Meccanica quantistica

Luciano Vanzo[1]

`Fisica Moderna`

corso di perfezionamento annuale

IPRASE Trentino

7 e 10 Settembre 2004

[1]vanzo@science.unitn.it

Indice

1 Teorie Fisiche **2**

2 Richiami di fisica classica, campi e particelle **3**
2.1 Dinamica Newtoniana 3
2.2 Il formalismo Lagrangiano 5
2.3 Il formalismo Hamiltoniano 7
2.4 Meccanica statistica 8
2.5 La radiazione 10
2.6 Interferenza e diffrazione 11

3 Difficoltà delle teorie classiche **14**
3.1 Calori specifici 15
3.2 La radiazione termica 17
3.3 La formula di Planck 18
3.4 L'effetto Compton 19
3.5 De Broglie e la diffrazione di elettroni 20

4 La meccanica quantistica **21**
4.1 Equazione di Schrödinger 21
4.2 Stati stazionari 25
4.3 Stati quantici 26
4.4 Le asserzioni statistiche della fisica quantistica 27
4.5 Digressione matematica: lo spazio L^2 31
4.6 Valori di aspettazione 34
4.7 Digressione matematica: operatori lineari 36
4.8 L'evoluzione temporale 41

5 Il formalismo generale **42**
5.1 I postulati 42
5.2 Statistica delle misure 44
5.3 Processi di misura 46
5.4 La determinazione dello stato 47
5.5 Relazioni di indeterminazione 48

A Equipartizione dell'energia **53**

B La distribuzione delle velocità **54**

C Equazioni di Maxwell **54**

1 Teorie Fisiche

Inizieremo il nostro corso mettendo in luce alcune caratteristiche che le teorie fisiche devono avere per essere considerate tali secondo il punto di vista della fisica. Innanzitutto si hanno i *sistemi fisici*. Un sistema fisico dovrebbe essere caratterizzato dalla lista di tutte le quantità effettivamente misurabili, dalle loro relazioni funzionali e dalle istruzioni di come dovrebbero essere misurate, e i loro valori letti o calcolati dalle indicazioni degli strumenti di misura. Detto questo, elenchiamo una serie di punti che ci sembrano importanti:

A. Una teoria fisica si può considerare come una teoria matematica in cui alcuni termini, non necessariamente tutti, hanno una interpretazione fisica, *di tipo operazionale*.

B Una interpretazione operazionale [1] è data in termini di operazioni eseguibili in laboratorio.

C. Ogni teoria si occupa di un particolare sistema fisico.

D. Alcuni termini rappresentano le *osservabili*, definite dalle operazioni fisiche (le misure) che hanno come scopo la determinazione di un numero (il risultato della misura). Altri rappresentano lo *stato* del sistema, definito dalle operazioni che si devono eseguire per prepararlo. Bisogna distinguere un sistema fisico da un sistema fisico in un dato stato.
Si misura sempre un osservabile in un dato stato.

E. La teoria deve essere in grado di fare previsioni sul risultato delle misure.

F. Le previsioni possono essere esatte (deterministiche) oppure di natura statistica. In questo caso bisogna ripetere la misura molte volte sul sistema preparato nello stesso stato, e determinare le frequenze di apparizione dei risultati.

G. La teoria deve essere in grado di calcolare lo stato in un istante t, se è noto lo stato all'istante iniziale t_0. In generale, questo si ottiene mediante la soluzione di un equazione differenziale, che determina la *dinamica* del sistema.

H. Bisogna ricordarsi che le teorie fisiche non forniscono una descrizione completa della realtà, e hanno sempre un campo di validità limitato.

2 Richiami di fisica classica, campi e particelle

[Ref.] H. Goldstein, *Meccanica classica*, Zanichelli, Bologna (1971).

Nella fisica classica si distinguono *particelle*, formanti la materia, e *radiazione*, perturbazione del campo elettromagnetico che si propaga con la velocità della luce. L'osservazione mostra che tra la materia e la radiazione esistono interazioni complicate, che però sono descritte da teorie sorprendentemente semplici e concise, delle quali l'elettrodinamica classica è il fondamento. Alcune inconsistenze di questa teoria si possono risolvere solo con l'introduzione della teoria della relatività ristretta; insieme, esse formano lo schema più evoluto raggiunto dalla fisica classica alla fine del secolo diciannovesimo, quello stesso schema che le nuove scoperte di fine secolo si appresteranno a mettere profondamente in crisi.

2.1 Dinamica Newtoniana

Un modello classico di riferimento per la struttura della materia è il seguente: N punti materiali (molecole monoatomiche) di massa m in un volume molto grande V rappresentano il nostro universo. Si hanno N variabili di posizione

$$(\vec{r_1}, \ldots, \vec{r_N}) \in R^{3N}$$

e N variabili di quantità di moto

$$(\vec{p_1}, \ldots, \vec{p_N}) \in R^{3N}, \qquad \vec{p}_k = m\vec{v}_k$$

dove $\vec{v}_k = \dot{\vec{r}}_k$ è la velocità della k-esima particella[1]. È sottinteso che si usano coordinate cartesiane, e che queste sono riferite a un sistema di riferimento inerziale. Il tempo è identico in tutti i sistemi riferimento inerziali. I problemi legati a questa interpretazione del continuo spazio-temporale riguardano la teoria della relatività ristretta, e non sono trattati in questo corso.
Le equazioni del moto sono le equazioni di Newton

$$m\dot{\vec{v}}_k = -\vec{\nabla}_k \Phi(\vec{r}_1, \ldots, \vec{r}_N) \tag{2.1}$$

dove $\vec{\nabla}_k$ indica il gradiente rispetto alla variabile $\vec{r}_k$ e l'energia potenziale ha un espressione della forma

$$\Phi(\vec{r}_1, \ldots, \vec{r}_N) = \sum_{1 \le i < k \le N} V(|\vec{r}_i - \vec{r}_k|) + \sum_{k=1}^{N} U(\vec{r}_k) \tag{2.2}$$

[1] Il punto sopra le variabili indica la derivata presa rispetto al tempo.

$U(\vec{r})$ è un potenziale esterno che tiene conto, ad esempio, del peso delle molecole, oppure rappresenta il potenziale di interazione con le pareti di un recipiente[2].
Il potenziale di interazione fra le molecole è definito invece dalla funzione $V(r)$, con $r = |\vec{r}| = (x^2 + y^2 + z^2)^{1/2}$. Al potenziale si richiedono alcune proprietà naturali

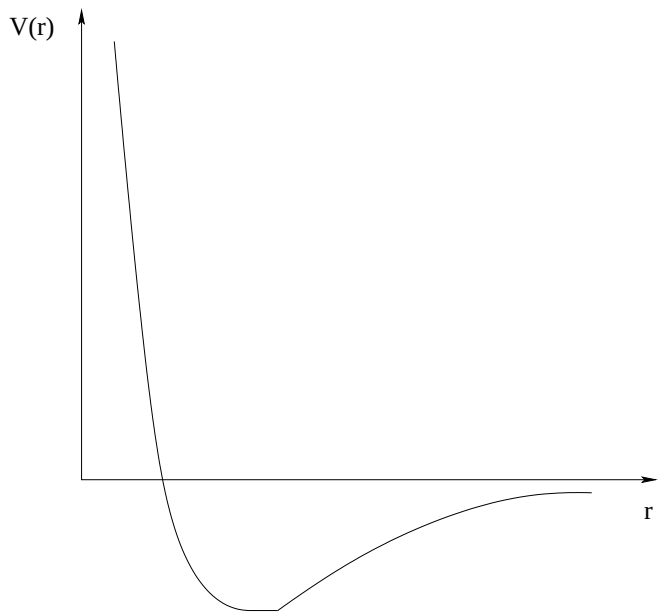
1. $V(r) \to 0$ per $r \to \infty$; sappiamo infatti che tra molecole elettricamente neutre a grandi distanze opera solo la forza gravitazionale, che è molto debole.

2. per qualche valore di r, $V(r)$ deve essere attrattivo, allo scopo di rendere conto degli stati condensati e delle forze di coesione.

3. a piccole distanze il potenziale deve essere repulsivo, per evitare l'accumulazione senza limite delle particelle in piccole regioni di spazio.

Dunque il grafico di $V(r)$ potrebbe assomigliare alla curva mostrata in figura

Figura 1: Il potenziale di interazione fra le molecole



In questo modo l'energia potenziale è limitata inferiormente, cioè esiste una costante C_N tale che

$$\Phi(\vec{r}_1, \ldots, \vec{r}_N) \geq -C_N$$

[2]Come nella teoria cinetica dei gas.

L'energia del sistema è data dalla nota espressione

$$H = \frac{1}{2m} \sum_{k=1}^{N} p_k^2 + \Phi(\vec{r}_1, \ldots, \vec{r}_N) \tag{2.3}$$

in cui[3] $p_k^2 = \vec{p}_k \cdot \vec{p}_k$.
Si ha qui una teoria fisica nel senso specificato sopra.
Gli **stati** sono matematicamente rappresentati dai punti[4]

$$P = (\vec{r}_1, \ldots, \vec{r}_N, \vec{p}_1, \ldots, \vec{p}_N) \in R^{3N}$$

P si chiama il *punto rappresentativo* e R^{3N} è il cosiddetto *spazio delle fasi* del sistema.
Le **osservabili** sono matematicamente rappresentate dalle funzioni reali definite sullo spazio delle fasi. Il risultato della misura di f nello stato rappresentato da P è il numero reale $f(P)$, ed è quindi univocamente determinato (si considerano funzioni almeno continue).
La **dinamica** è determinata dalle equazioni del moto di Newton (2.1), in cui si sceglie un punto rappresentativo come dato iniziale. Se il sistema fisico e lo stato sono noti (cioè è noto il punto che funge da dato iniziale), la teoria dà istruzioni non ambigue per rispondere a tutte le questioni mediante il calcolo.
La teoria si può estendere fino a includere le particelle elettricamente cariche e il campo elettromagnetico, ma per essere interamente consistente si devono introdurre i concetti propri della relatività di Einstein (a parte gli effetti di reazione che la radiazione ha sulle particelle che la producono).

2.2 Il formalismo Lagrangiano

Introduciamo la funzione di Lagrange

$$L(\vec{r}_1, \ldots, \vec{r}_N, \vec{v}_1, \ldots, \vec{v}_N) = \frac{1}{2} \sum_{k=1}^{N} m v_k^2 - \Phi(\vec{r}_1, \ldots, \vec{r}_N) \tag{2.4}$$

dove $\vec{v}_k = \dot{\vec{r}}_k$ sono le velocità delle particelle. La quantità di moto è

$$\vec{p}_k = \frac{\partial L}{\partial \vec{v}_k} = m\vec{v}_k$$

Con questa notazione si intende che una data componente della quantità di moto si ottiene derivando la Lagrangiana rispetto alla stessa componente

[3] Il prodotto scalare è $\vec{a} \cdot \vec{b} = a_x b_x + a_y b_y + a_z b_z$.

[4] Con R^n i matematici indicano l'insieme delle n-uple ordinate $(x^1, \ldots, x^n)$ di numeri reali.

della velocità. Si verifica subito che le equazioni del moto (2.1) si possono scrivere nel seguente modo

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \vec{v}_k}\right) - \frac{\partial L}{\partial \vec{r}_k} = 0 \tag{2.5}$$

Queste sono le equazioni di Eulero-Lagrange, le cui soluzioni godono della seguente proprietà notevole [2]: l'integrale (detto azione)

$$I = \int_{t_0}^{t_1} L(\vec{r}_1(t), \ldots, \vec{r}_N(t), \vec{v}_1(t), \ldots, \vec{v}_N(t))\, dt \tag{2.6}$$

assegna un numero reale ad ogni curva differenziabile definita nell'intervallo di tempo $[t_0, t_1]$. Si immagini ora di avere una soluzione $\vec{R}_k(t)$ delle equazioni di Eulero-Lagrange, e consideriamo una piccola variazione descritta dalla curva $\vec{r}_k(t) = \vec{R}_k(t) + \delta\vec{r}_k(t)$, soggetta però alle restrizioni

$$\delta\vec{r}_k(t_0) = 0, \qquad \delta\vec{r}_k(t_1) = 0$$

Queste variazione avranno l'effetto di indurre una piccola variazione del primo ordine nell'integrale di azione, ma risulta che in conseguenza delle equazioni (2.5) si ha sempre $\delta I = 0$. In altri termini, l'azione è stazionaria per piccole variazioni attorno alla soluzione, purché la curva variata passi negli istanti t_0 e t_1 per gli stessi punti per cui passa la soluzione $\vec{R}_k(t)$. È questo il principio di minima azione di Hamilton.
È possibile ottenere un formalismo più generale se si usano coordinate generalizzate arbitrarie $(q^1, \ldots, q^n)$, al posto delle coordinate cartesiane, purché l'assegnazione dei loro valori non sia soggetta a vincoli di nessun tipo. Si può avere $n < 3N$, perché fra le variabili originali $\vec{r}_k$ possono esistere vincoli che riducono il numero di variabili indipendenti. Ad esempio, per un corpo rigorosamente rigido si ha $n = 6$, perché bastano le tre coordinate del baricentro più tre coordinate angolari per specificare univocamente la posizione e l'orientazione nello spazio del corpo rigido. In questo caso i vincoli sono rappresentati dalle condizioni che fissano le distanze reciproche di tutti i punti del corpo. In un certo senso si può anche avere $n = \infty$, come nel caso della corda vibrante di lunghezza finita. Il numero n di coordinate generalizzate non più soggette a vincoli definisce *il numero dei gradi di libertà* del sistema considerato. Cosí si dirà che il corpo rigido è un sistema con sei gradi di libertà, mentre la corda vibrante è un sistema con infiniti gradi di libertà.
L'azione (2.6) diventa

$$I = \int_{t_0}^{t_1} L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n) dt$$

Il principio di minima azione di Hamilton stipula che sia $\delta I = 0$ per effetto di tutte le variazioni δq_k della traiettoria che si annullano negli istanti iniziale t_0 e finale t_1 (peraltro arbitrari). Si ottiene (al primo ordine nelle variazioni δq^k)

$$\delta I = \int_{t_0}^{t_1} dt \sum_k \left(\frac{\partial L}{\partial q^k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^k} \right) \delta q^k + \int_{t_0}^{t_1} \sum_k \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^k} \delta q^k \right) dt = 0$$

Il secondo termine (l'integrale di una derivata), coinvolge le variazioni negli estremi di integrazione e dunque non dà contributi. Resta

$$\delta I = \int_{t_0}^{t_1} dt \sum_k \left(\frac{\partial L}{\partial q^k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^k} \right) \delta q^k = 0$$

che deve valere per tutte le variazioni. Essendo queste linearmente indipendenti (per definizione di coordinate generalizzate) devono annullarsi i loro coefficienti, e si hanno cosí le equazioni di Eulero-Lagrange [2]

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^k} - \frac{\partial L}{\partial q^k} = 0 \tag{2.7}$$

Le grandezze

$$p_k = \frac{\partial L}{\partial \dot{q}^k} \tag{2.8}$$

sono detti *momenti coniugati* alle coordinate generalizzate. Le grandezze $(p, q) = (p_1, \ldots, p_n, q^1, \ldots, q^n)$ sono le coordinate dei punti di uno spazio detto *spazio delle fasi* del sistema. I punti stessi si chiamano *punti rappresentativi.*

2.3 Il formalismo Hamiltoniano

Supponiamo che sia possibile risolvere l'equazione (2.8) nelle velocità generalizzate, che diventano funzioni delle variabili di fase (p, q). In questo caso si può definire la funzione di Hamilton, o *Hamiltoniana*[5]

$$H(p, q, t) = \sum_k p_k \dot{q}^k - L(q, \dot{q}, t) \tag{2.9}$$

Variando le coordinate p_k, q^k si ottiene[6]

$$\delta H = \sum_k \left(\dot{q}^k \delta p_k - \frac{\partial L}{\partial q^k} \delta q^k \right) - \frac{\partial L}{\partial t}$$

[5]Qui supponiamo che la Lagrangiana possa anche dipendere esplicitamente dal tempo.
[6]Il termine $p_k \delta \dot{q}^k$ è cancellato dal termine $\delta \dot{q}^k \partial L / \partial \dot{q}^k$.

Ciò equivale a scrivere

$$\dot{q}^k = \frac{\partial H}{\partial p_k}, \qquad \dot{p}_k = \frac{\partial L}{\partial q^k} = -\frac{\partial H}{\partial q^k}, \qquad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \tag{2.10}$$

Queste si chiamano *equazioni canoniche di Hamilton* [2]. Esse determinano la traiettoria del punto rappresentativo nello spazio delle fasi,

$$t \to P_t = (p_1(t), \ldots, p_n(t), q^1(t), \ldots, q^n(t))$$

ossia la dinamica del sistema. Calcolando H dalla Lagrangiana (2.4), si vede che H rappresenta l'energia del sistema. In quel caso vale la formula

$$H(p,q) = \sum_{k=1}^{3N} \frac{p_k^2}{2m} + \Phi(q^1, \ldots, q^{3N}) \tag{2.11}$$

Dalle equazioni canoniche segue subito che

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}$$

e dunque l'Hamiltoniana è una costante del moto, quando non dipende esplicitamente dal tempo. Questa costante *definisce* l'energia del sistema anche quando la Lagrangiana non ha la forma semplice data dall'eq. (2.4).

2.4 Meccanica statistica

[Ref.] L. D. Landau e E. M. Lifchitz[7], *Fisica Statistica*, Editori Riuniti, (1978).

Una grammo molecola di una sostanza chimica qualsiasi contiene il numero enorme di circa $6,022 \cdot 10^{23}$ particelle (numero di Avogadro). Si vede bene che dal punto di vista della meccanica una mole è un sistema estremamente complicato: chi potrebbe assegnare $18,066 \cdot 10^{23}$ dati iniziali e risolvere lo stesso numero di equazioni differenziali ordinarie? Per sistemi di questo genere, è in generale possibile fare solamente asserzioni probabilistiche. Ma i metodi statistici non mettono in discussione i fondamenti della meccanica classica, ben diversamente da quanto avviene in meccanica quantistica. La fisica statistica inizia considerando un insieme molto numeroso di N punti rappresentativi, descriventi stati dello stesso sistema fisico ma corrispondenti a dati iniziali leggermente diversi. In realtà noi abbiamo un solo sistema,

[7] Nella traduzione Italiana questo autore è scritto Lifšits.

ma siccome non conosciamo quale degli N punti lo rappresenta (lo stato iniziale non è esattamente noto), è necessario ragionare su tutti gli N punti. Nel caso di un sistema macroscopico isolato e in condizioni stazionarie (cioè in equilibrio termodinamico), l'insieme statistico dei punti rappresentativi è descritto mediante una *funzione di distribuzione* $\rho(p,q)$ nello spazio delle fasi del sistema [3] (e che si può pensare proporzionale alla densità di questi punti). Se

$$\delta^n p \delta^n q = \delta p_1 \cdots \delta p_n \delta q^1 \cdots \delta q^n$$

rappresenta un piccolo volume dello spazio delle fasi, assegneremo la probabilità

$$\delta w = a\rho(p,q)\delta p^n \delta q^n, \qquad a^{-1} = \int \rho(p,q) d^n p d^n q$$

all'evento che $(p,q) \in \delta p^n \delta q^n$. Il valore di aspettazione delle osservabili in un insieme statistico, cioè la media del risultato di molte misure, è dunque

$$< f >= \frac{\int \rho(p,q) f(p,q) d^n p d^n q}{\int \rho(p,q) d^n p d^n q} \tag{2.12}$$

D'altronde le grandezze misurate dagli strumenti sensibili all'azione collettiva di moltissime molecole, che sono le grandezze macroscopiche (ad esempio la pressione di un gas), sono rappresentate più realisticamente da medie temporali del tipo

$$\bar{f} = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(p(t), q(t)) dt \tag{2.13}$$

prese lungo la traiettoria del punto rappresentativo del sistema. Si noti che per calcolare la media temporale occorre conoscere la soluzione delle equazioni del moto, il problema che i metodi statistici cercano appunto di evitare. Il problema importante è allora quello di sapere quando la media statistica dell'equazione (2.12) coincide con la media temporale (2.13), e questo fa parte di un ramo della fisica statistica chiamato *teoria ergodica*. Sotto certe ipotesi si dimostra che le due medie coincidono. In tal caso il problema fondamentale della fisica statistica diventa la determinazione della funzione di distribuzione $\rho(p,q)$. A tale proposito, si può tenere conto del fatto sperimentale che le proprietà macroscopiche dei sistemi in equilibrio dipendono solo dalla forma dell'Hamiltoniana e dal valore dell'energia E. Questo ci permette di scrivere che $\rho(p,q) = \rho(H(p,q))$, cioè che ρ dipende dal punto (p,q) unicamente attraverso H.

Se il sistema è in debole interazione con un termostato (avente lo scopo di mantenere costante la temperatura), si ha la formula fondamentale di Boltzmann-Gibbs [3]

$$\rho(p,q) = \exp\left[-\beta H(p,q)\right], \qquad \beta = \frac{1}{kT} \tag{2.14}$$

$k = 1,38 \cdot 10^{-16}$ erg/grado è la costante di Boltzmann e T è la temperatura assoluta. Questa distribuzione si chiama "la distribuzione canonica", ed è una delle formule più importanti della statistica.
La statistica classica, nella forma datale da L. Boltzmann e W. Gibbs, e la teoria cinetica dei gas da cui è storicamente derivata, sono teorie ammirevoli. Esse hanno permesso la spiegazione microscopica della termodinamica e reso praticamente indubitabile l'ipotesi della costituzione atomica della materia[8]. (Si veda, a questo proposito, il libro di J. Perrin [4].) In appendice sono descritte due importanti conseguenze della statistica classica.

2.5 La radiazione

[Ref.] E. M. Purcell, *Elettricità e magnetismo*, Collana "La Fisica di Berkeley", Vol. 1 e Vol. 2, Zanichelli, Bologna (1973).

[Ref.]: L. D. Landau e E. M. Lifchitz, *Teoria dei campi*, Editori Riuniti (1976).

La radiazione elettromagnetica è adeguatamente descritta dalle equazioni di Maxwell [6, 5]. Le variabili dinamiche fondamentali (nonché osservabili) sono le ampiezze dei campi elettrico e magnetico, $\vec{E}(\vec{r},t)$ e $\vec{B}(\vec{r},t)$ rispettivamente, che sono funzioni differenziabili delle variabili di posizione e del tempo. Si possono sempre derivare i campi da opportuni potenziali, in accordo con le formule

$$\vec{E} = -\vec{\nabla}\phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}$$

$$\vec{B} = \mathrm{rot}\,\vec{A}$$

I campi sono termini della teoria che hanno una interpretazione fisica di tipo operazionale, e sono dunque misurabili. Per contrasto, i potenziali scalare e vettore, ϕ e $\vec{A}$ rispettivamente, non godono di questa proprietà. Essi sono solo concetti ausiliari, usatiprincipalmente per ottenere le soluzioni delle equazioni di Maxwell in problemi specifici.

[8] L'altra prova stringente fu l'osservazione e la spiegazione teorica del moto Browniano.

Altre importanti osservabili sono la densità di energia nei campi[9]

$$u = \frac{1}{8\pi}(E^2 + B^2) \tag{2.15}$$

e il suo vettore di flusso, o vettore di Poynting, ossia la quantità di energia che fluisce per unità di tempo attraverso l'unità di superficie ortogonale al flusso

$$\vec{S} = \frac{c}{4\pi}\vec{E} \times \vec{B} \quad \text{(prodotto vettoriale)} \tag{2.16}$$

Dunque la quantità di energia che arriva per unità di tempo sopra una superficie irradiata $\mathcal{F}$ è

$$W = \int_{\mathcal{F}} \vec{S} \cdot \vec{n}\, d\sigma$$

dove $\vec{n}$ è il versore normale alla superficie.
La presenza nelle equazioni della velocità della luce nel vuoto, $c \simeq 2,99 \cdot 10^{10}$ cm/sec, indica la connessione della teoria elettromagnetica di Maxwell con la natura fisica della luce. La quale è appunto un campo elettromagnetico di alta frequenza (intorno ai 10^{15} cicli di oscillazione al secondo), piccola parte di un esteso spettro elettromagnetico che occupa tutta la gamma disponibile delle lunghezze d'onda (o delle frequenze) da zero all'infinito, ma che in pratica si estende dalle onde radio ($\lambda = 10^2 \div 10^5\, cm$), passando per le microonde ($\lambda = 10 \div 10^{-2}\, cm$), l'infrarosso ($\lambda = 10^{-2} \div 10^{-4}\, cm$), il visibile ($\lambda = 10^{-5}\, cm$), per giungere fino ai raggi X ($\lambda = 10^{-8} \div 10^{-10}\, cm$) e gamma ($\lambda > 10^{-10}\, cm$).
Si noti che la descrizione è radicalmente diversa da quella adottata nella meccanica. I punti dello spazio non si riferiscono più alle particelle, ma sono gli argomenti da cui dipendono i campi, e si ha un valore del campo in ogni punto ad ogni istante. I campi sono dunque funzioni (in genere differenziabili, per soddisfare le equazioni di Maxwell) e tali sono le osservabili, come il flusso di energia o la sua distribuzione nello spazio. Non appare nessuna discretezza, o atomicità, nei fenomeni che coinvolgono la radiazione, che invece si propaga secondo leggi affatto diverse che dalle particelle. A queste leggi particolari si dà il nome generico di *fenomeni ondulatori*.

2.6 Interferenza e diffrazione

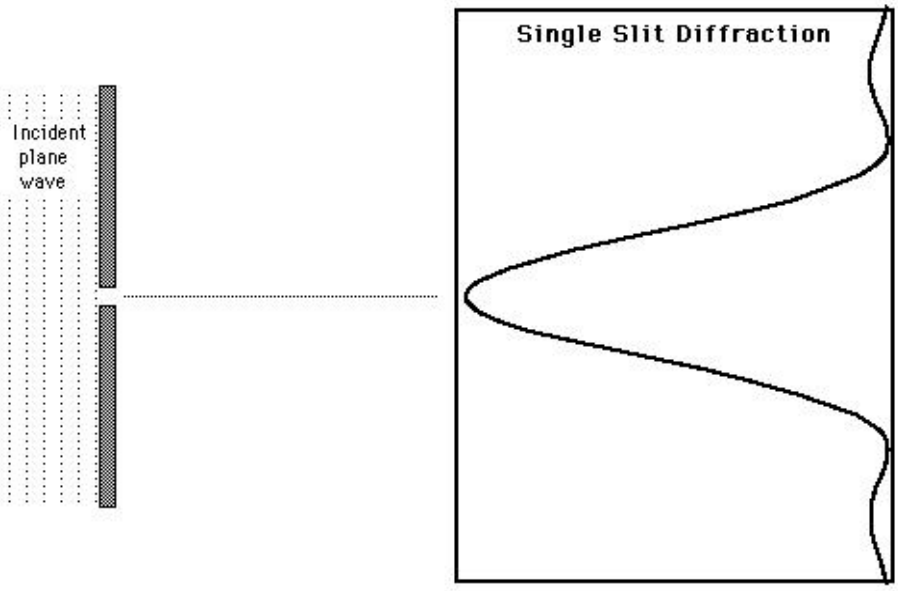
[Ref.] F. S. Crawford, *Onde e oscillazioni*, Collana "La Fisica di Berkeley", Vol. 3, Zanichelli, Bologna (1972).

[9] Come al solito, si è posto $E^2 = \vec{E} \cdot \vec{E}$ e per B^2 una espressione analoga.

[Ref.]: L. D. Landau e E. M. Lifschitz, *Teoria dei campi*, Editori Riuniti (1976).

La fig. [2] dà un'idea della figura di diffrazione prodotta dalla luce quando attraversa una fenditura stretta. La figura riporta l'intensità della luce nella posizione verticale dello schermo in funzione del parametro $D \sin\theta$, D essendo la larghezza della fenditura e θ l'angolo con la direzione orrizontale.

Figura 2: Diffrazione da una fenditura

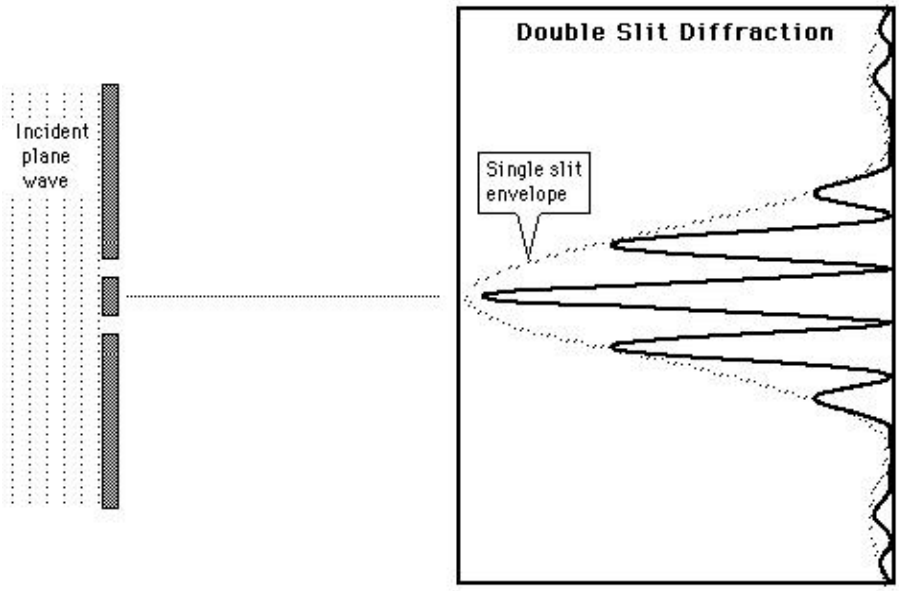


La formula dell'intensità secondo la teoria ondulatoria della luce, che è la teoria elettromagnetica di Maxwell, è [7, 5]

$$I_f(\theta) = I_{\max}\frac{\sin^2\Phi}{\Phi^2}, \qquad \Phi = \frac{2\pi D\sin\theta}{\lambda} \tag{2.17}$$

dove λ è la lunghezza d'onda della radiazione e $I_{\max}$ l'intensità nel massimo centrale. La larghezza angolare del picco è di ordine $\delta\theta \simeq \lambda/D$.
La fig. [3] presenta invece la diffrazione per due fenditure , nella quale si vede la modulazione dovuta alla diffrazione delle fenditure singole.
Come è noto, per ottenere la figura si considerano le fenditure come due sorgenti coerenti di radiazione, e si sommano i contributi di entrambe nei vari punti dello schermo. La geometria del problema è illustrata in fig. [4]

Figura 3: Diffrazione da due fenditure



Per illustrare la procedura, nel seguito ignoriamo la natura vettoriale del campo di radiazione (cioè gli effetti di polarizzazione) e scriviamo il campo in P

$$u(P) = \xi_0(P)e^{-i\omega t}\left(e^{i2\pi R_1/\lambda} + e^{i2\pi R_2/\lambda}\right)$$

dove $\xi_0(P)$ è il campo della fenditura singola. Qui supponiamo che la distanza del punto P sia grande rispetto alla distanza tra le fenditure (condizione di Fraunhofer); dalla figura si ha $R_2 = R_1 + d\sin\theta$, e quindi

$$u(P) = 2\xi_0(P)e^{i(kR_1+kd\sin\theta/2-i\omega t)}\cos\frac{1}{2}\Delta\phi$$

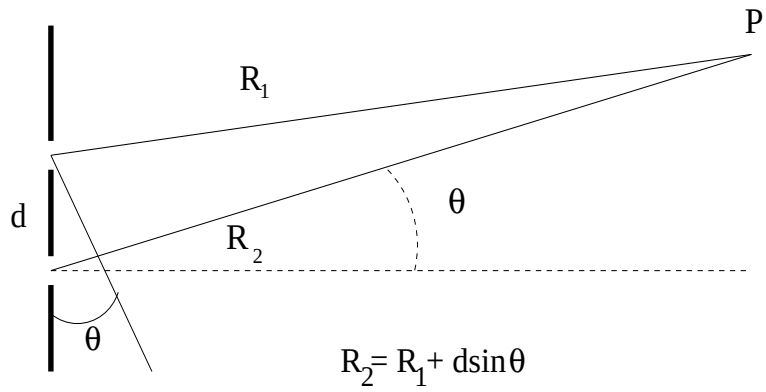
dove abbiamo introdotto il numero d'onde $k = 2\pi/\lambda$ e

$$\Delta\phi = \frac{2\pi d\sin\theta}{\lambda}$$

L'intensità è proporzionale al quadrato del modulo di $u(P)$; quindi si trova

$$I = I_f(\theta)\cos^2\left(\frac{1}{2}\Delta\phi\right)$$

Figura 4: Geometria della diffrazione



dove $I_f(\theta)$ è l'intensità della fenditura singola, data dalla (2.17). La larghezza angolare della figura di diffrazione è dell'ordine di λ/D.
L'osservazione sperimentale delle figure di diffrazione da parte di T. Young (1803/1804) mostra chiaramente la natura ondulatoria della luce. È interessante osservare che Young fece i suoi esperimenti proprio per rispondere alla questione, molto dibattuta nel suo tempo, se si dovesse considerare la luce come formata da particelle oppure no. Gli effetti collettivi delle particelle, come ad esempio il loro arrivo su di una emulsione fotografica, sono sempre additivi. Ma i fenomeni di interferenza mostrano che in alcune regioni gli effetti si sommano, mentre in altre si cancellano. È questo il comportamento che normalmente si ascrive alle onde; si potrebbe anche dire che i fenomeni di diffrazione e interferenza sono condizione necessaria per poter parlare di onde.
L'inclusione dei raggi X (W. Röntgen, 1895) nello spettro elettromagnetico, avvenne più tardi. Già nel 1905 era noto che i raggi X possono essere polarizzati, ma non rifratti, suggerendo che se fossero onde la loro lunghezza dovrebbe essere estremamente piccola (infatti $\lambda \simeq 10^{-10}$cm). In seguito M. von Laue (1912) produsse figure di diffrazione facendo passare i raggi X attraverso i cristalli (si veda il testo di M. Born [8] per una trattazione esauriente di questi argomenti).

3 Difficoltà delle teorie classiche

[Ref.] M. Born, *Fisica atomica*, Boringhieri (1968).

I fatti che imposero una revisione delle teorie classiche riguardano in parte i

calori specifici dei solidi e dei gas, la radiazione termica, gli spettri atomici e le proprietà particolari dei raggi X e della loro interazione con la materia, e in parte le sorprendenti proprietà della diffusione degli elettroni nei cristalli. Con i primi divenne chiaro che l'energia degli stati legati, atomi e molecole, può assumere solamente un insieme discreto di valori. Questo richiede che la radiazione elettromagnetica scambi l'energia con i sistemi atomici per quanti discreti, quei *fotoni* che furono successivamente introdotti da A. Einstein nella sua spiegazione dell'effetto fotoelettrico (Einstein, 1904). La meccanica classica non offre spiegazioni plausibili.
Con i secondi, si è visto che in alcune circostanze gli elettroni si comportano come se fossero onde, in altre come particelle. È importante osservare che le suddette circostanze sono definite dalle condizioni generali degli esperimenti, cioè dagli strumenti di misura impiegati e dalle procedure con le quali i fasci di elettroni sono stati effettivamente preparati. Anche qui la meccanica classica non offre spiegazioni plausibili.

3.1 Calori specifici

Un cristallo monoatomico con N atomi possiede $3N$ gradi di libertà di oscillazione indipendenti, dunque il teorema di equipartizione (vedi appendice) dà per il calore specifico a volume costante la legge di Dulong-Petit [8]

$$C_V = 3NK$$

o per una mole

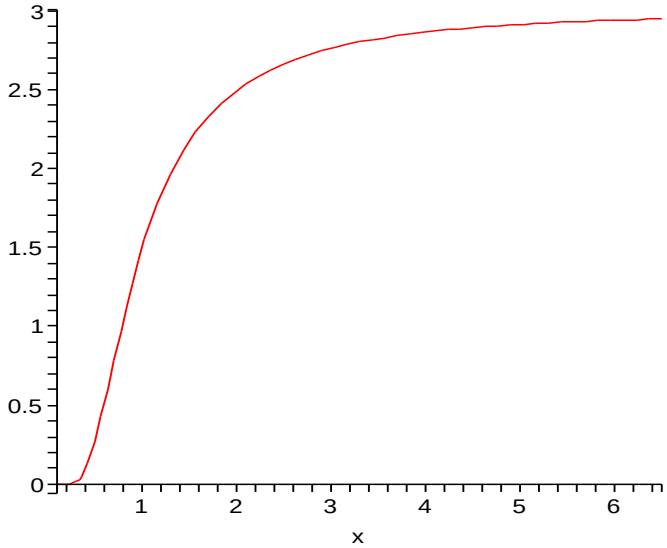
$$C_V = 3R \simeq 25 \text{ Joule}/K$$

La fig. [5] mostra una tipica curva di $C_V(T)/R$ per un metallo, e come si vede il risultato è in grossolano disaccordo con le misure di C_V.
Inoltre gli elettroni nei metalli dovrebbero contribuire a tutte le temperature, in conflitto con le osservazioni.
Le molecole biatomiche hanno tre gradi di libertà traslazionali, due rotazionali (la rotazione attorno all'asse di simmetria della molecola corrisponde a temperature di eccitazione enormi) e uno oscillatorio. Il calore specifico di una mole di gas biatomico dovrebbe essere

$$C_V = \frac{7R}{2} \simeq 7 \text{ calorie}/K$$

e mai in ogni caso inferiore a 6 calorie per grado. Invece a temperatura ambiente la maggior parte dei gas biatomici rivela solo 5 gradi di libertà, e il

Figura 5: Curva di calore specifico



calore specifico passa con continuità da $5R/2$ a $7R/2$, aumentando gradatamente T. A temperature sufficientemente basse poi, tutti i gas si comportano come se fossero monoatomici, con un calore specifico per mole pari a $3R/2$! In altri termini, i gradi di libertà rotazionali e oscillatori sono completamente inattivi oppure, come si dice, congelati. La meccanica classica è dunque grossolanamente in disaccordo con le osservazioni.
La teoria quantistica invece, assegna ad ogni oscillazione di frequenza ν un'energia pari a un multiplo intero del quanto di energia $E = h\nu$, dove

$$h = 6,62 \cdot 10^{-27} \text{erg} \cdot \text{sec}$$

è la costante di Planck. Le temperature necessarie per eccitare un tale oscillatore sono quindi dell'ordine di $T = h\nu/k$; questa temperatura è tipicamente molto alta, per l'idrogeno molecolare vale 6140 gradi, per l'azoto 3352, per l'ossigeno 2239, e cosí via. Si capisce dunque perché alla temperatura ambiente si osservano solo 5 gradi di libertà. Le temperature associate ai gradi di libertà rotazionali sono invece molto più basse, circa 85 gradi per l'idrogeno, 3 per l'azoto, 2 per l'ossigeno. Si capisce dunque perché a basse temperature (se il gas non liquefa) i gas biatomici si comportano come se fossero monoatomici.

3.2 La radiazione termica

Se le pareti di una cavità completamente vuota vengono portate e mantenute ad una data temperatura T, si genera nell'interno della cavità della radiazione elettromagnetica, e si raggiunge uno stato stazionario quando le pareti assorbono nell'unità di tempo una quantità di energia radiante uguale a quella che emettono. Sia $u(\nu, T)$ la densità spettrale dell'energia elettromagnatica nella cavità, cosicché $u(\nu, T)d\nu$ rappresenta l'energia per unità di volume contenuta nell'intervallo di frequenze $d\nu$; la notazione anticipa una speciale proprietà di questa funzione, scoperta da Kirchhoff nel 1859, che è il fatto di essere totalmente indipendente dalla forma e dai materiali delle pareti della cavità. G. R. Kirchhoff propose la determinazione di questa funzione universale, e la soluzione del problema è legata ai nomi di J. Stefan, L. Boltzmann, W. Wien e M. Planck. Tutti fecero un uso essenziale della natura elettromagnetica della radiazione. Secondo la teoria di Maxwell, il campo EM nella cavità cubica di lato L è equivalente ad un insieme infinito di oscillatori armonici con frequenze

$$\nu(\vec{n}) = \frac{c}{2L}\sqrt{n_x^2 + n_y^2 + n_z^2}$$

e ci sono due oscillatori per ogni frequenza permessa [8, 10]. I numeri (n_x, n_y, n_z) sono interi non negativi. Il numero di oscillatori con frequenza minore di ν è quindi uguale al doppio (due oscillatori per ogni frequenza) del volume di un ottavo della sfera definita dalla diseguaglianza

$$\sqrt{n_x^2 + n_y^2 + n_z^2} \leq \frac{2L\nu}{c}$$

cioè

$$N(\nu(\vec{n}) \leq \nu) = \frac{8\pi L^3}{3c^3}\,\nu^3$$

Il numero di oscillazioni elettromagnetiche proprie (modi normali) nell'intervallo di frequenza $d\nu$, in una cavità di volume $V = L^3$, si ottiene differenziando $N(\nu(\vec{n}) \leq \nu)$

$$dn_\nu = \frac{8\pi V}{c^3}\,\nu^2 d\nu$$

La meccanica statistica assegna energia media kT a ciascun oscillatore, e quindi la densità spettrale di energia nella cavità dovrebbe seguire la formula di Rayleigh-Jeans

$$u(\nu, T) = \frac{8\pi\nu^2}{c^3}\,kT$$

che però è in disaccordo con le misure per frequenze troppo alte o sufficientemente basse. L'assurdità della formula deriva anche dal fatto che essa implica densità di energia infinita nella cavità (l'integrale di $u(\nu, T)$ in $d\nu$ diverge per ogni T).

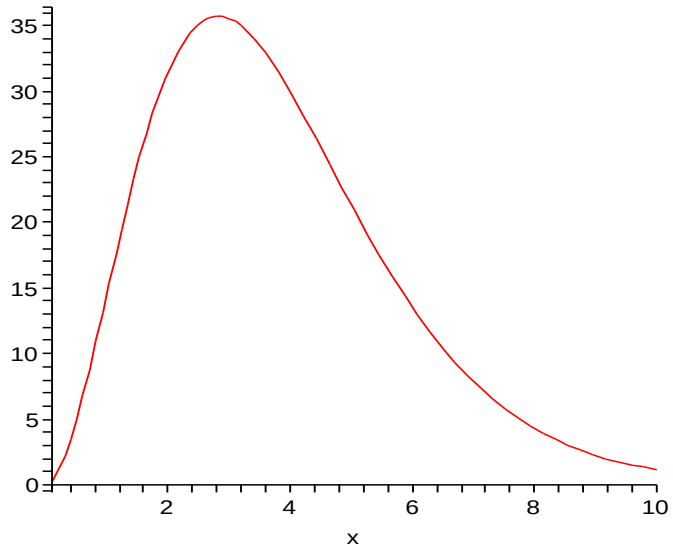
3.3 La formula di Planck

La formula corretta è stata proposta da Planck (1900), e si scrive [8]

$$u(\nu, T) = \frac{8\pi\nu^2}{c^3} \frac{h\nu}{e^{h\nu/kT} - 1} \tag{3.1}$$

e il suo grafico come funzione di ν si può vedere in Fig. [6]

Figura 6: La distribuzione di Planck



Si può ottenere la formula rinunciando all'idea che l'assorbimento e l'emissione di radiazione avvenga con continuità, attraverso lo scambio di quantità piccole a piacere. Piuttosto, lo scambio avviene per quanti indivisibili di energia $E = h\nu$. La derivazione corretta della formula si può solo ottenere con la meccanica quantistica, ed è arrivata 26 anni dopo con il lavoro di Dirac. Si vede che la formula differisce dal risultato classico di Rayleigh-Jeans

per il fatto che assegna all'oscillatore armonico l'energia media

$$< E >= \frac{h\nu}{e^{h\nu/kT} - 1}$$

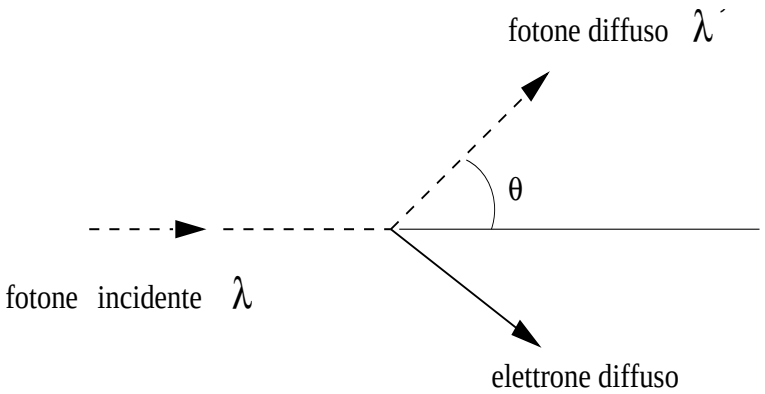
al posto di kT. Non è dunque vero che tutti gli oscillatori hanno in media la stessa energia: quelli con frequenza troppo grande ($h\nu \gg kT$) sono esponenzialmente soppressi. Il massimo della curva per una data temperatura si ha in corrispondenza della lunghezza d'onda (legge di spostamento di Wien)

$$\lambda\ (cm) = \frac{0,29}{T\ (^oK)}$$

3.4 L'effetto Compton

L'ipotesi dei quanti di luce è anche in perfetto accordo con l'effetto Compton-Simon (1923), che riguarda la diffusione di fotoni sugli elettroni. Il processo è illustrato in Fig. [7].

Figura 7: Effetto Compton



Si osserva una variazione istantanea della lunghezza d'onda in accordo con la formula di Compton [8]

$$\Delta\lambda = \lambda^{'} - \lambda = \frac{2h}{mc}\sin^2(\theta/2)$$

Si noti che $\Delta\lambda$ è indipendente dalla lunghezza d'onda. Questa relazione si deriva assumendo che

1. i fotoni abbiano energia e impulso connessi con la frequenza e il numero d'onda della radiazione incidente secondo le formule di Einstein-Planck

$$E = h\nu = \hbar\omega, \quad \vec{p} = \hbar\vec{k} = \frac{h}{\lambda}\vec{n} = \frac{h\nu}{c}\vec{n}$$

   dove $\hbar = h/2\pi$ e $\vec{n}$ è il versore nella direzione del moto del fotone

2. le leggi di conservazione dell'energia e dell'impulso siano valide.

Si noti che per i fotoni $E = c|\vec{p}|$, una relazione che è valida solo per le particelle di massa zero.
Secondo la descrizione classica, nel sistema di riferimento dove l'elettrone è istantaneamente in quiete la radiazione è diffusa con la stessa frequenza della radiazione incidente e l'impulso totale nell'irraggiamento è nullo; il processo di diffusione si accompagna quindi a un trasferimento continuo di impulso all'elettrone, che accelera progressivamente. La frequenza si sposta allora per effetto Doppler, e si determina che

$$\Delta\lambda = 2\lambda\frac{v}{c-v}\sin^2\frac{\theta}{2}$$

dove v è la velocità dell'elettrone. Si noti che $\Delta\lambda$ dipende da λ; il disaccordo con le osservazioni è evidente e classicamente incomprensibile.

3.5 De Broglie e la diffrazione di elettroni

Nel 1924 de Broglie propose che le relazioni di Einstein-Planck, $E = h\nu$ e $|\vec{p}| = h/\lambda$ fossero universalmente valide, dunque applicabili anche alle particelle come gli elettroni. La proposta contiene in sè l'idea rivoluzionaria che la materia, come la radiazione, abbia un carattere duale, secondo il quale dovrebbero manifestarsi proprietà ondulatorie per gli elettroni tutte le volte che le condizioni sperimentali lo permettono. Oggi sappiamo che è cosí non solo per gli elettroni, ma per tutte le particelle microscopiche successivamente scoperte [8] (ad esempio il neutrone, che non era noto al tempo di de Broglie, essendo stato identificato nel 1932).
La teoria di de Broglie trovò conferma nel 1927, quando Davisson e Germer osservarono per primi la diffrazione di elettroni riflessi o diffusi attraverso i cristalli [8], e le relazioni di de Broglie furono pienamente confermate. Molti anni dopo, nel 1961, C. Jönsson riuscí a ottenere la diffrazione usando un apparato di Young con due fenditure, ma non con un elettrone alla volta. Quest'ultimo fu realizzato da P. Merli et al. a Bologna nel 1976, ottenendo anche in questo caso una figura di diffrazione del tutto simile a quelle discusse sopra in relazione al comportamento della luce.

4 La meccanica quantistica

[Ref.] A. Messiah, *Meccanique quantique*, Vol. 1, Dunod (1958).

[Ref.] P. A. M. Dirac, *I principi della meccanica quantistica*, Boringhieri (1971).

Le esperienze di interferenza con particelle da una parte, e gli aspetti corpuscolari della radiazione dall'altra, rendono la maggior parte dei termini classici quali *particella*, *traiettoria*, *onda*, *diffrazione* e cosí via, del tutto inadeguati per descrivere in maniera completa i fenomeni fisici, eppure, nello stesso tempo, sono tutti egualmente necessari per descrivere e comunicare agli altri i risultati degli esperimenti. Ciò che serve è dunque un linguaggio teorico non ambiguo, una teoria; tale teoria è la *meccanica quantistica* (MQ) o, nel dominio dei fenomeni relativistici di alta energia, *la teoria quantistica dei campi* (TQC).
Le esperienze di interferenza e diffrazione con elettroni rendono anche urgenti i seguenti due quesiti (che non sono i soli naturalmente): che aspetto ha l'equazione a cui devono obbedire le onde di materia ? Qual'è l'interpretazione fisica operazionale dell'associazione di un'onda a ciascun tipo di particella, e viceversa ?
Nel seguito useremo con disinvoltura i termini classici solo allo scopo di comunicare i risultati della teoria, senza attribuire ai termini un diretto significato fisico. È un aspetto fondamentale della MQ, sottolineato con fermezza soprattutto da N. Bohr, che nessun termine classico può corrispondere a una descrizione completa della realtà. Esistono sempre termini *complementari* che sono egualmente necessari per la completezza della descrizione, ma che non possono essere usati contemporaneamente nelle medesime asserzioni, senza generare apparenti contraddizioni.

4.1 Equazione di Schrödinger

Il problema di determinare un'equazione per le onde di materia si può risolvere ricorrendo alle relazioni di de Broglie [9]. Consideriamo dapprima un'onda monocromatica piana

$$\Psi_k = e^{-i\omega t + i\vec{k}\cdot\vec{r}}$$

In accordo con de Broglie, Einstein e Planck, il carattere duale onda-particella si realizza associando alle caratteristiche dell'onda ω (frequenza angolare) e $\vec{k}$ (vettore numero d'onde), le caratteristiche della particella E (energia) e $\vec{p}$

(impulso o quantità di moto) secondo le relazioni ormai familiari

$$E = \hbar\omega, \qquad \vec{p} = \hbar\vec{k}$$

Ma per una particella libera si ha

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

e quindi deve valere la relazione di dispersione[10]

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

Come si verifica facilmente, questo significa che l'onda piana è soluzione dell'equazione di Schrödinger

$$i\hbar \frac{\partial \Psi_k}{\partial t} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2 \Psi_k}{\partial x^2} + \frac{\partial^2 \Psi_k}{\partial y^2} + \frac{\partial^2 \Psi_k}{\partial z^2}\right) = -\frac{\hbar^2}{2m}\nabla^2 \Psi_k \tag{4.1}$$

Abbiamo introdotto l'operatore differenziale di Laplace

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

Chiaramente, ci sono tre derivate seconde perchè ci sono tre dimensioni spaziali. Se si vuole semplificare la matematica si possono sopprimere due dimensioni (diciamo le dimensioni y e z). Allora l'equazione si semplifica un pò (ma la teoria matematica che ne tratta le proprietà si semplifica enormemente)

$$i\hbar \frac{\partial \Psi_k}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 \Psi_k}{\partial x^2} \tag{4.2}$$

Siccome non discuteremo le applicazioni dell'equazione, l'interpretazione fisica non soffrirà molto di questa riduzione di dimensioni.
È necessario che la sovrapposizione di onde piane sia ancora una possibile onda di materia, altrimenti non si saprebbe come determinare le figure di interferenza. Si possono allora formare i "pacchetti d'onde"

$$\psi(\vec{r}, t) = \int \hat{\psi}(\vec{k}) e^{-i\omega t + i\vec{k}\cdot\vec{r}} d^3k$$

[10] Cosí si chiama la relazione che lega la frequenza al numero d'onde.

e questi obbediranno ancora all'Eq. (4.1). Si noti che il secondo membro della (4.1) è semplicemente l'operatore differenziale che si ottiene dall'Hamiltoniana

$$H(p,q) = \frac{p^2}{2m}$$

attraverso la sostituzione

$$\vec{p} \to -i\hbar\vec{\nabla} \tag{4.3}$$

dove $\vec{\nabla}$ è l'operatore gradiente, che associa ad ogni funzione f il vettore con componenti $\partial f/\partial x$, $\partial f/\partial y$, $\partial f/\partial z$. Se la particella si muove in un potenziale esterno $V(\vec{r})$, l'Hamiltoniana classica è

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

e allora postuliamo semplicemente che l'equazione corretta sia

$$i\hbar\,\frac{\partial\psi}{\partial t} = H\,\psi \tag{4.4}$$

dove H denota adesso l'operatore differenziale

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r}) \tag{4.5}$$

Per un sistema di N particelle l'Hamiltoniano classico è

$$H = \sum_{k=1}^{N} \frac{p_k^2}{2m} + V(\vec{r}_1, \ldots, \vec{r}_N) \tag{4.6}$$

La sostituzione (4.3) (una per ogni $\vec{p}_k$) trasforma H in un operatore differenziale e l'equazione di Schrödinger è ancora la (4.4), eccetto che ψ è ora un'ampiezza d'onda in $3N$ dimensioni, vale a dire che

$$\psi = \psi(\vec{r}_1, \ldots, \vec{r}_N, t)$$

Si noti che non sarebbe consistente supporre che ψ prenda valori reali, perchè l'equazione di Schrödinger è complessa (contiene l'unità immaginaria i). Nel seguito denoteremo con $\overline{\psi}$ la funzione d'onda complessa coniugata di ψ e con $|\psi|^2 = \overline{\psi}\psi$ il suo modulo al quadrato.
Per il momento abbiamo acquisito un grado sufficiente di generalità, e adesso ci sembra opportuno fare una breve digressione per illustrare la grande potenza dell'idea di de Broglie e della sua formulazione matematica dovuta a E. Schrödinger.

Diffrazione di particelle

Il conto dell'interferenza di due fenditure strette che abbiamo illustrato nel caso della luce si adatta benissimo a onde qualsiasi. Si ottiene allora la figura di diffrazione con massimi a separazione angolare $\delta\theta \simeq \lambda/d = h/pd$. Si noti che è la grandezza $|\psi(P)|^2$ nei punti P dello schermo che determina l'intensità relativa dei massimi e dei minimi.

Stabilità degli atomi

L'Hamiltoniana di due particelle con carica opposta $\pm e$ si può mettere nella forma [2]

$$H = \frac{p^2}{2m} - \frac{e^2}{r}$$

dove m è la massa ridotta delle due particelle. Si vede che H non è limitata inferiormente. Se l'elettrone è delocalizzato in un volume di raggio r (raggio dell'atomo) allora la lunghezza d'onda è di ordine r e quindi la quantità di moto è $p = h/r$; questo dà

$$H = \frac{h^2}{2mr^2} - \frac{e^2}{r}$$

Si noti adesso che l'energia ha un minimo per

$$r = r_0 = \frac{h^2}{me^4}$$

che vale

$$E_0 = -\frac{me^4}{2h^2}$$

Vediamo dunque che per un'onda stazionaria non è possibile collassare nel centro, come farebbe una particella carica per emissione di radiazione elettromagnetica. Incidentalmente, il valore trovato per E_0 è quasi il risultato corretto (nel quale si ha semplicemente $\hbar$ al posto di h).
Si può anche dire, con maggior precisione, che la circonferenza di un'orbita circolare classica deve essere un multiplo intero della lunghezza d'onda

$$n\lambda = 2\pi r, \quad n = 1, 2, 3, \ldots$$

In tal caso $p = h/\lambda = n\hbar/r$ e l'energia assume i minimi nei valori $r_n = n^2\hbar^2/me^2$, per i quali si ha ($1\ eV$ (electron-volt) vale $1,6 \cdot 10^{-12}$ erg)

$$E_n = -\frac{me^4}{2\hbar^2 n^2} = -\frac{13,6\ eV}{n^2}$$

La lunghezza d'onda della radiazione emessa nelle transizioni atomiche $n \to m$ si può quindi scrivere nella forma (il fotone emesso deve avere energia $h\nu_{nm} = E_n - E_m = hc/\lambda_n - hc/\lambda_m$)

$$\frac{1}{\lambda_{nm}} = R\left(\frac{1}{n^2} - \frac{1}{m^2}\right), \quad R = 109678\ \mathrm{cm}^{-1}$$

e quindi rende immediatamente conto del *principio di combinazione di Ritz*, secondo il quale è possibile ordinare le righe spettrali dell'idrogeno in coppie tali che la somma dei rispettivi numeri d'onda è ancora una riga spettrale possibile.

4.2 Stati stazionari

Le funzioni d'onda della forma

$$\psi(\vec{r}, t) = e^{-iEt/\hbar}\Phi(\vec{r})$$

si chiamano *stati stazionari*, e soddisfano all'equazione di Schrödinger indipendente dal tempo [9, 10]

$$H\,\Phi = E\,\Phi \tag{4.7}$$

Nei cosiddetti stati legati, in cui le particelle sono confinate in regioni finite (atomi), si deve richiedere che $\Phi(\vec{r}) \to 0$ quando $|\vec{r}| \to \infty$. È anche necessario che Φ non abbia singolarità troppo severe in punti a distanza finita. Di norma ciò è possibile solo per certi valori discreti di E, che dunque sono i valori fisicamente permessi dell'energia nelle date condizioni.
Ad esempio, per una particella confinata in un segmento $[0, L]$ da due barriere infinite di potenziale, è ragionevole supporre valide le condizioni al contorno $\Phi(0) = \Phi(L) = 0$. L'equazione degli stati stazionari prende la forma

$$\frac{d^2\Phi}{dx^2} + \frac{2mE}{\hbar^2}\,\Phi = 0$$

Poniamo

$$k^2 = \frac{2mE}{\hbar^2}$$

Le soluzioni possibili sono $a \sin kx$ e $b \cos kx$, a e b costanti. Le condizioni al contorno rigettano il coseno e per il seno richiedono la condizione $\sin kL = 0$; quindi

$$k_n = \frac{n\pi}{L}, \qquad E_n = \frac{\pi^2 \hbar^2 n^2}{2mL^2}, \quad n = 1, 2, \ldots$$

Vediamo che la meccanica ondulatoria rende pienamente conto della possibilità che l'energia sia quantizzata.

4.3 Stati quantici

Non abbiamo ancora discusso una questione centrale della descrizione quantistica, e cioè quali siano i termini della teoria che corrispondono agli stati, e quali alle osservabili. Al punto in cui siamo non è ancora possibile dare una risposta esauriente perché non conosciamo l'interpretazione fisica della funzione d'onda, cioè non sappiamo in che modo dobbiamo usarla per ottenere informazioni fisiche sul sistema.
Ci limitiamo ad osservare il seguente fatto:

> se l'equazione di Schrödinger (4.4) è intesa sostituire completamente la descrizione classica, allora la funzione d'onda deve fornire una *descrizione completa* dello stato del sistema. Diremo dunque che $\Psi(\vec{r}_1, \ldots, \vec{r}_N, t)$ rappresenta matematicamente lo stato quantico del sistema (di N particelle) all'istante t.

La linearità dell'equazione di Schrödinger suggerisce che le combinazioni lineari di stati quantici rappresentino ancora stati quantici possibili: $a\phi + b\psi$ è un possibile stato quantico per ogni scelta di stati ϕ, ψ e numeri complessi a e b. Tuttavia è la funzione d'onda in un dato istante che rappresenta lo stato, e questa non è una soluzione dell'equazione di Schrödinger, ma svolge piuttosto il ruolo di dato iniziale. Quindi non è affatto evidente che la sovrapposizione di stati sia sempre possibile.
Tuttavia vedremo che l'insieme degli stati quantici può essere caratterizzato in maniera matematicamente rigorosa. Ma a questo arriveremo tra poco.
In linea di principio si possono usare coordinate generalizzate (p, q), e costruire l'Hamiltoniana quantistica da quella classica $H(p, q)$ mediante la sostituzione

$$p_k = -i\hbar \frac{\partial}{\partial q^k}, \quad k = 1, 2, \ldots, n$$

Gli stati sono allora le funzioni complesse $\psi(q^1, \ldots, q^n)$. Ci sono però due sorgenti di ambiguità:

(1) la regola fornisce operatori H diversi se si usano coordinate diverse. In altri termini, la regola non è invariante per trasformazioni delle coordinate generalizzate q^k

(2) Ci sono problemi a definire univocamente l'operatore

$$H\left(q^1, \ldots, q^n, -i\hbar \frac{\partial}{\partial q^1}, \ldots, -i\hbar \frac{\partial}{\partial q^n}\right)$$

in quanto le variabili q^k e p_k non commutano: definendo il commutatore $[A, B]$ di due operatori A e B come l'operatore tale che $[A, B]\psi = AB\psi - BA\psi$, $\forall\psi$, si ha

$$\left[q^k, -i\hbar \frac{\partial}{\partial q^k}\right] = i\hbar$$

Si può ovviare a queste difficoltà semplicemente permettendo trasformazioni a coordinate generalizzate solo dopo che un'Hamiltoniana quantistica ragionevole sia stata trovata. Il termine "ragionevole" si riferisce al fatto che devono esistere dei criteri generali ai quali un'Hamiltoniana quantistica deve obbedire. Siccome ogni ipotesi si giustifica dalle conseguenze, la scelta dell'operatore Hamiltoniano non appare essere un problema di natura concettuale, ma un'ipotesi tecnica. La sua reale giustificazione, nel dominio di fenomeni dove vale la MQ, è nei risultati da essa ottenuti (ad esempio lo spettro energetico), se siano o no conformi all'esperienza. Noi non ci soffermeremo oltre su questi problemi.

4.4 Le asserzioni statistiche della fisica quantistica

In linea di principio si può abbandonare l'idea intuitiva di particella e rimpiazzare la teoria classica con una teoria puramente ondulatoria. Le particelle sarebbero allora, in realtà, *pacchetti d'onde* di estensione piccola ma finita. Questi pacchetti si muovono in effetti secondo le leggi classiche del moto. Consideriamo ad esempio il pacchetto d'onde (in una sola dimensione)

$$\psi(x, t) = \int \hat{\psi}(k) e^{i(kx - \omega t)} dk$$

Se $|\hat{\psi}(k)|$ è una funzione con un massimo molto pronunciato in $k = k_0$, e sviluppiamo l'esponente attorno a k_0

$$kx - \omega t = k_0 x - \omega_0 t + \left(x - \frac{d\omega}{dk}_{|0} t\right)(k - k_0) + O((k - k_0)^2)$$

allora possiamo trascurare (in prima approssimazione) i termini quadratici in $k - k_0$, dopo di ché l'integrale diventa

$$\psi(x,t) = e^{i(k_0 x - \omega_0 t)} \int \hat{\psi}(k) \exp\left[i \left(x - \frac{d\omega}{dk}_{|0} t \right) (k - k_0) \right] dk$$

Quale che sia il valore dell'integrale, si vede che definisce una funzione della variabile

$$\xi = x - \frac{d\omega}{dk}_{|0} t$$

In altri termini

$$\psi(x,t) = e^{i(k_0 x - \omega_0 t)} \Phi(x - vt), \qquad v = \frac{d\omega}{dk}_{|0} \tag{4.8}$$

Le relazioni quantistiche $E = \hbar\omega$, $p = \hbar k$ implicano che

$$v = \frac{d\omega}{dk}_{|0} = \frac{dE}{dp}_{|0} = \frac{p_0}{m}$$

e quindi v è la velocità di una particella di massa m secondo la meccanica classica. Dalla (4.8) segue che il centro di $|\psi(x,t)|$ si propaga in effetti come una particella classica. Un calcolo più preciso mostra però che il pacchetto d'onde finisce per disperdersi, e dopo un certo tempo non è più possibile assimilarlo ad una particella.
Inoltre, le figure di diffrazione prodotte dalla diffusione degli elettroni sui cristalli rivelano una struttura discreta, se analizzate con maggior dettaglio. In particolare, diminuendo l'intensità del fascio si distinguono chiaramente le scintillazioni puntiformi sullo schermo, che si possono interpretare come gli impatti dei singoli elettroni.
Una descrizione puramente ondulatoria risulta dunque impossibile. Non si può neanche pensare che la funzione d'onda sia legata al numero di elettroni che arrivano sullo schermo per unità di tempo, perché le figure di diffrazione si possono produrre ripetendo molte volte l'esperimento con pochi elettroni per volta. Questo ci conduce ad una diversa interpretazione.

Statistica delle misure di posizione

Il caso della diffrazione di particelle discusso sopra è un tipico *test di natura statistica*: si ripete molte volte lo stesso esperimento, si misura ogni volta la stessa grandezza (la *posizione dell'evento di arrivo* dell'elettrone sullo schermo) e si fa la media sui risultati delle misure delle grandezze che interessano.

Ad esempio, la probabilità che l'evento di arrivo sia all'interno di un piccolo quadratino disegnato sullo schermo, è la media della quantità che assume il valore *uno* se l'evento cade nel quadratino, e *zero* se cade fuori.
La più naturale interpretazione fisica della funzione d'onda è dunque di natura statistica, e afferma che [8, 9, 10]

> **S1:** *la probabilità che una misura della posizione all'istante t nello stato descritto dalla funzione d'onda ψ, dia come risultato la terna di numeri reali (x, y, z), è proporzionale a $|\psi(x, y, z, t)|^2$.*
> *Più in generale, la probabilità che una misura delle posizioni all'istante t dia come risultato i $3n$ valori reali rappresentati dalle coordinate dei vettori $\vec{r}_k$, $k = 1, 2, \ldots, n$, è proporzionale a $|\psi(\vec{r}_1, \ldots, \vec{r}_n, t)|^2$.*

Questo richiede le condizioni

$$\int |\psi(\vec{r}, t)|^2 d^3\vec{r} = 1, \quad \int |\psi(\vec{r}_1, \ldots, \vec{r}_n, t)|^2 d^3r_1 \cdots d^3r_n = 1 \tag{4.9}$$

qualunque sia t, cosicché in entrambi i casi $|\psi|^2$ è più propriamente la probabilità per unità di volume dei valori di posizione, o densità di probabilità. È uno standard della letteratura quantistica indicare la radice quadrata degli integrali (4.9) con il simbolo $\|\psi\|$, che si legge *la norma di ψ*.
La probabilità di trovare i valori $(\vec{r}_1, \ldots, \vec{r}_n)$ nella regione Ω è dunque

$$P((\vec{r}_1, \ldots, \vec{r}_n) \in \Omega) = \int_\Omega |\psi(\vec{r}_1, \ldots, \vec{r}_n, t)|^2 d^3r_1 \cdots d^3r_n \tag{4.10}$$

Si noti anche che si deve supporre che le quantità (x, y, z) nel primo caso, o $\vec{r}_1, \ldots, \vec{r}_n$ nel secondo, siano simultaneamente misurabili, la qual cosa non sembra affatto ovvia.
Si noti però che queste affermazioni non si riferiscono alla *posizione dell'elettrone* nel senso classico del termine, ma ad una quantità più complicata, che è la *posizione dell'evento* del processo di interazione di quest'ultimo con lo schermo (nel ruolo di strumento di misura). Insomma, la posizione nello spazio non è un attributo delle particelle atomiche, ma del processo di interazione di queste con gli strumenti di misura. Bisognerebbe quindi pensare alla posizione come ad una *osservabile* che non ha un valore a priori, ma ne acquista uno in seguito a un processo di misura.

Statistica delle misure della quantità di moto

Torniamo al caso di una singola particella. Consideriamo la trasformata di Fourier-Plancherel

$$\widehat{\psi}(\vec{p}) = (2\pi\hbar)^{-3/2} \int \psi(\vec{r}) e^{-i\vec{p}\cdot\vec{r}/\hbar} d^3r \tag{4.11}$$

La corrispondenza $\psi \to \widehat{\psi}$ gode di due proprietà fondamentali: innanzitutto è invertibile, così se è nota $\widehat{\psi}$ si può ricostruire ψ con la formula di inversione

$$\psi(\vec{r}) = (2\pi\hbar)^{-3/2} \int \widehat{\psi}(\vec{p}) e^{i\vec{p}\cdot\vec{r}/\hbar} d^3p \tag{4.12}$$

E poi si ha la formula di Plancherel

$$\int \overline{\widehat{\phi}(\vec{p})}\, \widehat{\psi}(\vec{p}) d^3p = \int \overline{\phi(\vec{r})}\, \psi(\vec{r}) d^3r \tag{4.13}$$

In particolare

$$\int |\widehat{\psi}(\vec{p})|^2 d^3p = \int |\psi(\vec{r})|^2 d^3r$$

e quindi $\|\psi\| = 1$ se e solo se $\|\widehat{\psi}\| = 1$. In base a queste proprietà possiamo dire che

ψ e $\widehat{\psi}$ sono due descrizioni equivalenti dello stesso stato quantico.

Dunque anche $\widehat{\psi}$ ha una naturale interpretazione statistica, che però è naturalmente riferita alle variabili di impulso [9, 10]:

> **S2:** *la probabilità che una misura della quantità di moto all'istante t nello stato descritto dalla funzione d'onda ψ, o $\widehat{\psi}$, dia come risultato la terna di numeri reali (p_x, p_y, p_z), è proporzionale a $|\widehat{\psi}(p_x, p_y, p_z, t)|^2$.*
> *Più in generale, la probabilità che una misura degli impulsi all'istante t dia come risultato i $3n$ valori reali rappresentati dalle coordinate dei vettori $\vec{p}_k$, $k = 1, 2, \ldots, n$, è proporzionale a $|\widehat{\psi}(\vec{p}_1, \ldots, \vec{p}_n, t)|^2$.*

La trasformata di Fourier per più variabili è una semplice generalizzazione della formula (4.11)

$$\widehat{\psi}(\vec{p}_1, \ldots, \vec{p}_n, t) = \tag{4.14}$$
$$(2\pi\hbar)^{-3n/2} \int \psi(\vec{r}_1, \ldots, \vec{r}_n, t) \exp\left[-i(\vec{r}_1 \cdot \vec{p}_1 + \cdots + \vec{r}_n \cdot \vec{p}_n)/\hbar\right] d^3r_1 \cdots d^3r_n$$

Se $\|\psi\| = 1$ si ha

$$\|\widehat{\psi}\|^2 \equiv \int |\widehat{\psi}(\vec{p}_1, \ldots, \vec{p}_n, t)|^2 d^3p_1 \cdots d^3p_n = 1 \tag{4.15}$$

e dunque la funzione $|\widehat{\psi}(\vec{p}_1, \ldots, \vec{p}_n, t)|^2$ rappresenta più propriamente la densità di probabilità nello spazio degli impulsi. La probabilità di trovare i valori $(\vec{p}_1, \ldots, \vec{p}_n)$ nella regione $\widehat{\Omega}$ è

$$P((\vec{p}_1, \ldots, \vec{p}_n) \in \widehat{\Omega}) = \int_{\widehat{\Omega}} |\widehat{\psi}(\vec{p}_1, \ldots, \vec{p}_n, t)|^2 d^3p_1 \cdots d^3p_n \tag{4.16}$$

Si noti anche che si deve supporre che le quantità $\vec{p_1}, \ldots, \vec{p}_n$ siano simultaneamente misurabili, cosa che non è affatto ovvia.
Ripetiamo qui quanto si è detto per le misure di posizione. Le asserzioni statistiche non si riferiscono all'*impulso delle particelle* nel senso classico del termine, ma a quantità più complicate che si riferiscono agli *eventi* del processo di interazione del sistema fisico con gli strumenti di misura.
Insomma, la quantità di moto non andrebbe considerata come un attributo delle particelle atomiche, ma piuttosto come un attributo del processo di interazione di queste con gli strumenti di misura. Bisognerebbe quindi pensare alla quantità di moto come ad una *osservabile* che non ha un valore a priori, ma ne acquista uno in seguito a un processo di misura.
Poiché bisogna distinguere le osservabili dai loro valori, nel seguito indicheremo le quantità osservabili mediante simboli contrassegnati da una tilde. Così $\tilde{x}$ è l'osservabile che corrisponde alla componente x del vettore posizione, $\tilde{p}_x$ corrisponde alla componente x dell'impulso, $\widetilde{H}$ denota l'osservabile energia, et cetera.

4.5 Digressione matematica: lo spazio L^2

[Ref.] F. Riesz and B. Sz.-Nagy, *Functional analysis*, Frederick Ungar Publishing Co., NY (1972).

L'insieme di tutte le funzione d'onda a quadrato sommabile, tali cioè che l'integrale (4.9) (o l'integrale (4.15)) sia finito, è un ben noto spazio studiato dai matematici: lo spazio $L^2(R^n)$, dove n è il numero di variabili reali da cui dipende ψ. Per fissare le idee, indichiamo con $\|\psi\|$ la radice quadrata dell'integrale (4.9)

$$\|\psi\| = \left(\int |\psi(x_1, \ldots, x_n)|^2 d^n x \right)^{1/2}$$

che si chiama la norma di ψ. La dipendenza dal tempo non entra nella definizione dello spazio L^2, perciò nel seguito verrà omessa. Indicheremo con x le n variabili $(x_1, \ldots, x_n)$ (quindi $x \in R^n$) e useremo il simbolo dx per indicare la misura $d^n x$. Una barra sopra un simbolo indicherà la coniugazione complessa: così se $z = a + ib$ avremo $\bar{z} = a - ib$.
Valgono le proprietà

1. L^2 è uno spazio vettoriale complesso, ossia dati ψ e ϕ in L^2 e i numeri complessi a e b

$$a\psi + b\phi \in L^2$$

2. Nessun insieme finito $\{\phi_1, \ldots, \phi_m\}$ di elementi linearmente indipendenti genera tutto lo spazio L^2, cioè l'insieme degli elementi della forma $\sum_k a_k \phi_k$ è propriamente contenuto in L^2. Si descrive questa situazione dicendo che L^2 ha dimensione infinita.

   Gli elementi $\phi_1, \ldots, \phi_m$ si dicono linearmente indipendenti se l'equazione

$$\sum_{k=1}^{m} a_k \phi_k = 0$$

   è possibile solo quando $a_1 = a_2 = \cdots = a_m = 0$.

3. ad ogni coppia ϕ, ψ di elementi di L^2 è associato un numero complesso (ϕ, ψ), definito dall'integrale

$$(\phi, \psi) = \int \bar{\phi}(x)\psi(x)\, dx \tag{4.17}$$

   che si chiama *il prodotto scalare di ϕ e ψ*, e che ha le seguenti proprietà

   - $$(\phi, \psi) = \overline{(\phi, \psi)}$$

   - $$(\phi, a\psi_1 + b\psi_2) = a(\phi, \psi_1) + b(\phi, \psi_2)$$

   - $$(\phi, \phi) \geq 0$$

     ed è uguale a zero se e solo se $\phi = 0$.

- vale la diseguaglianza di Schwarz

$$|(\phi, \psi)| \leq \|\phi\| \|\psi\| \tag{4.18}$$

Il segno di uguaglianza vale se e solo se ϕ e ψ sono proporzionali.

Se $(\phi, \psi) = 0$ si dice che ϕ e ψ sono mutualmente ortogonali.

4. L^2 è uno spazio metrico completo: ogni sequenza infinita di elementi ϕ_n con la proprietà $\|\phi_n - \phi_m\| \to 0$ se $n, m \to \infty$, converge in L^2. Cioè esiste un unico $\phi \in L^2$ tale che $\|\phi - \phi_n\| \to 0$ per $n \to \infty$.

5. L^2 è uno spazio separabile, cioè ogni ϕ in L^2 è il limite di una sequenza *numerabile* di elementi $\phi_n \in L^2$, $n = 0, 1, \ldots$.

Sistemi ortonormali

Un sistema ortonormale è una sequenza di elementi ϕ_1, ϕ_2,..., in L^2, con le due proprietà

$$\forall n, \quad \|\phi_n\| = 1, \quad (\phi_n, \phi_m) = 0 \quad \text{se} \quad n \neq m$$

Il sistema ortonormale si dice *completo* se ogni ϕ in L^2 si può scrivere come serie

$$\phi = \sum_{\nu=1}^{\infty} c_\nu \phi_\nu$$

nel senso che

$$\|\phi - \sum_{\nu=1}^{N} c_\nu \phi_\nu\| \xrightarrow[N \to \infty]{} 0$$

In questo caso $c_\nu = (\phi_\nu, \phi)$ e vale il teorema di Parseval

$$\|\phi\|^2 = \sum_{\nu=1}^{\infty} |(\phi_\nu, \phi)|^2 \tag{4.19}$$

Spazi di Hilbert

Uno spazio metrico completo dotato di un prodotto scalare si chiama uno *spazio di Hilbert*. Se è anche separabile, allora si può mostrare che possiede un sistema ortonormale completo. Possiamo allora dire:

> Lo spazio degli stati di un sistema quantico è uno spazio di Hilbert separabile. I suoi elementi sono le funzioni complesse a quadrato sommabile sullo spazio delle configurazioni del sistema classico corrispondente; equivalentemente, sono le funzioni a quadrato sommabile sullo spazio degli impulsi.

Un'ovvia generalizzazione che si può menzionare è che le funzioni d'onda possono avere valori vettoriali, aventi un certo numero finito di componenti. In effetti, in meccanica quantistica queste funzioni d'onda sono necessarie per descrivere gli stati delle particelle dotate di *spin*, cioè di momento angolare intrinseco. La proiezione dello spin lungo un'asse qualunque è sempre quantizzato in multipli interi di $\hbar j$, dove i valori possibili per j sono $j = 0, 1/2, 1, 3/2, \ldots$. Se questo è il caso, si parla di particelle con "spin j". La funzione d'onda di una particella di spin j è precisamente un vettore con $2j + 1$ componenti.

4.6 Valori di aspettazione

Le medie dei risultati di molte misure di posizione o impulso, detti *valori di aspettazione*, seguono subito dalle asserzioni statistiche **S1**, **S2**. Se f e g sono funzioni reali di tre variabili reali, definiamo innanzitutto le osservabili $\widetilde{F} = f(\tilde{x}, \tilde{y}, \tilde{z})$ e $\widetilde{G} = g(\tilde{p}_x, \tilde{p}_y, \tilde{p}_z)$. Si misurano $\tilde{x}$, $\tilde{y}$ e $\tilde{z}$: se il risultato è la terna di numeri (x, y, z) il valore di $\widetilde{F}$ è $f(x, y, z)$. La definizione di $\widetilde{G}$ è del tutto analoga e non sarà ripetuta. Allora i valori di aspettazione sono

$$< \widetilde{F} >= \int f(x, y, z)|\psi(x, y, z)|^2 dxdydz \tag{4.20}$$

$$< \widetilde{G} >= \int g(p_x, p_y, p_z)|\widehat{\psi}(p_x, p_y, p_z)|^2 dp_x dp_y dp_z \tag{4.21}$$

Infatti i valori medi di una variabile casuale si ottengono facendo una lista dei suoi valori possibili moltiplicati per le rispettive probabilità, e sommando poi tutti i termini della lista. Nel caso di valori che variano con continuità si dovrà integrare sulla lista dei valori. Quindi le formule dei valori medi contengono un'idea intuitiva ma forse non ovvia: che i valori possibili di una coordinata di posizione o di impulso siano *tutti* i numeri reali. La generalizzazione a più particelle è immediata.
Vogliamo ora riscrivere la (4.21) in una forma che contiene solo la funzione d'onda ψ (si potrebbe anche mettere la (4.20) in una forma che contenga solo $\widehat{\psi}$). A questo scopo notiamo che, ad esempio

$$
\begin{aligned}
p_x\widehat{\psi}(\vec{p}) &= i\hbar\int\psi(\vec{r})\,\frac{\partial}{\partial x}\left(\exp(-i\vec{p}\cdot\vec{r}/\hbar)\right)d^3r \\
&= -i\hbar\int\frac{\partial\psi(\vec{r})}{\partial x}\exp(-i\vec{p}\cdot\vec{r}/\hbar)\,d^3r
\end{aligned}
\tag{4.22}
$$

come si vede dopo una integrazione per parti. Questa formula ci dice che se $\psi(\vec{r})$ è la trasformata di $\widehat{\psi}(\vec{p})$, quella di $p_x\widehat{\psi}(\vec{p})$ è $-i\hbar\partial\psi/\partial x$. Dall'Eq. (4.13) applicata a $p_x\widehat{\psi}(\vec{p})$ e $\widehat{\psi}(\vec{p})$, segue allora la formula importante

$$
<\tilde{p}_x>=\int\bar{\psi}(\vec{r})\left(-i\hbar\frac{\partial\psi(\vec{r})}{\partial x}\right)d^3r \tag{4.23}
$$

Non c'è nulla di speciale in p_x che non valga anche per p_y o p_z, quindi risulta ovvio che

$$
<\tilde{p}_y>=\int\bar{\psi}(\vec{r})\left(-i\hbar\frac{\partial\psi(\vec{r})}{\partial y}\right)d^3r \tag{4.24}
$$

$$
<\tilde{p}_z>=\int\bar{\psi}(\vec{r})\left(-i\hbar\frac{\partial\psi(\vec{r})}{\partial z}\right)d^3r \tag{4.25}
$$

Il caso più generale si ottiene sostituendo agli argomenti p_x, p_y e p_z di g gli operatori differenziali[11] $-i\hbar\partial/\partial x$, $-i\hbar\partial/\partial y$ e $-i\hbar\partial/\partial z$, dopo di ché la (4.21) diventa

$$
<\widetilde{G}>=\int\bar{\psi}(\vec{r})g\left(-i\hbar\frac{\partial}{\partial x},-i\hbar\frac{\partial}{\partial y},-i\hbar\frac{\partial}{\partial z}\right)\psi(\vec{r})\,d^3r \tag{4.26}
$$

Non si ha difficoltà a definire $g(-i\hbar\partial/\partial x,-i\hbar\partial/\partial y,-i\hbar\partial/\partial z)$ se g è un polinomio. Ad esempio, se $g=p_xp_yp_z$ si avrà

$$
g\left(-i\hbar\frac{\partial}{\partial x},-i\hbar\frac{\partial\psi}{\partial y},-i\hbar\frac{\partial}{\partial z}\right)\psi=-i^3\hbar^3\frac{\partial^3\psi}{\partial x\partial y\partial z}
$$

Si dimostra che $g(-i\hbar\partial/\partial x,-i\hbar\partial/\partial y,-i\hbar\partial/\partial z)$ può essere definita per una vasta classe di funzioni, che include tra l'altro le funzioni continue a tratti.
Viceversa, dai valori di aspettazione possiamo calcolare le probabilità. Infatti, se nella formula (4.20) f è la funzione caratteristica di una regione Ω, cioè

[11] Gli oggetti che seguono si chiamano operatori perché "operano" sulle funzioni trasformandole nelle loro derivate.

se $f(\vec{r}) = f_\Omega(\vec{r}) = 1$ se $\vec{r} \in \Omega$ e $f_\Omega(\vec{r}) = 0$ se $\vec{r} \notin \Omega$, il valore di aspettazione dell'osservabile $f_\Omega(\tilde{x}, \tilde{y}, \tilde{z})$ è la probabilità dell'evento $\vec{r} \in \Omega$, i.e.

$$\begin{aligned} P((x,y,z) \in \Omega)) &= < f_\Omega(\tilde{x}, \tilde{y}, \tilde{z}) > = \int f_\Omega(x,y,z) |\psi(x,y,z)|^2 d^3r \\ &= \int_\Omega |\psi(x,y,z)|^2 dxdydz \end{aligned} \tag{4.27}$$

e si ritrova la formula (4.10). Se è noto lo stato, si può anche calcolare la distribuzione statistica dei valori di una sola coordinata, ad esempio x. A questo scopo sia $e_I(\lambda)$ la funzione caratteristica di un intervallo I, $e_I(\lambda) = 1$ se $\lambda \in I$, $e_I(\lambda) = 0$ se $\lambda \notin I$. Possiamo allora considerare l'osservabile $e_I(\tilde{x})$, il cui valore è 1 se $x \in I$, 0 se $x \notin I$; il suo valore di aspettazione è la probabilità che $x \in I$, y e z avendo valori arbitrari. Vediamo che la formula (4.20) è tutto ciò che serve per rispondere a tutte le domande sull'osservabile di posizione.
Una situazione del tutto analoga si ha per la distribuzione statistica dei valori della quantità di moto. Ad esempio, possiamo scrivere per la probabilità che $p_x \in J$, la formula

$$P(p_x \in J) =< e_J(\tilde{p}_x) >= \int e_J(p_x) |\widehat{\psi}(p_x, p_y, p_z)|^2 dp_x dp_y dp_z \tag{4.28}$$

oppure (vedi la (4.26))

$$P(p_x \in J) = \int \bar{\psi}(x,y,z) e_J \left(-i\hbar \frac{\partial}{\partial x} \right) \psi(x,y,z)\, dxdydz \tag{4.29}$$

Come si è detto sopra, si dimostra che $e_J\left(-i\hbar\partial/\partial x\right)$ esiste.

4.7 Digressione matematica: operatori lineari

Sia $\mathcal{H}$ uno spazio di Hilbert separabile; gli elementi di $\mathcal{H}$ si chiamano *vettori*; un'applicazione

$$A : \mathcal{H} \to \mathcal{H}$$

che trasforma ϕ in $A\phi$ si chiama un *operatore lineare* se

$$A(a\phi + b\psi) = aA\psi + bA\psi, \quad \psi, \phi \in \mathcal{H}$$

dove a e b sono numeri complessi arbitrari; $A\phi$ si legge "A applicato a ϕ". Se esiste una costante C tale che

$$\|A\phi\| \leq C\|\phi\|$$

A si dice limitato[12]. Se $A\phi_n \to A\phi$ segue da $\phi_n \to \phi$ si dice che A è un operatore continuo[13]. Si ha il teorema [11]: *un A lineare è continuo se e solo se è limitato.*

Esempi

In connessione con la posizione e la quantità di moto, le formule (4.20) e (4.23)-(4.26) suggeriscono di introdurre gli operatori lineari in $L^2(R^3)$

$$\hat{x}: \ \phi(x,y,z) \to x\phi(x,y,z), \quad \hat{p_x}: \ \phi(x,y,z) \to -i\hbar\frac{\partial\phi(x,y,z)}{\partial x}$$

$$\hat{y}: \ \phi(x,y,z) \to y\phi(x,y,z), \quad \hat{p_y}: \ \phi(x,y,z) \to -i\hbar\frac{\partial\phi(x,y,z)}{\partial y}$$

$$\hat{z}: \ \phi(x,y,z) \to z\phi(x,y,z), \quad \hat{p_z}: \ \phi(x,y,z) \to -i\hbar\frac{\partial\phi(x,y,z)}{\partial z}$$

Questi operatori non sono limitati, ad esempio la norma $\|\hat{x}\phi\|$ può essere arbitrariamente grande anche se $\|\phi\| = 1$, e dunque non sono continui. Questi operatori non sono neanche definiti per tutti gli elementi di L^2: ad esempio non è detto che $x\phi(x)$ sia una funzione a quadrato sommabile, neanche se $\phi(x)$ lo è, e non è detto che $d\phi(x)/dx$ esista.

Operatori hermitiani

Nel considerare operatori non limitati si suppone però che il loro dominio di definizione, l'insieme degli elementi $\phi \in \mathcal{H}$ tali che $A\phi$ è ben definito, sia *denso* in $\mathcal{H}$, cioè che ogni elemento di $\mathcal{H}$ sia il limite di una successione di elementi del dominio di A. Allora si può definire l'aggiunto di A, indicato con A^*, come l'operatore tale che (si ricordi che $\mathcal{H}$ possiede un prodotto scalare per definizione)

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

[12]Ricordiamo che $\|\phi\|^2 = (\phi, \phi)$, dove (,) denota il prodotto scalare.

[13]Ricordiamo che $A\phi_n \to A\phi$ significa che $\|A\phi_n - A\phi\| \to 0$ per $n \to \infty$. Stesso significato per $\phi_n \to \phi$.

per ogni ϕ nel dominio $\mathcal{D}_A$ di A. Il dominio $\mathcal{D}_{A^*}$ di A^* può anche essere vuoto, o consistere della totalità di $\mathcal{H}$: in ogni caso, se $\mathcal{D}_{A^*} = \mathcal{D}_A$ e sul dominio comune $A = A^*$, l'operatore A si dice *autoaggiunto* o *hermitiano*. In tal caso

$$(A\phi, \psi) = (\phi, A\psi) = \overline{(A\psi, \phi)}$$

Vediamo che per gli operatori hermitiani $(A\phi, \phi)$ è un numero reale (in generale è complesso). È facile vedere che gli operatori di posizione $\hat{x}$, $\hat{y}$ e $\hat{z}$ sono hermitiani. Per una particella in una dimensione si avrebbe

$$(\phi, \hat{x}\psi) = \int \bar{\phi}(x) x \psi(x) dx = \int \overline{x\phi(x)} \psi(x) dx = (\hat{x}\phi, \psi)$$

Si ha poi

$$(\phi, \hat{p_x}\psi) = \int \bar{\phi}(x) \left(-i\hbar \frac{\partial \psi(x)}{\partial x} \right) dx = \int \overline{\left(-i\hbar \frac{\partial \phi(x)}{\partial x} \right)} \psi(x) dx = (\hat{p_x}\phi, \psi)$$

dove la seconda uguaglianza nasce da un'integrazione per parti, che è pemessa perché le funzioni d'onda tendono a zero all'infinito. Vediamo che anche le componenti della quantità di moto sono rappresentate da operatori hermitiani.

Autovalori e autovettori

Per gli operatori hermitiani, e per tutta la struttura della MQ, l'*equazione agli autovalori*

$$A\phi = \lambda\phi, \quad \lambda \in R, \quad A \text{ hermitiano} \tag{4.30}$$

è estremamente importante. I numeri λ per i quali la (4.30) ammette soluzioni si chiamano gli *autovalori* dell'operatore A e i vettori ϕ corrispondenti si chiamano *autovettori*. Si vede subito che gli autovalori devono essere numeri reali. Infatti $(\phi, A\phi) = \lambda(\phi, \phi)$ è reale per A hermitiano.
Date due soluzioni, $Af_1 = \lambda_1 f_2$ e $Af_2 = \lambda_2 f_2$, si ha

$$(\lambda_1 - \lambda_2)(f_2, f_1) = (f_2, Af_1) - (Af_2, f_1) = 0$$

perché A è supposto hermitiano. Quindi se i due autovalori sono distinti si deve avere $(f_1, f_2) = 0$, ed i corrispondenti autovettori sono ortogonali.
L'insieme degli autovalori di A si chiama *lo spettro dell'operatore* A; se questo consiste di un insieme numerabile di punti λ_1, λ_2, ..., si dice che lo spettro

è puramente *discreto*, o puntuale. È un fatto caratteristico degli operatori su spazi di Hilbert infinito dimensionali che lo spettro può anche avere una componente continua, o essere interamente continuo. Nel caso dello spettro discreto (può servire da esempio il caso dell'operatore

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

nel segmento $[0, L]$ della buca di potenziale infinita, studiato nella Sez. (4.2)), possono esistere più autovettori per ciascun autovalore. In questo caso si dice che l'autovalore è *degenere* e il sottospazio $\mathcal{M}_\nu$ generato dagli autovettori può anche avere dimensione infinita.
Un teorema fondamentale della teoria degli operatori sugli spazi di Hilbert garantisce che se lo spettro è discreto con autovalori λ_ν, allora gli autovettori ϕ_ν formano un sistema ortonormale completo in $\mathcal{H}$: cioè ogni $\phi \in \mathcal{H}$ ha lo sviluppo

$$\phi = \sum_\nu (\phi_\nu, \phi)\, \phi_\nu$$

eccetto che per gli autovalori degeneri la scelta delle autofunzioni in $\mathcal{M}_\nu$ non è univoca. In ogni caso si ha

$$A\phi = \sum_\nu \lambda_\nu (\phi_\nu, \phi)\, \phi_\nu$$

$$\|\phi\|^2 = \sum_\nu |(\phi_\nu, \phi)|^2$$

Operatori unitari

Dati due spazi di Hilbert $\mathcal{H}_1$ e $\mathcal{H}_2$, un operatore $U : \mathcal{H}_1 \to \mathcal{H}_2$ si dice *unitario* se U è isometrico e invertibile, vale a dire se

$$(U\phi, U\psi)_1 = (\phi, \psi)_2$$

dove $(\ ,\ )_a$, $a = 1, 2$, denota il prodotto scalare in $\mathcal{H}_a$. Si ha allora

$$U^*U = UU^* = 1, \quad U^{-1} = U^*$$

Se un tale operatore esiste si dice che $\mathcal{H}_1$ e $\mathcal{H}_2$ sono isomorfi. Dato che gli operatori unitari conservano il prodotto scalare, e dato che le predizioni statistiche della meccanica quantistica si mettono sempre nella forma di prodotti scalari (si veda la Sez. [5]), due spazi di Hilbert isomorfi danno luogo alla stessa descrizione fisica; si hanno dunque tante rappresentazioni matematiche fisicamente equivalenti della stessa teoria, inizialmente descritta in uno spazio di Hilbert $\mathcal{H}$, quanti sono gli spazi di Hilbert unitariamente isomorfi ad $\mathcal{H}$.

Operatori commutanti

Dati due operatori hermitiani A e B si definisce *commutatore* di A e B l'operatore

$$[A, B] = AB - BA$$

Se $[A, B] = 0$, un teorema di von Neumann [12] afferma che A e B sono funzioni di un operatore hermitiano C, ossia esistono funzioni F e G tali che $A = F(C)$ e $B = G(C)$. Dire che $A = F(C)$ ha il seguente significato: se C ha spettro discreto con autovalori λ_k e autovettori ϕ_k, si ha

$$A\phi_k = F(\lambda_k)\phi_k, \quad B\phi_k = G(\lambda_k)\phi_k$$

Questo definisce $F(A)$ e $G(B)$ per ogni ϕ perché gli autovettori formano un sistema ortonormale completo. Se lo spettro è continuo esiste una formula (dovuta a von Neumann) che generalizza la situazione appena descritta, ma che non trattiamo in questo corso per non appesantire troppo il testo.
Vediamo che gli operatori commutanti hanno un sistema completo ortonormale di autovettori in comune o, come si dice, sono *simultaneamente diagonalizzabili*. I teoremi enunciati valgono anche per un numero arbitrario di operatori reciprocamente commutanti.
Esempi di operatori commutanti sono gli operatori di posizione

$$[\hat{x}, \hat{y}] = [\hat{x}, \hat{z}] = [\hat{y}, \hat{z}] = 0$$

e di impulso

$$[\hat{p}_x, \hat{p}_y] = [\hat{p}_x, \hat{p}_z] = [\hat{p}_y, \hat{p}_z] = 0$$

Infatti, ad esempio, $\forall\psi$

$$\hat{x}\hat{y}\,\psi(x, y, z) = xy\,\psi(x, y, z) = yx\,\psi(x, y, z) = \hat{y}\hat{x}\,\psi(x, y, z)$$

e dunque $\hat{x}\hat{y} = \hat{y}\hat{x}$. Al contrario si ha

$$\hat{x}\hat{p}_x\psi = -i\hbar x\frac{\partial\psi}{\partial x} = -i\hbar\frac{\partial x\psi}{\partial x} + i\hbar\psi = \hat{p}_x\hat{x}\psi + i\hbar\psi$$

e quindi si ha

$$[\hat{x}, \hat{p}_x] = i\hbar \tag{4.31}$$

Questa è una delle più famose relazioni di commutazione della MQ. È ovvio che le coppie di operatori $\hat{y}$, $\hat{p}_y$ e $\hat{z}$, $\hat{p}_z$ soddisfano relazioni identiche.

Abbiamo appena accennato al fatto che è possibile definire le funzioni di un operatore A: si può definire la corrispondenza $u(\lambda) \to u(A)$, $u(\lambda)$ una funzione reale limitata, in modo tale da avere le seguenti proprietà naturali [11, 12]

1. se $u(\lambda)$ è una costante c, $u(A) = cI$ (I è l'operatore identità, $I\psi = \psi$).
2. la corrispondenza è lineare e omogenea: $(u_1 + u_2)(A) = u_1(A) + u_2(A)$ e $(cu)(A) = cu(A)$
3. la corrispondenza è moltiplicativa: se $u(\lambda) = u_1(\lambda)u_2(\lambda)$ si ha $u(A) = u_1(A)u_2(A)$. Questo mostra anche che tutte le funzioni di A commutano reciprocamente.
4. $u(A)$ è hermitiano: $(u(A)\phi, \psi) = (\phi, u(A)\psi)$
5. e di tipo positivo: su $u(\lambda) \geq 0$ allora $(u(A)\phi, \phi) \geq 0$ per ogni ϕ.

Queste proprietà si mostrano facilmente nel caso in cui A abbia spettro puramente discreto. Se lo spettro è continuo si veda la letteratura citata (o altra).
Un caso particolarmente importante di funzione di operatore è l'esponenziale

$$U(\lambda) = e^{-i\lambda A}, \quad A \text{ hermitiano}$$

Si noti la proprietà di gruppo, $U(\lambda)U(\mu) = U(\lambda + \mu)$, e l'unitarietà, $U(\lambda)^* = U(\lambda)^{-1}$. Formalmente si ha

$$i\frac{dU(\lambda)}{d\lambda} = AU(\lambda), \quad U(0) = 1 \tag{4.32}$$

ma si può dimostrare che la derivata esiste ed è uguale a quella scritta.

4.8 L'evoluzione temporale

Nella Sez. (4.1) abbiamo incontrato l'equazione di Schrödinger

$$i\hbar\frac{\partial\psi}{\partial t} = H\,\psi \tag{4.33}$$

dove H è l'operatore che corrisponde all'energia, ad esempio l'operatore (4.6). La funzione esponenziale appena descritta e l'equazione (4.32) mostrano subito che la soluzione dell'equazione di Schrödinger ha la forma

$$\psi(t) = U(t-t_0)\psi(t_0), \qquad U(t-t_0) = \exp\left(-iH(t-t_0)/\hbar\right) \tag{4.34}$$

dove $\psi(t_0)$ è la funzione d'onda dello stato iniziale. Dunque l'equazione determina in modo deterministico lo stato al tempo t se è dato lo stato a un tempo iniziale t_0: si ottiene lo stato all'istante t applicando l'operatore unitario $U(t-t_0)$ allo stato iniziale $\psi(t_0)$. In altri termini, l'equazione di Schrödinger determina la dinamica del sistema. Si vede anche che è l'operatore che rappresenta l'energia quello che "genera" l'evoluzione temporale, come avviene del resto in meccanica classica quando viene scritta nel formalismo Hamiltoniano.

5 Il formalismo generale

I risultati espressi dalle equazioni (4.20), (4.21), (4.23)-(4.24), (4.27)-(4.29), contenenti le principali asserzioni statistiche della MQ, hanno tutti la forma $(\psi, A\psi)$ per opportuni operatori hermitiani A, e stati normalizzati ψ. Ad esempio[14], con $\|\psi\| = 1$

$$< \tilde{x} >= (\psi, \hat{x}\psi)$$

$$< \tilde{p}_x >= \left(\psi, \frac{\hbar}{i}\frac{\partial\psi}{\partial x}\right) = (\psi, \hat{p}_x\psi)$$

Se $\|\psi\| \neq 1$, basta sostituire $\psi/\|\psi\|$ a ψ in tutte le formule: ad esempio il valore di aspettazione di $\tilde{x}$ in uno stato non normalizzato ψ è

$$< \tilde{x} >= (\psi, \hat{x}\psi)\|\psi\|^{-2} = \frac{(\psi, \hat{x}\psi)}{(\psi, \psi)}$$

Vediamo che gli stati ψ e $c\psi$, dove c è una costante, rappresentano dopotutto lo stesso stato. Ora in MQ si assume che la forma dei risultati trovati per la posizione e l'impulso valgano per tutte le quantità osservabili.

5.1 I postulati

Arriviamo in tal modo alla seguente generalizzazione, che forma la base assiomatica della meccanica quantistica.

[14] $\hat{x}$ e $\hat{p}_x$ sono stati definiti nelle sezione precedente, e sono operatori hermitiani non limitati.

P1. Lo spazio degli stati di un sistema quantico è uno spazio di Hilbert separabile $\mathcal{H}$. Se un vettore ϕ rappresenta uno stato, i suoi multipli complessi $a\phi$ rappresentano lo stesso stato.

P2. Ad ogni quantità fisica misurabile $\tilde{A}$, corrisponde nella teoria un operatore lineare hermitiano $A : \mathcal{H} \to \mathcal{H}$.

P3. Se si misura l'osservabile $\tilde{A}$ sullo stato rappresentato dal vettore ϕ, il valore di aspettazione dei risultati è dato da

$$< \tilde{A} >= (\phi, A\phi)\,, \quad \|\phi\| = 1$$

P4. Se $\tilde{A}$ è rappresentata dall'operatore A e $f(\lambda)$ è una funzione reale di una variabile reale, l'osservabile $f(\tilde{A})$ è rappresentata dall'operatore $f(A)$.

L'osservabile $f(\tilde{A})$ è così definita: si misura $\tilde{A}$; se il risultato è a il risultato di $f(\tilde{A})$ è $f(a)$. Se $f_I(\lambda)$ denota la funzione caratteristica dell'intervallo I, denotiamo l'operatore $f_I(A)$ che corrisponde all'osservabile $f_I(\tilde{A})$ con $E(I)$. Il valore di aspettazione di questo operatore in uno stato ψ

$$(E(I)\psi, \psi) = P(a \in I)$$

rappresenta la probabilità che in una misura di $\tilde{A}$ il risultato appartenga a I.

dispersione

I risultati di misure ripetute molte molte sullo stesso stato sono in generali "dispersi" attorno al valore di aspettazione. La *dispersione* dei risultati è definita dalla formula

$$\Delta A = \Big(< (\tilde{A} - < \tilde{A} >)^2 > \Big)^{1/2} = (\psi, (A - < \tilde{A} >)^2 \psi)^{1/2} \tag{5.1}$$

Sviluppando il quadrato si vede facilmente che

$$\Delta A = \Big(< \tilde{A}^2 > - < \tilde{A} >^2 \Big)^{1/2} \tag{5.2}$$

Inoltre $\Delta A \geq 0$.
Ci chiediamo adesso in quali stati la dispersione di un'osservabile è zero, cioè in quali stati l'osservabile $\tilde{A}$ ha un valore certo λ_0 (i.e. che si presenta con

probabilità uno) . Chiaramente si ha$< \tilde{A} >= \lambda_0$ e dunque la dispersione si annulla se e solo se

$$(\psi, (A - \lambda_0)^2 \psi) = ((A - \lambda_0)\psi, (A - \lambda_0)\psi) = \|(A - \lambda_0)\psi\|^2 = 0$$

e $A\psi = \lambda_0 \psi$. Un caso particolare era l'equazione di Schrödinger per gli stati stazionari. Vediamo che l'equazione agli autovalori ha un diretto significato fisico:

> *Gli autovettori di un operatore A sono gli stati quantici tali che una misura di $\tilde{A}$ dà con certezza l'autovalore corrispondente.*

5.2 Statistica delle misure

Si misura una osservabile $\tilde{A}$ nello stato ϕ:

1. quali sono i possibili valori del risultato ?
2. con quale probabilità si presentano ?

Assumiamo per semplicità che lo spettro di A sia discreto e non degenere, con autovettori ϕ_ν e autovalori λ_ν. Valgono le formule (v. la Sez. (4.7))

$$\phi = \sum_\nu (\phi_\nu, \phi)\, \phi_\nu$$

$$A\phi = \sum_\nu \lambda_\nu (\phi_\nu, \phi)\, \phi_\nu$$

Da **P3.** segue, con $c_\nu = (\phi_\nu, \phi)$,

$$< \tilde{A} >= (\phi, A\phi) = \sum_{\nu,\mu} (c_\nu \phi_\nu, c_\mu \lambda_\mu \phi_\mu) = \sum_{\nu,\mu} \bar{c}_\nu c_\mu \lambda_\mu (\phi_\nu, \phi_\mu)$$

sappiamo che $(\phi_\nu, \phi_\mu) = 0$ se $\nu \neq \mu$ e $\|\phi_\nu\| = 1$. Quindi

$$< \tilde{A} >= \sum_\nu \lambda_\nu |c_\nu|^2; \qquad \sum_\nu |c_\nu|^2 = 1 \tag{5.3}$$

Poco sopra abbiamo ottenuto il significato fisico degli autovettori di un'osservabile. Ora vediamo che

1. i soli valori che $\tilde{A}$ può prendere in un processo di misura sono gli autovalori del suo operatore A

2. la probabilità che il valore di $\tilde{A}$ sia l'autovalore λ_ν, se si misura sullo stato ϕ, è

$$P(\lambda_\nu) = |c_\nu|^2 = |(\phi_\nu, \phi)|^2 \tag{5.4}$$

Nel caso più generale in cui lo spettro è continuo, si ottiene un enunciato molto simile [9, 10, 12]

1. i soli valori che $\tilde{A}$ può prendere in un processo di misura appartengono allo spettro del suo operatore A

2. la probabilità che il valore a di $\tilde{A}$ sia nell'intervallo J del suo spettro, se si misura sullo stato ϕ, è

$$P(a \in J) = (E(J)\phi, \phi) \tag{5.5}$$

dove $E(J)$ è l'operatore di proiezione[15] introdotto sopra. Le formule (5.4) e (5.5) sembrano diverse, ma in realtà non lo sono, nel senso che anche la (5.4) si può scrivere come il valore medio di un opportuno operatore di proiezione. È stato mostrato da von Neumann che i postulati **P3.** e **P4.** sono matematicamente equivalenti all'unica asserzione

P5. La probabilità che in una misura sullo stato ϕ le osservabili $\tilde{A}_1, \ldots, \tilde{A}_N$, con operatori $A_1, \ldots, A_N$, prendano i valori nei rispettivi intervalli $J_1, \ldots, J_N$, è data dalla formula

$$P(a_i \in J_i) = (E(J_1) \cdots E(J_N)\phi, \phi) \tag{5.6}$$

dove si assume che $\|\phi\| = 1$.

La (5.6) include anche lo spettro discreto e rappresenta la più generale asserzione statistica possibile in meccanica quantistica, cioè in una teoria che include i postulati **P1.-P4.** L'asserzione ha due caratteristiche essenziali:

1. è statistica, e non causale, cioè non ci dice quali sono i valori delle osservabili, ma solo con quale probabilità si presentano

[15] Significa che $E(J)E(J) = E(J)$.

2. può essere enunciata solo per osservabili con operatori commutanti:

$$[A_i, A_k] = 0, \qquad i, k = 1, \ldots, N$$

e niente si può dire, in generale, sulla statistica comune di osservabili con operatori non commutanti.

Infatti, è evidente che l'ordine delle osservabili sia nella proposizione che nel primo membro della (5.6) è irrilevante, cosicché deve esserlo anche nel secondo membro: tutti gli operatori $E_k(J_k)$ devono commutare tra di loro, e ciò è possibile solo se commutano tra loro gli operatori A_k stessi. Ad esempio la questione "si misurano $\tilde{x}$ e $\tilde{p}_x$ nello stato ϕ: qual'è la probabilità che il risultato $x \in I$ e il risultato $p_x \in J$, dove I e J sono intervalli?" non ha significato in meccanica quantistica, perché $\hat{x}$ e $\hat{p}_x$ non commutano, mentre è perfettamente legittima in meccanica classica. Vogliamo discutere il significato fisico di questo fatto.

5.3 Processi di misura

Sia $\tilde{A}$ un'osservabile il cui operatore A possiede spettro discreto e non degenere. Supponiamo di misurare $\tilde{A}$ e di trovare il valore a^*. Ci chiediamo quale sia lo stato del sistema dopo la misura.
Osserviamo con von Neumann che nell'esperimento di Compton e Simon sulla diffusione dei fotoni da elettroni, sia il fotone che l'elettrone diffusi possono essere usati per determinare la posizione della collisione, oppure il momento trasferito all'elettrone. L'elettrone arriva al detector leggermente in ritardo rispetto al fotone (10^{-12} sec potrebbe essere un valore realistico), ma l'osservazione dell'uno o dell'altro conducono allo stesso valore delle quantità osservate. Generalizzando questo fatto, possiamo dire:

> *due misure della posizione, o di una componente della quantità di moto, eseguite in rapida successione danno lo stesso risultato.*

Quindi l'inferenza è che lo stato iniziale viene trasformato dalla misura in uno stato in cui $\tilde{A}$ ha il valore trovato a^* con probabilità uno, cioè dopo la misura il sistema viene a trovarsi in uno stato in cui a^* è certo. Questo stato dipende non soltanto dallo strumento di misura e da come viene usato, ma anche dal risultato della misura, perché il valore di $\tilde{A}$ in questo stato deve ancora essere uguale ad a^*.
Abbiamo appreso che in tal caso a^* deve essere un autovalore di A, e lo stato il corrispondente autovettore. Inoltre lo stato è univocamente determinato dalla misura di $\tilde{A}$, perché abbiamo supposto che gli autovalori sono non

degeneri. Se a^* è un autovalore degenere, lo stato è determinato in maniera incompleta, e si può solo dire che appartiene allo spazio generato dagli autovettori di A che hanno l'autovalore a^*. Abbiamo anche visto che la probabilità di a^* è $|(\phi, \phi^*)|^2$, se ϕ è lo stato iniziale e ϕ^* l'autovettore di A con autovalore a^*, nel caso non degenere. L'assunzione che il processo di misura riduce, o collassa, lo stato in un autostato dell'osservabile misurata si chiama *il postulato di riduzione* di von Neumann.
Per tornare al postulato **P5**, è chiaro che se gli operatori A_j commutano reciprocamente essi rappresentano osservabili misurabili contemporaneamente. Infatti per opportune funzioni f_j si ha $A_j = f_j(B)$, e quindi una misura dell'osservabile $\tilde{B}$ corrispondente a B è anche una misura simultanea di tutte le osservabili $\tilde{A}_j$. Si noti però che si deve assumere non solo che ad ogni osservabile corrisponde un operatore hermitiano, ma anche che ogni operatore hermitiano è l'operatore di un'osservabile, e ci sono situazioni in cui questo non è vero.
Ma supponiamo che $\tilde{A}$ e $\tilde{B}$ siano simultaneamente misurabili: cosa segue da questo ?
Come al solito supponiamo gli spettri discreti e non degeneri. Dopo la misura si ha uno stato in cui $\tilde{A}$ e $\tilde{B}$ hanno i valori misurati con certezza. Siano a^* e b^* questi valori, siano ϕ_n gli autovettori appartenenti ad A e ψ_m quelli appartenenti a B. Sia ϕ_* lo stato dopo la misura: allora

$$A\phi_* = a^*\phi_*, \qquad B\phi_* = b^*\psi_*$$

e quindi ϕ_* è un ϕ_n e un ψ_m allo stesso tempo, cioè ϕ_n è un ψ_m. Se lo stato era ϕ, la probabilità di a^* è $|(\phi, \phi_n)|^2$, e quindi se ϕ era un $\phi_{\bar{n}}$ il valore $a_{\bar{n}}$ è certo. Dunque per ogni $\bar{n}$ il vettore $\phi_{\bar{n}}$ è un $\psi_{\bar{m}}$: in altri termini, A e B possiedono un sistema ortonormale completo in comune. Segue subito che $[A, B] = 0$.
Si può estendere il risultato a molte osservabili $\tilde{A}_1, \ldots, \tilde{A}_N$: se sono simultaneamente misurabili i loro operatori commuteranno reciprocamente, $[A_i, A_k] = 0$.
Vediamo che il postulato **P5** è la più estesa asserzione che si può fare in una teoria che lo include. Senza la commutatività degli operatori, nulla si può dire dei risultati di misure simultanee, perché queste sono impossibili. Rimane però da capire qual'è la ragione fisica di questa impossibilità.

5.4 La determinazione dello stato

Le osservabili che possiedono operatori commutanti si dicono *osservabili compatibili* (terminologia dovuta a Dirac). Abbiamo notato che la misura di

un'osservabile che ha autovalori degeneri non produce uno stato univocamente determinato dal processo di misura e dal suo risultato, ma solo un elemento nello spazio di tutti gli autovettori appartenenti allo stesso autovalore. Per specificare completamente lo stato è in generale necessario misurare un certo numero di osservabili compatibili. Se la misura simultanea di un insieme di osservabili compatibili produce sempre un unico stato, diremo che l'insieme è *completo*. In meccanica classica tutte le osservabili sono compatibili, perciò misurandone una quantità sufficientemente elevata si riesce sempre a specificare lo stato in modo tale che le osservabili abbiano in esso un valore determinato. Sappiamo che per un sistema di N punti materiali, è sufficiente misurarne $6N$. Sappiamo anche che quando N è così grande che possiamo misurarne solo un numero molto minore di N (ad esempio per una mole di gas), la certezza delle nostre previsioni diventa molto minore, e si devono usare i metodi della meccanica statistica.
In meccanica quantistica non tutte le osservabili sono compatibili, ma solo quelle i cui operatori commutano (si veda la discussione fatta sopra). In questa teoria quindi, la massima precisione che si può ottenere è lo stato ψ risultante da una misura simultanea di un insieme completo di osservabili compatibili, e le osservabili che non commutano con i membri dell'insieme non hanno in ψ valori determinati. È questa la ragione formale del carattere statistico del postulato **P5** che è alla base dell'interpretazione fisica della MQ.
In molti casi poi, non si ha neanche una conoscenza completa dello stato, perché ad esempio non si riesce a misurare un insieme completo di osservabili compatibili. In questo caso la descrizione è meno dettagliata, e si sa solamente che lo stato è uno di un insieme di stati ortogonali $\{\phi_1, \ldots, \phi_n\}$, aventi probabilità w_1, w_2, $\ldots$, w_n, $\sum_k w_k = 1$. Allora i valori di aspettazione diventano

$$< \tilde{A} >= \sum_k w_k(\phi_k, A\phi_k)$$

La teoria che si occupa di queste "miscele" di stati quantici è la statistica quantistica. Von Neumann ha dimostrato che la descrizione mediante miscele di stati è la più generale che è permessa dai postulati della teoria.

5.5 Relazioni di indeterminazione

Due osservabili non commutanti $\tilde{A}$ e $\tilde{B}$ non possono essere misurate simultaneamente; abbiamo dedotto questo fatto dalla struttura matematica della meccanica quantistica, che associa operatori hermitiani alle osservabili, operatori che non sempre commutano. Tuttavia le ragioni fisiche non sono ancora

del tutto chiarite[16]. Intanto osserviamo che le dispersioni

$$\Delta A = \left(< \tilde{A}^2 > - < \tilde{A} >^2 \right)^{1/2}$$

$$\Delta B = \left(< \tilde{B}^2 > - < \tilde{B} >^2 \right)^{1/2}$$

non possono essere simultaneamente uguali a zero, sebbene ciascuna possa essere resa piccola a piacere. Dovrebbe allora esistere una formula che impedisce a ΔA e ΔB di essere piccole a piacere simultaneamente. Simili relazioni sono state scoperte da Werner Heisenberg (1927), e sono fondamentali per il significato fisico del formalismo e del fatto che osservabili non commutanti non siano misurabili contemporaneamente.
Le relazioni di indeterminazione di Heisenberg sono

$$\Delta A \Delta B \geq \frac{1}{2} | \, (\psi, [A, B]\psi) \, | = \frac{1}{2} | < [A, B] > | \tag{5.7}$$

Per la prova, osserviamo innanzitutto che la formula (5.7) non dice nulla se ψ è un autostato comune di A e B. Perciò supponiamo che questo non sia il caso, e consideriamo la quantità $2i \, \mathrm{Im}(A\psi, B\psi)$, per la quale si ha

$$2i \, \mathrm{Im}(A\psi, B\psi) = (A\psi, B\psi) - \overline{(A\psi, B\psi)}$$

Se ricordiamo la proprietà del prodotto scalare, $(f, g) = \overline{(g, f)}$, otteniamo

$$\begin{aligned} 2i \, \mathrm{Im}(A\psi, B\psi) &= (A\psi, B\psi) - (B\psi, A\psi) = (\psi, AB\psi) - (\psi, BA\psi) \\ &= (\psi, [A, B]\psi) \equiv a \end{aligned}$$

Abbiamo supposto $a \neq 0$; si ha dunque

$$1 = \frac{2i \, \mathrm{Im}(A\psi, B\psi)}{a} = \frac{|2i \, \mathrm{Im}(A\psi, B\psi)|}{|a|} \leq \frac{2}{|a|} |(A\psi, B\psi)| \leq \frac{2}{|a|} \|A\psi| \cdot \|B\psi\|$$

dove nell'ultimo passaggio si è usata la diseguaglianza di Schwarz (4.18). Vediamo che

$$\|A\psi| \cdot \|B\psi\| = (\psi, A^2\psi)^{1/2} (\psi, B^2\psi)^{1/2} \geq \frac{1}{2} |(\psi, [A, B]\psi)|$$

Se al posto di A si pone $A_1 = A - xI$ e al posto di B si pone $B_1 = B - yI$, con I l'operatore identità, il commutatore non cambia e la diseguaglianza

[16] L'impossibilità di costruire strumenti adeguati deve essere alla base dell'impossibilità di misurare osservabili non compatibili.

vale anche con A_1 e B_1. Ponendo allora $x =< \tilde{A} >$ e $y =< \tilde{B} >$, si ottiene la (5.7).
La forma più nota delle relazioni di indeterminazione di Heisenberg si ha se A è una componente del vettore posizione e B una componente della quantità di moto; in tal caso si ottiene

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}, \qquad \Delta y \Delta p_y \geq \frac{\hbar}{2}, \qquad \Delta z \Delta p_z \geq \frac{\hbar}{2} \tag{5.8}$$

Vediamo che le dispersioni delle osservabili interessate dalle relazioni di Heisenberg non possono essere entrambe piccole a piacere; più è accurata la misura di x (più è piccolo Δx) maggiore è l'indeterminazione di p_x (più è grande Δp_x), e viceversa. Osserviamo anche che le dispersioni delle osservabile nei vari stati quantici non hanno niente a che vedere con gli *errori delle misure*, dovute al fatto che gli strumenti non sono mai perfetti. Si può assumere che gli strumenti siano infinitamenti precisi, ed avere ancora il fenomeno della dispersione. Infatti lo stato quantico Ψ è genericamente una sovrapposizione di stati con valori definiti di una data osservabile $\tilde{A}$ e dunque, come abbiamo visto, $\tilde{A}$ non ha un valore definito in Ψ.
Bisogna comunque chiedersi se le relazioni di indeterminazione riflettono veramente le proprietà *fisiche* dei processi di misura.
Per illustrare questo punto, e capire la ragione fisica delle relazioni di Heisenberg, discutiamo ora il processo di misura di una coordinata di posizione, illustrato schematicamente in Fig. [8]. A questo scopo si può utilizzare un diaframma di apertura $\Delta y = d$, e uno schermo che registra gli arrivi delle particelle: l'arrivo sullo schermo significa che la coordinata y è determinata con l'incertezza $\Delta y = d$.
Ora l'attraversamento del diaframma è accompagnato da un effetto di diffrazione che è tanto più marcato quanto più è stretta l'apertura, cioè maggiore la precisione della misura. L'apertura angolare dovuta alla diffrazione è, come sappiamo, $\sin\theta \simeq \lambda/d$; quindi la componente p_y del momento può assumere qualunque valore fra $-p\sin\theta$ e $p\sin\theta$. In altri termini, c'è una incertezza in p_y di ordine $p\sin\theta$

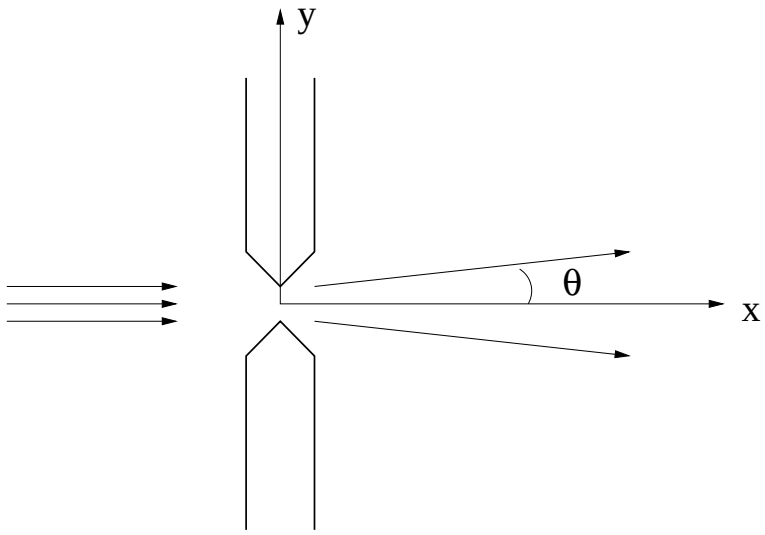
$$\Delta p_y \simeq p\sin\theta = \frac{h}{\lambda}\sin\theta = h/\Delta y$$

o, ciò che è lo stesso,

$$\Delta p_y \Delta y \simeq h$$

È importante verificare che il momento p_y non può veramente essere determinato con precisione migliore di quella data dalla relazione di incertezza.

Figura 8: Misura di posizione



In linea di principio una misura del momento trasferito al diaframma è sufficiente a determinare p_y. Notiamo però che la posizione del diaframma deve avere una imprecisione δy molto più piccola di Δy, affinché possa essere usato per una misura di posizione, e dunque l'incertezza nella componente P_y della quantità di moto del diaframma è[17]

$$\delta P_y \simeq h/\delta x \gg h/\Delta y \simeq \Delta p_y$$

Vediamo dunque che è impossibile determinare p_y con precisione migliore di quella data dalle relazioni di Heisenberg, perché l'uso del diaframma come strumento di misura della posizione rende assai indeterminato il suo impulso. Si può anche dire che il processo di misura di un'osservabile $\tilde{A}$ introduce una *perturbazione essenzialmente incontrollabile* nelle osservabili non compatibili con $\tilde{A}$. Che ci debba essere una perturbazione rilevante è chiaro dal fatto che le osservazioni dove è necessaria la MQ avvengono su scala microscopica. Anche l'osservazione di una palla in volo ne perturba il moto, per via della pressione di radiazione, ma si tratta di perturbazioni trascurabili. C'è da aspettarsi che non sia così però, se si illumina un elettrone. Per una discussione più dettagliata di questi problemi, rimando i lettori al testo di Messiah [9].

[17]Perciò il diaframma deve avere una grande massa affinché la sua velocità sia arbitrariamente piccola.

La discussione svolta illustra comunque un importante punto di principio: è necessario ammettere che anche gli strumenti di misura, per quanto macroscopici, obbediscano essi stessi alle relazioni di indeterminazione, e che la perturbazione da essi causata sia sempre sufficientemente forte che le relazioni di Heisenberg non cessino mai di essere soddisfatte. Se si scoprisse che non è così, la struttura della meccanica quantistica dovrebbe essere profondamente modificata.

Appendici

A Equipartizione dell'energia

Dalla forma dell'Hamiltoniana (2.11) segue che il contributo all'energia cinetica di ogni grado di libertà è

$$\frac{p_k^2}{2m} = \frac{1}{2} p_k \frac{\partial H}{\partial p_k}$$

Applicando la distribuzione canonica alla formula (2.12) dei valori di aspettazione si trova la formula (con una integrazione per parti)

$$< \frac{1}{2} p_k \frac{\partial H}{\partial p_k} >= \frac{1}{2} kT \tag{1.1}$$

Si vede che il risultato è lo stesso per tutti i gradi di libertà, in altri termini l'energia cinetica si equipartisce fra i vari gradi di libertà. Nel caso di sistemi composti di particelle identiche, il risultato segue da ovvi ragionamenti di simmetria. L'equipartizione è però valida anche se le particelle hanno masse diverse, o per sistemi composti di molecole poliatomiche che hanno gradi di libertà di rotazione e oscillazione.
Un grado di libertà oscillatorio ha l'energia potenziale bq^2, dove b è una costante. Questi contributi entrano additivamente nell'Hamiltoniana (2.11), e dunque ciascun grado di libertà oscillatorio contribuisce all'energia potenziale con un termine

$$b(q^j)^2 = \frac{1}{2} q^j \frac{\partial H}{\partial q^j}$$

Come prima, la media di questa espressione non dipende dall'indice j e vale $kT/2$.
Un grado di libertà rotazionale ha energia cinetica $I\omega^2/2$, dove I è il momento di inerzia del corpo rotante rispetto all'asse di rotazione, e ω è la velocità angolare. L'Hamiltoniana contiene il termine $p_\phi^2/2I$, dove p_ϕ è il momento coniugato all'angolo di rotazione e $\omega = \dot{\phi}$. Come prima, si trova che la media statistica del contributo vale ancora $kT/2$.
In conclusione, ogni grado di libertà cinetico contribuisce al calore specifico con $k/2$ ergs/K(grado Kelvin). Una mole contribuisce allora con $N_A k/2 = R/2$ ergs/K (la costante dei gas vale $R = 8,31 \cdot 10^7 \text{erg}/K$, e dunque $R/2$ è circa una caloria standard). Ogni grado di libertà oscillatorio contribuisce al calore specifico con $k/2 + k/2 = k$ ergs/K, e dunque una mole contribuisce con R ergs/K. E infine ogni grado di libertà rotazionale contribuisce con $k/2$ ergs/K, e dunque una mole con $R/2$ ergs/K.

B La distribuzione delle velocità

Per un sistema con Hamiltoniana (2.11) si ha

$$\exp\left[-\beta H(p,q)\right] = \prod_{k=1}^{3N} \exp\left[-\beta \frac{p_k^2}{2m}\right] \exp\left[-\beta \Phi(q^1, \ldots, q^{3N})\right]$$

Se $\vec{p}$ è la quantità di moto di una molecola, ne segue che il valor medio di una funzione $f(\vec{p})$ è

$$< f >= \frac{\int f(\vec{p}) e^{-\beta \vec{p}^2/2m} d^3\vec{p}}{\int e^{-\beta \vec{p}^2/2m} d^3\vec{p}} = \left(\frac{\beta}{2\pi m}\right)^{3/2} \int f(\vec{p}) e^{-\beta \vec{p}^2/2m} d^3\vec{p}$$

Si è usata la formula dell'integrale Gaussiano

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$$

Scritto per esteso, si vede dunque che

$$dw = \left(\frac{\beta}{2\pi m}\right)^{3/2} \exp\left(-\frac{\beta}{2m}(p_x^2 + p_y^2 + p_z^2)\right) dp_x dp_y dp_z \tag{2.1}$$

è la probabilità che le componenti della quantità di moto di una molecola siano nell'elemento di volume indicato dello spazio delle quantità di moto di una molecola. Questa formula non è influenzata dal potenziale ed è quindi valida anche per i gas reali, i liquidi e i solidi. La distribuzione delle velocità segue immediatamente; è la famosa distribuzione di Maxwell

$$dw = \left(\frac{\beta m}{2\pi}\right)^{3/2} \exp\left(-\frac{\beta m}{2}(v_x^2 + v_y^2 + v_z^2)\right) dv_x dv_y dv_z \tag{2.2}$$

C Equazioni di Maxwell

Come nel testo, indichiamo con $\vec{E}$ e $\vec{B}$ i campi elettrico e magnetico. Nel sistema di unità CGS razionalizzate, in cui il modulo della forza tra due cariche puntiformi poste a distanza di r cm l'una dall'altra si esprime con la legge di Coulomb nella forma

$$F = \frac{e_1 e_2}{r^2}$$

le equazioni fondamentali dell'elettrodinamica di Maxwell sono

$$\operatorname{div} \vec{E} = 4\pi\rho \tag{3.1}$$

$$\operatorname{rot} \vec{B} = \frac{1}{c}\frac{\partial \vec{E}}{\partial t} + \frac{4\pi}{c}\vec{J} \tag{3.2}$$

$$\operatorname{div} \vec{B} = 0 \tag{3.3}$$

$$\operatorname{rot} \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} \tag{3.4}$$

dove $\rho(\vec{r}, t)$ è la densità di carica elettrica e $\vec{J}(\vec{r}, t)$ la densità di corrente elettrica. Si noti l'assenza di una densità di cariche magnetiche e della relativa corrente: il magnetismo, in questa teoria, è interamente generato dalle correnti elettriche e dai campi elettrici variabili nel tempo, come vuole la legge sperimentale di induzione, scoperta da M. Faraday. Abbiamo indicato con $\operatorname{div} \vec{E}$ la *divergenza* di $\vec{E}$

$$\operatorname{div} \vec{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}$$

Analoga è l'espressione per $\operatorname{div} \vec{B}$. Con $\operatorname{rot} \vec{E}$, che si legge *rotore* di $\vec{E}$, si indica invece il campo di vettori le cui componenti sono

$$(\operatorname{rot} \vec{E})_x = \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}, \quad (\operatorname{rot} \vec{E})_y = \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}, \quad (\operatorname{rot} \vec{E})_z = \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}$$

Analoga è l'espressione di $\operatorname{rot} \vec{B}$.
È facile vedere che le equazioni ammettono soluzioni non banali anche nel vuoto, cioè in assenza di cariche e correnti, $\rho = 0$, $\vec{J} = 0$. Poniamo (onda monocromatica)

$$\vec{E}(t, \vec{r}) = \vec{a} \exp(-i\omega t + i\vec{k} \cdot \vec{r}), \qquad \vec{B}(t, \vec{r}) = \vec{b} \exp(-i\omega t + i\vec{k} \cdot \vec{r}) \tag{3.5}$$

Le caratteristiche dell'onda sono $\omega = 2\pi\nu$, dove ν è la frequenza in cicli per secondo, e $\vec{k}$, il vettore numero d'onde il cui modulo è $|\vec{k}| = 2\pi/\lambda$, dove $\lambda = c/\nu$ è la lunghezza d'onda e c la velocità della luce nel vuoto. La funzione $\phi = -i\omega t + i\vec{k} \cdot \vec{r}$ si chiama *la fase dell'onda.* Il vettore $\vec{k}$ è perpendicolare ai piani di fase costante e dà la direzione di propagazione dell'onda, nel senso che

i piani di fase costante si spostano nella direzione di $\vec{k}$ rimanendo ortogonali a $\vec{k}$. Per questo $v_\phi = \omega/|\vec{k}|$ si chiama *la velocità di fase* dell'onda.
I campi sono ovviamente funzioni reali, ma siccome le equazioni di Maxwell sono lineari e reali (nel senso che non hanno coefficienti complessi), le parti reali o immaginarie dei campi complessi sono anch'esse soluzioni delle equazioni. L'uso dei campi complessi è quindi permesso; si deve usare attenzione solo quando si calcolano espressioni non lineari nei campi, come ad esempio l'energia. È chiaro che $(\operatorname{Re} E)^2$ non è la stessa cosa di $\operatorname{Re} E^2$. A questo proposito si ricordi che se $z = x + iy$ è un numero complesso, la parte reale è per definizione x, e si scrive $\operatorname{Re} z = x$, mentre la parte immaginaria è y, e si scrive $\operatorname{Im} z = y$. Le potenze di z, z^n, si calcolano tenendo conto che $i^2 = -1$; ad esempio $z^2 = x^2 - y^2 + 2ixy$. La funzione esponenziale è definita dalla serie in campo complesso

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}, \qquad z = x + iy \tag{3.6}$$

e soddisfa la ben nota formula che la caratterizza: $\exp(z+u) = \exp(z)\exp(u)$. Il numero $e = \exp(1) = 2,718\ldots$ è irrazionale, e si chiama numero di Nepero. Si può notare che $\exp(n) = \exp(1)^n = e^n$ per ogni intero, e se $p = n/m$ si ha $\exp(p)^m = \exp(mp) = \exp(n) = e^n$, cosicché $\exp(p) = e^p$ per ogni numero razionale; questo motiva la definizione $e^z = \exp(z)$. In particolare, $e^z = e^x e^{iy}$, se $z = x + iy$.
Dalla definizione (3.6) segue la formula di Eulero[18]

$$e^{iy} = \cos y + i \sin y$$

In particolare, $e^{i\pi} = -1$, $e^{2i\pi} = 1$. Dunque se $\vec{a}$ è reale si ha

$$\operatorname{Re} \vec{E} = \vec{a} \cos(i\omega t - i\vec{k} \cdot \vec{r}), \quad \operatorname{Im} \vec{E} = \vec{a} \sin(i\omega t - i\vec{k} \cdot \vec{r})$$

Tornando alle (3.5), le equazioni con le divergenze dei campi danno le condizioni

$$\vec{k} \cdot \vec{a} = \vec{k} \cdot \vec{b} = 0$$

perciò i campi sono perpendicolari alla direzione di propagazione dell'onda: si dice che le onde elettromagnetiche sono "trasversali". Le equazioni di Maxwell con i rotori danno invece le condizioni

$$c\vec{k} \times \vec{a} = \omega \vec{b}, \quad c\vec{k} \times \vec{b} = -\omega \vec{a}$$

[18] Se uno si ricorda lo sviluppo in serie di potenze delle funzioni trigonometriche.

Queste equazioni hanno soluzioni non triviali (cioè non identicamente nulle) solo se vale la condizione $\omega^2 - c^2k^2 = 0$, e dunque si ha la relazione di dispersione caratteristica delle onde elettromagnetiche

$$\omega = c|\vec{k}| = c\sqrt{k_x^2 + k_y^2 + k_z^2}$$

Questo mostra che la velocità di fase delle onde elettromagnetiche è c, e dunque le ampiezze sono uguali in modulo ma ortogonali in direzione: $|\vec{a}| = |\vec{b}|$ e $\vec{a} \cdot \vec{b} = 0$.

Riferimenti bibliografici

[1] P. W. Bridgman, *La logica della fisica moderna*, Einaudi (1952).

[2] H. Goldstein, *Meccanica classica*, Zanichelli, Bologna (1971).

[3] L. D. Landau e E. M. Lifchitz *Fisica statistica*, Editori Riuniti, (1978).

[4] J. Perrin, *Gli atomi*, Editori Riuniti (1981).

[5] L. D. Landau e E. M. Lifchitz, *Teoria dei campi*, Editori Riuniti (1976).

[6] E. M. Purcell, *Elettricità e magnetismo*, Collana "La Fisica di Berkeley", Vol. 1 e Vol. 2, Zanichelli, Bologna (1973).

[7] F. S. Crawford, *Onde e oscillazioni*, Collana "La Fisica di Berkeley", Vol. 3, Zanichelli, Bologna (1972).

[8] M. Born, *Fisica atomica*, Boringhieri (1968).

[9] A. Messiah, *Meccanique quantique*, Vol. 1, Dunod (1958).

[10] P. A. M. Dirac, *I principi della meccanica quantistica*, Boringhieri (1971).

[11] F. Riesz and B. Sz.-Nagy, *Functional analysis*, Frederick Ungar Publishing Co., NY (1972).

[12] J. von Neumann, *Mathematical foundations of quantum mechanics*, Princeton University Press, (1955).