psi_\lambda(x) = \lambda\psi_\lambda(x)$$

si consideri la funzione $\chi(x) \equiv a\psi_\lambda(x)$. Risulta allora:

$$\hat{N}\chi(x) = \hat{N}a\psi_\lambda(x) = a^\dagger a a\psi_\lambda(x)$$

ma essendo $[a, a^\dagger] = 1$, si trova:

$$\hat{N}\chi(x) = a^\dagger a a\psi_\lambda(x) = (a a^\dagger - 1)a\psi_\lambda(x) = (a \underbrace{a^\dagger a}_{\hat{N}} - a)\psi_\lambda(x) = (\lambda - 1)a\psi_\lambda(x) = (\lambda - 1)\chi(x)$$

cioè se $\psi_\lambda(x)$ è autostato di \hat{N} relativo all'autovalore λ , allora $\hat{a}\psi_\lambda(x)$ è ancora autostato di \hat{N} , ma appartenente all'autovalore $(\lambda - 1)$:

$$\hat{a}\psi_\lambda(x) = \psi_{\lambda-1}(x)$$

In maniera analoga si dimostra che:

$$\hat{a}^\dagger\psi_\lambda(x) = \psi_{\lambda+1}(x)$$

Gli operatori \hat{a} e \hat{a}^\dagger fanno quindi passare da un autostato relativo ad un autovalore λ ad un autostato relativo ad un autovalore minore o maggiore di una unità rispettivamente, hanno quindi per autovalori $\lambda + k$ con $k \in \mathbb{Z}$. Essi sono pertanto denominati:⁶

\hat{a} = operatore di *distruzione*

\hat{a}^\dagger = operatore di *creazione*

Si noti che questi operatori non sono autoaggiunti, quindi *non sono osservabili*.

Naturalmente, questo discorso ha senso se $a\psi_\lambda(x) \neq 0$ e $a^\dagger\psi_\lambda(x) \neq 0$. Si consideri l'operatore di creazione:

$$\|a^\dagger\psi_\lambda(x)\| = \langle a^\dagger\psi_\lambda(x) | a^\dagger\psi_\lambda(x) \rangle = \langle \psi_\lambda(x) | a a^\dagger\psi_\lambda(x) \rangle = (\lambda + 1)\|\psi_\lambda(x)\| > 0 \quad (\lambda \geq 0)$$

dunque \hat{a}^\dagger non potrà mai dare 0 su uno stato. Nel caso dell'operatore di distruzione, siccome *deve* esistere un minimo dell'energia, allora *deve* esistere uno stato in cui questo ritorna 0:

$$\exists k: \hat{a}\psi_{\lambda-k}(x) = 0$$

ma questa relazione può essere verificata solo se l'autovalore è 0, quindi $\lambda = k$. Ne consegue che $\lambda = k$ è anche lo spettro di \hat{N} , i cui autovalori sono dunque gli interi non negativi. Questi sono allora anche gli autovalori dell'Hamiltoniana per la (5.4):

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right) \quad n \in \mathbb{N}_0$$

Questo operatore prende quindi il nome di **operatore numero**, in quanto restituisce il livello di energia (numero quantico) dell'oscillatore armonico.

La condizione di normalizzazione è $\langle \psi_n(x) | \psi_n(x) \rangle = 1$. Esprimendo l' n -simo autostato in termini dell'operatore di creazione:

$$\psi_n(x) = (\hat{a}^\dagger)^n \psi_0(x)$$

si ottiene:

$$|c_n|^2 \langle \psi_n(x) | \psi_n(x) \rangle = |c_n|^2 \langle (\hat{a}^\dagger)^n \psi_0(x) | (\hat{a}^\dagger)^n \psi_0(x) \rangle = |c_n|^2 \langle \psi_0(x) | \hat{a}^n (\hat{a}^\dagger)^n \psi_0(x) \rangle$$

⁶In quanto essi permettono di "creare" o "distruggere" un quanto di energia.

Dalla relazione (5.5) si ottiene:

$$\hat{a}^n (\hat{a}^\dagger)^n = \hat{a}^{n-1} \hat{a} \hat{a}^\dagger (\hat{a}^\dagger)^{n-1} = \hat{a}^{n-1} \hat{a}^\dagger \hat{a} (\hat{a}^\dagger)^{n-1} + \hat{a}^{n-1} (\hat{a}^\dagger)^{n-1}$$

questa scomposizione può essere applicata n volte sui prodotti $\hat{a} \hat{a}^\dagger$, ottenendo:

$$\hat{a}^n (\hat{a}^\dagger)^n = \hat{a}^{n-1} (\hat{a}^\dagger)^n \hat{a} + n \hat{a}^{n-1} (\hat{a}^\dagger)^{n-1}$$

ovvero:

$$|c_n|^2 \langle \psi_n(x) | \psi_n(x) \rangle = \langle \psi_0(x) | \hat{a}^{n-1} (\hat{a}^\dagger)^n \hat{a} \psi_0(x) \rangle + n \langle \psi_0(x) | \hat{a}^{n-1} (\hat{a}^\dagger)^{n-1} \psi_0(x) \rangle$$

ma il primo termine è nullo in quanto $\hat{a} \psi_0(x) = 0$ e per il secondo rimane:

$$|c_n|^2 \langle \psi_n(x) | \psi_n(x) \rangle = n \langle (\hat{a}^\dagger)^{n-1} \psi_0(x) | (\hat{a}^\dagger)^{n-1} \psi_0(x) \rangle = n |c_{n-1}|^2 \langle \psi_{n-1}(x) | \psi_{n-1}(x) \rangle$$

ovvero:

$$|c_n|^2 = n |c_{n-1}|^2$$

i coefficienti risultano quindi definiti per ricorrenza e siccome vale $|\psi_0(x)|^2 = 1$, risulta $|c_n|^2 = n!$ e potendo sempre essere scelti reali, si ricava la forma generale degli autostati dell'oscillatore armonico:

$$\psi_n(x) = \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n \psi_0(x)$$

5.3 Valori medi e stati coerenti

Si noti preliminarmente che a partire dalla definizione (5.3) degli operatori di creazione e distruzione si ricava facilmente:

$$\hat{q} = \sqrt{\frac{2\hbar}{m\omega}} \frac{a + a^\dagger}{2} \quad \hat{p} = \sqrt{2\hbar m\omega} \frac{a - a^\dagger}{2i}$$

Il valore medio di \hat{q} sullo stato n è dato da:

$$\langle \hat{q} \rangle_n = \langle \psi_n(x) | \hat{q} \psi_n(x) \rangle = \sqrt{\frac{\hbar}{2m\omega}} \langle \psi_n(x) | (a + a^\dagger) \psi_n(x) \rangle$$

ma come visto risulta:

$$\begin{aligned} \langle \psi_n(x) | \hat{a} \psi_n(x) \rangle &= \langle \psi_n(x) | \psi_{n-1}(x) \rangle = 0 & n: \hat{a} \psi_n(x) \neq 0 \\ \langle \psi_n(x) | \hat{a}^\dagger \psi_n(x) \rangle &= \langle \psi_n(x) | \psi_{n+1}(x) \rangle = 0 \end{aligned}$$

infatti autofunzioni relative ad autovalori diversi sono ortogonali. Ne consegue che $\langle \hat{q} \rangle_n = 0$ e analogamente $\langle \hat{p} \rangle_n = 0$.

Questo procedimento è particolarmente interessante in caso di calcoli più complessi:

$$\hat{q}^2 = \frac{\hbar}{2m\omega} (a + a^\dagger)^2 = \frac{\hbar}{2m\omega} (a^2 + a^\dagger a + a a^\dagger + a^{\dagger 2})$$

ma $(a a^\dagger - a^\dagger a) = 1$, per cui:

$$\hat{q}^2 = \frac{\hbar}{2m\omega} (a^2 + a^{\dagger 2} + 2a^\dagger a + 1) = \frac{\hbar}{2m\omega} (a^2 + a^{\dagger 2} + 2\hat{N} + 1)$$

Nel momento in cui si calcola il valore medio, gli stati $a \psi_n(x)$ e $a^\dagger \psi_n(x)$ risultano ortogonali a $\psi_n(x)$ e quindi non danno contributo:

$$\langle \hat{q}^2 \rangle_n = \frac{\hbar}{2m\omega} \langle \psi_n(x) | (a^2 + a^{\dagger 2} + 2\hat{N} + 1) \psi_n(x) \rangle = \frac{\hbar}{2m\omega} (2n + 1)$$

Si noti che risulta:

$$\langle \hat{q}^2 \rangle_n = \frac{\hbar}{2m\omega} (2n+1) = \frac{\hbar\omega}{m\omega^2} \left(n + \frac{1}{2} \right) = \frac{E_n}{m\omega^2}$$

Un discorso simile può essere fatto per il valore medio di \hat{p} :

$$\hat{p}^2 = -\frac{\hbar m\omega}{2} (a^2 + a^\dagger a + a a^\dagger + a^{\dagger 2}) = -\frac{\hbar m\omega}{2} (a^2 + a^{\dagger 2} - 2a a^\dagger - 1) = -\frac{\hbar m\omega}{2} (a^2 + a^{\dagger 2} - 2\hat{N} - 1)$$

da cui:

$$\langle \hat{p}^2 \rangle_n = \hbar m\omega \left(n + \frac{1}{2} \right) = mE_n$$

ne segue naturalmente che:

$$\left\langle \frac{\hat{p}^2}{2m} \right\rangle = \frac{1}{2} E_n = \left\langle \frac{1}{2} \hbar k \hat{q}^2 \right\rangle$$

ovvero, il valore medio dell'energia cinetica e dell'energia potenziale coincidono.

Si consideri ora l'equazione agli autovalori per l'operatore di distruzione:

$$\hat{a}\psi(x) = \lambda\psi(x)$$

Nello spazio delle configurazioni risulta ($y = x/l$):

$$\begin{aligned} \hat{a} &= y + \frac{d}{dy} \\ \hat{a}^\dagger &= y - \frac{d}{dy} \end{aligned}$$

per cui l'equazione agli autovalori diventa:

$$\left(y + \frac{d}{dy} \right) \psi(y) = \lambda\psi(y)$$

da cui:

$$\frac{d}{dy}\psi(y) = (\lambda - y)\psi(y) \quad \Rightarrow \quad \psi(y) = Ae^{\lambda y - y^2/2}$$

Questa funzione appartiene a \mathcal{L}^2 per ogni $\lambda \in \mathbb{C}$, infatti a infinito il termine predominante è sempre l'esponenziale negativo. *Questi stati sono chiamati stati coerenti, sono gli autostati dell'operatore di distruzione e formano un sistema completo. Essi rappresentano delle gaussiane centrate in $\sqrt{2} \Re(\lambda)$.*

Procedendo in maniera analoga per l'operatore di creazione si trova:

$$a^\dagger \psi(y) = \lambda\psi(y) \quad \Rightarrow \quad \psi(y) = Ae^{-\lambda y + y^2/2}$$

e quindi l'operatore di creazione non ha mai autostati.

Si considerino ora i valori medi di \hat{q} e \hat{p} calcolati su un autostato coerente:

$$\langle \hat{q} \rangle_\lambda = \sqrt{\frac{\hbar}{2m\omega}} \langle \psi_\lambda(x) | (a + a^\dagger) \psi_\lambda(x) \rangle = \sqrt{\frac{\hbar}{2m\omega}} (\lambda + \lambda^*) = \sqrt{\frac{2\hbar}{m\omega}} \Re(\lambda) = \sqrt{\frac{2\hbar}{m\omega}} |\lambda| \sin \theta$$

infatti risulta:

$$\langle \psi_\lambda(x) | a^\dagger \psi_\lambda(x) \rangle = \langle \psi_\lambda(x) | a \psi_\lambda(x) \rangle^*$$

per cui:

$$\langle \hat{q} \rangle_\lambda \propto \Re(\lambda)$$

in maniera analoga:

$$\langle \hat{p} \rangle_\lambda = \sqrt{\frac{\hbar m\omega}{2}} \langle \psi_\lambda(x) | \frac{(a - a^\dagger)}{i} \psi_\lambda(x) \rangle = \sqrt{\frac{\hbar m\omega}{2}} \frac{\lambda - \lambda^*}{i} = \sqrt{2\hbar m\omega} \frac{1}{i} \Im(\lambda) = \sqrt{2\hbar m\omega} \frac{1}{i} |\lambda| \cos \theta$$

ovvero:

$$\langle \hat{p} \rangle_\lambda \propto \Im \mathbf{m}(\lambda)$$

Per i valori quadratici si ha, applicando la relazione (5.5):

$$\begin{aligned} \langle \hat{q}^2 \rangle_\lambda &= \frac{\hbar}{2m\omega} \langle \psi_\lambda(x) | (a^2 + a^{\dagger 2} + aa^\dagger + a^\dagger a) \psi_\lambda(x) \rangle = \frac{\hbar}{2m\omega} \langle \psi_\lambda(x) | (a^2 + a^{\dagger 2} + 2a^\dagger a + 1) \psi_\lambda(x) \rangle = \\ &= \frac{\hbar}{2m\omega} (\lambda^2 + \lambda^{*2} + 2\lambda^* \lambda + 1) = \frac{\hbar}{2m\omega} (\lambda + \lambda^*)^2 + \frac{\hbar}{2m\omega} = \frac{2\hbar}{m\omega} \Re \mathbf{e}(\lambda)^2 + \frac{\hbar}{2m\omega} \end{aligned}$$

e

$$\begin{aligned} \langle \hat{p}^2 \rangle_\lambda &= -\frac{\hbar m\omega}{2} \langle \psi_\lambda(x) | (a^2 + a^{\dagger 2} - aa^\dagger - a^\dagger a) \psi_\lambda(x) \psi_\lambda(x) \rangle = \\ &= -\frac{\hbar m\omega}{2} \langle \psi_\lambda(x) | (a^2 + a^{\dagger 2} - 2a^\dagger a - 1) \psi_\lambda(x) \rangle = -\frac{\hbar m\omega}{2} (\lambda^2 + \lambda^{*2} - 2\lambda^* \lambda - 1) = \\ &= -\frac{\hbar m\omega}{2} (\lambda - \lambda^*)^2 + \frac{\hbar m\omega}{2} = -2\hbar m\omega \Im \mathbf{m}(\lambda)^2 + \frac{\hbar m\omega}{2} \end{aligned}$$

Lo scarto quadratico medio è dato allora da:

$$\begin{aligned} (\Delta q)^2 &= \langle \hat{q}^2 \rangle_\lambda - \langle \hat{q} \rangle_\lambda^2 = \frac{2\hbar}{m\omega} \cancel{\Re \mathbf{e}(\lambda)^2} + \frac{\hbar}{2m\omega} - \frac{2\hbar}{m\omega} \cancel{\Re \mathbf{e}(\lambda)^2} = \frac{\hbar}{2m\omega} \\ (\Delta p)^2 &= \langle \hat{p}^2 \rangle_\lambda - \langle \hat{p} \rangle_\lambda^2 = -2\hbar m\omega \cancel{\Im \mathbf{m}(\lambda)^2} + \frac{\hbar m\omega}{2} + 2\hbar m\omega \cancel{\Im \mathbf{m}(\lambda)^2} = \frac{\hbar m\omega}{2} \end{aligned}$$

ovvero:

$$(\Delta q)^2 (\Delta p)^2 = \frac{\hbar}{2m\omega} \frac{\hbar m\omega}{2} = \frac{\hbar^2}{4} \quad \Rightarrow \quad \Delta x \Delta p = \frac{\hbar}{2}$$

quindi, *gli stati coerenti realizzano la minima indeterminazione possibile permessa dal principio di Heisemberg. Gli stati coerenti realizzano quindi anche il comportamento più simile possibile a quello classico.*

Il momento angolare

6.1 Il momento angolare in meccanica quantistica

Come è noto, il momento angolare ha una stretta relazione con le rotazioni. Classicamente, esso è definito come:

$$\vec{L} = \vec{r} \times \vec{p}$$

si tratta quindi propriamente di uno pseudovettore.¹

Esplicitando le componenti, si ha:

$$\begin{aligned} l_x &= yp_z - zp_y \\ l_y &= zp_x - xp_z \\ l_z &= xp_y - yp_x \end{aligned}$$

ovvero in maniera più concisa:

$$l_i = \varepsilon_{ijk} q_j p_k$$

A queste quantità si può associare un operatore hermitiano, ad esempio l_x :

$$\hat{l}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y$$

Questo operatore è evidentemente hermitiano, in quanto compaiono sempre prodotti $\hat{q}\hat{p}$ lungo direzioni diverse e pertanto commutano:

$$\hat{l}_x^\dagger = \hat{y}^\dagger \hat{p}_z^\dagger - \hat{z}^\dagger \hat{p}_y^\dagger = \hat{p}_z \hat{y} - \hat{p}_y \hat{z} = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y = \hat{l}_x$$

Per potere effettuare però misure contemporanee delle varie componenti del momento angolare occorre che queste commutino fra loro, condizione che non è soddisfatta. Si prenda infatti la componente lungo y : $\hat{l}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z$ e si consideri il commutatore con \hat{l}_x , tenendo conto che $\hat{z}\hat{p}_y$ e $\hat{x}\hat{p}_z$ commutano:

$$\begin{aligned} [\hat{l}_x, \hat{l}_y] &= [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z] = [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] - [\hat{y}\hat{p}_z, \hat{x}\hat{p}_z] - [\hat{z}\hat{p}_y, \hat{z}\hat{p}_x] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z] = \\ &= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z] = \hat{y}\hat{p}_z \hat{z}\hat{p}_x - \hat{z}\hat{p}_x \hat{y}\hat{p}_z + \hat{z}\hat{p}_y \hat{x}\hat{p}_z - \hat{x}\hat{p}_z \hat{z}\hat{p}_y = \\ &= \hat{y}\hat{p}_x [\hat{p}_z, \hat{z}] + \hat{p}_y \hat{x} [\hat{z}, \hat{p}_z] = i\hbar(\hat{p}_y \hat{x} - \hat{y}\hat{p}_x) = i\hbar(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x) \equiv i\hbar \hat{l}_z \end{aligned}$$

Ne consegue quindi:

$$[\hat{l}_x, \hat{l}_y] = i\hbar \hat{l}_z$$

¹Il prodotto vettoriale definisce un oggetto indipendente dal sistema di riferimento scelto. Tale indipendenza è però relativa alle rotazioni e alle traslazioni del sistema, ma non vale per riflessioni speculari. Gli oggetti che hanno questa proprietà si chiamano *pseudovettori*.

Lo stesso calcolo può essere ripetuto per tutte le componenti, ottenendo la relazione:

$$[\hat{l}_a, \hat{l}_b] = i\hbar\varepsilon_{abc}\hat{l}_c$$

Questa relazione fondamentale dice che le componenti del momento angolare non commutano fra loro e pertanto non hanno una base comune di autovettori.² Il momento angolare ha le dimensioni di \hbar .

In generale, se il commutatore risulta nullo su uno stato, allora le componenti relative del momento angolare sono misurabili contemporaneamente, altrimenti no. La regola di commutazione pregiudica solo l'esistenza di una base di autovettori comune.

Si consideri ora (\hat{u}_i sono i versori lungo la direzione i):

$$\begin{aligned}\vec{l} &= l_x\hat{u}_x + l_y\hat{u}_y + l_z\hat{u}_z \\ l^2 &= l_x^2 + l_y^2 + l_z^2 = l_a l_a\end{aligned}$$

ne segue allora che:

$$[\hat{l}^2, \hat{l}_b] = [\hat{l}_a\hat{l}_a, \hat{l}_b] = \hat{l}_a[\hat{l}_a, \hat{l}_b] + [\hat{l}_a, \hat{l}_b]\hat{l}_a = i\hbar\varepsilon_{abc}\hat{l}_a\hat{l}_c + i\hbar\varepsilon_{abc}\hat{l}_c\hat{l}_a = i\hbar\varepsilon_{abc}(\hat{l}_a\hat{l}_c + \hat{l}_c\hat{l}_a)$$

Ora, $\hat{l}_a\hat{l}_c \neq \hat{l}_c\hat{l}_a$, ma il termine $\hat{l}_a\hat{l}_c + \hat{l}_c\hat{l}_a$ nel suo complesso è simmetrico, pertanto moltiplicato per la quantità antisimmetrica ε_{abc} fornisce zero. Risulta quindi:

$$[\hat{l}^2, \hat{l}_b] = 0 \quad \forall b$$

Il modulo quadro del momento angolare commuta con una sua componente, possono quindi essere misurati contemporaneamente ed esiste una base di autovettori in comune. Ha senso quindi ricercare la soluzione del problema:

$$\begin{cases} \hat{l}^2\psi(r, \theta, \phi) &= \lambda \psi(r, \theta, \phi) \\ \hat{l}_a\psi(r, \theta, \phi) &= m \psi(r, \theta, \phi) \end{cases}$$

CASO DI PIÙ PARTICELLE.

Nel caso il sistema sia composto da più particelle risulta:

$$\vec{L} = \sum_i \vec{l}_i = \sum_i \vec{r}_i \times \vec{p}_i$$

siccome i diversi momenti angolari \vec{l}_i commutano fra loro, per la risultante \vec{L} si può fare lo stesso discorso delle singole componenti:

$$[L_a, L_b] = i\hbar\varepsilon_{abc}L_c$$

e siccome vale sempre $L^2 = L_a^2 + L_b^2 + L_c^2$ segue anche che $[L^2, L_i] = 0$, pertanto si può impostare lo stesso problema agli autovalori, che fornirà lo stesso risultato riferito al vettore momento angolare totale.

Ne consegue che tutte le proprietà che saranno studiate in questo capitolo si possono applicare ugualmente bene al momento angolare totale di un sistema.

²Queste regole di commutazione hanno un analogo in meccanica classica:

$$\{l_a, l_b\} = \varepsilon_{abc}l_c$$

Naturalmente, in meccanica classica il fatto che le parentesi di Poisson siano non nulle non implica la non misurabilità contemporanea delle componenti del momento angolare, ma solo che non se ne può scegliere più di una componente nel ruolo di variabile canonica.

È interessante comunque sottolineare l'analogia fra variabili canoniche classiche e osservabili simultanei.

6.2 Autofunzioni di l^2 e l_z : le armoniche sferiche

Consideriamo quindi il problema impostato alla fine del paragrafo precedente:

$$\begin{cases} \hat{l}^2 \psi(r, \theta, \phi) = \lambda \psi(r, \theta, \phi) \\ \hat{l}_z \psi(r, \theta, \phi) = m \psi(r, \theta, \phi) \end{cases}$$

Nel seguito di questo paragrafo sarà generalmente omessa la notazione di operatore $\hat{\quad}$, inoltre sarà scelta come componente del momento angolare quella lungo z , essenzialmente per ragioni storiche. Infine, gli autovalori saranno resi numeri puri introducendo il fattore \hbar :

$$\begin{cases} l^2 \psi(r, \theta, \phi) = \lambda \hbar^2 \psi(r, \theta, \phi) \\ l_z \psi(r, \theta, \phi) = m \hbar \psi(r, \theta, \phi) \end{cases}$$

L'operatore momento angolare è dato da:

$$\vec{l} = \vec{r} \times \vec{p} \quad \rightarrow \quad \hat{l} = -i\hbar \vec{r} \times \vec{\nabla}$$

e in coordinate polari risulta:

$$\vec{\nabla} = \frac{\partial}{\partial r} \hat{u}_r + \frac{1}{r} \frac{\partial}{\partial \theta} \hat{u}_\theta + \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \hat{u}_\phi$$

ne consegue che il momento angolare in coordinate polari è espresso da:

$$\vec{l} = -i\hbar \left(\frac{\partial}{\partial \theta} \hat{u}_\theta - \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \hat{u}_\phi \right)$$

siccome poi risulta $\hat{u}_z \cdot \hat{u}_\varphi = -\sin \theta$, vale ($l_z = \hat{u}_z \cdot \vec{l}$):

$$l_z = -i\hbar \frac{\partial}{\partial \phi} \quad (6.1)$$

Analogamente, $\hat{l}^2 = \vec{l} \cdot \vec{l} = l_x^2 + l_y^2 + l_z^2$, e usando lo stesso procedimento utilizzato per ricavare l_z si trova:³

$$l^2 = -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{\cos \theta}{\sin \theta} \frac{\partial^2}{\partial \theta} \right)$$

perché a differenza di quella cartesiana, la base sferica ha i versori che dipendono dal punto considerato. Questa forma può essere riscritta come:

$$l^2 = -\hbar^2 \left[\frac{\partial}{\partial \cos \theta} \left(\sin^2 \theta \frac{\partial}{\partial \cos \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] \quad (6.2)$$

e rappresenta quindi l'equazione agli autovalori per il quadrato del momento angolare.

Si consideri ora l'equazione relativa alla componente z . Si noti preliminarmente che essendo possibile trovare delle autofunzioni comuni a l_z e l^2 , allora queste autofunzioni sono fattorizzabili come $\Psi(\theta, \phi) = Y(\theta) \cdot \psi(\phi)$.⁴

$$-i\hbar \frac{\partial}{\partial \phi} \psi(\phi) = m\psi(\phi) \quad \Rightarrow \quad \Psi(\theta, \phi) = e^{im\phi} \cdot Y(\theta) \quad \forall m \in \mathbb{C}$$

³In maniera alternativa, queste relazioni si possono ricavare esplicitamente notando che valgono le relazioni:

$$\begin{cases} x = r \sin \theta \cos \phi \\ y = r \sin \theta \sin \phi \\ z = r \cos \theta \end{cases}$$

e ricordando le relazioni di trasformazione:

$$\frac{\partial}{\partial x_i} = \frac{\partial r}{\partial x_i} \frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x_i} \frac{\partial}{\partial \theta} + \frac{\partial \phi}{\partial x_i} \frac{\partial}{\partial \phi}$$

⁴Ci si occupa per il momento solo della forma di questa soluzione e non della sua normalizzazione.

questa equazione ammette in linea di principio soluzioni per ogni $m \in \mathbb{C}$, tuttavia la richiesta che la soluzione abbia senso fisicamente porta a limitarne i possibili valori. In particolare, la richiesta che l'operatore l_z sia un osservabile, ovvero hermitiano, conduce alla condizione:

$$\langle \psi_1(\phi) | l_z \psi_2(\phi) \rangle = \langle l_z \psi_1(\phi) | \psi_2(\phi) \rangle \quad \psi(\phi) = e^{im\phi}$$

ovvero, esplicitando il prodotto scalare:

$$-i\hbar \int_0^{2\pi} \psi_1^*(\phi) \frac{\partial}{\partial \phi} \psi_2(\phi) d\phi = -i\hbar \psi_1^*(\phi) \psi_2(\phi) \Big|_0^{2\pi} + \hbar \int_0^{2\pi} \left(-i \frac{\partial}{\partial \phi} \psi_1(\phi) \right)^* \psi_2(\phi) d\phi$$

affinché l'operatore sia hermitiano deve quindi risultare:

$$-i\hbar \psi_1^*(\phi) \psi_2(\phi) \Big|_0^{2\pi} = 0$$

e quindi:

$$\psi_1^*(2\pi) \psi_2(2\pi) - \psi_1^*(0) \psi_2(0) = 0$$

ovvero, le funzioni devono essere periodiche con periodo 2π . Questo implica che:⁵

$$\boxed{\psi(\phi) = e^{im\phi} \quad m \in \mathbb{Z}_0}$$

Si consideri ora la (6.2). Effettuando la sostituzione:

$$\cos \theta = x \quad \rightarrow \quad x \in [-1, 1]$$

tenendo conto inoltre che $\frac{\partial}{\partial \phi} = l_z$, questa equazione diviene:

$$\begin{aligned} - \left\{ \frac{d}{dx} \left[(1-x^2) \frac{d}{dx} \right] + \frac{m^2}{1-x^2} \right\} Y(x) &= \lambda Y(x) \\ (1-x^2) \frac{d^2}{dx^2} Y(x) - 2x \frac{d}{dx} Y(x) + \left(\lambda - \frac{m}{1-x^2} \right) Y(x) &= 0 \end{aligned}$$

Per il teorema di Poincaré, che stabilisce che le soluzioni hanno una singolarità nei punti in cui i coefficienti si annullano o hanno singolarità, si devono escludere i punti $x = \pm 1, 0$. Ponendo quindi:⁶

$$Y(x) = (1-x^2)^{|m|/2} f(x)$$

si trova per $f(x)$:

$$(1-x^2)^{|m|/2} \{ (1-x^2) f''(x) - 2x(|m|+1) f'(x) + [\lambda - |m|(|m|+1)] \} = 0$$

siccome si sono esclusi i punti di discontinuità e quelli dove si annullano i coefficienti, la relazione di sopra diventa:

$$(1-x^2) f''(x) - 2x(|m|+1) f'(x) + [\lambda - |m|(|m|+1)] = 0$$

Ora, si può certamente trovare una soluzione nella forma di una serie di potenze, per cui ponendo:

$$f(x) = \sum_{n=0}^{\infty} a_n x^{n+\alpha}$$

e sostituendo (con la posizione $\eta \equiv \lambda - |m|(|m|+1)$):

$$\sum_{n=0}^{\infty} [(1-x^2)(n+\alpha)(n+\alpha-1) a_n x^{n+\alpha-2} - 2x(|m|+1)(n+\alpha) a_n x^{n+\alpha-1} + \eta a_n x^{n+\alpha}] = 0$$

⁵Volendo essere precisi al massimo, si noti che non è la $\psi(\phi)$ da sola che deve essere osservabile, ma il prodotto fattorizzato. Di conseguenza m potrebbe in linea di principio avere valori seminteri, questo caso è però comunque escluso nel momento in cui si richiede che l'operatore l_z sia ben definito.

⁶Il teorema di Fuchs garantisce che questa posizione sia lecita.

e svolgendo i prodotti che coinvolgono le x :

$$\sum_{n=0}^{\infty} [(n+\alpha)(n+\alpha-1)a_n x^{n+\alpha-2} - (n+\alpha)(n+\alpha-1)a_n x^{n+\alpha} - 2(|m|+1)(n+\alpha)a_n x^{n+\alpha} + \eta a_n x^{n+\alpha}] = 0$$

I coefficienti di $x^{n+\alpha-2}$ devono essere nulli per $n=0$ e $n=1$, per cui:

$$\begin{cases} \alpha(\alpha-1) = 0 \\ (\alpha+1)\alpha = 0 \end{cases} \Rightarrow \alpha = 0$$

per cui:

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} + \sum_{n=0}^{\infty} [-n(n-1)a_n x^n - 2n(|m|+1)a_n x^n + \eta a_n x^n] = 0$$

infatti, la derivata seconda fa sparire i primi due termini dello sviluppo in serie. Mandando quindi $n \rightarrow n-2$ in questo pezzo della sommatoria, questa può partire da 0 e mettendo tutto insieme si ricava:⁷

$$\begin{aligned} \sum_{n=0}^{\infty} [(n+2)(n+1)a_{n+2}x^n - n(n-1)a_n x^n - 2n(|m|+1)a_n x^n + \eta a_n x^n] &= 0 \\ \sum_{n=0}^{\infty} \{(n+2)(n+1)a_{n+2} - [n(n-1)a_n + 2n(|m|+1)a_n - \eta] a_n\} x^n &= 0 \end{aligned}$$

e quindi la relazione ricorsiva:

$$a_{n+2} = \frac{n(n-1) + 2n(|m|+1) - \eta}{(n+2)(n+1)} a_n$$

Occorre ora studiare la convergenza di questa serie. Risulta:

$$\lim_{n \rightarrow \infty} \frac{a_{n+2}}{a_n} = 1$$

per cui il raggio di convergenza è 1. Questo significa che si può scrivere:

$$\sum_{n=0}^{\infty} a_n x^n \approx P(x) + a_N x^{2N} \sum_{k=0}^{\infty} x^{2k}$$

con x^{2k} serie geometrica di ragione x^2 , quindi:

$$\sum_{n=0}^{\infty} a_n x^n \approx P(x) + \frac{a_N x^{2N}}{1-x^2}$$

pertanto in $x = \pm 1$ diverge.

Siccome le autofunzioni devono essere a quadrato sommabile in $[-1, 1]$, la serie deve ridursi ad un polinomio e quindi arrestarsi. Questo accade se si verifica la relazione:

$$a_{n+2} = 0 \quad \Rightarrow \quad n(n-1) + 2n(|m|+1) - \eta = 0$$

ovvero, ricordando la definizione di η :

$$\begin{aligned} n(n-1) + 2n(|m|+1) - \lambda + |m|(|m|+1) &= 0 \rightarrow \\ \lambda = n(n-1) + 2n + 2n|m| + |m|(|m|+1) &\rightarrow \\ = n(n+1) + 2n|m| + |m|(|m|+1) &\rightarrow \\ = (n+|m|)(n+|m|+1) & \end{aligned}$$

⁷La tecnica è del tutto analoga a quella utilizzata per l'oscillatore armonico.

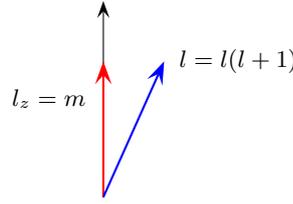
deve dunque risultare:

$$\lambda = l(l+1) \quad l \equiv n + |m|$$

gli autovalori di \hat{l}^2 e \hat{l}_z risultano quindi:

$$\boxed{\begin{cases} \hat{l}^2 \psi_{lm}(r, \theta, \phi) = l(l+1)\hbar^2 \psi_{lm}(r, \theta, \phi) \\ \hat{l}_z \psi_{lm}(r, \theta, \phi) = m\hbar \psi_{lm}(r, \theta, \phi) \end{cases} \quad \begin{cases} m \in \mathbb{Z}_0 \\ l = n + |m| \\ |m| \leq l \end{cases} \quad (6.3)}$$

Si noti un fatto importante: $l_z^2 < l^2$ tranne quando $l = 0$. Questa è una differenza notevole rispetto al caso classico: significa che la componente z del momento angolare non è mai lunga quanto il vettore stesso e quindi *il momento angolare non è mai allineato esattamente lungo una direzione prestabilita*:



Questa caratteristica è una conseguenza del principio di indeterminazione. Il fatto che il momento angolare non può mai essere orientato lungo una direzione precisa implica che anche se si riesce a stabilire la sua componente lungo z , nulla si può dire per quanto riguarda le due altre componenti l_x e l_y .

L'implicazione diretta è che *il momento angolare deve precedere intorno alla direzione scelta*. Si consideri infatti il valore medio della componente l_x sull'autostato $\psi_{lm}(r, \theta, \phi)$ con autovalore $l_z = m$:

$$\begin{aligned} \langle l_x \rangle_m &= \langle \psi_{lm}(r, \theta, \phi) | l_x \psi_{lm}(r, \theta, \phi) \rangle = \\ &= \frac{1}{i\hbar} \langle \psi_{lm}(r, \theta, \phi) | [l_y, l_z] \psi_{lm}(r, \theta, \phi) \rangle = \\ &= \frac{1}{i\hbar} \langle \psi_{lm}(r, \theta, \phi) | (l_y l_z - l_z l_y) \psi_{lm}(r, \theta, \phi) \rangle = \\ &= \frac{m}{i} (\langle \psi_{lm}(r, \theta, \phi) | l_y \psi_{lm}(r, \theta, \phi) \rangle - \langle \psi_{lm}(r, \theta, \phi) | l_y \psi_{lm}(r, \theta, \phi) \rangle) = 0 \end{aligned}$$

Un discorso analogo può essere fatto per il valore medio di l_y , si ricava pertanto:

$$\begin{cases} \langle l_x \rangle_m = 0 \\ \langle l_y \rangle_m = 0 \end{cases}$$

Siccome il valore medio della componente x e y del momento angolare è nullo e la componente lungo z è inferiore alla lunghezza totale, il momento angolare deve precedere.

ARMONICHE SFERICHE

Le autofunzioni contemporanee dell'operatore l^2 e l_z si indicano con la notazione $Y_l^m(\theta, \phi)$ e sono chiamate **armoniche sferiche**. Risulta quindi:⁸

$$Y_l^m(\theta, \phi) = P_l^m(\cos \theta) e^{im\phi}$$

⁸Si faccia attenzione che scritte così non sono normalizzate.

Esse sono estremamente importanti in meccanica quantistica e pertanto saranno trattate più approfonditamente.

I polinomi $P_l^0(\cos\theta)$ sono detti **polinomi di Legendre** e $P_l^m(\cos\theta)$ sono i cosiddetti **polinomi associati di Legendre**.

I polinomi di Legendre rappresentano l'ortogonalizzazione nell'intervallo $[-1, 1]$ della serie di potenze $\{1, x, x^2, \dots\}$ rispetto alla norma di \mathcal{L}^2 :

$$\int_{-1}^1 P_n(x)P_m(x)dx = \frac{2}{n+1}\delta_{nm}$$

e quindi concettualmente possono ricavarsi con il procedimento di ortonormalizzazione di Gram-Schmidt. Esiste tuttavia una formula ricorsiva, detta di Rodrigues, che permette di calcolarli più semplicemente:

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l$$

I polinomi associati di Legendre sono definiti a partire dai polinomi di Legendre secondo la regola:

$$P_l^m(x) = (1-x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_l^0(x) = \frac{1}{2^l l!} (1-x^2)^{\frac{m}{2}} \frac{d^{l+m}}{dx^{l+m}} (x^2 - 1)^l \quad |m| < l$$

Una volta normalizzate le armoniche sferiche hanno quindi la forma:

$$Y_l^m(\theta, \phi) = (-1)^m \sqrt{\frac{2l+1}{4\pi} \frac{(l-|m|)!}{(l+|m|)!}} P_l^{|m|}(\cos\theta) e^{im\phi} \quad |m| < l$$

e soddisfano dunque la proprietà di ortonormalità ($d\Omega = \sin\theta d\theta d\phi$):

$$\int Y_l^{m' *}(\theta, \phi) Y_l^m(\theta, \phi) d\Omega = \delta_{l'l'} \delta_{mm'}$$

e di completezza:

$$\sum_{l=0}^{+\infty} \sum_{m=-l}^{m=l} Y_l^{m*}(\theta', \phi') Y_l^m(\theta, \phi) = \frac{\delta(\theta - \theta') \delta(\phi - \phi')}{\sin\theta}$$

Le armoniche sferiche hanno importanti proprietà di simmetria. Sotto inversione completa di coordinate hanno parità pari a l :

$$Y_l^m(\pi - \theta, \phi - \pi) = (-1)^l Y_l^m(\theta, \phi)$$

mentre per inversione nel piano \widehat{xy} hanno parità pari a m :

$$Y_l^m(\theta, \phi + \pi) = (-1)^m Y_l^m(\theta, \phi)$$

infine, sotto inversione della coordinata z hanno parità pari a $l + m$:

$$Y_l^m(\pi - \theta, \phi) = (-1)^{l+m} Y_l^m(\theta, \phi)$$

Le prime armoniche sferiche sono:

- per $l = 0$:

$$Y_0^0(\theta, \phi) = \sqrt{\frac{1}{4\pi}}$$

- per $l = 1$:

$$Y_1^{-1} = \sqrt{\frac{3}{8\pi}} \sin\theta \cdot e^{-i\phi} \quad Y_1^0 = \sqrt{\frac{3}{4\pi}} \cos\theta \quad Y_1^1 = -\sqrt{\frac{3}{8\pi}} \sin\theta \cdot e^{i\phi}$$

- per $l = 2$:

$$Y_2^{-2} = \sqrt{\frac{15}{32\pi}} \sin^2 \theta \cdot e^{-2i\phi} \qquad Y_2^1 = -\sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta \cdot e^{i\phi}$$

$$Y_2^0 = \sqrt{\frac{15}{16\pi}} (3 \cos^2 \theta - 1)$$

$$Y_2^{-1} = \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta \cdot e^{-i\phi} \qquad Y_2^2 = \sqrt{\frac{15}{32\pi}} \sin^2 \theta \cdot e^{2i\phi}$$

EQUIVALENZA DELLE DIREZIONI NELLO SPAZIO

È interessante sottolineare che la quantizzazione del momento angolare non rompe comunque l'isotropia dello spazio. In altri termini anche se una componente del momento angolare lungo una determinata direzione risulta quantizzata e nulla si può dire per quanto riguarda le altre, tutte le direzioni dello spazio risultano ancora equivalenti.

Per mostrare questo, si consideri il caso di $l = 1$. Se si fissa una direzione z , allora esistono tre autostati di \hat{l}_z relativi agli autovalori $+1, 0, -1$, con autofunzioni rispettivamente (si considera qui la sola parte angolare):

$$\begin{aligned} \psi_+(\theta, \phi) &= Y_1^1(\theta, \phi) = N_+ \sin \theta e^{i\phi} \\ \psi_0(\theta, \phi) &= Y_1^0(\theta, \phi) = N_0 \cos \theta \\ \psi_-(\theta, \phi) &= Y_1^{-1}(\theta, \phi) = N_- \sin \theta e^{-i\phi} \end{aligned}$$

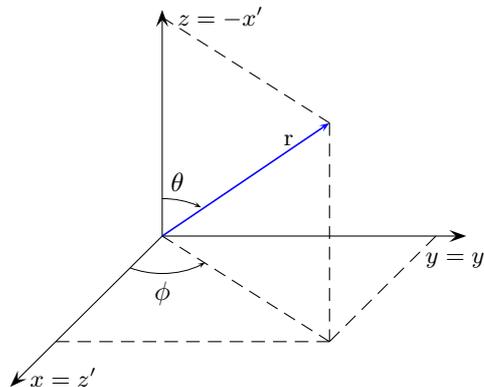
con $N_+ = -N_- = N$. Si supponga ora di trovarsi nell'autostato $l_z = 0$. L'operatore \hat{l}_x non ha valore definito in questo stato e si possono pertanto solo associare delle probabilità al possibile risultato di una misura. Essendo soggetto alle stesse limitazioni di \hat{l}_z , l'autostato di \hat{l}_x relativo a $l_z = 0$ si deve quindi scrivere come una combinazione lineare delle tre autofunzioni di \hat{l}_z (denominate con χ) relative ai possibili esiti di una misurazione:

$$\psi_0(\theta, \phi) = \alpha \chi_{1,1}(\theta, \phi) + \beta \chi_{1,0}(\theta, \phi) + \gamma \chi_{1,-1}(\theta, \phi)$$

dove $|\alpha|^2, |\beta|^2$ e $|\gamma|^2$ rappresentano evidentemente le probabilità di ottenere i tre risultati di una misura di l_x .

Per trovare gli autostati di \hat{l}_x , basta semplicemente ruotare il sistema di riferimento di $\pi/2$ intorno all'asse y , ovvero applicare le trasformazioni:

$$\begin{cases} x \rightarrow -z = x' \\ y \rightarrow y = y' \\ z \rightarrow x = z' \end{cases}$$



Torna ora utile notare che:⁹

$$\begin{aligned} N_0 \frac{z}{r} &= N_0 \cos \theta = Y_1^0(\theta, \phi) \\ N \frac{x}{r} &= N \sin \theta \cos \phi = \frac{1}{2} \sin \theta (e^{i\phi} + e^{-i\phi}) = N [Y_1^1(\theta, \phi) + Y_1^{-1}(\theta, \phi)] \end{aligned}$$

Ora, con la trasformazione di coordinate riportata sopra, si ricava:

$$N_0 \frac{z}{r} \rightarrow -N_0 \frac{x'}{r} = -N_0 [Y_1^1(\theta', \phi') + Y_1^{-1}(\theta', \phi')]$$

dove θ' e ϕ' indicano le coordinate angolari nel nuovo sistema di riferimento. Nel nuovo sistema di riferimento quindi c'è una uguale probabilità di trovare 1 o -1 come risultato di una misura di l_x e quindi $\langle l_x \rangle = 0$.

Ne consegue anche che per uno stato generico risulta:

$$\psi(\theta, \phi) = \sum_{n=-1}^1 c_n Y_1^n(\theta, \phi) = \sum_{k=-1}^1 c_k Y_1^k(\theta', \phi')$$

che rappresenta formalmente il fatto che *la quantizzazione non rompe l'isotropia dello spazio*.

6.3 Gli operatori gradino L_+ e L_-

Si seguirà ora un percorso analogo a quello presentato nel paragrafo §5.2 per definire precisamente lo spettro degli operatori momento angolare. Si considerino quindi gli operatori:

$$\begin{cases} L_+ = l_x + il_y \\ L_- = l_x - il_y \end{cases} \Rightarrow L_- = L_+^\dagger \quad (6.4)$$

Entrambi questi operatori commutano con l^2 ,⁽¹⁰⁾ quindi se agiscono su una autofunzione non ne cambiano l'autovalore:

$$l^2 L_+ \psi_{lm}(r, \theta, \phi) = L_+ l^2 \psi_{lm}(r, \theta, \phi) = l(l+1) \hbar^2 L_+ \psi_{lm}(r, \theta, \phi)$$

in termini formali, $L_+ \psi_{lm}(r, \theta, \phi)$ appartiene allo stesso sottospazio di L^2 . L_+ invece non commuta con l_z :

$$[l_z, L_+] = [l_z, l_x + il_y] = i\hbar l_y + \hbar l_x = \hbar(l_x + il_y) = \hbar L_+$$

Quindi risulta:

$$\begin{cases} [l_z, L_+] = \hbar L_+ \\ [l_z, L_-] = -\hbar L_- \end{cases}$$

Quello che è interessante è il comportamento di questi operatori su un'autofunzione di l_z :

$$\begin{aligned} l_z L_+ \psi_{lm}(r, \theta, \phi) &= (l_z L_+ - L_+ l_z + L_+ l_z) \psi_{lm}(r, \theta, \phi) = \\ &= [l_z, L_+] \psi_{lm}(r, \theta, \phi) + L_+ l_z \psi_{lm}(r, \theta, \phi) = \\ &= \hbar L_+ \psi_{lm}(r, \theta, \phi) + m \hbar L_+ \psi_{lm}(r, \theta, \phi) = \\ &= (m+1) \hbar \psi_{lm}(r, \theta, \phi) \end{aligned}$$

⁹Si utilizza banalmente la formula di Eulero:

$$\begin{cases} e^{i\phi} = \cos \phi + i \sin \phi \\ e^{-i\phi} = \cos \phi - i \sin \phi \end{cases}$$

da cui si ricava immediatamente:

$$e^{i\phi} + e^{-i\phi} = 2 \cos \phi$$

¹⁰Infatti l^2 commuta con ciascuna componente.

cioè l'operatore L_+ agendo su un autostato di l_z di autovalore m restituisce l'autostato di l_z di autovalore $(m+1)$. Ne consegue anche che se m è autovalore, anche $(m+1)$ è autovalore.

In maniera del tutto analoga si trova:

$$\begin{aligned} l_z L_- \psi_{lm}(r, \theta, \phi) &= (l_z L_- - L_- l_z + L_- l_z) \psi_{lm}(r, \theta, \phi) = \\ &= [l_z, L_-] \psi_{lm}(r, \theta, \phi) + L_- l_z \psi_{lm}(r, \theta, \phi) = \\ &= -\hbar L_- \psi_{lm}(r, \theta, \phi) + m \hbar L_- \psi_{lm}(r, \theta, \phi) = \\ &= (m-1) \hbar \psi_{lm}(r, \theta, \phi) \end{aligned}$$

e quindi l'operatore L_- agendo su un autostato di l_z di autovalore m restituisce l'autostato di l_z di autovalore $(m-1)$. Ne consegue anche che se m è autovalore, anche $(m-1)$ è autovalore.

Questo particolare comportamento degli operatori L_+ e L_- giustifica il loro nome di **operatori gradino**. Come anticipato all'inizio di questo paragrafo, è facile verificare guardando direttamente la forma di questi operatori che questi non sono hermitiani ma che l'uno è l'aggiunto dell'altro. La loro combinazione è invece autoaggiunta, infatti:

$$\begin{aligned} L_- L_+ &= l_x^2 + l_y^2 + i(l_x l_y - l_y l_x) = l_x^2 + l_y^2 - \hbar l_z = l^2 - \hbar l_z - l_z^2 \\ L_+ L_- &= l_x^2 + l_y^2 - i(l_x l_y - l_y l_x) = l_x^2 + l_y^2 + \hbar l_z = l^2 + \hbar l_z - l_z^2 \end{aligned} \quad (6.5)$$

Siccome queste due combinazioni sono autoaggiunte, e supponendo di lavorare con autofunzioni normalizzate, deve risultare:

$$\begin{aligned} \langle L_- L_+ \rangle &= \langle \psi_{lm}(r, \theta, \phi) | L_- L_+ \psi_{lm}(r, \theta, \phi) \rangle = \|L_+ \psi_{lm}(r, \theta, \phi)\| \geq 0 \\ \langle L_+ L_- \rangle &= \langle \psi_{lm}(r, \theta, \phi) | L_+ L_- \psi_{lm}(r, \theta, \phi) \rangle = \|L_- \psi_{lm}(r, \theta, \phi)\| \geq 0 \end{aligned}$$

ma $L_- L_+ \psi_{lm}(r, \theta, \phi) = \hbar^2[l(l+1) - m - m^2]$ e $L_+ L_- \psi_{lm}(r, \theta, \phi) = \hbar^2[l(l+1) + m - m^2]$ per cui, tenendo conto che $l > 0$, si arriva alle importanti relazioni:

$$\begin{cases} l(l+1) \geq m(m+1) & m > 0 \\ l(l+1) \geq m(m-1) & m < 0 \end{cases}$$

ora, nel caso di m negativo, risulta $m(m-1) \rightarrow -m(-m+1) = |m|(|m|+1)$, per cui le relazioni precedenti si possono riassumere nell'unica:

$$l(l+1) \geq |m|(|m|+1) \quad \Rightarrow \quad |m| \leq l$$

Di conseguenza, il valore di m deve essere limitato in alto e in basso. In altri termini, il processo di innalzamento e abbassamento dell'autovalore m non può continuare indefinitamente e devono esistere due particolari valori m_+ e m_- per i quali risulta:

$$\begin{aligned} \exists \quad \psi_{lm_+}(r, \theta, \phi) : L_+ \psi_{lm_+}(r, \theta, \phi) &= 0 \\ \exists \quad \psi_{lm_-}(r, \theta, \phi) : L_- \psi_{lm_-}(r, \theta, \phi) &= 0 \end{aligned}$$

Essendo possibile agire con l'operatore L_+ sullo stato $\psi_{lm_-}(r, \theta, \phi)$ fino a raggiungere lo stato $\psi_{lm_+}(r, \theta, \phi)$ sul quale fornisce zero, si ricava banalmente che:

$$m_+ = m_- + n \quad n \in \mathbb{N}_0 \quad (6.6)$$

Dalle posizioni (6.5) si ricava anche banalmente:

$$\begin{aligned} l^2 &= -\hbar L_- L_+ + l_z - l_z^2 \\ l^2 &= \hbar L_+ L_- + l_z - l_z^2 \end{aligned}$$

applicando quindi la prima di queste all'autostato del momento angolare relativo all'autovalore m_+ si ottiene:

$$l^2\psi_{lm_+}(r, \theta, \phi) = (L_-L_+ + l_z - l_z^2)\psi_{lm_+}(r, \theta, \phi) = \hbar^2 m_+(m_+ + 1)$$

e la seconda all'autostato relativo all'autovalore m_- :

$$l^2\psi_{lm_-}(r, \theta, \phi) = (L_+L_- + l_z - l_z^2)\psi_{lm_-}(r, \theta, \phi) = \hbar^2 m_-(m_- - 1)$$

siccome l'autovalore di l^2 è $l(l+1)$, dalle relazioni precedenti si ricavano le condizioni:

$$\begin{cases} l = m_+ \\ l = -m_+ - 1 \end{cases} \quad \text{oppure} \quad \begin{cases} l = -m_- \\ l = m_- - 1 \end{cases}$$

e siccome per definizione $m_+ \geq m_-$, l'unica scelta consistente con le due condizioni è porre:

$$\boxed{l = m_+ = -m_- \geq 0}$$

questa relazione unita alla (6.6) permette di ricavare l'altra relazione fondamentale:

$$2m_+ = n \rightarrow m_+ = \frac{n}{2} = l \quad n \in \mathbb{N}_0$$

e siccome ogni valore di m differisce per un intero da $m_+ = -m_- = l$, si ricava anche che:

$$2m \in \mathbb{Z}$$

Infine, siccome si è visto che vale:

$$-\frac{n}{2} \leq m \leq \frac{n}{2}$$

risulta anche che il momento angolare l può assumere valori interi o seminteri e la degenerazione della componente m del momento angolare rispetto a l vale $2l + 1$.

In realtà, i valori seminteri del momento angolare sono esclusi da considerazioni di carattere fisico-matematico. Se il valore $l = 1/2$ fosse possibile, allora si avrebbe:

$$\psi_{\frac{1}{2}, \frac{1}{2}}(r, \theta, \phi) \propto \sqrt{\sin \theta} e^{i\frac{\phi}{2}}$$

Dalla forma di L_- in coordinate sferiche:

$$L_- = -\hbar e^{i\phi} \left(\frac{\partial}{\partial \theta} - i \cot \theta \frac{\partial}{\partial \phi} \right)$$

si ricava che:

$$L_- \psi_{\frac{1}{2}, \frac{1}{2}}(r, \theta, \phi) \propto \frac{\hbar}{2} \cot \theta \sqrt{\sin \theta} e^{-i\frac{\phi}{2}}$$

che non è proporzionale a $\psi_{\frac{1}{2}, -\frac{1}{2}}(r, \theta, \phi)$, che invece è dato da:

$$\psi_{\frac{1}{2}, -\frac{1}{2}}(r, \theta, \phi) \propto \sqrt{\sin \theta} e^{-i\frac{\phi}{2}}$$

la funzione $\cot \theta \sqrt{\sin \theta} e^{i\frac{\phi}{2}}$ così ricavata inoltre non appartiene a \mathcal{L}^2 secondo la misura $\sin \theta d\theta d\phi$, ne consegue che i valori seminteri di l non sono accettabili.

La questione può essere vista sotto un altro aspetto notando che l'operatore differenziale che agisce sulle funzioni deve essere ad un solo valore ed hermitiano: e questo esclude appunto la possibilità di valori di l seminteri. D'altronde, una volta effettuata una rotazione di 2π intorno all'asse z la funzione d'onda deve raccordarsi in fase e pertanto può variare solo per numeri interi.

L'utilità degli operatori gradino si rivela in pieno nel calcolare determinate quantità. Ad esempio, in termini di questi operatori il valore medio di l_x è dato da:

$$\begin{aligned} \langle l_x \rangle &= \frac{1}{2} \langle \psi_{lm}(r, \theta, \phi) | (L_+ + L_-) \psi_{lm}(r, \theta, \phi) \rangle = \\ &= \langle \psi_{lm}(r, \theta, \phi) | \psi_{l, m+1}(r, \theta, \phi) \rangle + \langle \psi_{lm}(r, \theta, \phi) | \psi_{l, m-1}(r, \theta, \phi) \rangle = 0 \end{aligned}$$

data l'ortogonalità delle autofunzioni del momento angolare. In maniera perfettamente analoga si dimostra che $\langle l_y \rangle = 0$. Per quanto riguarda invece i valori medi di $\langle l_x^2 \rangle$ e $\langle l_y^2 \rangle$, ricordando la definizione di operatorio gradino (6.4), si ricava:

$$l_x^2 = \frac{1}{2} (L_+^2 + L_-^2 + L_+L_- + L_+L_-) \quad l_y^2 = -\frac{1}{2i} (L_+^2 + L_-^2 - L_+L_- - L_+L_-)$$

tenendo conto che $\langle L_+^2 \rangle = \langle L_-^2 \rangle = 0$, si ricava:

$$\langle l_x^2 \rangle = \langle l_y^2 \rangle = \frac{1}{4} \langle L_+L_- - L_+L_- \rangle$$

A partire da questi risultati, si ricava che:

$$\langle l^2 \rangle = \langle l_x^2 \rangle + \langle l_y^2 \rangle + \langle l_z^2 \rangle = 2\langle l_x^2 \rangle + \langle l_z^2 \rangle$$

da cui discende immediatamente che:

$$\langle l_x^2 \rangle = \frac{1}{2} [\langle l^2 \rangle - \langle l_z^2 \rangle] = \frac{\hbar^2}{2} [l(l+1) - m^2]$$

e risulta sempre diverso da zero tranne che nel caso $l = 0$.

Gli operatori gradino permettono anche di ricavare esplicitamente gli autostati relativi agli autovalori $|m| = l$ in maniera semplice. In base alle leggi di trasformazione delle coordinate polari:

$$\begin{cases} x = r \sin \theta \cos \phi \\ y = r \sin \theta \sin \phi \\ z = r \cos \theta \end{cases}$$

ed alla regola di trasformazione:

$$\frac{\partial}{\partial x_i} = \frac{\partial r}{\partial x_i} \frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x_i} \frac{\partial}{\partial \theta} + \frac{\partial \phi}{\partial x_i} \frac{\partial}{\partial \phi}$$

si ricava la forma esplicita degli operatori componenti del momento angolare:

$$\begin{aligned} l_x &= i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) \\ l_y &= i\hbar \left(-\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) \\ l_z &= -i\hbar \frac{\partial}{\partial \phi} \end{aligned}$$

in particolare, dalla forma di l_x e l_y e ricordando la formula di Eulero si ricava:

$$L_+ = l_x + il_y = \hbar e^{i\phi} \left(\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right)$$

e per l'autostato della parte angolare $\chi_{l,l}(\theta, \phi) = \chi(\theta)e^{il\phi}$ vale:

$$L_+\chi_{l,l}(\theta, \phi) = \hbar e^{i(l+1)\phi} \left[\frac{\partial}{\partial \theta} \chi(\theta) - l \cot \theta \chi(\theta) \right] = 0$$

ovvero:

$$\frac{\partial}{\partial \theta} \chi(\theta) = l \cot \theta \chi(\theta) \quad \Rightarrow \quad \frac{d\chi(\theta)}{\chi(\theta)} = l \cot \theta d\theta$$

che ha per soluzione $\chi(\theta) = (\sin \theta)^l$, quindi:

$$\chi_{l,l}(\theta, \phi) = N(\sin \theta)^l e^{il\phi} = Y_l^l(\theta, \phi)$$

e in maniera del tutto analoga:

$$\chi_{l,-l}(\theta, \phi) = N(\sin \theta)^l e^{-il\phi} = Y_l^{-l}(\theta, \phi)$$

6.4 Lo spin

A pagina 22 (§2.4) è stato descritto l'esperimento di Stern-Gerlach.¹¹ Come anticipato in quella sede, il risultato davvero sorprendente si ebbe quando fu utilizzato lo stato fondamentale dell'atomo di idrogeno, con $l = 0$.¹² Questi atomi non hanno momento angolare orbitale, di conseguenza nessuna corrente orbitale e pertanto non si dovrebbe osservare alcuna deviazione in un apparato alla Stern e Gerlach.

L'esperimento mostra invece un fascio diviso in due. L'interpretazione di questo risultato è altamente problematica, perché secondo la teoria quantistica, nello stato $l = 0$ si deve osservare un solo fascio e anche nel primo stato eccitato, con $l = 1$, si dovrebbero osservare tre fasci.¹³ L'esistenza di due fasci con la stessa intensità risulta quindi inspiegabile con le leggi della meccanica quantistica illustrate finora.

La soluzione di questo problema fu proposta da due studenti, Goudsmit e Uhlenbeck. Essi suggerirono "semplicemente" che l'elettrone stesso avesse un *momento angolare intrinseco*, che chiamarono **spin**. In questo approccio, Goudsmit e Uhlenbeck immaginarono che l'elettrone ruotasse su sé stesso, in analogia al moto di rotazione dei pianeti. Assumendo che il momento angolare associato all'elettrone valga $\hbar/2$, in base alle leggi di quantizzazione del momento angolare si trova che i soli valori possibili per la componente z sono $\hbar/2$ e $-\hbar/2$, perché come visto la componente z del momento angolare può variare solo per valori interi di \hbar .

Si noterà che questo è in realtà un argomento imposto *ad hoc*. La ragione fondamentale per cui la componente z del momento angolare può variare solo per unità \hbar è legata al fatto che le funzioni d'onda angolari devono raccordarsi in fase dopo un giro di 2π . Ora si trova una funzione d'onda che deve contenere un fattore $e^{i\phi/2}$, il quale dopo un giro completo di 2π vale -1 . Detto in termini pittorici, significa che dopo un giro completo un elettrone non mostra la stessa "faccia", ma che occorrono due giri affinché si presenti come all'inizio.

Difficoltà maggiori sorgono qualora si cerchi di costruire un modello che tenga conto del fatto che l'elettrone ruota. In linea di principio, una sfera carica che ruota crea delle correnti e non presenta problemi particolari. Tuttavia, una volta applicato questo modello all'elettrone si scopre che le regioni equatoriali dovrebbero ruotare a velocità superiori a quelle della luce per generare il momento angolare osservato.

Queste difficoltà sono tutt'altro che banali ed in effetti furono risolte intorno al 1930 da P.A.M. Dirac, il quale mostrò che impostando la teoria quantistica in maniera relativisticamente invariante, sia lo spin dell'elettrone sia le *matrici di spin* – che Pauli dovette introdurre a mano per poter trattare il momento magnetico intrinseco dell'elettrone – emergono in maniera naturale e spontanea. Come detto nella nota a pagina 23, la teoria di Dirac è tuttavia completamente al di là di questo corso.

Vale la pena notare che la questione dello spin è però molto importante, infatti è grazie ad esso che si riescono a spiegare le proprietà della tavola periodica degli elementi, ma anche questo è oggetto di corsi più approfonditi.

Si passerà quindi ora a delineare un approccio formale al problema dello spin, seguendo la strada delle matrici di Pauli.

Per via sperimentale si sa che il momento intrinseco dell'elettrone si comporta come un momento angolare e vale $s = 1/2$, $s_z = \pm 1/2$. Ne consegue che gli autovalori dell'operatore di spin devono essere:

$$\begin{cases} s(s+1)\hbar^2 = \frac{3}{4}\hbar^2 \\ s_z\hbar = \frac{\hbar}{2} \end{cases}$$

Siccome s è fissato, è sufficiente riferirsi alla sola componente z , si parla quindi di stati con spin $1/2$ o spin $-1/2$. Questo significa che per descrivere completamente lo stato di un elettrone non è

¹¹ Si può consultare sull'argomento l'eccellente articolo al sito:
http://galileo.phys.virginia.edu/classes/252/Angular_Momentum/Angular_Momentum.html.

¹² Phipps and Taylor, *Phys. Rev.* **29**, 309.

¹³ È appena il caso di notare a questo punto del corso che per $l = 1$ i valori possibili di l_z sono tre: $-1, 0, 1$.

sufficiente fornire la sua funzione d'onda, ma occorre specificare anche il valore di un grado di libertà aggiuntivo che è stato chiamato appunto *spin*.

Lo stato di un elettrone risulta quindi completamente descritto se in ogni punto dello spazio si aggiungono due ampiezze di probabilità che definiscono lo stato dello spin. In pratica, la funzione d'onda di un elettrone deve avere una forma del tipo:

$$\psi(\vec{r}, s) = \alpha\psi_+(\vec{r}) + \beta\psi_-(\vec{r})$$

dove $|\psi_+(\vec{r})|^2$ e $|\psi_-(\vec{r})|^2$ indicano rispettivamente la probabilità di trovare l'elettrone nella posizione \vec{r} con spin "su" ($s_z = \hbar/2$) o spin "giù" ($s_z = -\hbar/2$).

Siccome sono possibili solo due stati di spin, questo può essere descritto da uno spazio vettoriale a due dimensioni la cui base è quindi composta da due vettori. In questo contesto, la funzione d'onda $\psi(\vec{r}, s)$ di un elettrone può essere indicata come un vettore a due componenti:

$$\psi(\vec{r}, s) = \begin{pmatrix} \psi_+(\vec{r}) \\ \psi_-(\vec{r}) \end{pmatrix}$$

Lo spazio degli stati di un elettrone è quindi rappresentato dal prodotto tensoriale $\mathcal{H} = \mathcal{L}^2 \otimes \mathcal{S}$, ovvero dello spazio \mathcal{L}^2 studiato finora e dello spazio \mathcal{S} a due dimensioni dello spin. Il prodotto scalare in questo spazio è definito dalla relazione:

$$\langle \psi(\vec{r}) | \varphi(\vec{r}) \rangle = \int [\psi_+^*(\vec{r})\varphi_+(\vec{r}) + \psi_-^*(\vec{r})\varphi_-(\vec{r})] d\vec{r}$$

Gli operatori di spin commutano evidentemente con tutti gli operatori di \mathcal{L}^2 (posizione, impulso, energia, ecc.), in quanto relativi a due spazi vettoriali diversi.

Una funzione d'onda scritta in termini di un vettore a due componenti prende il nome di **spinore**.¹⁴ In particolare, fissata una base ortonormale χ_+ e χ_- in questo spazio bidimensionale, ad esempio lungo la direzione z , la parte di spin della funzione d'onda risulta espressa dalla combinazione lineare:

$$\psi(s) = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

dove i coefficienti dello sviluppo possono dipendere dal vettore posizione \vec{r} . In questo caso la probabilità di misurare un determinato spin dipende dalla posizione della particella.

La base ortonormale scelta χ_+ e χ_- è costituita evidentemente da due autovettori, deve quindi risultare:

$$\begin{aligned} \hat{s}_z\chi_+ &= \frac{\hbar}{2}\chi_+ & \text{ovvero} & \hat{s}_z \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \hat{s}_z\chi_- &= -\frac{\hbar}{2}\chi_- & \text{ovvero} & \hat{s}_z \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

I due autovettori relativi a valori di spin diversi sono evidentemente ortonormali, per cui risulta anche:

$$\begin{aligned} \langle \chi_+ | \hat{s}_z\chi_+ \rangle &= \frac{\hbar}{2} \\ \langle \chi_{\pm} | \hat{s}_z\chi_{\mp} \rangle &= 0 \\ \langle \chi_- | \hat{s}_z\chi_- \rangle &= -\frac{\hbar}{2} \end{aligned}$$

¹⁴Per questa definizione occorre in realtà anche che le due componenti trasformino tra loro per rotazioni, ma questo è il caso come sarà mostrato fra breve.

Da queste relazioni discende immediatamente che la rappresentazione matriciale dell'operatore \hat{s}_z è estremamente semplice:

$$\hat{s}_z = \begin{pmatrix} \frac{\hbar}{2} & 0 \\ 0 & -\frac{\hbar}{2} \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \equiv \frac{\hbar}{2} \sigma_z$$

Ora, è stato detto sopra che questo grado di libertà aggiuntivo dell'elettrone si comporta a tutti gli effetti come un momento angolare. Risulta quindi naturale ricalcarne in qualche modo il formalismo, e in particolare introdurre gli analoghi degli operatori gradino. Questi operatori agiranno quindi nel seguente modo sugli autovettori dello spazio di spin \mathcal{S} :

$$\begin{aligned} \hat{s}_+ \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \hbar \begin{pmatrix} 0 \\ 0 \end{pmatrix} && (\hat{s}_+ \chi_+ = 0) \\ \hat{s}_+ \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= \hbar \begin{pmatrix} 1 \\ 0 \end{pmatrix} && (\hat{s}_+ \chi_- = \hbar \chi_+) \end{aligned}$$

ovvero, in termini di rappresentazione matriciale:

$$\hat{s}_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad ; \quad \hat{s}_- = \hat{s}_+^\dagger = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

In analogia al momento angolare, se si suppone che valgano le relazioni:

$$\begin{aligned} \hat{s}_x &= \frac{1}{2}(s_+ + s_-) \\ \hat{s}_y &= -i\frac{1}{2}(s_+ - s_-) \end{aligned}$$

si può ricavare la forma esplicita degli operatori \hat{s}_x e \hat{s}_y :

$$\hat{s}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \equiv \frac{\hbar}{2} \sigma_x \quad ; \quad \hat{s}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \equiv \frac{\hbar}{2} \sigma_y$$

Le tre matrici σ_x , σ_y e σ_z sono le cosiddette **matrici di Pauli**.

Le regole di commutazione di queste matrici sono facilmente trovate:

$$\begin{aligned} \sigma_x \cdot \sigma_y &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = i\sigma_z \\ \sigma_y \cdot \sigma_x &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = i \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -i\sigma_z \end{aligned} \tag{6.7}$$

da cui:

$$[\sigma_x, \sigma_y] = \sigma_x \cdot \sigma_y - \sigma_y \cdot \sigma_x = 2i\sigma_z$$

Se si effettua il calcolo per tutte le altre combinazioni¹⁵ si ricava la regola generale di commutazione per le matrici di Pauli:

$$[\sigma_j, \sigma_k] = 2i\varepsilon_{jkl}\sigma_l$$

Queste matrici sono inoltre unitarie in quanto $\sigma_i^2 = \mathbb{I}$ e, come si può vedere direttamente dalla (6.7), soddisfano la relazione $\sigma_i \cdot \sigma_j + \sigma_j \cdot \sigma_i = 0$ con $i \neq j$, ovvero la proprietà di anticommutazione:¹⁶

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}$$

Queste matrici insieme alla matrice identità costituiscono una base nello spazio delle matrici 2×2 .

¹⁵Calcolo che non sarà eseguito qui in quanto *veramente* banale e identico a quello mostrato.

¹⁶Sinteticamente, si dice che queste matrici soddisfano un'algebra di Clifford, una struttura che generalizza i numeri complessi e i quaternioni:

$$\sigma_i \cdot \sigma_j = \delta_{ij} \cdot \mathbb{I} + i\varepsilon_{ijk}\sigma_k$$

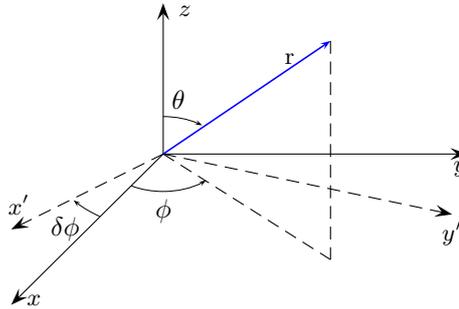
Tenendo infine conto della relazione $\hat{s}_i = \frac{\hbar}{2}\sigma_i$ si ricavano le regole di commutazione degli operatori di spin:

$$[\hat{s}_i, \hat{s}_j] = i\hbar\varepsilon_{ijk}\hat{s}_k$$

ovvero soddisfano le relazioni di commutazione del momento angolare. Questo non deve sorprendere e non deve trarre in inganno, perché nelle ipotesi iniziali si era immesso “a mano” proprio il fatto che lo spin si comportasse come un momento angolare quantistico. Questa relazione, quindi, può al massimo confermare che l’apparato formale dello spin costruito qui è coerente con queste premesse.

6.5 Il momento angolare come generatore infinitesimo delle rotazioni

Si consideri una rotazione infinitesima di un angolo $\delta\phi$ intorno all’asse z :



Come è noto, le leggi di trasformazione di una generica rotazione α intorno all’asse z sono date da:

$$\begin{cases} x' = x \cos \alpha - y \sin \alpha \\ y' = x \sin \alpha + y \cos \alpha \\ z' = z \end{cases}$$

ovvero in forma matriciale:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R_z(\alpha) \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad ; \quad R_z(\alpha) = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (6.8)$$

Le matrici per le altre due rotazioni, intorno all’asse x e y , sono date rispettivamente da:

$$R_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix} \quad ; \quad R_y(\alpha) = \begin{pmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{pmatrix} \quad (6.9)$$

Si supponga ora che uno stato sia descritto dalla funzione d’onda $\psi(r, \theta, \phi)$. Una rotazione infinitesima di un angolo $\delta\phi$ intorno all’asse z può essere espressa nella forma seguente:¹⁷

$$\psi(r, \theta, \phi + \delta\phi) \simeq \psi(r, \theta, \phi) + \delta\phi \left. \frac{\partial}{\partial \phi} \psi(r, \theta, \phi) \right|_{r, \theta, \phi} + O(2)$$

ma ricordando la forma dell’operatore \hat{l}_z , si ricava che:

$$\hat{l}_z = -i\hbar \frac{\partial}{\partial \phi} \Rightarrow \frac{\partial}{\partial \phi} = \frac{i}{\hbar} \hat{l}_z$$

per cui:

$$\psi(r, \theta, \phi + \delta\phi) \simeq \psi(r, \theta, \phi) + \frac{i}{\hbar} \hat{l}_z \psi(r, \theta, \phi) \delta\phi = \left(\mathbb{I} - \frac{i}{\hbar} \delta\phi \hat{l}_z \right) \psi(r, \theta, \phi)$$

¹⁷Si tratta di una banale espansione in serie di Taylor sulla variabile ϕ .

Un modo alternativo per mostrare questa relazione è il seguente. Si consideri la funzione d'onda $\psi(x, y, z)$. Una generica rotazione intorno all'asse z , descritta dall'operatore matrice $R_z(\alpha)$, fornisce:

$$R_z(\alpha)\psi(x, y, z) = \psi(x \cos \alpha + y \sin \alpha, -x \sin \alpha + y \cos \alpha, z)$$

Una rotazione infinitesima di un angolo $\delta\phi$ intorno allo stesso asse z può essere pertanto espressa come:¹⁸

$$R_z(\delta\phi)\psi(x, y, z) \simeq \psi(x + y\delta\phi, -x\delta\phi + y, z) \simeq \psi(x, y, z) + \delta\phi \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right) \psi(x, y, z)$$

ma $\left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right)$ è proprio la componente l_z lungo z del momento angolare.

In definitiva, l'operatore di rotazione infinitesima intorno a z è dato da:

$$R_z(\delta\phi)\psi(x, y, z) = \underbrace{\left(\mathbb{I} - \frac{i}{\hbar} \delta\phi \hat{l}_z \right)}_{\substack{\text{operatore} \\ \text{infinitesimo} \\ \text{di rotazione}}} \psi(x, y, z)$$

la componente i -sima \hat{l}_i del momento angolare è quindi il generatore infinitesimo delle rotazioni intorno all'asse i .

Una rotazione finita può essere ottenuta come una serie di N (al limite: infinite) rotazioni infinitesime in sequenza. La rotazione di un angolo finito α può quindi essere ottenuta come $N\delta\alpha \equiv N \frac{\alpha}{N}$. In sostanza quindi una rotazione finita di un angolo α si può scrivere come un limite sulle rotazioni infinitesime:

$$\psi(r, \theta, \phi + \alpha) = \lim_{N \rightarrow \infty} \left(\mathbb{I} + \frac{i}{\hbar} \frac{\alpha}{N} \hat{l}_z \right)^N \psi(r, \theta, \phi) = e^{i\alpha \hat{l}_z / \hbar} \psi(r, \theta, \phi)$$

ne consegue che l'operatore $e^{i\alpha \hat{L} / \hbar}$ è il generatore delle rotazioni finite.

Effettivamente, si può anche dimostrare che le rotazioni soddisfano le stesse regole di commutazione del momento angolare. Si considerino ad esempio le due matrici di rotazione intorno agli assi x e y date in (6.9). Siccome si considerano rotazioni infinitesime, si possono sviluppare queste matrici arrestandosi al secondo ordine:

$$R_x(\epsilon) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \epsilon^2/2 & -\epsilon \\ 0 & \epsilon & 1 - \epsilon^2/2 \end{pmatrix} \quad ; \quad R_y(\epsilon) = \begin{pmatrix} 1 - \epsilon^2/2 & 0 & \epsilon \\ 0 & 1 & 0 \\ -\epsilon & 0 & 1 - \epsilon^2/2 \end{pmatrix}$$

da cui, tenendo conto solo dei termini fino a ϵ^2 e trascurando di termini $O(3)$:

$$R_x(\epsilon)R_y(\epsilon) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \epsilon^2/2 & -\epsilon \\ 0 & \epsilon & 1 - \epsilon^2/2 \end{pmatrix} \cdot \begin{pmatrix} 1 - \epsilon^2/2 & 0 & \epsilon \\ 0 & 1 & 0 \\ -\epsilon & 0 & 1 - \epsilon^2/2 \end{pmatrix} = \begin{pmatrix} 1 - \epsilon^2/2 & 0 & \epsilon \\ \epsilon^2 & 1 - \epsilon^2/2 & -\epsilon \\ -\epsilon & \epsilon & 1 - \epsilon^2 \end{pmatrix}$$

$$R_y(\epsilon)R_x(\epsilon) = \begin{pmatrix} 1 - \epsilon^2/2 & 0 & \epsilon \\ 0 & 1 & 0 \\ -\epsilon & 0 & 1 - \epsilon^2/2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \epsilon^2/2 & -\epsilon \\ 0 & \epsilon & 1 - \epsilon^2/2 \end{pmatrix} = \begin{pmatrix} 1 - \epsilon^2/2 & \epsilon^2 & \epsilon \\ 0 & 1 - \epsilon^2/2 & -\epsilon \\ -\epsilon & \epsilon & 1 - \epsilon^2 \end{pmatrix}$$

per cui risulta:

$$R_x(\epsilon)R_y(\epsilon) - R_y(\epsilon)R_x(\epsilon) = \begin{pmatrix} 0 & -\epsilon^2 & 0 \\ \epsilon^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

e tenendo presente che la matrice di rotazione intorno all'asse z (6.8) approssimata al primo ordine in ϵ fornisce:

$$R_z(\epsilon) = \begin{pmatrix} 1 & -\epsilon & 0 \\ \epsilon & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

¹⁸Essendo una rotazione infinitesima $\sin \delta\phi \simeq \delta\phi$ e $\cos \delta\phi \simeq 1$.

si ricava:

$$R_x(\epsilon)R_y(\epsilon) - R_y(\epsilon)R_x(\epsilon) = R_z(\epsilon^2) - \mathbb{I}$$

Se si associano quindi degli operatori di rotazione infinitesima fatti come:

$$\begin{cases} \hat{R}_x(\epsilon) &= \mathbb{I} + \frac{i}{\hbar}\epsilon\hat{l}_x - \frac{\epsilon^2}{\hbar^2}\hat{l}_x^2 \\ \hat{R}_y(\epsilon) &= \mathbb{I} + \frac{i}{\hbar}\epsilon\hat{l}_y - \frac{\epsilon^2}{\hbar^2}\hat{l}_y^2 \end{cases}$$

risulta:

$$[\hat{R}_x(\epsilon), \hat{R}_y(\epsilon)] = -\frac{\epsilon^2}{\hbar^2}[\hat{l}_x, \hat{l}_y] = -i\frac{\epsilon^2}{\hbar^2}\hat{l}_z = (\hat{R}_z(\epsilon) - \mathbb{I})\hat{l}_z$$

6.6 Addizione di momenti angolari

Sia ora:

$$\vec{J} = \vec{J}_1 + \vec{J}_2$$

L'operatore vettoriale che caratterizza il momento angolare di un sistema, somma di due momenti angolari. Si assuma che i due momenti angolari del sistema \vec{J}_1 e \vec{J}_2 siano compatibili, ovvero che una misura dell'uno non influenzi il valore dell'altro. Formalmente, questo significa che i due operatori commutano fra loro.

Le varie componenti del momento angolare sono additive, per cui risulta:

$$(J_x, J_y, J_z) = (J_{1x}, J_{1y}, J_{1z}) + (J_{2x}, J_{2y}, J_{2z}) = (J_{1x} + J_{2x}, J_{1y} + J_{2y}, J_{1z} + J_{2z})$$

Naturalmente, il modulo quadro non è additivo:

$$J^2 = (\vec{J}_1 + \vec{J}_2)^2 = J_1^2 + J_2^2 + 2\vec{J}_1 \cdot \vec{J}_2 = J_1^2 + J_2^2 + 2(J_{1x}J_{2x} + J_{1y}J_{2y} + J_{1z}J_{2z})$$

Si vuole studiare lo spettro risultante del momento angolare totale, ovvero degli operatori J^2 e \hat{J}_z .

È bene notare che si tratta comunque di momenti angolari, quindi esiste una base di autovettori comuni a \hat{J} e \hat{J}_z e per gli autovalori simultanei valgono le stesse regole espone fin qui. In particolare, nell'autospazio in comune V_j gli autovalori j di \hat{J}^2 e m di \hat{J}_z soddisfano la stessa regola $-j \leq m \leq j$. Questo autospazio prende il nome di *multipletto*.

Risulta importante stabilire le relazioni che sussistono fra gli autovalori e le autofunzioni del momento angolare totale e degli addendi.

Si considerino due momenti angolari addendi e il momento angolare che risulta dalla loro somma. Si denotino gli operatori e i numeri quantici relativi ai moduli quadri e alle componenti z con la notazione usata fin qui.

Fissati i due numeri quantici relativi ai moduli quadri dei momenti addendi, j_1 e j_2 , il numero quantico relativo al modulo quadro del momento risultante deve appartenere all'intervallo $[|j_1 - j_2|, j_1 + j_2]$.

Fissato il numero quantico relativo al modulo quadro del momento angolare totale, il numero quantico relativo alla componente z del momento angolare totale può assumere i valori appartenenti all'intervallo $[-j, j]$.

Questo teorema esprime semplicemente la proprietà triangolare, secondo la quale la lunghezza del lato di un triangolo è sempre compreso tra la somma degli altri due e la loro differenza. Questo vale evidentemente anche per i moduli dei due vettori addendi e della loro somma, in quanto i vettori formano un triangolo. Questa è però una dimostrazione intuitiva, in quanto j è il numero quantico associato al modulo quadro del momento angolare risultante e non al suo modulo. La dimostrazione rigorosa è più complessa e viene riportata qui di seguito.

DIMOSTRAZIONE RIGOROSA DELLA REGOLA DI SOMMA DEI MOMENTI ANGOLARI

Si noti innanzitutto che l'autospazio risultante della somma dei momenti angolari è la somma diretta degli autospazi dei momenti angolari addendi. Quindi, se le autofunzioni che formano una base comune di \hat{J}_1^2 e \hat{J}_{1z} appartengono ad uno spazio di Hilbert \mathcal{H}_1 e le autofunzioni che formano una base comune di \hat{J}_2^2 e \hat{J}_{2z} appartengono ad un'altro spazio di Hilbert \mathcal{H}_2 , allora le autofunzioni che formano una base comune di \hat{J}^2 e \hat{J}_z appartengono allo spazio di Hilbert $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ prodotto tensoriale dei due.¹⁹

Inoltre, se si considera che il momento angolare è il generatore infinitesimo delle rotazioni, allora la somma di due momenti angolari può essere vista come l'applicazione di due distinte rotazioni al sistema. Passando da queste rotazioni all'operatore unitario che le rappresenta, si vede che questo corrisponde alla somma dei due addendi.

Formalmente, questo si può indicare nel modo seguente:

$$\begin{array}{ll} |j_1, m_1\rangle & \text{Autostato simultaneo di } \hat{J}_1^2 \text{ e } \hat{J}_{1z} \\ |j_2, m_2\rangle & \text{Autostato simultaneo di } \hat{J}_2^2 \text{ e } \hat{J}_{2z} \\ |j_1, m_1\rangle \otimes |j_2, m_2\rangle = |j_1, m_1; j_2, m_2\rangle & \text{Autostato simultaneo di } \hat{J}^2 \text{ e } \hat{J}_z \end{array} \quad (6.10)$$

Naturalmente, per il momento angolare addendo \hat{J}_1 , una volta fissato il numero quantico j_1 il numero quantico azimutale m_{1z} può assumere i $(2j_1 + 1)$ valori $\{-j_1, -j_1 + 1, \dots, j_1 - 1, j_1\}$. Ne consegue che i corrispondenti $(2j_1 + 1)$ autovettori sono degeneri rispetto all'autovalore $j_1(j_1 + 1)\hbar^2$ di \hat{J}_1^2 , ma distinti rispetto all'autovalore $m_1\hbar$ di \hat{J}_{1z} , con $m_1 \in \{-j_1, -j_1 + 1, \dots, j_1 - 1, j_1\}$. Lo stesso discorso può essere fatto per \hat{J}_2 .

Si consideri ora la base determinata da \hat{J} e \hat{J}_z . Siccome la componente z del momento angolare è additiva, gli autovettori di \hat{J}_z sono espressi dai prodotti tensoriali $\hat{J}_{1z} \otimes \hat{J}_{2z}$ degli autovettori delle singole componenti.²⁰ Gli autovalori saranno quindi dati di conseguenza dalla somma $m = m_1 + m_2$.

Questo discorso non può essere esteso al momento angolare totale \hat{J} , perché l'autostato relativo non risulta additivo: $\hat{J}^2 \neq \hat{J}_1^2 + \hat{J}_2^2$. Rimane comunque il fatto che l'autospazio del momento risultante è la somma diretta dei due autospazi dei momenti addendi e questo implica che ogni autostato comune di \hat{J}^2 e \hat{J}_z si possa scrivere come prodotto tensoriale $\langle j_1, m_1 | \times \langle j_2, m_2 |$ dei due autostati di \hat{J}_1^2 e \hat{J}_{1z} e di \hat{J}_2^2 e \hat{J}_{2z} .

Si consideri ora il caso di j_1 e j_2 fissati. Ci si vuole limitare a studiare l'autospazio generato dagli autovettori che abbiano tutti i valori possibili di m_1 e m_2 . Formalmente, questo autospazio è indicato con:

$$\{\langle j_1, m_1 | \otimes \langle j_2, m_2 | \}_{-j_1 \leq m_1 \leq j_1; -j_2 \leq m_2 \leq j_2}^{j_1; j_2}$$

che, siccome i valori di j sono fissati, può essere indicata sinteticamente con $\langle m_1, m_2 |$. Si tratta di un insieme di autovettori comuni di \hat{J} e \hat{J}_z e siccome $-j_1 \leq m_1 \leq j_1$ e $-j_2 \leq m_2 \leq j_2$ la dimensione di questo autospazio è $(2j_1 + 1) \times (2j_2 + 1)$. Si consideri ora un determinato autovalore $m\hbar$ di \hat{J}_z . Poiché vale l'additività e risulta $m = m_1 + m_2$, ne consegue che l'operatore \hat{J}_z è degenero in quanto diverse coppie di \hat{J}_1, \hat{J}_2 (ovvero m_1 e m_2) danno origine allo stesso autovalore m .

Si chiami questo autospazio degenero $W_m = \{|m_1, m_2\rangle\} : m_1 + m_2 = m$ fissato.

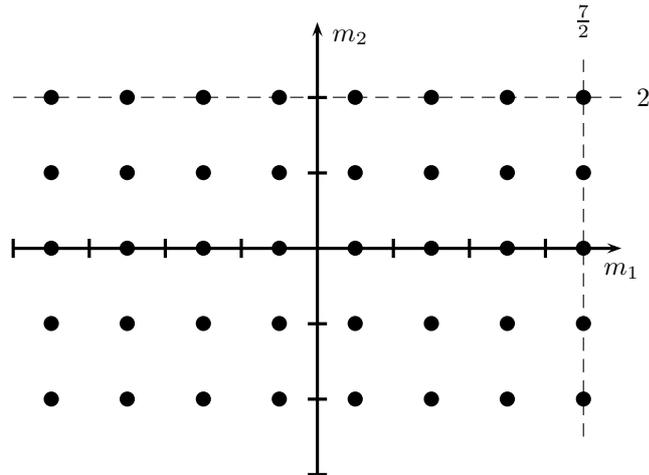
Evidentemente, quando entrambi m_1 e m_2 assumono il massimo valore non c'è degenerazione, perché esiste una sola scelta possibile e quindi $m = m_1 + m_2 = j_1 + j_2$. L'autovettore di \hat{J}_z relativo all'autovalore massimo è quindi unico.

L'autospazio W_m relativo all'autovalore $m = j_1 + j_2 - 1$ ha evidentemente dimensione 2 perché è generato da due autovettori ortogonali: $|j_1 - 1, j_2\rangle$ e $|j_1, j_2 - 1\rangle$.

Per proseguire nello studio della dimensione del sottospazio W_m è meglio a questo punto seguire un approccio grafico. Si mettano in un piano cartesiano i valori di m_1 sulle ascisse e i valori di m_2 sulle ordinate. In questo modo si viene a costituire un reticolo finito di passo unitario costituito da un rettangolo di dimensioni $(2j_1 + 1) \times (2j_2 + 1)$. Si considerino ad esempio i due momenti angolari $j_1 = 7/2$ e $j_2 = 2$:

¹⁹Ciò deriva senz'altro dall'aver assunto i due momenti angolari compatibili, per cui i due spazi di Hilbert non si "sovrappongono".

²⁰Questo deriva dall'applicazione del metodo di separazione delle variabili.

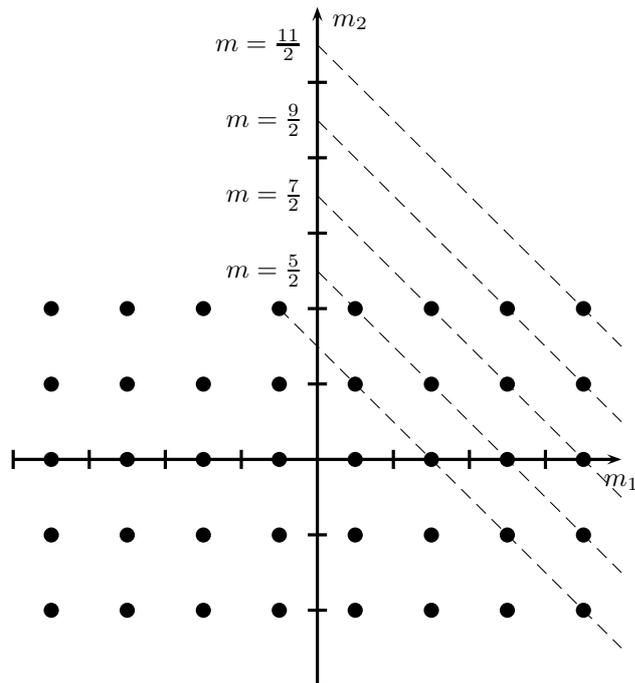


Si noti che siccome j_1 è semidispari il valore $m = 0$ non è permesso, valore che invece è possibile per j_2 intero.

Il problema della degenerazione in questo modo può essere risolto abbastanza facilmente. L'equazione $m_1 + m_2 = m$ può infatti essere riscritta come:

$$m_2 = -m_1 + m$$

che rappresenta una retta di pendenza 45° ⁽²¹⁾ che interseca l'asse delle ordinate in m :



Si vede quindi chiaramente che quando m ha il valore massimo ($m = j_1 + j_2$) non esiste degenerazione. Man mano che m decresce di una unità, la degenerazione aumenta a sua volta di una unità fino a diventare massima quando $m = -|j_1 - j_2|$. Il valore della degenerazione vale qui $[2 \min(j_1, j_2) + 1]$.

Da questo punto in poi, diminuendo ancora di una unità il valore di m la degenerazione diminuisce di una unità alla volta fino a quando m raggiunge il minimo valore possibile ($m = -j_1 - j_2$), dove la degenerazione si annulla in quanto esiste una sola coppia di valori che può dare origine a questo valore.

²¹Il coefficiente angolare è -1.

Si consideri ora l'operatore \hat{J}^2 e l'autospazio W_m introdotto poco sopra. Si considerino gli operatori gradino associati al momento angolare totale:

$$\begin{aligned}\hat{J}_+ &\equiv \hat{L}_x + i\hat{L}_y \\ \hat{J}_- &\equiv \hat{L}_x - i\hat{L}_y\end{aligned}$$

Se si applica l'operatore di innalzamento all'autospazio W_m questo si trasforma nell'autospazio W_{m+1} in quanto per definizione tutti i vettori di W_m sono autovettori di \hat{J}_z relativi all'autovalore m , e quindi l'innalzamento li fa divenire autovettori relativi all'autovalore $m+1$.

D'altra parte, per valori $|j_1 - j_2| \leq m \leq j_1 + j_2$ la dimensione²² di W_{m+1} è di un'unità più piccola di W_m . La conseguenza è che tutti gli autospazi W_m devono contenere un vettore – lo si denoti con $|\psi\rangle$ – che viene trasformato dall'operatore \hat{J}_+ nel vettore nullo: $\hat{J}_+|\psi\rangle = |0\rangle$.

Ricordando ora la relazione (6.5) sugli operatori gradino, si ricava:

$$||0\rangle|^2 = 0 = |\hat{J}_+|\psi\rangle|^2 = \langle\psi| \hat{J}_- \hat{J}_+ |\psi\rangle = \langle\psi| \hat{J}^2 - \hat{J}_z (J_z + 1) |\psi\rangle = \langle\psi| \hat{J}^2 |\psi\rangle - m(m+1)|\psi\rangle = 0$$

ovvero:

$$\langle\psi| \hat{J}^2 |\psi\rangle = m(m+1)|\psi\rangle$$

Ma $|\psi\rangle$ è un autovettore di \hat{J}_z relativo all'autovalore m e di \hat{J}^2 relativo ad un certo autovalore $j(j+1)$ tale che $-j \leq m \leq j$. Ne consegue quindi che:

$$j(j+1) = m(m+1) \quad \Rightarrow \quad j = m$$

L'autovettore dell'autospazio W_m annichilato dall'operatore \hat{J}_+ è quello relativo all'autovalore j , ovvero quello per cui vale $j = m$.

È utile esprimere questo risultato anche in una forma diversa. Solo gli autospazi W_m che sono formati da autovettori di \hat{J}_z con un numero quantico m tale che $|j_1 - j_2| \leq m \leq j_1 + j_2$ contengono un elemento che è anche autovettore di \hat{J}^2 con numero quantico $j = m$. Gli autospazi invece con m non compreso fra $|j_1 - j_2|$ e $j_1 + j_2$ non contengono un tale autovettore. Questo dimostra che questi autovettori particolari soddisfano la relazione:

$$|j_1 - j_2| \leq j \leq j_1 + j_2$$

Questa relazione prende il nome di *relazione triangolare* ed esprime il fatto che i due vettori addendi formano un parallelogramma e che il vettore risultante è sempre inferiore alla somma dei due addendi e superiore alla loro differenza.

Incidentalmente, si dimostra facilmente che non esistono altri numeri quantici. Infatti, negli autospazi W_m con $m \notin [|j_1 - j_2|, j_1 + j_2]$ non esistono autovalori annichilati da \hat{J}_+ e quindi nessuno ha numero quantico $m = j$.

Sia j_1 che j_2 possono essere interi o seminteri. Come si è visto, il massimo valore che può assumere il numero quantico j è $j_1 + j_2$ e può variare solo per unità. Ne consegue che:

$$\begin{aligned}j_1 \text{ e } j_2 \text{ entrambi interi o semidispari} &\quad \Rightarrow \quad j \text{ è intero} \\ j_1 \text{ e } j_2 \text{ uno intero e l'altro semidispari} &\quad \Rightarrow \quad j \text{ è semidispari}\end{aligned}$$

da cui:

$$j + j_1 + j_2 \text{ è sempre intero}$$

All'inizio di questo discorso la base simultanea di \hat{J} e \hat{J}_z era stata indicata in (6.10) come:

$$|j_1, m_1\rangle \otimes |j_2, m_2\rangle = |j_1, m_1; j_2, m_2\rangle = |m_1, m_2\rangle$$

L'ultima notazione è giustificata dal fatto che dare i valori di m_1 e m_2 permette di indentificare completamente lo stato.²³ Questa rappresentazione indica esplicitamente il fatto che questi autovettori sono ottenuti come prodotti tensoriali degli autovettori dei singoli autospazi.

²²Si tratta in pratica del livello di degenerazione appena discusso.

²³A parte ovviamente la degenerazione dello stesso.

Se si considera invece il momento angolare totale, si è appena visto che esiste una base di autovettori comuni a \hat{J} e \hat{J}_z relativa agli autovalori j e m , e rappresentata quindi da:

$$\{|j, m\rangle\}_{-j \leq m \leq j}^{j=0 \rightarrow n}$$

dove n rappresenta il valore massimo che può assumere il momento totale e che dipende dalla massima energia del particolare sistema.²⁴ Questi autovettori sono naturalmente anche autovettori di \hat{J}_1^2 e \hat{J}_2^2 .

Esistono dunque due modi di rappresentare gli autovettori comuni di \hat{J}^2 e \hat{J}_z . Si tratta in realtà di due basi ortonormali dello stesso spazio ed è quindi possibile passare da una rappresentazione all'altra tramite una matrice di trasformazione. Gli elementi di questa matrice sono chiamati *coefficienti di Clebsch-Gordan* e sono evidentemente dati dal prodotto scalare tra gli autovettori delle due possibili rappresentazioni della base comune di \hat{J}^2 e \hat{J}_z :

$$C_{j,m}^{m_1, m_2} = \langle j_1, m_1; j_2, m_2 | j, m \rangle$$

Ulteriori approfondimenti sulla manipolazione dei momenti angolari riguardano generalmente corsi superiori.

²⁴La notazione di questo valore con n richiama l'atomo di idrogeno, come si vedrà nel prossimo capitolo.

L'atomo di idrogeno

L'atomo di idrogeno può essere rappresentato formalmente come il problema di due particelle che interagiscono tramite un potenziale centrale. L'hamiltoniana generica di un sistema di questo tipo è data dalla forma generale:

$$H = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + V(|\vec{r}_1 - \vec{r}_2|) \quad (7.1)$$

La dipendenza di questo potenziale dalla differenza $\vec{r}_1 - \vec{r}_2$ riflette il principio di azione e reazione (legato all'omogeneità dello spazio) e il modulo di questa differenza deriva dalla conservazione del momento angolare (legato a sua volta all'isotropia dello spazio).

7.1 Equazione di Schrödinger per l'atomo di idrogeno

Siccome il problema è a simmetria radiale, l'Hamiltoniana è invariante per rotazioni.

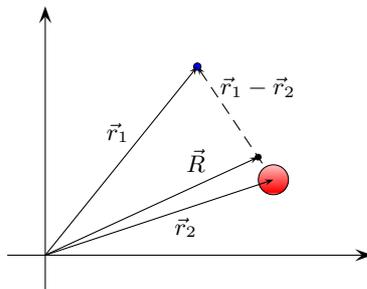
L'Hamiltoniana del sistema si può, tramite una trasformazione di coordinate, decomporre in una parte che descrive il moto del centro di massa e in una parte che descrive il moto intorno al centro di massa. Questa risulta essere una trasformazione canonica, che in meccanica quantistica significa dire che gli operatori associati alle coordinate commutano.

Denotando infatti con \vec{R}, \vec{P} le coordinate relative al baricentro e con $\vec{r}_1, \vec{r}_2, \vec{p}_1, \vec{p}_2$ quelle relative ai due corpi, si vede che risulta:

$$\begin{aligned} [r_i, P_j] &= (i\hbar - i\hbar) \delta_{ij} \\ [r_i, p_j] &= \left(\frac{m_2}{M} i\hbar + \frac{m_1}{M} i\hbar \right) \delta_{ij} = i\hbar \delta_{ij} \end{aligned}$$

Queste trasformazioni di coordinate sono definite dalle relazioni:

$$\begin{cases} \vec{R} = \frac{m_1 \vec{r}_1 + m_2 \vec{r}_2}{m_1 + m_2} \\ \vec{P} = \vec{p}_1 + \vec{p}_2 \end{cases} \quad \begin{cases} \vec{r} = \vec{r}_1 - \vec{r}_2 \\ \vec{p} = \frac{m_2 \vec{r}_1 + m_1 \vec{r}_2}{m_1 + m_2} \end{cases}$$



Definendo la *massa ridotta* come $\mu = \frac{m_1 m_2}{m_1 + m_2}$ e la massa totale come $M = m_1 + m_2$, l'Hamiltoniana (7.1) diventa:

$$H = \underbrace{\frac{P^2}{2M}}_{\text{Centro di massa}} + \underbrace{\frac{p^2}{2\mu} + V(r)}_{H_r = \text{Sistema legato}}$$

Si noti che in questo caso specifico risulta $m_P \simeq 1836 m_e$ e quindi la massa ridotta vale circa $0.999 m_e$. Il centro di massa coincide quindi praticamente con il protone e il moto intorno al centro di massa è a tutti gli effetti pratici il moto dell'elettrone intorno al nucleo.

Il moto del centro di massa può quindi effettivamente essere separato da quello del sistema "ridotto", in quanto le due Hamiltoniane commutano fra loro. Ne consegue che la funzione d'onda totale del sistema può essere fattorizzata nel moto del baricentro e quello della particella:

$$\psi_E(\vec{r}_1, \vec{r}_2) = e^{i\frac{\vec{P} \cdot \vec{R}}{\hbar}} \cdot \psi_\varepsilon(|\vec{r}_1 - \vec{r}_2|)$$

dove:

$$E = \frac{P^2}{2M} + \varepsilon \quad H_r \psi_\varepsilon = \varepsilon \psi_\varepsilon$$

Naturalmente, non tutte le autofunzioni di questo sistema sono fattorizzabili in questo modo, ma quelle fattorizzabili costituiscono una base. Le autofunzioni $e^{i\frac{\vec{P} \cdot \vec{R}}{\hbar}}$ non sono normalizzate, e questo dipende dal fatto che non è stato scelto alcun valore particolare per il momento angolare L . Una volta normalizzate, si ritrova un pacchetto d'onda in \vec{R} :

$$\psi(\vec{R}, \vec{r}) = \int_0^{+\infty} A(\vec{P}) e^{i\frac{\vec{P} \cdot \vec{R}}{\hbar}} d\vec{P} \cdot \psi_\varepsilon(\vec{r})$$

L'integrale fornisce il pacchetto d'onda in \vec{R} , e mostra che il baricentro del sistema si comporta come una particella libera di energia $E = \frac{\hbar^2 k^2}{2M}$.¹

La parte interessante è quindi quella che descrive il moto intorno al baricentro che, per quanto notato sopra, coincide praticamente con il moto dell'elettrone intorno al nucleo.

L'equazione di Schrödinger per il sistema ridotto si scrive quindi:

$$\left[\frac{\vec{p}^2}{2\mu} + V(r) \right] \psi(\vec{r}) = \varepsilon \psi(\vec{r})$$

ovvero, esplicitando gli operatori:

$$\left[-\frac{\hbar}{2\mu} \nabla^2 + V(r) \right] \psi(\vec{r}) = \varepsilon \psi(\vec{r})$$

Ora, si mostrerà che in un moto in un potenziale centrale la parte radiale e angolare sono facilmente separabili.²

In primo luogo, data la simmetria centrale del problema si può esplicitare l'equazione di Schrödinger in coordinate polari (r, θ, ϕ) . In queste coordinate l'operatore laplaciano si scrive:

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2}$$

Ricordando la forma del momento angolare data in (6.2) e notando che:

$$\frac{\partial}{\partial \cos \theta} \left(\sin^2 \theta \frac{\partial}{\partial \cos \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2}$$

si ricava che:

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{l^2}{\hbar^2 r^2}$$

¹Si noti per inciso che questo pacchetto non è un'autofunzione dell'Hamiltoniana.

²Non dovrebbe essere una sorpresa, in quanto c'è una forte analogia con il caso classico e dipende sostanzialmente dall'invarianza per rotazioni del sistema.

dove la parte espressa come momento angolare dipende esclusivamente dalle coordinate angolari. L'Hamiltoniana dell'atomo di idrogeno³ assume quindi la forma:

$$H = -\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \underbrace{\frac{\hat{l}^2}{2\mu r^2}}_{\text{Potenziale efficace}} + V(r) \quad (7.2)$$

Il termine $\hat{l}^2/2\mu r^2$ compare anche nella trattazione classica del potenziale centrale e rappresenta un termine di potenziale "centrifugo" che tende ad allontanare la particella dall'origine. Questo è dovuto al fatto che il sistema è in rotazione (se $\hat{l}^2 > 0$) e si sta osservando la sola parte radiale. Il problema può quindi essere studiato introducendo un *potenziale efficace*, dove il secondo termine tende a spostare la massa verso r crescenti. Il problema resta un problema unidimensionale, ma in un potenziale effettivo diverso da $V(r)$.

In meccanica quantistica si possono fare alcune ulteriori osservazioni interessanti.

In primo luogo, notando che nella (7.2) l'Hamiltoniana dipende solo dalla coordinata radiale e dal quadrato del momento angolare \hat{l}^2 , si ricava che:

$$[H, \hat{l}^2] = 0$$

Dalla forma (6.1) di \hat{l}_z e dal fatto il quadrato del momento angolare commuta con una sua componente, si vede che risulta:

$$[H, \hat{l}_z] = 0$$

Il quadrato del momento angolare rappresenta quindi un buon numero quantico e può essere utilizzato per classificare gli stati dell'atomo di idrogeno.

Siccome \hat{l}^2 e \hat{l}_z si possono assegnare a priori ed è noto un sistema di autofunzioni comune, la soluzione si può scrivere in una forma fattorizzata:

$$\psi(r, \theta, \phi) = R(r)Y_l^m(\theta, \phi)$$

e l'equazione di Schrödinger (7.2), una volta indicata l'energia del sistema con E , assume la forma:

$$-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left[r^2 \frac{\partial}{\partial r} R(r) \right] Y_l^m(\theta, \phi) + R(r) \frac{\hat{l}^2}{2\mu r^2} Y_l^m(\theta, \phi) + V(r) R(r) Y_l^m(\theta, \phi) = E R(r) Y_l^m(\theta, \phi)$$

che, dividendo ambedue i membri per $R(r)Y_l^m(\theta, \phi)$, fornisce:

$$-\frac{\hbar^2}{2\mu} \frac{1}{r^2 R(r)} \frac{\partial}{\partial r} \left[r^2 \frac{\partial}{\partial r} R(r) \right] + \frac{1}{Y_l^m(\theta, \phi)} \frac{\hat{l}^2}{2\mu r^2} Y_l^m(\theta, \phi) + V(r) = E$$

e moltiplicando per $2\mu r^2/\hbar^2$ e riarrangiando i termini:

$$\frac{1}{R(r)} \frac{d}{dr} \left[r^2 \frac{d}{dr} R(r) \right] - \frac{2\mu r^2}{\hbar} [V(r) - E] = \frac{1}{Y_l^m(\theta, \phi)} \frac{1}{\hbar^2} \hat{l}^2 Y_l^m(\theta, \phi)$$

Il membro a sinistra dipende solo da r mentre quello a destra solo da θ e ϕ , pertanto entrambi devono essere uguali ad una costante. Ma la soluzione dell'equazione a destra è quella del momento angolare ed è già stata trovata in (6.3):

$$\hat{l}^2 Y_l^m(\theta, \phi) = \hbar^2 \cdot \text{cost.} \cdot Y_l^m(\theta, \phi) \quad \Rightarrow \quad \text{cost.} \equiv l(l+1)$$

L'equazione per la parte radiale assume quindi la forma definitiva:

$$-\frac{\hbar}{2\mu} \frac{1}{r^2} \frac{d}{dr} \left[r^2 \frac{dR_{nl}(r)}{dr} \right] + \left[V(r) + \frac{\hbar l(l+1)}{2\mu r^2} \right] R_{nl}(r) = E_{nl} R_{nl}(r)$$

³Si tratta in effetti della *componente nel baricentro*, ma nel seguito ci si riferirà a questa parte come all'*atomo di idrogeno*.

dove si è notato esplicitamente il fatto che l'energia dipende in generale da l (perché il potenziale efficace dipende da questo numero quantico) e il fatto che gli stati legati devono presentare una quantizzazione dei livelli energetici, denotata con un indice n .

La funzione d'onda totale per il sistema atomo di idrogeno è quindi fornita da:

$$\boxed{\psi_{nlm}(r, \theta, \phi) = R_{nl}(r)Y_l^m(\theta, \phi)}$$

Si passerà quindi ora a studiare la soluzione $R_{nl}(r)$ della parte radiale.

7.2 Soluzione dell'equazione radiale

Si è visto quindi che l'equazione da risolvere per la parte radiale è la seguente:

$$-\frac{\hbar}{2\mu} \frac{1}{r^2} \frac{d}{dr} \left[r^2 \frac{dR_{nl}(r)}{dr} \right] + \left[V(r) + \frac{\hbar l(l+1)}{2\mu r^2} \right] R_{nl}(r) = E_{nl} R_{nl}(r) \quad r \in [0, +\infty)$$

Affinché l'Hamiltoniana ridotta sia ben definita sulle funzioni a quadrato sommabile in \mathbb{R}^3 si deve richiedere che l'operatore $\hat{p} = -i\hbar \frac{1}{r} \frac{\partial}{\partial r} r$ sia hermitiano sulle funzioni di r a quadrato sommabile in $[0, +\infty)$ e sulla misura $d\mu = r^2 dr$, il che equivale a chiedere:

$$\int_0^{+\infty} R^*(r) \hat{p} R(r) r^2 dr = \int_0^{+\infty} [\hat{p} R^*(r)] R(r) r^2 dr$$

il che implica $\lim_{r \rightarrow 0} rR(r) = 0$, ovvero, con la comoda posizione $u(r) \equiv rR(r)$:

$$\lim_{r \rightarrow 0} u(r) = 0$$

Devono quindi essere escluse dal dominio di hermiticità di \hat{p} tutte le funzioni che divergono nell'origine come $1/r$, anche se sono a quadrato sommabile. Questa posizione può essere interpretata dicendo che per evitare che la particella penetri nella regione $r < 0$ bisogna porre una barriera di potenziale infinita in $r = 0$, il che implica $u(r) = 0$.

Si passi ora a specificare il caso dell'atomo di idrogeno. In questo caso, il potenziale è quello coulombiano ed è espresso pertanto da:⁴

$$V(r) = -\frac{e^2}{r} \quad \longrightarrow \quad V_{\text{eff.}} = -\frac{e^2}{r} + \frac{l(l+1)}{r^2}$$

Si vede che per r grandi il termine predominante è quello coulombiano. L'andamento delle soluzioni $u(r)$ per $r \rightarrow +\infty$ è dettato dall'equazione:

$$\frac{d^2}{dr^2} u(r) + \frac{2\mu E}{\hbar} u(r) = 0$$

che ha per soluzione la combinazione di esponenziali:

$$u(r) \underset{r \rightarrow \infty}{\simeq} A e^{+i\sqrt{2\mu E/\hbar^2} r} + B e^{-i\sqrt{2\mu E/\hbar^2} r}$$

Se $E > 0$ le radici sono reali e gli esponenziali sono oscillanti. Non si tratta di funzioni a quadrato sommabile, ma di onde piane: in questo caso non ci sono stati legati. Se invece risulta $E < 0$ l'esponenziale negativo diverge per $r \rightarrow +\infty$ e quindi questa soluzione deve essere scartata per motivi fisici ($B = 0$).⁵

Con la posizione fatta $rR(r) = u(r)$ ed esplicitando la forma del potenziale, l'equazione di Schrödinger diventa:

$$\frac{d^2}{dr^2} u(r) + \left[\frac{2\mu E}{\hbar^2} + \frac{2\mu e^2}{\hbar^2 r} - \frac{l(l+1)}{r^2} \right] u(r) = 0 \quad (7.3)$$

⁴In unità gaussiane, in cui $4\pi\epsilon_0 \equiv 1$.

⁵Si noti che l'esponenziale negativo è immaginario e se $E < 0$ l'argomento è immaginario anche lui: di conseguenza, questo esponenziale diventa reale positivo.

Visto che la soluzione all'infinito deve decrescere come un esponenziale, è sensato fare la seguente assunzione sulla soluzione dell'equazione:

$$u(r) = e^{-\varepsilon r} f(r) \quad \varepsilon^2 = -\frac{2\mu E}{\hbar^2}$$

Notando che:

$$\begin{aligned} \frac{d^2}{dr^2} u(r) &= \frac{d}{dr} \left[-\varepsilon e^{-\varepsilon r} f(r) + e^{-\varepsilon r} \frac{d}{dr} f(r) \right] = \\ &= \varepsilon^2 e^{-\varepsilon r} f(r) - \varepsilon e^{-\varepsilon r} \frac{d}{dr} f(r) - \varepsilon e^{-\varepsilon r} \frac{d}{dr} f(r) + e^{-\varepsilon r} \frac{d^2}{dr^2} f(r) = \\ &= e^{-\varepsilon r} \left[\frac{d^2}{dr^2} f(r) - 2\varepsilon \frac{d}{dr} f(r) + \varepsilon^2 f(r) \right] \end{aligned}$$

ricordando che $\varepsilon \equiv -\frac{2\mu E}{\hbar^2}$ e notando che $r_0 = \frac{\mu e^2}{\hbar^2}$ (raggio di Bohr), l'equazione (7.3) si riscrive come:

$$\frac{d^2}{dr^2} f(r) - 2\varepsilon \frac{d}{dr} f(r) + \left[\cancel{\varepsilon^2} + \frac{2}{r_0 r} - \frac{l(l+1)}{r^2} + \cancel{\varepsilon^2} \right] f(r) = 0 \quad (7.4)$$

La funzione $f(r)$ può essere ricercata nella forma di una serie di potenze, di cui dovrà essere studiata la convergenza. Si vedrà che questa serie in effetti diverge come un esponenziale e quindi per poter avere senso fisico si deve ridurre ad un polinomio. Questo permetterà di ricavare lo spettro dei valori possibili di E . Si ponga dunque:

$$f(r) = \sum_{n=0}^{\infty} a_n r^{n+\alpha} \quad f(0) = 0 \Rightarrow \alpha > 0 \quad (7.5)$$

e sostituendo nella (7.4) si ricava:

$$\sum_{n=0}^{+\infty} \left[(n+\alpha)(n+\alpha-1) a_n r^{n+\alpha-2} - 2\varepsilon(n+\alpha) a_n r^{n+\alpha-1} + \frac{2}{r_0} a_n r^{n-\alpha-1} - l(l+1) a_n r^{n+\alpha-2} \right] = 0$$

L'idea di fondo è quella di trovare una formula di ricorrenza per i coefficienti. Si isoli innanzitutto il primo coefficiente, che è quello che corrisponde a $n=0$:

$$[\alpha(\alpha-1) - l(l+1)] a_0 r^{\alpha-2}$$

successivamente, nei termini rimanenti della serie si mandi $n \rightarrow n+1$ in modo da avere ancora una sommatoria che parte da $n=0$:

$$[\alpha(\alpha-1) - l(l+1)] a_0 r^{\alpha-2} + \sum_{n=0}^{+\infty} \left\{ [(n+\alpha+1)(n+\alpha) - l(l+1)] a_{n+1} - \left[2\varepsilon(n+\alpha) - \frac{2}{r_0} \right] a_n \right\} r^{n+\alpha-1}$$

Siccome è presente una derivata seconda, il primo termine deve essere nullo,⁶ e questo si può verificare in due modi:

$$\begin{cases} a_0 & = 0 \\ \alpha(\alpha-1) - l(l+1) & = 0 \end{cases}$$

Il caso $a_0 = 0$ è naturalmente da escludere in quanto, essendo i coefficienti definiti per ricorrenza, questo implicherebbe che tutti i coefficienti siano nulli: la soluzione si riduce quindi alla soluzione banale.⁷

⁶Perché il primo termine è una potenza di r^0 ed è quindi costante.

⁷Si noti per inciso che la soluzione banale – essendo nulla dappertutto – è la soluzione che “descrive” l'assenza di sistema fisico.

Il secondo caso permette ancora due possibilità:

$$\alpha(\alpha - 1) - l(l + 1) = 0 \quad \Rightarrow \quad \begin{cases} \alpha = -l \\ \alpha = l + 1 \end{cases}$$

La prima possibilità, α negativo, è esclusa dalla (7.5),⁸ resta quindi solo la possibilità $\alpha = l + 1$.

Un modo alternativo di vedere la forma della soluzione radiale

Esiste un modo forse più lineare per ritrovare le posizioni fatte sulla forma radiale delle autofunzioni dell'atomo di idrogeno. Si parta dall'equazione (7.3). Per $r \rightarrow 0$ questa equazione assume la forma:

$$\frac{d^2}{dr^2}u(r) + \frac{2\mu e^2}{\hbar^2 r}u(r) - \frac{l(l+1)}{r^2}u(r) = 0$$

La soluzione certamente non può divergere nell'origine⁹, per cui si può ricercare una soluzione nella forma $u(r) = Cr^s$. Sostituendo questa forma nell'equazione scritta qui sopra, si ricava:

$$s(s-1) + 2s - l(l+1) = 0$$

che ammette per soluzioni $s = l$ e $s = -(l+1)$. Siccome la funzione non può divergere nell'origine, la soluzione fisicamente accettabile è $s = l$.

Per $r \rightarrow +\infty$ l'equazione (7.3) assume invece la forma:

$$\frac{d^2}{dr^2}u(r) + \frac{2\mu E}{\hbar^2}u(r) = 0$$

che ha evidentemente soluzioni del tipo $u(r) = e^{\pm \varepsilon r}$, dove la soluzione con il segno positivo deve essere escluso sempre per ragioni fisiche. Questa soluzione infatti diverge all'infinito.

Ne consegue che una soluzione fisicamente accettabile per la parte radiale dell'Hamiltoniana dell'atomo di idrogeno deve avere la forma:

$$u(r) = r^l e^{-\varepsilon r} f(r)$$

A questo punto, si ricerca la forma della $f(r)$ in termini di serie di potenze rientrando nel discorso principale.

Dovendosi annullare i coefficienti delle potenze $r^{n+\alpha-1}$,⁽¹⁰⁾ deve sussistere la seguente relazione di ricorrenza fra i coefficienti:

$$a_{n+1} = \frac{2\varepsilon(n+\alpha) - \frac{2}{r_0}}{(n+\alpha+1)(n+\alpha) - l(l+1)} a_n$$

Occorre ora studiare la convergenza di questa serie. In primo luogo, visto che vale:

$$\lim_{n \rightarrow +\infty} \frac{a_{n+1}}{a_n} = 0$$

si deduce che il raggio di convergenza della serie è infinito e quindi la soluzione è ben valida nell'intervallo $r \in [0, +\infty)$. Tuttavia, questa serie si comporta a infinito come un esponenziale. Infatti considerando che il termine $2/r_0$ e il termine $l(l+1)$ diventano trascurabili per alti n si ricava:

$$\lim_{n \rightarrow +\infty} \frac{a_{n+1}}{a_n} \simeq \frac{2\varepsilon}{n}$$

e ricordando che l'espansione in serie dell'esponenziale è data da:

$$e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$$

⁸Questa costituisce dal punto di vista matematico una delle due soluzioni indipendenti. Quella accettabile fisicamente è l'altra.

⁹Non può in realtà divergere in nessun punto dell'intervallo $[0, +\infty)$ per essere fisicamente accettabile.

¹⁰Si ricordi che le potenze sono linearmente indipendenti.

si ricava che la soluzione, a meno di potenze di r , si comporta all'infinito come $f(r) \simeq e^{2\varepsilon r}$, e quindi $u(r) \simeq e^{\varepsilon r}$. Questa funzione non appartiene a \mathcal{L}^2 . Per un generico valore di ε ¹¹ quindi la soluzione non è a quadrato sommabile. Se però esiste una condizione per cui:

$$\exists n: a_n \neq 0 \quad \Rightarrow \quad a_{n+1} = 0$$

la serie si riduce ad un polinomio e la soluzione diventa *fisicamente accettabile*. Questa condizione è data da:

$$2\varepsilon(n + \alpha) - \frac{2}{r_0} = 0 \quad \Rightarrow \quad \varepsilon = \frac{1}{r_0(n + \alpha)} \quad (7.6)$$

Ricordando le posizioni fatte su ε e r_0 si ricava:

$$-\frac{2\mu E_n}{\hbar^2} = \frac{\mu^2 e^2}{\hbar^4} \frac{1}{(n + l + 1)^2}$$

ovvero in definitiva:

$$E_n = -\frac{\mu e^2}{2\hbar^2} \frac{1}{(k + 1)^2} \quad k \equiv n + l \geq 0$$

e una volta fissato k (ovvero l'energia) deve risultare $0 \leq l \leq n - 1$:¹²

$$E_n = -\frac{\mu e^2}{2\hbar^2} \frac{1}{n^2} \quad k \equiv n \geq 1$$

Il numero n prende il nome di **numero quantico principale**.

Si passi ora a studiare la degenerazione dei livelli energetici. Come si vede, l'energia dell'atomo di idrogeno dipende dal solo numero quantico n , ne consegue che una volta fissata l'energia sono ancora possibili tutti i valori del momento angolare $0 \leq l \leq n - 1$.

Ricordando che la somma dei primi n interi è data da:

$$\sum_{i=1}^n i = \frac{n(n-1)}{2}$$

e considerando che ogni valore di l è a sua volta $2l + 1$ volte degenere (a causa della componente z del momento angolare associata al numero quantico m), si ricava che la degenerazione del livello n è data da:

$$d = \sum_{l=0}^{n-1} (2l + 1) = \sum_{l=1}^n (2l + 1) = 2 \sum_{l=1}^n l + n = \frac{2n(n-1)}{2} + n = n^2 - n + n = n^2$$

ovvero che *la degenerazione del livello n è data da n^2* . La degenerazione rispetto a l è una cosiddetta *degenerazione accidentale* in quanto dipende specificamente dalla forma del potenziale coulombiano (in particolare dal fatto che è l'unico potenziale che dipende dall'inverso di una potenza della distanza a permettere orbite chiuse), mentre la degenerazione rispetto a m dipende dalla simmetria centrale del potenziale, ovvero in ultima analisi dal fatto che tutte le direzioni dello spazio sono equivalenti per quanto riguarda l'energia.

Si passerà ora a discutere più in dettaglio le autofunzioni dell'atomo di idrogeno.

7.3 Autofunzioni dell'atomo di idrogeno

I polinomi definiti dalla condizione (7.6) sono chiamati *polinomi associati di Laguerre*, sono indicati con $L_p^q(x)$ e definiti dalla relazione:

$$L_p^q(x) = (-1)^q \frac{d^q}{dx^q} L_{p+q}(x)$$

¹¹L'unico parametro che si può far variare in questa equazione è l'energia contenuta nel termine ε .

¹²Si noti che questi valori coincidono effettivamente con quelli trovati da Bohr.

dove i polinomi $L_{p+q}(x)$ sono i *polinomi (semplici) di Laguerre*, definiti dalla relazione (detta di *Rodrigues*):

$$L_p(x) = e^x \frac{d^p}{dx^p} (e^{-x} x^p)$$

Si vede subito che $L_p^0(x) = L_p(x)$ e $L_0^q = q!$. Questi polinomi rappresentano l'ortogonalizzazione delle potenze x^n sul dominio $x \in [0, +\infty)$ secondo la misura e^{-x} e hanno p zeri. I primi quattro polinomi sono i seguenti:

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= 1 - x \\ L_2(x) &= \frac{1}{2}x^2 - 2x + 1 \\ L_3(x) &= \frac{1}{6}(-x^3 + 9x^2 - 18x + 6) \end{aligned}$$

Ricordando le posizioni fatte sulla funzione $u(r) = e^{-\varepsilon r} f(r)$, $R_{nl}(r) = r \cdot u(r)$, $\varepsilon^2 = -\frac{2\mu E}{\hbar^2}$, $r_0 = \frac{\mu e^2}{\hbar^2}$ e l'espansione (7.5) con $\alpha = l + 1$ si trova per la parte radiale della soluzione:

$$R_{nl}(r) = N_{nl} r^l e^{-\frac{r}{nr_0}} L_{n-l-1}^{2l+1} \left(\frac{2r}{nr_0} \right)$$

dove N_{nl} è un coefficiente di normalizzazione ricavabile imponendo la condizione:

$$\int_0^{+\infty} r^2 |R_{nl}(r)|^2 dr = 1$$

e che permette di giungere alla forma:

$$R_{nl}(r) = \frac{1}{r_0^{3/2}} \frac{2}{n^2} \sqrt{\frac{(n-l-1)!}{(n+l)!}} \left(\frac{2r}{nr_0} \right)^l e^{-\frac{r}{nr_0}} L_{n-l-1}^{2l+1} \left(\frac{2r}{nr_0} \right)$$

Ricordando che le armoniche sferiche sono già normalizzate, le autofunzioni complete dell'atomo di idrogeno sono date da:

$$\psi_{nlm}(r, \theta, \phi) = \frac{1}{r_0^{3/2}} \frac{2}{n^2} \sqrt{\frac{(n-l-1)!}{(n+l)!}} \left(\frac{2r}{nr_0} \right)^l e^{-\frac{r}{nr_0}} L_{n-l-1}^{2l+1} \left(\frac{2r}{nr_0} \right) \cdot Y_l^m(\theta, \phi)$$

È interessante notare che il modulo quadro della parte radiale $|R_{nl}(r)|^2$ rappresenta la *densità di probabilità di trovare l'elettrone ad una determinata distanza r dal nucleo* ed è quindi legata alle dimensioni dell'atomo, mentre il modulo quadro delle armoniche sferiche $|Y_l^m(\theta, \phi)|^2$ risulta legato alla *densità di probabilità di trovare l'elettrone in un punto dello spazio identificato dalle coordinate (θ, ϕ) indipendentemente dalla distanza*.

I valori di $|R_{nl}(r)|^2$ permettono quindi di identificare quali sono i raggi¹³ delle orbite dell'atomo di idrogeno:

$$\langle r^k \rangle = \int_0^{+\infty} r^{2+k} |R_{nl}(r)|^2 dr$$

da cui si ricava il valore medio per l'orbita:

$$\langle r \rangle = \frac{r_0}{2} [3n^2 - l(l+1)]$$

si vede quindi come questo dipenda sia da n che da l , a differenza del valore trovato da Bohr per il quale non c'è dipendenza da l . La dimensione globale dell'atomo risulta in ogni caso determinata essenzialmente dal numero quantico principale n .

¹³Inteso nel senso di *valori più probabili*.

Le prime autofunzioni radiali hanno la forma esplicita:

$$\begin{aligned}
 R_{10}(r) &= 2r_0^{-3/2} e^{-r/r_0} \\
 R_{20}(r) &= \frac{1}{2\sqrt{2}} r_0^{-3/2} \left(2 - \frac{r}{r_0} \right) e^{-r/2r_0} \\
 R_{21}(r) &= \frac{1}{6\sqrt{2}} r_0^{-3/2} \frac{r}{r_0} e^{-r/2r_0} \\
 R_{30}(r) &= 2(3r_0)^{-3/2} \left[1 - \frac{2r}{3r_0} - \frac{2r^2}{27r_0^2} \right] e^{-r/3r_0} \\
 R_{31}(r) &= \frac{4\sqrt{2}}{3} (3r_0)^{-3/2} \left[\frac{r}{r_0} - \frac{r^2}{6r_0^2} \right] e^{-r/3r_0} \\
 R_{32}(r) &= \frac{2\sqrt{2}}{27\sqrt{5}} (3r_0)^{-3/2} \left[\frac{r}{r_0} \right]^2 e^{-r/3r_0}
 \end{aligned}$$

Lo stato fondamentale è definito da $n = 1$ e $l = 0$, si tratta dell'unico caso in cui non c'è degenerazione in energia. La probabilità $4r_0^{-3} r^2 e^{-2r/r_0}$ decade in maniera esponenziale per r crescenti e la corrispondente armonica sferica è costante¹⁴, quindi la distribuzione della probabilità è a simmetria sferica. Essa inoltre tende a 0 per $r \rightarrow 0$.

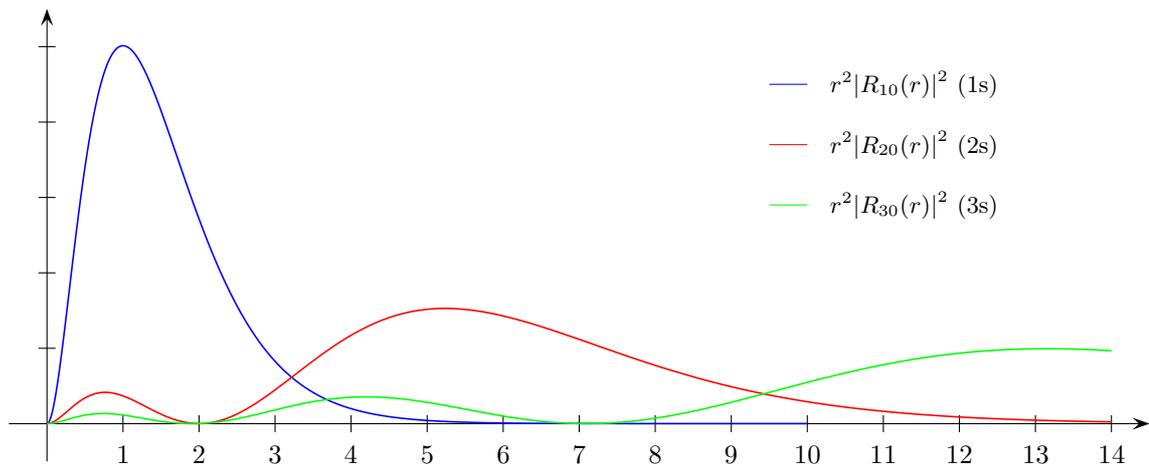
In particolare, le densità di probabilità delle funzioni radiali in prossimità del nucleo si comportano come $\approx r^{l+1}$: quindi quanto più grande è il valore del momento angolare l , tanto più queste funzioni vanno a zero rapidamente in prossimità del nucleo. Questo è coerente con quanto detto sopra, ovvero che il momento angolare si comporta come una sorta di “potenziale centrifugo” tendente ad allontanare l'elettrone dal nucleo. A valori grandi di r invece la probabilità è dettata dal termine dell'esponenziale decrescente, che assicura uno smorzamento molto veloce della probabilità di trovare l'elettrone lontano dal nucleo e rende così il sistema atomico sostanzialmente ben localizzato nello spazio.¹⁵

Nella figura seguente si mostrano le prime tre densità di probabilità radiali per $l = 0$, che sono quelle più intuitive da visualizzare in quanto hanno simmetria sferica. Come si vede dalla figura, il livello fondamentale è molto stretto intorno al valore del raggio di Bohr r_0 , mentre all'aumentare di n la probabilità di trovare l'elettrone si “sparpaglia” via via sempre più. Come è stato detto più sopra, la funzione radiale del livello n presenta n poli (punti in cui la probabilità si annulla) e quindi n massimi. Il massimo maggiore si allontana via via dal nucleo, questo insieme al fatto che i massimi sono separati da un guscio sferico in cui la probabilità di trovare l'elettrone è nulla dà ragione del concetto di “livello energetico”. È da notare tuttavia che anche per $n > 1$ esiste una probabilità non nulla di trovare l'elettrone vicino al nucleo, probabilità che decresce in ogni caso velocemente all'aumentare del livello di energia.

Nella notazione spettroscopica i diversi livelli sono identificati indicando il numero quantico principale seguito da una lettera che indica il valore del momento angolare orbitale: s ($l = 0$), p ($l = 1$), d ($l = 2$), f ($l = 3$), e così via di seguito. I livelli dell'atomo di idrogeno vengono quindi identificati con $1s$ (livello fondamentale), $2s$ (primo livello eccitato con $l = 0$), $2p$ (primo livello eccitato con $l = 1$), $3s$, $3p$, $3d$, ...

¹⁴Si ricordi che $Y_0^0(\theta, \phi) = \frac{1}{4\pi}$.

¹⁵Esiste un eccellente applet JAVA sul sito del Davidson College che permette di visualizzare le autofunzioni e le densità di probabilità per l'atomo di idrogeno e che può essere trovato al link <http://webphysics.davidson.edu/Applets/Hydrogenic/>.



7.3.1 Sullo spin dell'elettrone

L'esperienza dimostra che questa trattazione non è sufficiente a spiegare tutte le proprietà sperimentali dell'atomo di idrogeno. In effetti, secondo la teoria fin qui esposta le righe dello spettro devono essere singole per via della degenerazione dei livelli energetici rispetto al momento angolare che "collassano" su un solo livello identificato da n . In realtà, sperimentalmente si rileva che queste linee non sono singole ma sono sdoppiate o presentano dei multipletti. La spiegazione di questo fenomeno risiede nel fatto che questa trattazione non ha tenuto conto dello spin dell'elettrone, che può assumere due valori distinti (si veda il paragrafo §6.4).

Allo spin è associato quindi un altro numero quantico indicato tradizionalmente con s ⁽¹⁶⁾. Questo momento angolare interagisce con il campo magnetico del nucleo e dà origine a due diversi valori per l'energia che splittano le linee spettrali in due. Ogni trattazione completa della struttura atomica non può prescindere da questa quantità, e infatti nello studio della teoria atomica si utilizza il momento angolare totale $\vec{J} = \vec{L} + \vec{s}$. Questa evoluzione è ancora una volta oggetto di corsi superiori.

7.3.2 Sulla stabilità degli autostati

Al lettore attento potranno sorgere a questo punto due dubbi:

- non bisogna dimenticare che una carica in moto irraggia. Nulla di quanto detto finora spiega perché nello stato fondamentale l'elettrone non irraggia e rimanga stabilmente nel suo stato,
- dal punto di vista quantistico, invece, nulla spiega perché un atomo di idrogeno in uno stato eccitato ($n > 1$) debba decadere verso lo stato fondamentale. L'elettrone che si trovi in un livello con $n > 1$ si trova comunque in un autostato e pertanto non dovrebbe evolvere, ma rimanere indefinitamente in questo stato,
- se si ammette tuttavia che un autostato possa non essere stabile e si permette il decadimento verso livelli energetici inferiori, ancora una volta nulla spiega perché l'atomo di idrogeno nel suo stato fondamentale sia stabile e non possa ulteriormente decadere, facendo precipitare l'elettrone sul nucleo.

Si tratta di due domande acute, ma la cui risposta purtroppo non può essere data nell'ambito di questo corso perché richiede la conoscenza della QED (*elettrodinamica quantistica*). È possibile dare qui solo un abbozzo della risposta, trascurando i dettagli tecnici.

Il punto essenziale è che si deve tenere conto dell'interazione dell'elettrone con le fluttuazioni quantistiche del vuoto (oltre che dell'interazione dell'elettrone con il campo elettromagnetico generato da lui stesso e dal nucleo), interazione che può essere schematizzata come una emissione /

¹⁶In origine, Pauli propose di utilizzare un numero quantico aggiuntivo a n, l, m per spiegare questo comportamento. Goudsmit e Uhlenbeck proposero di utilizzare il momento di spin dell'elettrone, si veda a questo proposito la parte finale del paragrafo §2.4 e il paragrafo §6.4.

assorbimento continuo di fotoni. Qualora si tenga in considerazione questo meccanismo, si riesce a dimostrare che per i livelli eccitati dell'atomo il bilancio è a favore dell'emissione di fotoni, ovvero i fotoni emessi sono più di quelli assorbiti. Questo significa che l'elettrone mediamente perde energia (irraggia) e questo lo porta a decadere ai livelli energetici più bassi. Nello stato fondamentale, invece, si riesce a dimostrare che il numero di fotoni emessi e assorbiti si bilancia perfettamente, per cui in questo caso non c'è emissione netta di energia e l'elettrone può rimanere indefinitamente nel suo stato.

Teoria delle perturbazioni

In questo capitolo si affronterà un argomento fondamentale per approfondire lo studio della meccanica quantistica, la cosiddetta *teoria perturbativa*.

I sistemi fisici quantistici del mondo reale sono descritti da equazioni differenziali la cui risoluzione pratica è di grande difficoltà, se non addirittura impossibile. Infatti, i sistemi fisici che si possono risolvere in maniera analiticamente esatta sono estremamente pochi: l'oscillatore armonico, l'atomo di idrogeno e la particella in una scatola.

In tutti gli altri casi occorre individuare delle tecniche di calcolo che pur non permettendo di ottenere una soluzione analitica esatta del problema, consentano almeno di ottenere delle soluzioni che descrivono i sistemi reali nel modo più corretto possibile. Queste soluzioni non sono quindi *esatte*, ma permettono di descrivere il sistema con un sufficiente grado di accuratezza. I risultati ottenuti con questi metodi sono generalmente espressi in termini di serie infinite, che convergono abbastanza rapidamente alla soluzione corretta (e ignota in forma analitica) man mano che ci si arresta ad ordini superiori dello sviluppo.

Concettualmente, *i metodi perturbativi si basano sul fatto che è possibile introdurre nella Hamiltoniana che descrive il sistema libero alcuni termini di potenziale perturbativi, cioè abbastanza piccoli rispetto al termine dell'Hamiltoniana libera da essere considerati come una piccola perturbazione e potere quindi essere sviluppati in serie di potenze.*

È questo il caso, ad esempio, di un atomo di idrogeno immerso in un campo elettrico (*effetto Stark lineare*) o in un campo magnetico (*effetto Zeeman*) e dell'accoppiamento dell'elettrone con il campo magnetico generato da lui stesso nel moto intorno al nucleo (*accoppiamento spin-orbita o L-S*). Il caso più eclatante è rappresentato forse dall'elettrodinamica quantistica (QED), in cui l'interazione elettrone-fotone è trattata perturbativamente.

Giusto per dare un'idea corretta, occorre dire che esistono anche situazioni in cui la teoria perturbativa non può essere utilizzata perché le interazioni in gioco non sono trascurabili. È appena il caso di accennare che questa è la situazione che si presenta qualora si prenda in considerazione l'interazione forte fra quark e gluoni (QCD) a bassa energia.

Ritornando nell'ambito di questo corso, nel seguito saranno descritti i tre approcci possibili: il caso in cui il potenziale perturbativo è costante nel tempo, il caso in cui questo dipende dal tempo ed infine il metodo variazionale.¹

8.1 Perturbazioni indipendenti dal tempo

Si consideri un sistema fisico descritto da una Hamiltoniana H_0 , il cui spettro si suppone completamente noto. Questo sistema fisico è allora detto essere noto con precisione.

¹A rigor di logica questo non è un metodo perturbativo propriamente detto, ma lo si può far rientrare in questa categoria per comodità di studio.

Si supponga ora di voler studiare un sistema fisico che può essere descritto da una Hamiltoniana della forma:

$$H = H_0 + \epsilon H_1$$

con ϵ parametro adimensionale piccolo e H_1 un operatore di cui sono noti gli elementi di matrice.

Nel caso pratico, accade spesso che il sistema in studio sia descritto da una Hamiltoniana H non trattabile, nel senso che non sia facile o possibile diagonalizzare, ma che sia possibile separare in una parte "nota" H_0 (chiamata *Hamiltoniana imperturbata*) ed una parte non trattabile moltiplicata appunto per un parametro adimensionale.²

Si denoti con $|\varphi_n\rangle$ l'autostato relativo all'Hamiltoniana imperturbata con autovalore dell'energia $E_n^{(0)}$ e con $|\psi_l\rangle$ una funzione d'onda (autostato sconosciuto) dell'Hamiltoniana completa H relativa all'autovalore E_l . Risulta quindi:

$$\begin{aligned} H_0|\varphi_n\rangle &= E_n^{(0)}|\varphi_n\rangle \\ H|\psi_l\rangle &= (H_0 + \epsilon H_1)|\psi_l\rangle = E_l|\psi_l\rangle \end{aligned} \quad (8.1)$$

Per poter proseguire nello studio dello spettro dell'Hamiltoniana completa, è *necessario* fare alcune ipotesi.

- Ipotesi di sviluppabilità degli autovalori. Si suppone che ogni autovalore E_l sia scrivibile come uno sviluppo in serie di potenze di ϵ :

$$E_l = \sum_{n=0}^{\infty} \epsilon^n E_l^{(n)} = E_l^{(0)} + \epsilon E_l^{(1)} + \epsilon^2 E_l^{(2)} + \dots$$

Considerare l'autovalore imperturbato come primo termine di questo sviluppo in serie è sensato perché per $\epsilon = 0$ l'Hamiltoniana H si riduce a quella imperturbata H_0 e così quindi deve fare l'autovalore;

- H e H_0 hanno lo stesso dominio di autoaggiuntezza. Siccome le $|\varphi_n\rangle$ sono una base di autofunzioni di H_0 , questa ipotesi permette di dire che le $|\varphi_n\rangle$ sono una base anche per H e che quindi l'autofunzione $|\psi_l\rangle$ dell'Hamiltoniana completa si può scrivere come combinazione lineare di queste autofunzioni:

$$|\psi_l\rangle = \sum_{n=0}^{\infty} a_{nl} |\varphi_n\rangle$$

questa ipotesi sarà in effetti utilizzata più avanti. Per ora si farà l'ipotesi che l'autostato sconosciuto $|\psi_l\rangle$ sia esprimibile anch'esso in termini di una serie in ϵ intorno all'autostato imperturbato:

$$|\psi_l\rangle = \sum_{n=0}^{\infty} \epsilon^n |\psi_l^{(n)}\rangle = |\psi_l^{(0)}\rangle + \epsilon |\psi_l^{(1)}\rangle + \epsilon^2 |\psi_l^{(2)}\rangle + \dots$$

ancora una volta, considerare il primo termine dello sviluppo come l'autofunzione imperturbata, indicato qui con $|\psi_l^{(0)}\rangle$, è sensato perché per $\epsilon = 0$ l'Hamiltoniana si riduce a quella imperturbata e così devono fare le autofunzioni.

Questi sviluppi in serie prendono il nome di **serie perturbative**.

Per il momento, si aggiunga anche l'ipotesi che H_0 abbia solo spettro discreto e tutti gli autovalori siano *non degeneri*. Questa ipotesi sarà rimossa in una fase successiva dello studio.

CASO NON DEGENERE

Innanzitutto, l'equazione (8.1) può essere riscritta nella forma:

$$(H_0 + \epsilon H_1) \sum_{n=0}^{\infty} \epsilon^n |\psi_l^{(n)}\rangle = \sum_{n=0}^{\infty} \epsilon^n E_l^{(n)} \sum_{n=0}^{\infty} \epsilon^n |\psi_l^{(n)}\rangle$$

²Concretamente, questo parametro si può tirare fuori moltiplicando e dividendo opportunamente.

le potenze sono indipendenti e ortogonali, quindi per essere valida questa relazione devono essere uguali i coefficienti delle potenze uguali di ϵ . In questo modo si ottengono le relazioni per le approssimazioni ai vari ordini dello sviluppo perturbativo, di cui si mostrano di seguito solo i primi tre termini:

$$\text{Ordine 0} \quad H_0|\psi_i^{(0)}\rangle = E_i^{(0)}|\psi_i^{(0)}\rangle \quad (8.2a)$$

$$\text{Ordine 1} \quad H_0|\psi_i^{(1)}\rangle + H_1|\psi_i^{(0)}\rangle = E_i^{(1)}|\psi_i^{(0)}\rangle + E_i^{(0)}|\psi_i^{(1)}\rangle \quad (8.2b)$$

$$\text{Ordine 2} \quad H_0|\psi_i^{(2)}\rangle + H_1|\psi_i^{(1)}\rangle = E_i^{(0)}|\psi_i^{(2)}\rangle + E_i^{(1)}|\psi_i^{(1)}\rangle + E_i^{(2)}|\psi_i^{(0)}\rangle \quad (8.2c)$$

La relazione all'ordine zero (8.2a) rappresenta evidentemente il sistema imperturbato, coerentemente con le posizioni fatte.

Nella relazione della correzione al primo ordine (8.2b) bisogna isolare il valore di $E_i^{(1)}$, che rappresenta appunto il primo ordine di correzione da apportare al valore imperturbato dell'energia. Per fare questo, è sufficiente moltiplicare scalarmente a sinistra per $\langle\psi_i^{(0)}|$, ovvero proiettare l'equazione sullo stato imperturbato del livello l :

$$\langle\psi_i^{(0)}|H_0|\psi_i^{(1)}\rangle + \langle\psi_i^{(0)}|H_1|\psi_i^{(0)}\rangle = \langle\psi_i^{(0)}|E_i^{(1)}|\psi_i^{(0)}\rangle + \langle\psi_i^{(0)}|E_i^{(0)}|\psi_i^{(1)}\rangle$$

Gli autovalori dell'energia sono dei numeri e le autofunzioni imperturbate sono ortonormali, per cui l'equazione precedente si riduce alla:

$$\langle\psi_i^{(0)}|H_0|\psi_i^{(1)}\rangle + \langle\psi_i^{(0)}|H_1|\psi_i^{(0)}\rangle = E_i^{(1)} + E_i^{(0)}\langle\psi_i^{(0)}|\psi_i^{(1)}\rangle$$

Si noti ora che l'Hamiltoniana imperturbata è autoaggiunta, per cui vale:

$$\langle\psi_i^{(0)}|H_0|\psi_i^{(1)}\rangle = \langle\psi_i^{(1)}|H_0|\psi_i^{(0)}\rangle^* = E^{(0)*}\langle\psi_i^{(1)}|\psi_i^{(0)}\rangle^* = E^{(0)}\langle\psi_i^{(0)}|\psi_i^{(1)}\rangle$$

per cui:

$$\cancel{E^{(0)}\langle\psi_i^{(0)}|\psi_i^{(1)}\rangle} + \langle\psi_i^{(0)}|H_1|\psi_i^{(0)}\rangle = E_i^{(1)} + \cancel{E_i^{(0)}\langle\psi_i^{(0)}|\psi_i^{(1)}\rangle}$$

e quindi la correzione all'energia al primo ordine per il livello l è data da:

$$\boxed{E_i^{(1)} = \langle\psi_i^{(0)}|H_1|\psi_i^{(0)}\rangle}$$

ovvero, *la correzione al primo ordine perturbativo dell'energia è dato dal valore di aspettazione del termine perturbativo sugli stati imperturbati corrispondenti.*

Per ricavare la correzione al primo ordine dell'autofunzione, è più pratico seguire il cosiddetto **metodo di Rayleigh-Schrödinger**, che consiste nel considerare l'autofunzione come uno sviluppo in serie in termini delle autofunzioni dell'Hamiltoniana imperturbata. Si denotino queste autofunzioni con $|\varphi_n\rangle$, l'equazione (8.2b) diventa:

$$\begin{aligned} H_0 \sum_{n=0}^{\infty} a_{nl}|\varphi_n\rangle + H_1|\varphi_l\rangle &= E_i^{(1)}|\varphi_l\rangle + E_i^{(0)} \sum_{n=0}^{\infty} a_{nl}|\varphi_n\rangle \quad \rightarrow \\ \sum_{n=0}^{\infty} a_{nl}H_0|\varphi_n\rangle + H_1|\varphi_l\rangle &= E_i^{(1)}|\varphi_l\rangle + E_i^{(0)} \sum_{n=0}^{\infty} a_{nl}|\varphi_n\rangle \end{aligned}$$

proiettando questa equazione sullo stato imperturbato $|\varphi_m\rangle$ e considerando l'ortonormalità di queste autofunzioni:

$$\begin{aligned} \sum_{n=0}^{\infty} a_{nl}\langle\varphi_m|H_0|\varphi_n\rangle + \langle\varphi_m|H_1|\varphi_l\rangle &= E_i^{(1)}\langle\varphi_m|\varphi_l\rangle + E_i^{(0)} \sum_{n=0}^{\infty} a_{nl}\langle\varphi_m|\varphi_n\rangle \quad \rightarrow \\ E_n^{(0)} \sum_{n=0}^{\infty} a_{nl} \underbrace{\langle\varphi_m|\varphi_n\rangle}_{\delta_{mn}} + \langle\varphi_m|H_1|\varphi_l\rangle &= E_i^{(1)} \underbrace{\langle\varphi_m|\varphi_l\rangle}_{\delta_{ml}} + E_i^{(0)} \sum_{n=0}^{\infty} a_{nl} \underbrace{\langle\varphi_m|\varphi_n\rangle}_{\delta_{mn}} \quad \rightarrow \\ E_m^{(0)} a_{ml} + \langle\varphi_m|H_1|\varphi_l\rangle &= E_i^{(1)} \delta_{ml} + E_i^{(0)} a_{ml} \end{aligned}$$

questa equazione scritta per $m = l$ permette di ritrovare la correzione al primo ordine dell'energia. La parte interessante ora è quella che permette di trovare la correzione all'autofunzione, per $m \neq l$ si ritrova infatti:

$$\left[E_m^{(0)} - E_l^{(0)} \right] a_{ml} = \langle \varphi_l | H_1 | \varphi_m \rangle$$

ovvero:

$$a_{ml} = \frac{\langle \varphi_l | H_1 | \varphi_m \rangle}{E_m^{(0)} - E_l^{(0)}}$$

La correzione al primo ordine risulta quindi:

$$\begin{aligned} E_n &= E_n^{(0)} + \epsilon \langle \varphi_n | H_1 | \varphi_n \rangle \\ |\psi_n\rangle &= |\varphi_n\rangle + \epsilon \sum_{\substack{l=0 \\ l \neq n}}^{+\infty} \frac{\langle \varphi_l | H_1 | \varphi_n \rangle}{E_n^{(0)} - E_l^{(0)}} |\varphi_l\rangle \end{aligned}$$

Da questa relazione si possono dedurre due cose importanti. La prima è che *la correzione al primo ordine del livello n è ortogonale all'autostato imperturbato corrispondente*: $\langle \varphi_n | \psi_n^{(1)} \rangle = 0$, la seconda è che la condizione di validità per l'applicazione del metodo perturbativo è data da:

$$|\langle \varphi_l | H_1 | \varphi_n \rangle| \ll |E_n^{(0)} - E_l^{(0)}|$$

ovvero che *la perturbazione deve essere piccola rispetto alla separazione dei livelli di energia*.

Dalla relazione (8.2c) si può ricavare la correzione dell'energia al secondo ordine, sempre proiettando sullo stato imperturbato $|\psi_l^{(0)}\rangle$ in modo da isolare $E_l^{(2)}$:

$$\langle \psi_l^{(0)} | H_0 | \psi_l^{(2)} \rangle + \langle \psi_l^{(0)} | H_1 | \psi_l^{(1)} \rangle = E_l^{(0)} \langle \psi_l^{(0)} | \psi_l^{(2)} \rangle + \cancel{E_l^{(1)} \langle \psi_l^{(0)} | \psi_l^{(1)} \rangle} + E_l^{(2)} \langle \psi_l^{(0)} | \psi_l^{(0)} \rangle$$

Il primo termine per l'hermiticità dell'Hamiltoniana imperturbata equivale a:

$$\langle \psi_l^{(0)} | H_0 | \psi_l^{(2)} \rangle = \langle \psi_l^{(2)} | H_0 | \psi_l^{(0)} \rangle^* = E_l^{(0)*} \langle \psi_l^{(2)} | \psi_l^{(0)} \rangle^* = E_l^{(0)} \langle \psi_l^{(0)} | \psi_l^{(2)} \rangle$$

e quindi in definitiva:

$$E_l^{(2)} = \langle \psi_l^{(0)} | H_1 | \psi_l^{(1)} \rangle$$

La forma della $\psi_l^{(1)}$ è proprio la correzione al primo ordine calcolata sopra:

$$E_l^{(2)} = \langle \psi_l^{(0)} | H_1 | \sum_{\substack{n=0 \\ n \neq l}}^{+\infty} \frac{\langle \varphi_n | H_1 | \varphi_l \rangle}{E_l^{(0)} - E_n^{(0)}} |\varphi_n\rangle \rangle \equiv \langle \varphi_l | H_1 | \sum_{\substack{n=0 \\ n \neq l}}^{+\infty} \frac{H_{nl}}{E_l^{(0)} - E_n^{(0)}} |\varphi_n\rangle \rangle = \sum_{\substack{n=0 \\ n \neq l}}^{+\infty} \frac{H_{nl}}{E_l^{(0)} - E_n^{(0)}} \langle \varphi_l | H_1 | \varphi_n \rangle$$

dove si è indicato con la notazione H_{nl} l'elemento di matrice $\langle \varphi_n | H_1 | \varphi_l \rangle$. La correzione al secondo ordine dell'energia è data pertanto da:

$$E_n^{(2)} = \sum_{\substack{l=0 \\ l \neq n}}^{+\infty} \frac{|H_{ln}|^2}{E_n^{(0)} - E_l^{(0)}}$$

Per ricavare la correzione al secondo ordine delle autofunzioni si userà ancora l'approccio di Rayleigh-Schrödinger espandendo in serie di autofunzioni imperturbate questo termine correttivo:

$$|\psi_l^{(2)}\rangle = \sum_{m=0}^{\infty} b_{ml} |\varphi_m\rangle$$

la tecnica è sempre la stessa: si sostituisce questa espansione nella relazione (8.2c), che quindi diventa:

$$\sum_{n=0}^{\infty} b_{nl} H_0 |\varphi_n\rangle + H_1 |\psi_l^{(1)}\rangle = E_l^{(0)} \sum_{n=0}^{\infty} b_{nl} |\varphi_n\rangle + E_l^{(1)} |\psi_l^{(1)}\rangle + E_l^{(2)} |\varphi_l\rangle$$

in seguito si tiene conto della forma della correzione al primo ordine dell'autofunzione e si fa agire l'Hamiltoniana H_0 sulle autofunzioni imperturbate $|\varphi_n\rangle$:

$$\sum_{n=0}^{\infty} b_{nl} E_n^{(0)} |\varphi_n\rangle + \sum_{\substack{n=0 \\ n \neq l}}^{+\infty} \frac{H_{nl}}{E_l^{(0)} - E_n^{(0)}} H_1 |\varphi_n\rangle = E_l^{(0)} \sum_{n=0}^{\infty} b_{nl} |\varphi_n\rangle + E_l^{(1)} \sum_{\substack{n=0 \\ n \neq l}}^{+\infty} \frac{H_{nl}}{E_l^{(0)} - E_n^{(0)}} |\varphi_n\rangle + E_l^{(2)} |\varphi_l\rangle$$

Si proietta quindi sull'autostato imperturbato $|\varphi_m\rangle$ per isolare il coefficiente dello sviluppo b_{ml} :

$$\begin{aligned} \sum_{n=0}^{\infty} b_{nl} E_n^{(0)} \langle \varphi_m | \varphi_n \rangle + \sum_{\substack{n=0 \\ n \neq l}}^{+\infty} \frac{H_{nl}}{E_l^{(0)} - E_n^{(0)}} \langle \varphi_m | H_1 | \varphi_n \rangle = \\ = E_l^{(0)} \sum_{n=0}^{\infty} b_{nl} \langle \varphi_m | \varphi_n \rangle + E_l^{(1)} \sum_{\substack{n=0 \\ n \neq l}}^{+\infty} \frac{H_{nl}}{E_l^{(0)} - E_n^{(0)}} \langle \varphi_m | \varphi_n \rangle + E_l^{(2)} \langle \varphi_m | \varphi_l \rangle \end{aligned}$$

Si ipotizza ora che $m \neq l$ (la posizione $m = l$ conduce al termine correttivo dell'energia, come per il primo ordine) e si esplicita il termine correttivo $E_l^{(1)} = H_{ll}$. Tenendo poi conto che le autofunzioni imperturbate sono ortonormali (ovvero che $\langle \varphi_m | \varphi_l \rangle = \delta_{ml}$), si ottiene:

$$E_m^{(0)} b_{ml} + \sum_{\substack{n=0 \\ n \neq l}}^{+\infty} \frac{H_{nl} H_{mn}}{E_l^{(0)} - E_n^{(0)}} = E_l^{(0)} b_{ml} + \frac{H_{ll} H_{ml}}{E_l^{(0)} - E_m^{(0)}}$$

che può essere riarrangiata nella forma:

$$\begin{aligned} \left[E_m^{(0)} - E_l^{(0)} \right] b_{ml} &= \frac{H_{ll} H_{ml}}{E_l^{(0)} - E_m^{(0)}} - \sum_{\substack{n=0 \\ n \neq l}}^{+\infty} \frac{H_{nl} H_{mn}}{E_l^{(0)} - E_n^{(0)}} \rightarrow \\ b_{ml} &= \frac{H_{ll} H_{ml}}{\left[E_l^{(0)} - E_m^{(0)} \right] \left[E_m^{(0)} - E_l^{(0)} \right]} - \sum_{\substack{n=0 \\ n \neq l}}^{+\infty} \frac{H_{nl} H_{mn}}{\left[E_l^{(0)} - E_n^{(0)} \right] \left[E_m^{(0)} - E_l^{(0)} \right]} \rightarrow \\ b_{ml} &= \sum_{\substack{n=0 \\ n \neq l}}^{+\infty} \frac{H_{nl} H_{mn}}{\left[E_l^{(0)} - E_n^{(0)} \right] \left[E_l^{(0)} - E_m^{(0)} \right]} - \frac{H_{ll} H_{ml}}{\left[E_l^{(0)} - E_m^{(0)} \right]^2} \end{aligned}$$

per cui la correzione al secondo ordine dell'autofunzione è espressa dalla formula:

$$|\psi_l^{(2)}\rangle = \sum_{\substack{m=0 \\ m \neq l}}^{+\infty} \sum_{\substack{n=0 \\ n \neq l}}^{+\infty} \frac{H_{nl} H_{mn}}{\left[E_l^{(0)} - E_n^{(0)} \right] \left[E_l^{(0)} - E_m^{(0)} \right]} |\varphi_m\rangle - \sum_{\substack{m=0 \\ m \neq l}}^{+\infty} \frac{H_{ll} H_{ml}}{\left[E_l^{(0)} - E_m^{(0)} \right]^2} |\varphi_m\rangle$$

da cui si ricava la correzione completa al secondo ordine dell'energia e delle autofunzioni:

$$E_n = E_n^{(0)} + \epsilon \langle \varphi_n | H_1 | \varphi_n \rangle + \epsilon^2 \sum_{\substack{l=0 \\ l \neq n}}^{+\infty} \frac{|H_{ln}|^2}{E_n^{(0)} - E_l^{(0)}}$$

$$|\psi_n\rangle = |\varphi_n\rangle + \epsilon \sum_{\substack{l=0 \\ l \neq n}}^{+\infty} \frac{H_{ln}}{E_n^{(0)} - E_l^{(0)}} |\varphi_l\rangle +$$

$$+ \epsilon^2 \left[\sum_{\substack{m=0 \\ m \neq n}}^{+\infty} \sum_{\substack{l=0 \\ l \neq n}}^{+\infty} \frac{H_{ln} H_{ml}}{[E_n^{(0)} - E_l^{(0)}][E_n^{(0)} - E_m^{(0)}]} |\varphi_m\rangle - \sum_{\substack{m=0 \\ m \neq n}}^{+\infty} \frac{H_{nn} H_{mn}}{[E_n^{(0)} - E_m^{(0)}]^2} |\varphi_m\rangle \right]$$

CASO DEGENERE

Come si è visto, quando la perturbazione tende a 0, l'autostato perturbato $|\psi_l\rangle$ del livello l deve tendere all'autofunzione dell'Hamiltoniana imperturbata corrispondente $|\varphi_l\rangle$. Quando lo stato a cui deve tendere l'autofunzione perturbata è degenere sorge però il problema che l'autostato imperturbato non è ben definito, trattandosi di un sottospazio a dimensione > 1 . In questo caso è come se l'autofunzione perturbata "non sapesse" a quale stato ridursi e la procedura descritta in precedenza perde di validità. D'altra parte, siccome esistono in questo caso stati $|\varphi_l\rangle \neq |\varphi_n\rangle$ certamente diversi ma che corrispondono allo stesso autovalore, formule che contengano un termine $E_l^{(0)} - E_n^{(0)}$ a denominatore perdono evidentemente di significato a meno che gli elementi H_{ln} non siano nulli per $|\varphi_l\rangle$ appartenenti al sottospazio $S(E^{(0)})$ dell'autovalore $E^{(0)}$ e diversi da $|\varphi_n\rangle$, ovvero a meno che la perturbazione non sia diagonale nel sottospazio $S(E^{(0)})$. Questa è la chiave per risolvere la questione.

Innanzitutto è necessario aggiungere un indice, che sarà indicato con r , per denotare gli stati degeneri. Si supponga inoltre che l'indice di degenerazione sia α . Vale allora:

$$H_0 |\varphi_{nr}\rangle = E_n^{(0)} |\varphi_{nr}\rangle \quad r \in \{1, \dots, \alpha\}$$

queste autofunzioni sono certamente ortonormali alle $\{\varphi_{ms}\}$ del sottospazio relativo ad un altro autovalore ed è anche possibile sceglierle ortonormali fra loro.³

L'ipotesi di espansione in termini di ϵ è ancora valida e quindi anche le equazioni (8.2a), (8.2b) e (8.2c), a patto di aggiungere opportunamente una sommatoria sull'indice di degenerazione r . Per l'ordine 0 si ottiene quindi:

$$H_0 |\psi_{lr}^{(0)}\rangle = E_l^{(0)} |\psi_{lr}^{(0)}\rangle$$

che permette di concludere banalmente che $E_{lr}^{(0)} \equiv E_l^{(0)}$, ovvero che le autofunzioni $|\psi_{lr}^{(0)}\rangle$ sono relative tutte allo stesso autovalore e che le stesse sono quindi esprimibili come combinazioni lineari delle autofunzioni del sottospazio degenere:

$$|\psi_{lr}^{(0)}\rangle = \sum_{s=0}^{\alpha} c_{rs} |\varphi_{ls}\rangle$$

Per conservare l'analogia al caso non degenere, si deve esprimere la $|\psi_{lr}^{(1)}\rangle$ in termini delle autofunzioni degeneri dell'Hamiltoniana imperturbata:

$$|\psi_{lr}^{(1)}\rangle = \sum_{n=0}^{+\infty} \sum_{s=0}^{\alpha} c_{lr,ns} |\varphi_{ns}\rangle$$

che sostituite nell'equazione (8.2b) del primo ordine, permettono di scrivere:

$$H_0 \sum_{n=0}^{+\infty} \sum_{s=0}^{\alpha} c_{lr,ns} |\varphi_{ns}\rangle + H_1 \sum_{s=0}^{\alpha} c_{rs} |\varphi_{ls}\rangle = E_l^{(0)} \sum_{n=0}^{+\infty} \sum_{s=0}^{\alpha} c_{lr,ns} |\varphi_{ns}\rangle + E_{lr}^{(1)} \sum_{s=0}^{\alpha} c_{rs} |\varphi_{ls}\rangle$$

³È senz'altro sempre possibile ortonormalizzarle utilizzando il metodo di Gram-Schmidt.

Ora $H_0|\varphi_{ns}\rangle = E_n^{(0)}|\varphi_{ns}\rangle$, per cui si ha:

$$\sum_{n=0}^{+\infty} \sum_{s=0}^{\alpha} c_{lr,ns} E_n^{(0)} |\varphi_{ns}\rangle + H_1 \sum_{s=0}^{\alpha} c_{rs} |\varphi_{ls}\rangle = E_l^{(0)} \sum_{n=0}^{+\infty} \sum_{s=0}^{\alpha} c_{lr,ns} |\varphi_{ns}\rangle + E_{lr}^{(1)} \sum_{s=0}^{\alpha} c_{rs} |\varphi_{ls}\rangle$$

ovvero:

$$\sum_{n=0}^{+\infty} \sum_{s=0}^{\alpha} c_{lr,ns} [E_n^{(0)} - E_l^{(0)}] |\varphi_{ns}\rangle + \sum_{s=0}^{\alpha} c_{rs} [H_1 - E_{lr}^{(1)}] |\varphi_{ls}\rangle = 0$$

Si può ora proiettare questa equazione sullo stato $|\varphi_{lu}\rangle$:

$$\sum_{n=0}^{+\infty} \sum_{s=0}^{\alpha} c_{lr,ns} [E_n^{(0)} - E_l^{(0)}] \langle \varphi_{lu} | \varphi_{ns} \rangle + \sum_{s=0}^{\alpha} c_{rs} \langle \varphi_{lu} | [H_1 - E_{lr}^{(1)}] | \varphi_{ls} \rangle = 0$$

Tenendo presente che nel sottospazio degenere risulta $\langle \varphi_{lu} | \varphi_{ls} \rangle = \delta_{us}$,⁴ mentre per autovalori diversi risulta certamente $\langle \varphi_{lu} | \varphi_{ns} \rangle = 0$, la prima sommatoria sparisce. Infatti, per $l \neq n$ il prodotto scalare è nullo e per $l = n$ è il termine $E_n^{(0)} - E_l^{(0)}$ che si annulla. Quindi l'equazione precedente si riduce alla seguente:

$$\sum_{s=0}^{\alpha} c_{rs} [\langle \varphi_{lu} | H_1 | \varphi_{ls} \rangle - E_{lr}^{(1)} \delta_{ls}] = 0$$

questa formula rappresenta in realtà un sistema di equazione nelle α incognite $c_{r1}, \dots, c_{r\alpha}$, per avere quindi una soluzione diversa da quella banale bisogna imporre che il determinante del sistema sia nullo:

$$\det [\langle \varphi_{lu} | H_1 | \varphi_{ls} \rangle - E_{lr}^{(1)} \delta_{ls}] = 0 \quad (8.3)$$

il termine fra parentesi rappresenta infatti una matrice con indici di riga e colonna l e s . Questo sistema permette di ricavare le α correzioni $E_{l1}^{(1)}, \dots, E_{l\alpha}^{(1)}$ al primo ordine da apportare a ogni autovalore degenere dell'Hamiltoniana imperturbata. Si noti che questa ha la forma di un'equazione secolare ed infatti questo procedimento equivale in pratica a *diagonalizzare la perturbazione nel sottospazio degenere*.

In linea di principio la perturbazione riesce a risolvere la degenerazione almeno parzialmente, in altre parole ad ogni autovalore degenere corrispondono più correzioni diverse al primo ordine che si riducono a zero – quindi al valore imperturbato – al tendere a zero del parametro ϵ . Se le soluzioni sono tutte distinte, allora la perturbazione rimuove completamente la degenerazione.

Riassumendo, quando l'Hamiltoniana è degenere, la strada da seguire per applicare la teoria delle perturbazioni è la seguente:

- Si sceglie una base di autofunzioni φ_{ns} di H_0 ,
- Si scrive l'elemento di matrice $\langle \varphi_{nr} | H_1 | \varphi_{ns} \rangle$,
- Si diagonalizza questa matrice nel sottospazio degenere risolvendo l'equazione (8.3),
- Le soluzioni di questa equazione rappresentano le correzioni al primo ordine dell'energia,
- Gli autovettori diagonalizzati corrispondenti rappresentano le autofunzioni imperturbate da cui partire,
- Con queste autofunzioni, le correzioni successive al primo ordine hanno la stessa forma del caso non degenere.

⁴Perché si è assunto che gli autovettori nel sottospazio degenere siano ortonormalizzati.

Nella pratica, il fatto che l'Hamiltoniana sia degenere significa che esistono degli osservabili che commutano con essa e che permettono di distinguere i vari autostati degeneri. La ricerca di una base di autofunzioni per diagonalizzare la perturbazione può essere quindi guidata da motivazioni fisiche, si può cioè scegliere una base di autofunzioni simultanea di H_0 e degli altri osservabili che commutano sia con H_0 che con la perturbazione H_1 . In questo caso, infatti, la perturbazione risulta già diagonale nel sottospazio degenere. Questo non significa però che H_1 è diagonale in tutto lo spazio di Hilbert, che sarebbe in contraddizione con il fatto che H_0 e H_1 non commutano.

Una nota conclusiva sulla teoria delle perturbazioni indipendenti dal tempo prima di passare ad un esempio concreto.

In generale si dovrebbe applicare sempre la teoria per autovalori degeneri, il caso di autovalore non degenere è in effetti un caso particolare della teoria più generale. Inoltre la correzione in entrambi i casi si può considerare come l'autovalore del termine di perturbazione. Si può sostenere questa affermazione in maniera formale, ma se ne può anche fornire un'argomentazione intuitiva. La correzione all'energia è certamente un'osservabile e dunque deve essere indipendente dalla rappresentazione. Se si rappresentassero gli stati nella base di autofunzioni del termine di perturbazione, sia nel caso di correzione per stati non degeneri che nel caso di correzione per stati degeneri, l'espressione trovata per le correzioni coinciderebbe con la definizione di autovalori del termine di perturbazione.

Ma se la correzione deve essere indipendente dalla rappresentazione questo vuol dire che anche nelle altre rappresentazioni (per esempio nella rappresentazione degli autostati dell'Hamiltoniana imperturbata utilizzata) si stanno calcolando in maniera implicita gli autovalori del termine di perturbazione.

D'altra parte sono poche le caratteristiche di un operatore che sono indipendenti dalla rappresentazione usata, e fra queste ci sono gli autovalori.

In conclusione, *la teoria delle perturbazioni conduce in ogni caso alla ricerca degli autovalori del termine di perturbazione (diagonalizzazione).*

8.1.1 Un esempio concreto: La correzione relativistica all'atomo di idrogeno

In questa sezione si tratterà in concreto un caso di perturbazione: la correzione relativistica al moto dell'elettrone nell'atomo di idrogeno (*correzione di Thomas*).

La trattazione dell'atomo di idrogeno fatta nel capitolo §7 è *in linea generale* adeguata, esistono però ancora una serie di effetti importanti (e pertanto di correzioni) che devono essere presi in conto. Questi effetti sono relativamente piccoli per l'atomo di idrogeno, ma sono importanti per i cosiddetti *atomi idrogenoidi*, ovvero atomi con $Z > 1$ e un solo elettrone. In questi casi, i termini correttivi diventano via via più importanti man mano che aumenta Z . Il primo di questi termini, e l'unico che sarà preso in considerazione in questo corso, è quello che tiene conto della forma relativistica dell'energia cinetica.⁵

$$E_{\text{cin.}} = \sqrt{m^2c^4 + p^2c^2} - mc^2$$

La velocità dell'elettrone è bassa da un punto di vista relativistico, quindi questo termine si può espandere in serie in termini di v/c , il rapporto fra la velocità dell'elettrone e quella della luce. Qualche semplice manipolazione permette di esprimere questa energia cinetica in una forma facilmente utilizzabile in termini di teoria perturbativa:

$$E_{\text{cin.}} = \sqrt{m^2c^4 + p^2c^2} - mc^2 = mc^2 \sqrt{1 + \frac{p^2c^2}{m^2c^4}} - mc^2$$

ricordando che $p = mv$, sotto radice si ritrova proprio il termine in v/c ricercato:

$$\frac{p^2c^2}{m^2c^4} = \frac{p^2}{m^2c^2} = \frac{m^2 v^2}{m^2 c^2} \ll 1$$

⁵La massa qui è denotata con m per comodità, si tenga bene a mente però che in tutto quello che riguarda gli atomi si intende sempre *massa ridotta*.

l'energia cinetica ha quindi una forma del tipo $(1+x)^n$ con $n = 1/2$ e $x = p^2 c^2 / m^2 c^4$, che per $x \ll 1$ si può espandere in serie:

$$(1+x)^n \simeq 1 + ax + \frac{a(a-1)}{2} x^2$$

di conseguenza, trascurando i termini superiori al secondo ordine, si ricava:

$$E_{\text{cin.}} \simeq mc^2 \left(1 + \frac{1}{2} \frac{p^2}{mc^2} - \frac{1}{8} \frac{p^4}{m^4 c^2} \right) - mc^2 = \frac{p^2}{2m} - \frac{p^4}{8m^3 c^2}$$

Si vede quindi che il termine di correzione relativistico è, appunto, un termine del secondo ordine.

Ricordando che la forma dell'Hamiltoniana imperturbata è data da:⁶

$$H_0 = \frac{p^2}{2m} - \frac{e^2}{r}$$

si vede immediatamente che la forma del termine perturbativo è esprimibile in questi termini:

$$H_1 = \frac{1}{2mc^2} \left(H_0 + \frac{e^2}{r} \right)^2$$

Questa forma permette di vedere che anche la correzione relativistica è invariante per rotazioni e quindi commuta con il momento angolare. Come base di autofunzioni può pertanto essere scelta proprio quella dell'atomo di idrogeno, e la correzione al livello fondamentale è quindi data dal termine:

$$\begin{aligned} E_0^{(1)} &= -\frac{1}{2mc^2} \langle \psi_{nlm} | \left(\frac{p^2}{2m} \right)^2 | \psi_{nlm} \rangle = \\ &= -\frac{1}{2mc^2} \langle \psi_{nlm} | H_0^2 + \frac{e^2}{r} H_0 + H_0 \frac{e^2}{r} + \frac{e^4}{r^2} | \psi_{nlm} \rangle = \\ &= -\frac{1}{2mc^2} \langle \psi_{nlm} | H_0^2 + 2E_n^{(0)} \frac{e^2}{r} + \frac{e^4}{r^2} | \psi_{nlm} \rangle \end{aligned}$$

dove $E_n^{(0)} = -me^4 / 2n^2 \hbar^2$ sono naturalmente i livelli imperturbati dell'atomo di idrogeno. Il calcolo di questi integrali può essere certamente fatto per forza bruta, esiste però un metodo più furbo e economico di ottenere il risultato.

Si consideri il termine $\langle \psi_{nlm} | \frac{e^2}{r} | \psi_{nlm} \rangle$, che rappresenta una funzione omogenea di grado -1. Si noti che in un autostato dell'Hamiltoniana risulta $\langle H \rangle = \langle T \rangle - \langle e^2/r \rangle$. Il teorema del Viriale permette inoltre di stabilire che $2\langle T \rangle = \langle e^2/r \rangle$, per cui si giunge alla relazione:

$$\begin{aligned} \langle H_0 \rangle &= \frac{1}{2} \langle \frac{e^2}{r} \rangle - \langle \frac{e^2}{r} \rangle \quad \Rightarrow \quad \langle H_0 \rangle = -\frac{1}{2} \langle \frac{e^2}{r} \rangle \quad \Rightarrow \quad \langle \frac{e^2}{r} \rangle \\ &= -2\langle H_0 \rangle \quad \Rightarrow \quad \langle \psi_{nlm} | \frac{e^2}{r} | \psi_{nlm} \rangle = -2E_n^{(0)} \end{aligned}$$

Il valore di $\langle e^4/r^2 \rangle$ può essere ricavato con argomenti analoghi (*Teorema di Feynman-Hellman*), ricavando:

$$\langle \frac{e^4}{r^2} \rangle = \frac{8n}{2l+1} E_n^{(0)2}$$

Il teorema del viriale in meccanica quantistica

⁶Si utilizza qui il sistema di unità elettrostatiche per comodità di notazione. Al momento opportuno le costanti saranno ripristinate correttamente.

Il *teorema del viriale* afferma che in un sistema fisico in cui l'energia potenziale sia una funzione omogenea di grado n nelle coordinate,⁷ il valore medio dell'energia cinetica T e dell'energia potenziale U sono legati dalla relazione:

$$2\langle T \rangle = n\langle U \rangle$$

Infatti, notando che vale $\langle \psi | [q_i p_i, H] | \psi \rangle = 0$, si ricava:

$$\begin{aligned} 0 &= \langle \psi | [q_i p_i, H] | \psi \rangle = \langle \psi | q_i [p_i, H] | \psi \rangle + \langle \psi | [q_i, H] p_i | \psi \rangle \\ &= \langle \psi | q_i [p_i, T(p_i)] | \psi \rangle + \langle \psi | [q_i, U(q_i)] p_i | \psi \rangle \end{aligned}$$

avendo tenuto conto del fatto che:

$$[q_i, T(p_i)] = [p_i, U(q_i)] = 0$$

Dalle relazioni sul commutatore (4.3) segue che:

$$\begin{cases} \frac{\partial T(p_i)}{\partial p_i} p_i = 2T \\ \frac{\partial U(q_i)}{\partial q_i} q_i = nU \end{cases}$$

ovvero:

$$-i\hbar \langle \psi | nU | \psi \rangle = i\hbar \langle \psi | 2T | \psi \rangle$$

cioè il teorema del viriale.

La correzione relativistica al primo ordine è data pertanto da:

$$\begin{aligned} \Delta E_{nl}^{(0)} &= -\frac{1}{2mc^2} \langle \psi_{nlm} | H_0^2 + 2E_n^{(0)} \frac{e^2}{r} + \frac{e^4}{r^2} | \psi_{nlm} \rangle = \\ &= -\frac{1}{2mc^2} \left[E_n^{(0)2} + 2E_n^{(0)} \cdot (-2E_n^{(0)2}) + \frac{8n}{l+1} E_n^{(0)2} \right] = \\ &= -\frac{1}{2mc^2} \left[E_n^{(0)2} - 4E_n^{(0)2} + \frac{8n}{l+1} E_n^{(0)2} \right] = \\ &= -\frac{1}{2mc^2} \left[\frac{8n}{l+1} - 3 \right] E_n^{(0)2} \end{aligned}$$

ricordando che $E^{(0)} = -mc^4/2n^2\hbar^2$ e $\alpha \equiv e^2/c\hbar$ (*costante di struttura fine*), il termine di correzione si può riscrivere nella forma:

$$\Delta E_n^{(0)} = E_n^{(0)} \frac{\alpha^2}{4n^2} \left[\frac{8n}{l+1} - 3 \right]$$

Come si vede, questa correzione dipende dal numero quantico l , per cui permette di distinguere livelli appartenenti allo stesso numero quantico principale n che altrimenti sono degeneri. Il risultato è che anche in assenza di "perturbazioni esterne" – per esempio un campo elettrico o magnetico applicato all'atomo – i livelli energetici risultano in realtà divisi in una serie di linee sottili, presentano cioè quella che si definisce una *struttura fine* delle righe spettrali. Siccome l'entità di questo effetto dipende in ultima istanza dal valore della costante α , questo giustifica il suo nome di *costante di struttura fine* e la piccolezza dell'effetto è legato proprio al piccolo valore di questa costante.⁸

Si noti che questo effetto dipende in realtà dal prodotto $Z\alpha$, per cui in atomi idrogenoidi con $Z \gg 1$ questo termine non è più sufficientemente piccolo da giustificare l'approccio perturbativo.

⁷Si definisce *omogenea di grado n* una funzione tale che quando si moltiplica per un certo numero $\alpha > 0$ ogni sua variabile, il suo valore si calcola moltiplicando per α^n la funzione calcolata negli argomenti originari (cioè senza α). La definizione formalmente corretta è la seguente: sia $\alpha, k \in \mathbb{R}$, $\alpha > 0$, una funzione $f(x_1, \dots, x_n)$ definita su un cono di \mathbb{R}^n si definisce *positivamente omogenea di grado k* se:

$$\forall (x_1, \dots, x_n) \quad f(\alpha x_1, \dots, \alpha x_n) = \alpha^k f(x_1, \dots, x_n)$$

Se questa relazione vale per ogni valore di α , la funzione è detta *omogenea di grado n* .

⁸Si ricordi che vale $\alpha \simeq 1/137$.

8.2 Perturbazioni dipendenti dal tempo

Nel caso in cui il termine perturbativo H_1 dipende dal tempo si parla di *teoria delle perturbazioni dipendenti dal tempo*. Si tratta in questo caso di un sistema non isolato in cui non si conserva l'energia, per cui il problema da risolvere non sarà quello di determinare le autofunzioni (che fra l'altro cambiano da un istante all'altro), piuttosto quello di trovare la probabilità che un sistema nello stato $\psi(\vec{r}, t)$ all'istante t si trovi nello stato $\psi'(\vec{r}', t')$ all'istante t' .

Se l'Hamiltoniana del sistema è esprimibile come $H(t) = H_0 + \epsilon H_1(t)$ o, secondo un'altra notazione molto diffusa, come:

$$H(t) = H_0 + \epsilon W(t)$$

ovvero come un'Hamiltoniana imperturbata (che si suppone nota) e un termine dipendente dal tempo, il problema si può ricondurre al problema di valutare la probabilità di transizione fra un autostato di H_0 all'istante iniziale t verso un altro autostato di H_0 all'istante finale t' . Se il termine $\epsilon W(t)$ è piccolo, si può impostare uno sviluppo perturbativo di questa probabilità di transizione. Per fare ciò è necessario però introdurre anche l'ipotesi di *approssimazione adiabatica*: ovvero che la perturbazione vari lentamente rispetto ai tempi caratteristici del sistema in modo che abbia senso la scrittura:⁹

$$H(t)\psi(\vec{r}, t) = E(t)\psi(\vec{r}, t)$$

Si suppone inoltre che la perturbazione sia limitata nel tempo (*ipotesi adiabatica*), ovvero che:

$$\lim_{t \rightarrow -\infty} W(t) = \lim_{t \rightarrow +\infty} W(t) = 0 \quad (8.4)$$

La funzione d'onda che descrive lo stato $\psi(\vec{r}, t)$ in questo caso deve ovviamente soddisfare l'equazione di Schrödinger:

$$i\hbar \frac{\partial}{\partial t} |\psi(\vec{r}, t)\rangle = [H_0 + \epsilon W(t)] |\psi(\vec{r}, t)\rangle$$

Questa funzione d'onda può essere sviluppata in termini di autofunzioni $|\varphi_m(\vec{r})\rangle$ dell'Hamiltoniana imperturbata:

$$|\psi(\vec{r}, t)\rangle = \sum_{m=0}^{+\infty} \langle \psi(\vec{r}, t) | \varphi_m(\vec{r}) \rangle |\varphi_m(\vec{r})\rangle \equiv \sum_{n=0}^{+\infty} c_n(t) |\varphi_n(\vec{r})\rangle \quad (8.5)$$

sviluppo che può essere sostituito nell'equazione di Schrödinger:

$$i\hbar \sum_{m=0}^{+\infty} \frac{d}{dt} c_m(t) |\varphi_m(\vec{r})\rangle = \sum_{m=0}^{+\infty} c_m(t) [H_0 + \epsilon W(t)] |\varphi_m(\vec{r})\rangle$$

Proiettando ora sullo stato $|\varphi_n\rangle$, sfruttando le proprietà di ortonormalità delle autofunzioni imperturbate e il fatto che H_0 è diagonale nella sua base:

$$i\hbar \sum_{m=0}^{+\infty} \frac{d}{dt} c_m(t) \langle \varphi_n | \varphi_m \rangle = \sum_{m=0}^{+\infty} c_m(t) \langle \varphi_n | H_0 | \varphi_m \rangle + \epsilon \sum_{m=0}^{+\infty} c_m(t) \langle \varphi_n | W(t) | \varphi_m \rangle \rightarrow$$

$$i\hbar \frac{d}{dt} c_n(t) = E_n c_n(t) + \epsilon \sum_{m=0}^{+\infty} c_m(t) W_{nm}(t) \quad (8.6)$$

dove la dipendenza dalle coordinate spaziali \vec{r} è stata omessa e si è denotato con $W_{nm}(t)$ l'elemento di matrice $\langle \varphi_n | W(t) | \varphi_m \rangle$ per alleggerire la notazione. Se sul sistema non agisse alcuna perturbazione, il termine $W_{nm}(t)$ sarebbe nullo e l'equazione si ridurrebbe ad una separazione delle variabili, fornendo la nota soluzione:

$$c_n(t) = b_n e^{-i \frac{E_n}{\hbar} t}$$

con b_n coefficienti costanti ricavabili a partire dalle condizioni iniziali. Nel caso sul sistema agisca una perturbazione dipendente dal tempo, si fa l'ipotesi che il sistema continui ad essere descritto

⁹Questo equivale a dire che si possono trascurare le probabilità di transizione fra un livello e l'altro del sistema e che i livelli di energia variano lentamente nel tempo.

dalla stessa equazione, ma con i coefficienti $b_n \equiv b_n(t)$ dipendenti dal tempo (che corrisponde in pratica ad utilizzare l'approssimazione adiabatica):

$$c_n(t) = b_n(t)e^{-i\frac{E_n}{\hbar}t}$$

che sostituita nella (8.6) fornisce:

$$\begin{aligned} i\hbar \frac{d}{dt} \left[b_n(t)e^{-i\frac{E_n}{\hbar}t} \right] &= E_n b_n(t)e^{-i\frac{E_n}{\hbar}t} + \epsilon \sum_{m=0}^{+\infty} W_{nm}(t)b_m(t)e^{-i\frac{E_m}{\hbar}t} \rightarrow \\ i\hbar \frac{d}{dt} b_n(t) + E_n b_n(t)e^{-i\frac{E_n}{\hbar}t} &= E_n b_n(t)e^{-i\frac{E_n}{\hbar}t} + \epsilon \sum_{m=0}^{+\infty} W_{nm}(t)b_m(t)e^{-i\frac{E_m}{\hbar}t} \end{aligned}$$

Eliminando il termine comune $e^{-i\frac{E_m}{\hbar}t}$ (che fra l'altro non si annulla mai) e ponendo $\omega_{nm} \equiv (E_n - E_m)/\hbar$ (frequenza di Bohr) si giunge alla forma:

$$i\hbar \frac{d}{dt} b_n(t) = \epsilon \sum_{m=0}^{+\infty} W_{nm}(t)b_m(t)e^{-i\omega_{nm}t} \quad (8.7)$$

Si noti che se la perturbazione si annulla, i coefficienti $b_n(t)$ diventano indipendenti dal tempo in quanto descrivono la soluzione della ben nota equazione di Schrödinger dipendente dal tempo ma con una Hamiltoniana indipendente dal tempo. Ne consegue che *la dipendenza dal tempo della perturbazione è tutta contenuta nei coefficienti $b_n(t)$* .

L'ipotesi adiabatica permette di stabilire che per $t \rightarrow -\infty$ i coefficienti $b_n(t)$ devono diventare costanti e noti, perché lo stato iniziale del sistema si suppone noto¹⁰. Per comodità si può supporre che lo stato iniziale sia un'autostato dell'energia E_i , per cui:

$$\lim_{t \rightarrow -\infty} b_n(t) = \delta_{ni}$$

Durante la perturbazione, i coefficienti cambieranno e saranno diversi da zero, in quanto la perturbazione provoca delle transizioni dallo stato iniziale verso altri stati. Dopo che la perturbazione ha agito (quindi per $t \rightarrow +\infty$), il sistema si troverà in uno stato E_f , ed i coefficienti $|b_n(+\infty)|^2$ rappresenteranno le probabilità di trovarsi nei vari stati finali. Questi coefficienti prendono quindi il nome di *ampiezze di transizione*.

Le equazioni (8.7) rappresentano un sistema di infinite equazioni differenziali al primo ordine in t nelle funzioni $b_n(t)$. Questo sistema può essere risolto in modo relativamente semplice in maniera iterativa se si fa l'ipotesi che le ampiezze di transizione si possano espandere in serie di potenze di ϵ , ovvero che valga la relazione:

$$b_n(t) = \sum_{k=0}^{+\infty} b_n^{(k)}(t)\epsilon^k \quad (8.8)$$

Si può quindi inserire questo sviluppo nella (8.7) ottenendo:

$$\sum_{k=0}^{+\infty} \frac{d}{dt} b_n^{(k)}(t)\epsilon^k = \frac{1}{i\hbar} \sum_{m=0}^{+\infty} \sum_{k=0}^{+\infty} \epsilon^{k+1} b_m^{(k)}(t)W_{nm}(t)e^{-i\omega_{nm}t}$$

ed eguagliando le potenze corrispondenti in ϵ si ricavano le relazioni di ricorrenza:

$$\begin{aligned} \frac{d}{dt} b_n^{(0)}(t) &= 0 \\ \frac{d}{dt} b_n^{(1)}(t) &= \frac{1}{i\hbar} \sum_{m=0}^{+\infty} b_m^{(0)}(t)W_{nm}(t)e^{-i\omega_{nm}t} \\ &\dots \\ \frac{d}{dt} b_n^{(k)}(t) &= \frac{1}{i\hbar} \sum_{m=0}^{+\infty} b_m^{(k-1)}(t)W_{nm}(t)e^{-i\omega_{nm}t} \end{aligned} \quad (8.9)$$

¹⁰E quindi sono noti i coefficienti costanti dello sviluppo in serie di autovalori.

Dalla prima di queste si vede subito che i coefficienti all'ordine 0 sono costanti nel tempo:

$$\frac{d}{dt}b_n^{(0)}(t) = 0 \quad \Rightarrow \quad b_n^{(0)}(t) \equiv b_n = \text{costante}$$

questo è coerente, in quanto l'ordine $\epsilon = 0$ corrisponde all'Hamiltoniana imperturbata e quindi lo sviluppo in autostati è quello indipendente dal tempo. I coefficienti all'ordine 0 valgono quindi $b_n = \delta_{ni}$ dove i indica l'autostato iniziale.

Sostituendo il valore di $b_n^{(0)}$ nell'equazione al primo ordine si ottiene:

$$\frac{d}{dt}b_n^{(1)}(t) = \frac{1}{i\hbar}\delta_{ni}W_{nm}(t)e^{-i\omega_{nm}t} = \frac{1}{i\hbar}W_{ni}(t)e^{-i\omega_{ni}t}$$

che può essere immediatamente integrata rispetto al tempo:

$$b_n^{(1)}(t) = \frac{1}{i\hbar} \int_{-\infty}^t W_{ni}(t')e^{-i\omega_{ni}t'} dt'$$

dove si è fatto partire l'integrale da $t = -\infty$, tenendo conto che $b_n^{(1)}(t)$ si deve annullare per $t \rightarrow -\infty$ e che la perturbazione non ha ancora cominciato ad agire (per il principio di causalità, il coefficiente può dipendere solo dalle condizioni iniziali e da $H_1(t')$, $t' < t$). Analogamente, se si aggiunge il termine 0 e si estende l'integrazione fino a $t = +\infty$ si ottiene l'*ampiezza di transizione completa al primo ordine*:

$$b_n^{(1)}(t) = \delta_{ni} + \epsilon \frac{1}{i\hbar} \int_{-\infty}^{+\infty} W_{ni}(t')e^{-i\omega_{ni}t'} dt'$$

questo integrale rappresenta la *trasformata di Fourier dell'elemento di matrice della perturbazione calcolata alla frequenza di Bohr ω_{ni} della transizione*. Per avere senso, l'integrale deve convergere e la convergenza è assicurata proprio dall'approssimazione adiabatica. Inoltre, se la perturbazione dura un tempo finito fra $t = 0$ e $t = t_0$, allora l'integrale è di fatto limitato a questo intervallo, perché per $t < 0$ e per $t > t_0$ l'integrando¹¹ è nullo.

In definitiva, dopo che è cessata la perturbazione la probabilità di transizione dallo stato iniziale i allo stato n è data al primo ordine da:

$$\mathcal{P}_{i \rightarrow n} = |b_n^{(1)}(+\infty)|^2$$

mentre la probabilità di restare nello stato i è data evidentemente da:

$$\mathcal{P}_i = 1 - \sum_{\substack{n=1 \\ n \neq i}}^{+\infty} |b_n^{(1)}(+\infty)|^2$$

L'insieme di equazioni (8.9) permette di ricavare per ricorrenza il termine al secondo ordine (e più in generale tutti gli altri). Per ricavare questo termine conviene tuttavia cambiare la rappresentazione, e passare da quella di Schrödinger S a quella di Dirac o di *interazione* I . Per fare questo occorre effettuare la trasformazione unitaria sui vettori di stato e sugli operatori:

$$|\psi(t)\rangle_I = e^{i\frac{H_0}{\hbar}t} |\psi(t)\rangle_S \quad \hat{O}_I(t) = e^{i\frac{H_0}{\hbar}t} \hat{O}_S(t) e^{-i\frac{H_0}{\hbar}t}$$

e applicando questa trasformazione allo sviluppo in termini di autofunzioni (8.5) si ricava:¹²

$$|\psi(t)\rangle_I = \sum_{n=0}^{+\infty} b_n(t) |\varphi_n\rangle$$

per cui i coefficienti dello sviluppo (8.5) sono proprio i coefficienti di sviluppo dello stato nella rappresentazione di interazione.

¹¹Ovvero: $W_{ni}(t')$.

¹²Si applica l'operatore $e^{i\frac{H_0}{\hbar}t}$ a entrambi i membri e si nota che $c_n(t) = b_n(t)e^{i\frac{H_0}{\hbar}t}$

La rappresentazione di interazione

La *rappresentazione di interazione* o di *Dirac* è una rappresentazione della meccanica quantistica che si colloca a metà strada fra la *rappresentazione di Schrödinger* considerata finora e quella di *Heisemberg*. La rappresentazione di Schrödinger, trattata nel paragrafo §4.9, prevede che l'evoluzione nel tempo sia dettata dall'operatore temporale (4.6) che agisce sullo stato del sistema, mentre gli osservabili sono indipendenti dal tempo. Nella rappresentazione di Heisemberg, invece, sono i vettori di stato ad essere costanti nel tempo mentre gli osservabili evolvono temporalmente secondo la $\hat{O}(t) = e^{i\frac{H_0}{\hbar}t}\hat{O}e^{-i\frac{H_0}{\hbar}t}$.

Nella rappresentazione di interazione, sia il vettore di stato che l'operatore osservabile evolvono nel tempo, anche se in maniera diversa. Questa rappresentazione, definita rispetto alla rappresentazione di Schrödinger dalle equazioni (8.2) è particolarmente utile in questo contesto perché permette di "scaricare" la dipendenza temporale dell'Hamiltoniana H_0 sugli operatori, mentre l'evoluzione temporale degli stati risulta determinata dalla sola H_1 .

Risulta quindi chiaro che in assenza del termine perturbativo $W_I(t)$, la rappresentazione di interazione si riduce a quella di Heisemberg,¹³ e l'equazione del moto risulta:

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle_I = i\hbar\frac{\partial}{\partial t}e^{i\frac{H_0}{\hbar}t}|\psi(t)\rangle_S = e^{i\frac{H_0}{\hbar}t}(H - H_0)|\psi(t)\rangle_S = W_I(t)|\psi(t)\rangle_I$$

dove con $W_I(t)$ si è indicato il termine perturbativo nella rappresentazione di interazione.

Sostituendo lo sviluppo in autofunzioni nella rappresentazione di interazione nell'equazione del moto appena scritta si ricava:

$$i\hbar\frac{\partial}{\partial t}\sum_{m=0}^{+\infty}b_m(t)|\varphi_m\rangle = \sum_{m=0}^{+\infty}b_m(t)W_I(t)|\varphi_m\rangle \quad \rightarrow \quad i\hbar\sum_{m=0}^{+\infty}\frac{d}{dt}b_m(t)|\varphi_m\rangle = \sum_{m=0}^{+\infty}b_m(t)W_I(t)|\varphi_m\rangle$$

e proiettando come al solito sull'autostato $|\varphi_n\rangle$:

$$i\hbar\sum_{m=0}^{+\infty}\frac{d}{dt}b_m(t)\underbrace{\langle\varphi_n|\varphi_m\rangle}_{\delta_{nm}} = \sum_{m=0}^{+\infty}b_m(t)\langle\varphi_n|W_I(t)|\varphi_m\rangle \quad \rightarrow \quad i\hbar\frac{d}{dt}b_n(t) = \sum_{m=0}^{+\infty}b_m(t)\langle\varphi_n|W_I(t)|\varphi_m\rangle$$

ovvero:

$$b_n(t) = \frac{1}{i\hbar}\sum_{m=0}^{+\infty}b_m(t)(W_{nm})_I(t)$$

dove con $(W_{nm})_I(t)$ si è indicato l'elemento di matrice $\langle\varphi_n|W_I(t)|\varphi_m\rangle$.¹⁴ Introducendo lo sviluppo (8.8) si ritrovano le stesse equazioni iterative (8.9), scritte però nella rappresentazione di interazione:

$$\frac{d}{dt}b_n^{(k)}(t) = \frac{1}{i\hbar}\sum_{m=0}^{+\infty}b_m^{(k-1)}(t)(W_{nm})_I(t)$$

dove la soluzione al primo ordine è evidentemente ancora:

$$b_n^{(1)}(t) = \frac{1}{i\hbar}\int_{-\infty}^t(W_{ni})_I(t_1)dt_1$$

La correzione al secondo ordine si ottiene facilmente in funzione di quella al primo ordine:

$$\frac{d}{dt}b_n^{(2)}(t) = \frac{1}{i\hbar}\sum_{m=0}^{+\infty}b_m^{(1)}(t)(W_{nm})_I(t) = \frac{1}{(i\hbar)^2}\int_{-\infty}^t\sum_{m=0}^{+\infty}(W_{nm})_I(t)(W_{mi})_I(t_2)dt_2$$

¹³Evidentemente, se l'evoluzione degli stati è determinata dal solo termine perturbativo, se questo si annulla gli stati sono indipendenti dal tempo e quindi l'evoluzione temporale è contenuta nei soli operatori: la rappresentazione si riduce a quella di Heisemberg.

¹⁴Si noti che questa equazione coincide con la (8.7), in quanto:

$$\langle\varphi_n|W_I(t)|\varphi_m\rangle = e^{i\omega_{nm}t}\langle\varphi_n|W_S(t)|\varphi_m\rangle$$

Questo è perfettamente logico, in quanto si tratta della stessa interazione, solamente scritta in due rappresentazioni diverse.

integrando sul tempo:

$$b_n^{(2)}(t) = \frac{1}{(i\hbar)^2} \int_{-\infty}^t \int_{-\infty}^{t_1} \sum_{m=0}^{+\infty} (W_{nm})_I(t_1) (W_{mi})_I(t_2) dt_2 dt_1$$

ovvero:

$$b_n^{(2)}(t) = \frac{1}{(i\hbar)^2} \int_{-\infty}^t \int_{-\infty}^{t_1} \langle \varphi_n | W_I(t_1) W_I(t_2) | \varphi_i \rangle dt_2 dt_1$$

Questa relazione può facilmente essere iterata per ricavare il k -simo termine correttivo:

$$b_n^{(k)}(t) = \frac{1}{(i\hbar)^k} \int_{-\infty}^t \int_{-\infty}^{t_1} \cdots \int_{-\infty}^{t_{k-1}} \langle \varphi_n | W_I(t_1) W_I(t_2) \cdots W_I(t_k) | \varphi_i \rangle dt_k \cdots dt_2 dt_1$$

Ricordando che vale $\lim_{t \rightarrow -\infty} |\psi(t)_I\rangle = |\varphi_i\rangle$ e che vale:

$$\lim_{t \rightarrow +\infty} b_n(t) = \lim_{t \rightarrow +\infty} \langle \varphi_n(t) | \psi(t) \rangle_I = \langle \varphi_n(t) | U_I(t) | \varphi_i(t) \rangle$$

dove $U_I(t)$ è l'operatore di evoluzione temporale nella rappresentazione di interazione, allo sviluppo perturbativo dei coefficienti $b_n(t)$ si può far corrispondere lo sviluppo perturbativo di $U_I(t)$:

$$U_I(t) = \mathbb{I} + \sum_{k=1}^{+\infty} \frac{1}{(i\hbar)^k} \int_{-\infty}^t \int_{-\infty}^{t_1} \cdots \int_{-\infty}^{t_{k-1}} W_I(t_1) W_I(t_2) \cdots W_I(t_k) dt_k \cdots dt_2 dt_1$$

che prende il nome di **serie di Dyson** e i cui termini sono ordinati secondo le potenze della perturbazione.

L'interpretazione fisica di questa serie è abbastanza interessante. Essa mostra che l'evoluzione temporale di un sistema sottoposto a perturbazione $W(t)$ può avvenire in diversi modi:

- il sistema evolve senza intervento della perturbazione $W(t)$, in questo caso l'evoluzione è guidata dall'Hamiltoniana imperturbata. Questo corrisponde al primo termine della serie (ordine zero),
- il sistema evolve liberamente fino all'istante t_1 , quando la perturbazione agisce e successivamente al termine della perturbazione evolve liberamente,
- il sistema evolve liberamente fino all'istante t_1 , quando la perturbazione agisce e termina, successivamente evolve liberamente fino all'istante t_2 quando la perturbazione agisce di nuovo e al termine di questa seconda perturbazione evolve liberamente,
- i termini successivi rappresentano la perturbazione che interagisce a diverse riprese, l'ordine k -simo rappresenta la perturbazione che interviene in k istanti successivi

La probabilità di arrivare allo stato finale $|\varphi_n\rangle$ partendo dallo stato iniziale $|\varphi_i\rangle$ è data in meccanica quantistica dalla somma su tutte le possibilità (si ricordi che *se un evento può accadere in diversi modi indistinguibili, allora occorre sommare le ampiezze relative ai vari modi*), il che dà origine proprio alla serie di Dyson.

8.3 Metodo variazionale

Si tratta di un metodo che permette di avere una stima del valore più basso dell'autovalore dell'Hamiltoniana di un sistema. Esso si basa sul **Teorema di Ritz** che afferma che *l'autovalore più basso di un operatore hermitiano è minore o uguale al valore medio dell'operatore calcolato su uno stato generico, e che l'uguaglianza vale se e solo se lo stato coincide con l'autostato corrispondente* ovvero, detto $\mathcal{H}_{\text{sist.}}$ lo spazio di Hilbert del sistema, che:

$$\forall \psi_n \in \mathcal{H}_{\text{sist.}} \quad \langle \psi_n | H | \psi_n \rangle \geq E_0$$

Questo teorema è ragionevolmente ovvio in quanto per definizione il valore medio non può mai essere inferiore al valore minimo, se ne può però naturalmente dare una dimostrazione più formale.

È chiaro che se si cerca il minimo del funzionale $\langle \psi | \hat{H} | \psi \rangle$ dell'Hamiltoniana, spazzando tutto lo spazio di Hilbert $\mathcal{H}_{\text{sist.}}$ si troverebbe proprio il valore E_0 . In pratica, far variare il funzionale sull'intero spazio di Hilbert è troppo complicato dal punto di vista dei calcoli, per cui si sceglie normalmente un sottoinsieme di questo spazio parametrizzato tramite alcuni parametri differenziabili reali α_i , $i = 1, \dots, N$. La scelta del sottospazio prende il nome di *ansatz* ed è molto importante perché determina quale minimo relativo viene individuato dal metodo variazionale. Il valore dell'energia E_0 è infatti il minimo assoluto del funzionale sullo spazio di Hilbert totale, mentre la scelta di un sottospazio permetterà evidentemente di trovare solo un minimo relativo, che può coincidere o meno con quello assoluto (coinciderà nel caso l'autofunzione sia compresa nel sottospazio scelto). Questa procedura permette di passare da un problema di ricerca del minimo di un funzionale:¹⁵

$$E_0(\alpha_i) = \langle \psi_{\alpha_i} | H | \psi_{\alpha_i} \rangle$$

ad un problema algebrico di ricerca del minimo di una funzione che richiede la risoluzione del sistema di equazioni:

$$\frac{\partial}{\partial \alpha_i} E_0(\alpha_i) = 0$$

Sebbene certamente più semplice, questo non è necessariamente un problema facile da risolvere, in quanto si ricerca (in linea di principio) un minimo globale e la derivazione dei $E_0(\alpha_i)$ rispetto a α_i può non essere sufficiente. Se tuttavia le funzioni ψ_{α_i} sono espresse come combinazioni lineari di altre funzioni aventi α_i come coefficienti, allora il minimo è unico e facilmente determinabile. Questo metodo è noto come *metodo di Ritz*¹⁶ L'ansatz è quindi fondamentale sia per semplificare abbastanza i calcoli sia per ottenere una buona approssimazione per eccesso dell'energia dello stato fondamentale. Non esistono regole precise per questa scelta, ma come principi guida si possono considerare le simmetrie del problema, dei parametri che rappresentino i fattori di scala del problema o la semplicità di calcolare il valore di $E_0(\alpha_i)$ a partire dai parametri.

Un metodo particolarmente semplice consiste nel considerare lo stato incognito come combinazione lineare delle autofunzioni normalizzate dell'Hamiltoniana imperturbata e troncare lo sviluppo ad un numero finito di termini. Si noti che mentre la base completa di infinite autofunzioni spazia su tutto lo spazio di Hilbert $\mathcal{H}_{\text{sist.}}$, troncare lo sviluppo significa limitarsi ad un sottospazio: questa posizione rappresenta quindi l'ansatz e i parametri variazionali sono i coefficienti dello sviluppo:

$$|\psi\rangle = \sum_{i=1}^N \alpha_i |\varphi_i\rangle$$

In questo caso, quindi, l'energia è data dalla forma:

$$E_0(\alpha) = \langle \psi | H | \psi \rangle = \sum_{i=1}^N \alpha_i^* \alpha_j \langle \varphi_i | H | \varphi_j \rangle \equiv \sum_{i=1}^N \alpha_i^* \alpha_j H_{ij}$$

dove la condizione di ortonormalità della base di autofunzioni implica $\sum_{i=1}^N |\alpha_i|^2 = 1$. Il funzionale da minimizzare utilizzando il *metodo dei moltiplicatori di Lagrange*¹⁷ è quindi:

$$E_0^{(\alpha)} = E_0(\alpha) - \lambda \left(\sum_{i=1}^N |\alpha_i|^2 - 1 \right)$$

¹⁵ Per la precisione, si tratta della ricerca di un estremo.

¹⁶ Esistono altri metodi non lineari, come il *metodo di Hartree-Fock*, che hanno un numero limitato di minimi e sono quindi ragionevolmente semplici da risolvere, ma questi sono l'oggetto di corsi superiori.

¹⁷ Sostanzialmente, il metodo dei moltiplicatori di Lagrange si basa sul fatto che data una funzione $f(\vec{x})$ ed il vincolo $g(\vec{x}) = c$, allora esiste un valore λ (detto *moltiplicatore di Lagrange*) per il quale il funzionale:

$$\mathcal{F}(\vec{x}_0, \lambda) = f(\vec{x}_0) + \lambda[g(\vec{x}_0) - c]$$

è stazionario nel punto di massimo \vec{x}_0 del problema vincolato originario. Questa rappresenta una *condizione necessaria*.

Calcolando la variazione rispetto a α_i^* si ricava:

$$\frac{\partial}{\partial \alpha_i^*} E_0 = 0 \quad \Rightarrow \quad H_{ij} \alpha_j = \lambda \alpha_i$$

ovvero l'equazione di Schrödinger proiettata sullo spazio lineare generato dagli stati $|\varphi_i\rangle$. La soluzione di questo sistema sarà una stima per eccesso dell'autovalore ricercato.

È a questo punto interessante fare un paio di osservazioni. È sempre possibile allargare la base di autofunzioni utilizzata nello sviluppo troncato della $|\psi\rangle$, quindi in linea di principio è possibile approssimare l'autovalore vero E_0 con il livello di precisione voluto. Per N finito si avrà sempre un'approssimazione che sarà tanto migliore quanto meglio si sceglie la base di partenza. Un criterio pratico può essere quello di verificare che l'allargamento della base non conduce ad un miglioramento del risultato. Inoltre, una buona scelta della base di partenza si riflette anche sulla rapidità di convergenza del metodo.

Questo metodo può essere migliorato utilizzando delle funzioni parametriche: in questo caso si risolve il problema per un valore del parametro generico e poi si minimizza rispetto a questo parametro. Un caso limite di questo approccio è quello utilizzato nel *metodo di Hartree-Fock* per trattare gli atomi a più elettroni: in questo caso, le funzioni di base stesse sono calcolate utilizzando un metodo variazionale assumendo diverse semplificazioni, dopo di che queste funzioni vengono utilizzate nel metodo variazionale di Ritz descritto qui. Lo studio di questo approccio è generalmente riservato a corsi superiori.

È da notare che mentre minimizzando il funzionale l'energia tende effettivamente al valore reale, non è possibile garantire che per le funzioni d'onda accada lo stesso.¹⁸

Si concluderà questo paragrafo notando che il metodo variazionale può in realtà essere esteso anche ai livelli eccitati, a patto di conoscere la funzione d'onda dello stato fondamentale: in questo caso l'ansatz deve essere ortogonale alla funzione dello stato fondamentale.

¹⁸Questo è stato dimostrato praticamente effettuando i calcoli su un oscillatore armonico, i cui stati sono noti analiticamente.

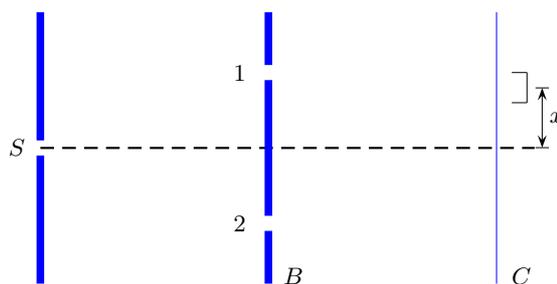
L'interpretazione di Feynman e i *Path Integral*

NOTA: Questa appendice riguarda un argomento che non fa parte nel corso principale, ma che riveste secondo l'autore un'interesse particolare soprattutto per l'approccio concettuale di Feynman alla meccanica quantistica. Il ragionamento sviluppato in questa appendice può risultare senz'altro utile per chiarire alcuni aspetti concettuali della meccanica quantistica che possono risultare non particolarmente evidenti nella teoria "tradizionale".

A.1 Impostazione concettuale

Il concetto di partenza dell'approccio seguito qui è quello di probabilità. La meccanica quantistica non ne modifica la definizione, che rimane quella classica: ovvero, si dice ancora che il risultato di un esperimento ha una probabilità \mathcal{P} se ripetendo molte volte l'esperienza si ottiene una frazione di risultati corrispondenti a quello atteso pari a \mathcal{P} .

Quello che cambia in meccanica quantistica è invece il modo con cui vengono calcolate queste probabilità e questo cambiamento ha conseguenze importanti quando si trattano sistemi di dimensioni atomiche. Si consideri il noto esperimento delle due fenditure. Una sorgente S emette una particella, ad esempio un elettrone, che incide su uno schermo B . Questo schermo ha due fenditure, attraverso le quali l'elettrone può passare e giungere su un piano C , dove è disposto un rivelatore che può scorrere lungo il piano in tutte le posizioni.



La sorgente può essere resa abbastanza debole e il rivelatore abbastanza sensibile da ricevere e contare una sola particella per volta. Il rivelatore registrerà sempre un solo elettrone alla volta, mai una mezza particella, l'elettrone o arriva intero o non arriva sul rivelatore. Se ci sono due rivelatori, solo uno rivelerà l'elettrone e mai i due contemporaneamente¹. Questo è quello che permette di

¹Ovviamente i dettagli reali dell'esperimento sono parecchio più complessi, ma il concetto è identico.

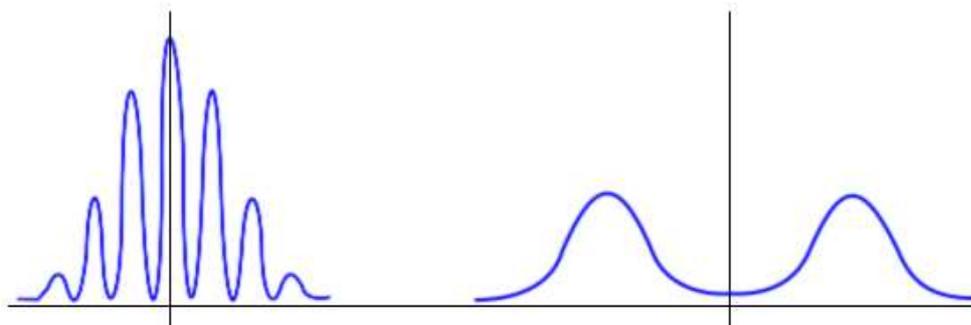
identificare l'elettrone come una *particella*. Il rivelatore registra quindi il passaggio di un elettrone emesso dalla sorgente attraverso uno dei fori dello schermo fino al rivelatore stesso.

Il passo successivo è di misurare il numero di elettroni che arrivano in diverse posizioni a distanza x rispetto all'asse della sorgente, in altre parole si cerca di determinare sperimentalmente la probabilità che un elettrone venga rivelato nella posizione x . Questa curva è abbastanza complessa e presenta diversi massimi e minimi.

Per trovare la legge fisica che spiega questo comportamento si può supporre, basandosi sulla natura particellare dell'elettrone, che:

1. ogni elettrone che raggiunge lo schermo S sia passato attraverso uno dei due fori nello schermo B e che quindi
2. la probabilità di arrivare in un punto x dello schermo sia la somma della probabilità \mathcal{P}_1 di arrivare in x attraverso il foro 1 e \mathcal{P}_2 di arrivare in x attraverso il foro 2.

Questa assunzione è facilmente verificabile sperimentalmente: basta chiudere il foro 2 e misurare la probabilità di arrivare in x tramite il foro 1 e rifare lo stesso con l'altro foro. La somma di queste due probabilità però non corrisponde a quella dell'esperimento originario, ovvero $\mathcal{P}(x) \neq \mathcal{P}_1(x) + \mathcal{P}_2(x)$, quindi evidentemente *la probabilità di arrivare in x con entrambi i fori aperti non è la somma delle probabilità di arrivare in x passando dal foro 1 più la probabilità di arrivarci attraverso il foro 2.*



I due fori sono aperti

I fori sono aperti alternativamente

La curva $\mathcal{P}(x)$ è in realtà familiare, si tratta infatti della figura di interferenza che si ha quando un'onda parte dalla sorgente S e attraversa i due fori per fare interferenza sullo schermo C . Quindi, in analogia al caso ondulatorio, si può dire che la legge $\mathcal{P}(x)$ è data dal modulo quadro di una certa quantità complessa $\varphi(x)$ che viene chiamata *ampiezza di probabilità* di arrivare in x . Questa $\varphi(x)$ è la somma di due contributi $\varphi_1(x)$ e $\varphi_2(x)$ che rappresentano rispettivamente l'ampiezza di probabilità di passare attraverso il foro 1 e 2. In altri termini, la legge di probabilità nel caso entrambi i fori siano aperti deve essere descritta dalla posizione:

$$\begin{cases} \varphi(x) = \varphi_1(x) + \varphi_2(x) \\ \mathcal{P}(x) = |\varphi(x)|^2 \end{cases} \quad \begin{cases} \mathcal{P}_1(x) = |\varphi_1(x)|^2 \\ \mathcal{P}_2(x) = |\varphi_2(x)|^2 \end{cases}$$

In sostanza, si *calcola l'intensità (ovvero il modulo quadro dell'ampiezza) dell'onda che arriverebbe sullo schermo e si interpreta questa come la probabilità di trovare la particella*. È notevole che questa doppia interpretazione dell'esperienza (particelle o onda) non conduca in realtà a contraddizione, a patto di considerare con cura quello che è possibile dire in un esperimento.

La proposizione "l'elettrone passa per il foro 1 o per il foro 2 per arrivare in x ", perfettamente lecita per una particella, implica che queste siano due classi disgiunte ovvero che l'una escluda l'altra. In questo caso la frequenza di arrivo in x è data sicuramente dalla somma delle probabilità individuali $\mathcal{P}_1(x)$ e $\mathcal{P}_2(x)$, ma questo è escluso dall'esperienza.

Si può certamente concepire un esperimento leggermente modificato in modo da individuare per quale foro è passato l'elettrone.² Se si fa questo, si trova inevitabilmente che l'elettrone è passato attraverso un foro o l'altro. Se ora si combinano i risultati di questo esperimento, si trova effettivamente che gli elettroni passati attraverso il foro 1 si distribuiscono come ci si attende in questo caso, e lo stesso per gli elettroni passati attraverso il foro 2. Ma in questo esperimento, però, i due gruppi di elettroni rappresentano sicuramente due classi disgiunte e pertanto ci si aspetta che la loro probabilità totale sia data da $\mathcal{P}(x) = \mathcal{P}_1(x) + \mathcal{P}_2(x)$. Questo è effettivamente il caso che si osserva.

Questo può apparire sorprendente. Se non si osservano gli elettroni durante il tragitto fra la sorgente e il rivelatore la probabilità di trovarlo in x è data da $\mathcal{P}(x) = |\varphi_1(x) + \varphi_2(x)|^2$, se invece li si osserva la probabilità è data da $\mathcal{P}(x) = \mathcal{P}_1(x) + \mathcal{P}_2(x)$. La risposta a questa situazione anomala è che il processo di misura, in sostanza fatto da fotoni che interagiscono con gli elettroni, altera il moto di questi ultimi, in particolare altera le probabilità di arrivare in x . Si può mostrare che qualunque processo di misura interagisce con il sistema in due modi: o in maniera tale da perturbarlo pesantemente o in maniera tale da non permettere di estrarre sufficiente informazione. Questo è la base su cui Heisenberg ha impostato il suo principio. Egli notò per primo che la meccanica quantistica imponeva dei limiti alla precisione con cui può essere condotto un esperimento perché l'interazione fra l'operazione di misura e il sistema cambia necessariamente lo stato stesso del sistema. Questa limitazione deve riguardare ogni sistema fisico e di misura, e finora non è stata scoperta alcuna eccezione al principio di indeterminazione. Nell'ambito dell'approccio seguito qui, il principio può essere espresso dicendo che *ogni tentativo di determinare la reale scelta presa dal sistema in una situazione in cui sono possibili diverse alternative distrugge l'interferenza fra le alternative*.

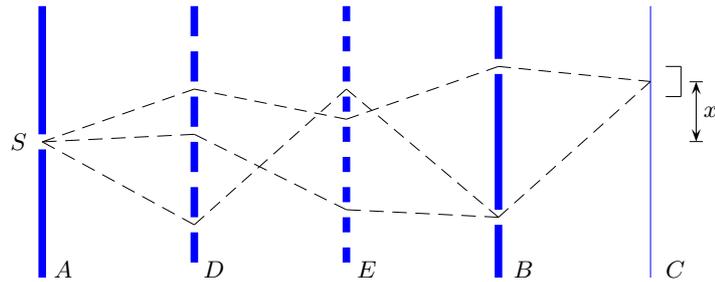
Per risolvere l'impasse logica creata da questa situazione, ci sono due modi. Nel primo caso, l'approccio tradizionale, si dice che la premessa (1.) non è valida cioè che *non è vero che l'elettrone passa per un foro o per l'altro*. Nel secondo caso, l'approccio seguito da Feynmann, invece si invalida la premessa (2.) ovvero si dice che *la probabilità di un risultato non si calcola sommando sulle probabilità delle alternative possibili*. Prima di continuare, occorre precisare il significato corretto da attribuire al termine **alternativa**. Il primo significato in cui si può intendere questo termine è nel senso di *alternative mutuamente escludenti*: è il caso in cui uno dei due fori è chiuso o esiste un apparato che permetta di sapere esattamente in quale foro è passato l'elettrone. Il secondo significato è quello di *alternative interferenti*, analogo a quello in ottica: è il caso in cui i fori sono entrambi aperti e non viene determinato attraverso quel dei due è passato l'elettrone. I due casi richiedono di combinare le probabilità in maniera diversa, come visto.

Questo modo di vedere le cose rende più chiara anche un'altra questione. Si supponga di predisporre un esperimento in cui si determina attraverso quale foro sia passato l'elettrone, ma di non utilizzare i dati raccolti per costruire la distribuzione di probabilità. In questo caso, la regola di combinazione deve comunque seguire quella delle alternative mutuamente escludenti: in altre parole, è sufficiente la possibilità di avere informazioni sul sistema per distruggere la figura di interferenza.

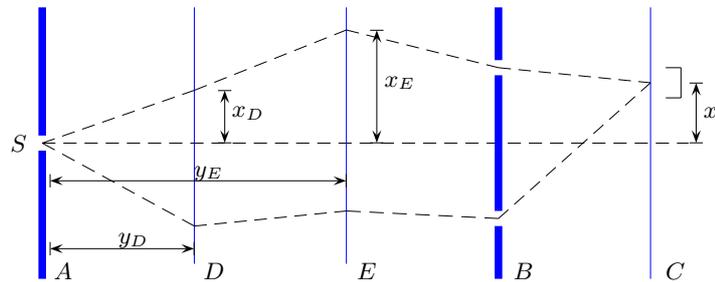
Si è visto quindi che ad ogni evento in natura può essere associata una ampiezza di probabilità. Quando si rivela un elettrone emesso da S sul rivelatore, esiste una ampiezza associata all'elettrone che attraversa il foro 1, $\varphi_1(x)$ ed una associata all'elettrone che attraversa il foro 2, $\varphi_2(x)$. All'intero evento si deve quindi associare una ampiezza di probabilità totale, ed il quadrato di questa è interpretata come la probabilità che l'evento abbia luogo. Se si interrompe l'evento in corso, ad esempio misurando attraverso quale foro passa l'elettrone, si disturba la costruzione dell'ampiezza totale in quanto osservare la particella in un determinato stato esclude che possa trovarsi in un altro. Di conseguenza, le ampiezze non si possono più sommare.

L'ampiezza di probabilità di un evento è data dalla somma delle ampiezze di tutte le alternative che possono dare origine all'evento. Si consideri ad esempio una particella che va da A a B . Questa particella potrà seguire diversi percorsi nello spazio e nel tempo, e ad ogni percorso si potrà associare una ampiezza di probabilità da sommare. Si consideri ancora l'esperimento con i due fori, ma si aggiungano un paio di schermi supplementari, denotati con D e E , ognuno dei quali ha una serie di fori che sarà indicata con D_1, D_2, \dots e E_1, E_2, \dots . Si supponga per comodità il moto dell'elettrone nel piano xy .

²Non ci si dilungherà sui dettagli di questi esperimenti, di cui è importante qui solo ritenere i risultati.



Come si vede, ci sono ora diverse alternative per le quali un elettrone può raggiungere il rivelatore in x . Ognuno di questi percorsi ha una sua ampiezza che deve essere sommata a quella di tutti gli altri. Si supponga ora di creare dei fori sempre più fitti e sempre più piccoli, fino a far scomparire completamente gli schermi.



Il percorso dell'elettrone sarà allora parametrizzato dall'altezza x_D a cui passa in y_D , dove si trova lo schermo-non-esistente D , e analogamente per x_E e y_E . Ad ogni coppia di altezze corrisponde un'ampiezza e siccome il principio di sovrapposizione vale sempre, si devono prendere gli integrali di queste ampiezze su tutti i valori possibili di x_D e x_E .

A questo punto è semplice immaginare di collocare sempre più "non-schermi" fra la sorgente S e il rivelatore. In questo modo, il percorso dell'elettrone viene definito con sempre più precisione fino ad arrivare alla conclusione che questo percorso è una semplice funzione dell'altezza $x(y)$. Applicando il principio di sovrapposizione, si arriva all'integrale esteso alle ampiezze di tutti i cammini possibili.

È possibile un ulteriore raffinamento di questa idea. Non solo è possibile pensare a tutti i possibili cammini nello *spazio*, ma anche nel *tempo* specificando in quale istante l'elettrone passa in tutti i punti del percorso. Il cammino bidimensionale è quindi completamente determinato una volta fornite due funzioni $x(t)$ e $y(t)$. L'ampiezza totale di arrivare in un determinato punto è la somma (o l'integrale) dell'ampiezza associata ad ogni possibile percorso.

A.2 Formulazione matematica

Se una particella parte all'istante t_a dal punto x_a per arrivare nell'istante t_b nel punto x_b , si dice classicamente che la particella è andata da a a b secondo la funzione $x(t)$ che ha la proprietà $x(t_a) = x_a$ e $x(t_b) = x_b$. In meccanica quantistica si ha piuttosto una ampiezza di probabilità di andare da x_a a x_b , denotata con $K(b, a)$ e chiamata *kernel*. Questa ampiezza è la somma di tutte le ampiezze associate a tutti i percorsi possibili fra x_a e x_b : *ogni possibile traiettoria contribuisce alla probabilità finale*. Questa è la grande differenza rispetto alla meccanica classica, dove esiste una sola specifica traiettoria che la particella segue per andare da x_a a x_b , definita *traiettoria classica* e denotata nel seguito con $\bar{x}(t)$.

Nella meccanica classica, la condizione che permette di determinare il percorso effettivo di una particella fra i molti possibili è detto *principio di minima azione*. Questo principio afferma che esiste una certa quantità S calcolabile per ogni percorso che una volta minimizzata³ fornisce proprio il

³In realtà si tratta di trovare un estremo e non necessariamente un minimo.

moto effettivo $\bar{x}(t)$. In termini più formali, S resta invariato al primo ordine se il percorso $\bar{x}(t)$ è leggermente modificato. In meccanica classica, la quantità S , l'azione appunto, è espressa in termini della lagrangiana del sistema:

$$S = \int_{t_a}^{t_b} L(\dot{x}, x, t) dt$$

Non ci si dilungherà qui nella dimostrazione, ma è noto che applicando una variazione a questa quantità e richiedendo che si annulli al primo ordine si ritrovano le equazioni del moto di Eulero-Lagrange. È da notare però che mentre *in meccanica classica è importante la forma dell'integrale d'azione e non tanto il suo valore specifico nell'estremo, in meccanica quantistica entrambi sono importanti*. Questa differenza si spiega con il fatto che in meccanica classica occorre sapere la forma del funzionale lungo un insieme di percorsi che si discostano tra loro per determinare quello che minimizza l'azione, mentre in meccanica quantistica si è visto che tutti i percorsi devono contribuire allo stesso modo alla costruzione dell'ampiezza di probabilità.

Il punto essenziale è che i diversi "percorsi quantistici" contribuiscono in ugual modo ma con fasi diverse e la fase di questi contributi è costituita dall'azione in unità di \hbar . Riassumendo, la probabilità di andare da x_a a x_b è data dal modulo quadro dell'ampiezza $\mathcal{P}(b, a) = |K(b, a)|^2$. Questa ampiezza è a sua volta data dalla somma di tutte le ampiezze di tutti i percorsi $\varphi[x(t)]$:

$$K(b, a) = \sum_{\substack{\text{tutti i percorsi} \\ \text{possibili} \\ \text{da } x_a \text{ a } x_b}} \varphi[x(t)]$$

e il contributo ai percorsi ha una fase proporzionale all'azione S :

$$\varphi[x(t)] \propto e^{(i/\hbar)S[x(t)]}$$

dove l'azione ha ancora la forma dei sistemi classici, ovvero in termini di lagrangiana. La costante di proporzionalità deve essere scelta in modo da normalizzare opportunamente l'ampiezza $K(b, a)$.

Potrebbe non risultare chiaro come dal fatto che quantisticamente tutti i percorsi contribuiscano sullo stesso piano ma con fasi diverse si arrivi al fatto che classicamente un solo percorso diventa quello importante. Il caso classico corrisponde in ogni caso a situazioni in cui le dimensioni e le masse sono così grandi che l'azione è enorme in termini di unità \hbar . Questo significa che la fase del contributo S/\hbar è un angolo molto grande. La parte reale (o immaginaria) di $\varphi[x(t)]$ è il coseno (o il seno) di questo angolo e può avere un segno positivo o negativo. Piccoli cambiamenti nel percorso della particella producono in queste condizioni enormi variazioni della fase e di conseguenza oscillazioni molto veloci del coseno (o del seno) intorno a valori negativi o positivi. Il contributo totale tende quindi ad annullarsi perché se un percorso dà un contributo positivo molto grande, esisterà un percorso infinitamente vicino (su scala classica) che darà un contributo molto grande ma negativo, in modo che il contributo netto sia nullo.

Nelle vicinanze di $\bar{x}(t)$ un piccolo cambiamento nel percorso non produce variazioni di S , almeno al primo ordine. Tutti i contributi nelle vicinanze (in senso quantistico) sono quasi in fase e quindi non si cancellano. Di conseguenza, solo i percorsi nelle immediate vicinanze di $\bar{x}(t)$ sono importanti e nel limite classico diventa sufficiente considerare solo il percorso $\bar{x}(t)$. In questo modo, le leggi della meccanica classica possono venir fuori dalla meccanica quantistica. Si noti che questo significa che le traiettorie contribuiscono all'azione finché restano nei dintorni di \hbar . In altri termini questo implica che la traiettoria classica non può essere definita con una precisione arbitraria inferiore a questo termine, restando quindi indeterminata di una quantità dell'ordine di \hbar .

Queste considerazioni sono valide finché si parla dell'azione nel limite classico, dove $S \gg \hbar$. Quando l'azione S è comparabile con \hbar , come avviene a livello atomico, tutte le traiettorie devono essere prese in considerazione e nessuna di loro assume un'importanza preponderante sulle altre.

Il numero di percorsi da prendere in conto per calcolare la somma è di un elevato ordine di infinito e non è evidente quale misura può essere attribuita allo spazio dei percorsi. Nell'integrazione di Riemann si dice sostanzialmente che l'area sottesa da una curva è data dalla somma delle ordinate, ovvero si prende un sottoinsieme delle ordinate spaziate di ϵ e si pone $A \simeq \sum_i f(x_i)$. Facendo tendere ϵ a zero, si dice che il limite di questa somma rappresenta l'area. Ora, questo limite non esiste per

come è scritto sopra in quanto occorre un fattore che normalizzi questa somma rispetto all'intervallo ϵ , che nel caso dell'integrazione di Riemann corrisponde proprio a ϵ stesso. Il limite corretto per l'area sottesa dalla curva è dato quindi da $A = \lim_{\epsilon \rightarrow 0} [\epsilon \sum_i f(x_i)]$. Si tratta ora di applicare questo ragionamento alla somma sui percorsi (la cosiddetta *somma sulle storie*).

Si consideri un sottoinsieme dei percorsi definiti nel modo seguente. Si divida l'intervallo della variabile indipendente tempo t in intervalli di larghezza ϵ . Questo fornisce un insieme di valori $\{t_i\}$ spazati di ϵ da t_a a t_b . Per ogni istante t_i si sceglie un determinato punto x_i e il percorso si ottiene unendo tutti questi punti con una linea retta. È possibile definire una somma su tutti i percorsi prendendo un integrale multiplo su tutti i valori di x_i per $1 \leq i \leq N - 1$, ovvero:

$$\begin{cases} N\epsilon &= t_b - t_a \\ \epsilon &= t_{i+1} - t_i \end{cases} \quad \begin{cases} t_0 &= t_a \\ x_0 &= x_a \end{cases} \quad \begin{cases} t_N &= t_b \\ x_N &= x_b \end{cases}$$

L'equazione risultante è:

$$K(b, a) = \int \cdots \int \varphi[x(t)] dx_1 dx_2, \cdots dx_{N-1}$$

Dove l'integrazione non è estesa a x_0 e x_N perché gli estremi x_a e x_b sono considerati fissi. In analogia al caso dell'integrazione secondo Riemann, l'ampiezza è via via più rappresentativa dell'intero insieme dei percorsi possibili man mano che si riduce l'intervallo ϵ . Tuttavia, come nel caso di Riemann, questo limite non esiste a meno di introdurre un fattore di normalizzazione, che deve dipendere da ϵ . La ricerca di questo fattore è generalmente parecchio complessa, però fortunatamente per tutti i casi di rilevanza pratica si riesce a trovarlo. Una volta introdotto questo fattore A il limite esiste e quindi si può scrivere:

$$\begin{cases} K(b, a) &= \lim_{\epsilon \rightarrow 0} \frac{1}{A} \int \cdots \int e^{(i/\hbar)S[b,a]} \frac{dx_1}{A} \frac{dx_2}{A} \cdots \frac{dx_{N-1}}{A} \\ S[b, a] &= \int_{t_a}^{t_b} L(\dot{x}, x, t) dt \end{cases} \quad (\text{A.1})$$

e l'azione S rappresenta un integrale di linea preso sulla traiettoria rettilinea passante attraverso i punti x_i .

Ci sono naturalmente svariati modi per identificare un sottoinsieme di percorsi possibili, e quello definito qui sopra non è quello matematicamente più semplice da affrontare. Ad esempio, la lagrangiana può dipendere da diverse variabili fra cui l'accelerazione: i percorsi costruiti con la procedura illustrata sopra presentano però delle discontinuità nei punti (x_i, t_i) , dove l'accelerazione diventa quindi infinita. In questo caso specifico, si può effettuare la sostituzione $\ddot{x} = \frac{1}{\epsilon^2}(x_{i+1} - 2x_i + x_{i-1})$,⁴ ma possono esserci certamente altri casi per cui l'applicazione del metodo esposto sopra si rivela complesso e artificioso. È una situazione che a ben vedere si presenta anche per l'integrazione secondo Riemann su alcuni domini particolari, che conduce a nuove definizioni di integrazione come quella di Lebegue.

La necessità di ridefinire il concetto di integrazione o di modificare la procedure di scelta del sottoinsieme di percorsi possibili non invalida però il concetto di integrazione. Il concetto di integrazione sui percorsi è infatti indipendente dalla particolare implementazione, e si può quindi indicare questa integrazione su tutti i percorsi possibili con una notazione più generale:

$$K(b, a) = \int_a^b e^{(i/\hbar)S[b,a]} \mathcal{D}x(t)$$

che viene chiamato *path integral*.

Resta ancora da definire la regola per combinare diversi eventi in successione. Si supponga che t_c sia un istante di tempo compreso fra t_a e t_b , allora l'azione associata ad un qualunque percorso fra x_a e x_b è data da:

$$S[b, a] = S[b, c] + S[c, a]$$

⁴Si noti che questa sostituzione coincide con il dire che si sostituisce la derivata seconda con le differenze finite centrali.

Questa posizione è giustificata dal fatto che l'azione è definita come un integrale e che la lagrangiana non dipende da derivate superiori alla velocità.⁵ Questa regola permette di scrivere:

$$K(b, a) = \int_a^b e^{(i/\hbar)(S[b,c]+S[c,a])} \mathcal{D}x(t)$$

Ogni percorso può essere diviso in due parti, la prima fra x_a e $x(t_c)$ e la seconda fra $x(t_c)$ e x_b . Una volta fatto questo, è possibile integrare su tutti i percorsi fra x_a e x_c , quindi su tutti i percorsi fra x_c e x_b e infine su tutti i possibili punti x_c . Nel primo passo, la parte di azione $S[b, c]$ è costante, per cui l'ampiezza si scrive:

$$K(b, a) = \int_{x_c} \int_c^b e^{(i/\hbar)S[b,c]} K(c, a) \mathcal{D}x(t) dx_c$$

dove si è effettuata l'integrazione su tutti i possibili percorsi fra c e b e poi su tutti i possibili punti x_c . Il passo successivo è quindi integrare dal punto arbitrario x_c e il punto finale b :

$$K(b, a) = \int_{x_c} K(b, c)K(c, a) dx_c$$

Si ha quindi la regola generale: *le ampiezze di eventi in istanti successivi di tempo si moltiplicano*. Naturalmente, è possibile dividere ulteriormente il percorso da x_c a x_a in due, introducendo un altro punto in t_d e di conseguenza un percorso da x_c a d e da d a x_a . Il kernel totale ha allora la forma:

$$K(b, a) = \int_{x_c} \int_{x_d} K(b, c)K(c, d)K(d, a) dx_c dx_d \quad (\text{A.2})$$

Fisicamente, si sta dicendo che la particella per andare da a a b passa per il punto c e d e per prendere in considerazione tutti i percorsi possibili bisogna integrare su tutti i valori possibili di x_c e x_d . Si noti che nella teoria relativistica non è possibile definire l'ampiezza di un percorso come $e^{iS/\hbar}$, ma la regola di composizione delle ampiezze è sempre valida ed è sempre possibile associare un'ampiezza ad ogni percorso possibile.

Questo modo di calcolare le probabilità implica seguire il moto della particella in dettaglio. Spesso però non interessa sapere come si è arrivati in un punto (x, t) in dettaglio, ma solo all'ampiezza di probabilità di trovarsi in (x, t) senza alcun particolare sulla storia passata. Si può quindi definire $\psi(x, t)$ l'ampiezza totale di arrivare nel punto (x, t) da una qualunque situazione nel passato, e questa ampiezza ha esattamente le stesse caratteristiche di quelle viste finora ovvero che la probabilità di trovare la particella al punto x nell'istante t è data da $|\psi(x, t)|^2$. Questa ampiezza si definisce *funzione d'onda*, ma si tratta in definitiva solo di una questione di denominazione. Si può quindi dire che un sistema si *trova nello stato* $\psi(x, t)$, per dire che il sistema è descritto dall'ampiezza di probabilità $\psi(x, t)$.

Quindi, il kernel $K(x_2, t_2; x_1, t_1) \rightarrow \psi(x_2, t_2)$ è effettivamente una funzione d'onda. Questo fornisce però più informazioni, infatti permette di determinare anche che la particella è arrivata in (x_2, t_2) venendo da (x_1, t_1) : si tratta di un'informazione che non è generalmente di interesse nello studio del sistema e di cui non è utile tenere traccia. Ecco perché si considera solo la $\psi(x_2, t_2)$.

Siccome si tratta sempre di ampiezze, la regola di composizione (A.2) vale anche per le funzioni d'onda e quindi in particolare vale per tutti i punti (x_1, t_1) e si può scrivere:⁶

$$\psi(x_2, t_2) = \int_{-\infty}^{+\infty} K(x_2, t_2; x_1, t_1)\psi(x_1, t_1)dx_1$$

che fisicamente può essere letta nel seguente modo: *l'ampiezza totale di arrivare in (x_2, t_2) è data dalla somma (o dall'integrale) di tutte le possibile ampiezze di arrivare in (x_1, t_1) moltiplicata l'ampiezza di andare da (x_1, t_1) a (x_2, t_2)* . In altri termini, gli effetti della storia passata della particella

⁵In caso contrario, si sarebbe dovuto specificare il valore della velocità e eventualmente delle derivate superiori al punto x_c .

⁶Si noti incidentalmente che questa equazione giustifica il nome di *kernel* per la funzione $K(b, a)$: si tratta infatti di un'equazione integrale di Fredholm omogenea in $\psi(x, t)$ di cui $K(x, t; x', t')$ costituisce il kernel.

possono essere espressi da una sola funzione. *Se si trascura tutto quello che si conosce di una particella eccetto la sua funzione d'onda ad un istante t , allora è possibile calcolare qualunque cosa a partire da questo istante particolare. Tutti gli effetti della storia passata sul futuro di una particella possono essere condensati in una sola funzione d'onda.*

A.3 Le relazioni di de Broglie e di Planck

In questo paragrafo si mostrerà come è possibile ricavare le relazioni di de Broglie e di Planck nell'approccio esposto qui.

Si consideri una particella libera con lagrangiana data da $L = \frac{1}{2}m\dot{x}^2$. Dalle relazioni (A.1) e dalla forma del fattore di normalizzazione:

$$A = \left(\frac{2\pi i \hbar \epsilon}{m} \right)^{1/2}$$

si ricava che l'ampiezza ha la forma:

$$K(b, a) = \left(\frac{2\pi i \hbar \epsilon}{m} \right)^{-N/2} \lim_{\epsilon \rightarrow 0} \int \dots \int e^{\frac{i m}{2\hbar \epsilon} \sum_{i=1}^N (x_i - x_{i-1})^2} dx_1 \dots dx_{N-1}$$

che rappresentano una serie di integrali gaussiani. Poiché l'integrale di una gaussiana è ancora una gaussiana, le integrazioni possono essere effettuate l'una dopo l'altra e quindi prendere il limite. Il risultato che si ottiene è:

$$K(b, a) = \left[\frac{2\pi i \hbar (t_b - t_a)}{m} \right]^{-1/2} e^{\frac{i m (x_b - x_a)^2}{2\hbar (t_b - t_a)}}$$

Si consideri per semplicità una particella che parta all'istante $t = 0$ dal punto $x = 0$. Il kernel allora diventa:

$$K(x, t, 0, 0) = \left(\frac{2\pi i t}{m} \right)^{-1/2} e^{\frac{i m x^2}{2\hbar t}}$$

Si consideri il comportamento di questa ampiezza a t fissato. La parte reale presenta delle oscillazioni sempre più rapide man mano che x cresce. Se x è sufficientemente grande, e quindi sono già avvenute molte oscillazioni, allora la distanza fra due nodi successivi è praticamente costante (almeno per le oscillazioni immediatamente successive) e quindi l'ampiezza si comporta come una funzione sinusoidale di lunghezza d'onda lentamente variabile. Cambiando x proprio di una lunghezza d'onda λ la fase dell'ampiezza deve variare di 2π , ovvero:

$$2\pi = \frac{m(x + \lambda)^2}{2\hbar t} - \frac{m x^2}{2\hbar t} = \frac{m x \lambda}{\hbar t} + \frac{m \lambda^2}{2\hbar t}$$

e trascurando λ^2 rispetto a $x\lambda$, in quanto $x \gg \lambda$, si trova:

$$\lambda = \frac{2\pi \hbar}{m(x/t)}$$

classicamente, una particella in moto ha una velocità di x/t e una quantità di moto pari a $m x/t$. Dal punto di vista della meccanica quantistica, quando il moto può essere descritto adeguatamente assegnando il momento classico alla particella, allora l'ampiezza di probabilità varia nello spazio con lunghezza d'onda:

$$\lambda = \frac{h}{p}$$

Si consideri ora la variazione nel tempo del kernel. In questo caso sia l'ampiezza che la frequenza decrescono nel tempo. Si supponga che t sia abbastanza grande da poter trascurare le variazioni in ampiezza del kernel. In questo caso, in analogia al caso precedente, il periodo è costante da

un'oscillazione all'altra e denotato questo con T , si ha che quando è trascorso un tempo T la fase deve essere variata di 2π , quindi:

$$2\pi = \frac{mx^2}{2\hbar t} - \frac{mx^2}{2\hbar(t+T)} = \frac{mx^2}{2\hbar t^2} \frac{T}{1+T/t}$$

assumendo che $t \gg T$ e denotando con $\omega = 2\pi/T$ la frequenza angolare questa equazione diventa:

$$\omega \simeq \frac{m}{2\hbar} \left(\frac{x}{t}\right)^2$$

ma $\frac{1}{2}m(x/t)^2$ è l'energia cinetica classica di una particella, per cui nelle condizioni in cui una particella può essere descritta adeguatamente dalla meccanica classica si ottiene:

$$E = \hbar\omega$$

A.4 L'equazione di Schrödinger

Si è visto che i *path integral* hanno degli integrandi che comprendono l'esponenziale dell'azione, il che conduce a delle equazioni integrali. Anche se in linea teorica è possibile risolvere queste equazioni, normalmente si cerca di ridurre lo studio alla risoluzione di equazioni differenziali. In effetti, tutta la meccanica quantistica si basa su una equazione differenziale, l'equazione di Schrödinger. In questo paragrafo si mostrerà quindi come è possibile ricavare l'equazione di Schrödinger a partire dall'approccio seguito qui.

L'idea di base è che la relazione $S[b, a] = S[b, c] + S[c, a]$ vale per punti qualsiasi, quindi anche per punti infinitamente vicini. Quindi t_c può essere solo un infinitesimo più grande di t_a : questo permette di legare il valore del path integral a quello di un punto infinitamente vicino e di passare così ad una formulazione differenziale.

Si è già visto che si può definire una funzione d'onda tale che:

$$\psi(x_2, t_2) = \int_{-\infty}^{+\infty} K(x_2, t_2; x_1, t_1) \psi(x_1, t_1) dx_1$$

che fornisce la funzione d'onda all'istante t_2 a partire da quella all'istante t_1 . Per ottenere l'equazione differenziale corrispondente, occorre prendere il tempo t_2 infinitamente vicino a t_1 : $t_2 = t_1 + \epsilon$. Il kernel $K(2, 1)$ dell'equazione è proporzionale all'esponenziale di $i/\hbar S[2, 1]$. Nel caso l'intervallo temporale sia infinitesimo si può fare l'approssimazione:

$$S = \int_{t_a}^{t_b} L(\dot{x}, x, t) dt \simeq \epsilon L(\dot{x}, x)$$

e usare per la lagrangiana l'approssimazione, valida fino al primo ordine in ϵ :

$$L(\dot{x}, x) \simeq L\left(\frac{x-y}{\epsilon}, \frac{x+y}{2}\right)$$

avendo denotato con x e y due punti vicini. Si noti che questa approssimazione corrisponde in pratica a discretizzare perché si prende il rapporto incrementale invece della derivata rispetto al tempo e la media dei due punti infinitamente vicini al posto della dipendenza da x . Sostituendo si ricava:

$$\psi(x, t + \epsilon) = \int_{-\infty}^{+\infty} \frac{1}{A} e^{i\epsilon/\hbar L\left(\frac{x-y}{\epsilon}, \frac{x+y}{2}\right)} \psi(y, t) dy$$

Si consideri ora il caso di una particella che si muove in un potenziale $V(x, t)$. Risulta allora:

$$L = \frac{1}{2}m\dot{x}^2 - V(x, t) = m\frac{(x-y)^2}{\epsilon} - \epsilon V\left(\frac{x+y}{2}, \epsilon t\right)$$

da cui:

$$\psi(x, t + \epsilon) = \int_{-\infty}^{+\infty} \frac{1}{A} e^{i\epsilon/\hbar m \frac{(x-y)^2}{\epsilon}} \cdot e^{-i\epsilon/\hbar V(\frac{x+y}{2}, \epsilon t)} \psi(y, t) dy$$

Il primo termine esponenziale di questo integrale contiene la quantità $\frac{(x-y)^2}{\epsilon}$. Quando x è lontano da y questo integrale è molto grande e quindi oscilla rapidamente. Siccome gli altri fattori non hanno grandi variazioni, il contributo totale dell'integrale è piccolo. Quindi solo quando x è vicino a y si hanno dei contributi importanti, si può quindi porre $y = x + \eta$ con η molto piccolo. Con questa sostituzione si ottiene:

$$\psi(x, t + \epsilon) = \int_{-\infty}^{+\infty} \frac{1}{A} e^{\frac{i m \eta^2}{2 \hbar \epsilon}} \cdot e^{-\frac{i \epsilon}{\hbar} V(\frac{x+\eta}{2}, t)} \psi(x + \eta, t) d\eta$$

La fase del primo esponenziale varia di un radiante quando $\eta \simeq \sqrt{\epsilon \hbar / m}$, quindi la maggior parte dell'integrale viene da contributi che sono di questo ordine di grandezza.

In queste ipotesi è possibile espandere la $\psi(x, t)$ in serie e fermarsi al primo ordine in ϵ , che implica il secondo ordine in η . Inoltre, il termine $V(\frac{x+\eta}{2}, t)$ può essere approssimato con $\epsilon V(x, t)$, in quanto in questo modo si commette un errore di ordine superiore in ϵ .

Espandendo quindi il termine a sinistra al primo ordine in ϵ e quello a destra al primo ordine in ϵ e al secondo in η si ottiene:

$$\psi(x, t) + \epsilon \frac{\partial}{\partial t} \psi(x, t) = \int_{-\infty}^{+\infty} \frac{1}{A} e^{\frac{i m \eta^2}{2 \hbar \epsilon}} \left[1 - \frac{i \epsilon}{\hbar} V(x, t) \right] \left[\psi(x, t) + \eta \frac{\partial}{\partial x} \psi(x, t) + \frac{1}{2} \eta^2 \frac{\partial^2}{\partial x^2} \psi(x, t) \right] d\eta$$

Il termine principale di questa equazione fornisce (si noti che $\psi(x, t)$ non dipende dalla variabile di integrazione η e quindi si può portare fuori integrale):

$$\psi(x, t) = \psi(x, t) \cdot \frac{1}{A} \int_{-\infty}^{+\infty} e^{\frac{i m \eta^2}{2 \hbar \epsilon}} d\eta$$

che permette di dedurre che:

$$\frac{1}{A} \int_{-\infty}^{+\infty} e^{\frac{i m \eta^2}{2 \hbar \epsilon}} d\eta = 1 \quad \Rightarrow \quad A = \left(\frac{2\pi i \hbar \epsilon}{m} \right)^{1/2}$$

e questo è in realtà il modo per calcolare la costante A , ovvero imporre che l'equazione all'ordine zero in ϵ sia corretta. In caso contrario, non esiste il limite per ϵ tendente a zero del path integral originario.

Per il calcolo del primo ordine in ϵ occorre guardare l'equazione:

$$\epsilon \frac{\partial}{\partial t} \psi(x, t) = \frac{1}{A} \int_{-\infty}^{+\infty} \left[\frac{1}{2} \eta^2 e^{\frac{i m \eta^2}{2 \hbar \epsilon}} \frac{\partial^2}{\partial x^2} - \frac{i \epsilon}{\hbar} e^{\frac{i m \eta^2}{2 \hbar \epsilon}} V(x, t) - \frac{i \epsilon}{\hbar} \eta e^{\frac{i m \eta^2}{2 \hbar \epsilon}} V(x, t) \right] \psi(x, t) d\eta$$

ovvero:

$$\begin{aligned} \epsilon \frac{\partial}{\partial t} \psi(x, t) = & \frac{1}{A} \int_{-\infty}^{+\infty} \eta^2 e^{\frac{i m \eta^2}{2 \hbar \epsilon}} \frac{\partial^2}{\partial x^2} \psi(x, t) d\eta - \frac{1}{A} \frac{i \epsilon}{\hbar} \int_{-\infty}^{+\infty} \eta e^{\frac{i m \eta^2}{2 \hbar \epsilon}} \frac{\partial}{\partial x} \psi(x, t) d\eta - \\ & - \frac{1}{A} \frac{i \epsilon}{\hbar} \int_{-\infty}^{+\infty} e^{\frac{i m \eta^2}{2 \hbar \epsilon}} V(x, t) \psi(x, t) d\eta \end{aligned}$$

Tenendo conto che i fattori $\psi(x, t)$ e $V(x, t)$ possono essere portati fuori dell'integrale, che è in η , ci si riduce a calcolare gli integrali:

$$\int_{-\infty}^{+\infty} \eta e^{\frac{i m \eta^2}{2 \hbar \epsilon}} d\eta \quad ; \quad \int_{-\infty}^{+\infty} \eta^2 e^{\frac{i m \eta^2}{2 \hbar \epsilon}} d\eta$$

Il primo è nullo,⁷ mentre il secondo integrando ripetutamente per parti risulta:

$$\int_{-\infty}^{+\infty} \eta^2 e^{\frac{im\eta^2}{2\hbar\epsilon}} d\eta = \frac{i\hbar\epsilon}{m}$$

che sostituito nell'equazione fornisce (semplificando anche il fattore ϵ che compare ora in tutti i termini):

$$\frac{\partial}{\partial t}\psi(x, t) = \frac{i\hbar\epsilon}{m} \frac{\partial^2}{\partial x^2}\psi(x, t) - \frac{i}{\hbar}V(x, t)\psi(x, t)$$

ovvero:

$$i\hbar \frac{\partial}{\partial t}\psi(x, t) = \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x, t) \right] \psi(x, t)$$

che è proprio l'equazione di Schrödinger per una particella libera in un potenziale $V(x, t)$.

È giunto il momento di tirare qualche conclusione. L'approccio seguito qui è consistito nel cambiare il modo con cui calcolare le probabilità di un evento (ma non la definizione stessa di probabilità), ipotizzando che su scala atomica non esista un percorso "privilegiato" nel senso della meccanica classica (formalizzato nel principio di minima azione), ma che ogni possibile percorso dia un contributo identico nel calcolo della probabilità finale dell'evento. La formalizzazione di questa idea ha condotto alla definizione del *path integral* (o *somma sulle storie*), dove si è visto che il contributo dei percorsi differisce solo per un fattore di fase, definito in termini dell'azione classica. L'applicazione di questo metodo ha permesso di ritrovare l'ipotesi di de Broglie e Planck nel limite in cui l'azione quantistica può essere lecitamente approssimata dall'azione classica e ha mostrato come l'equazione di Schrödinger possa essere vista come l'approssimazione al primo ordine dell'approccio con i path integral. Questo mostra la completa equivalenza fra l'impostazione di Feynmann e la meccanica quantistica tradizionale.

≈≈ FINE ≈≈

⁷Per rendersene conto, basta osservare che si tratta di una funzione simmetrica pari ($e^{\frac{im\eta^2}{2\hbar\epsilon}}$) moltiplicata per una funzione simmetrica dispari (η) e integrata da $-\infty$ a $+\infty$.