

TRATTATO GENERALE DI FISICA
A CURA DEL CONSIGLIO NAZIONALE DELLE RICERCHE

ENRICO PERSICO

FONDAMENTI
DELLA
MECCANICA ATOMICA



NICOLA ZANICHELLI EDITORE
BOLOGNA 1936-XIV

ERRATA CORRIGE

L'EDITORE, ADEMPIUTI I DOVERI,
ESERCITERÀ I DIRITTI SANCITI DALLE LEGGI

Copyright 1936-XIV by Casa Ed. N. Zanichelli

pag. 9	riga 16 :	$e = 47,668 \cdot 10^{-10}$	leggasi	$e = 4,7668 \cdot 10^{-10}$
» 37	titolo :	I livelli e nergetici	»	I livelli energetici
» 39	riga 17 :	$\tilde{\nu} = R \left(-1 - \frac{1}{n^2} \right);$	»	$\tilde{\nu} = R \left(1 - \frac{1}{n^2} \right);$
» 40	» 29 :	sono quadruple di quelle della serie di Lyman dell'idrogeno, e quindi cadono nel campo della luce visibile anzichè nell'ultravioletto:	»	, per n pari, coincidono con quelle della serie di Balmer dell'idrogeno:
» 52	» 4 :	analizata	»	analizzata
» 102	» 3 :	dovremmo	»	dovremo
» 130	» 7 :	equazione fondamentale	»	equazione caratteristica
» 275	» 6 :	$0,9174 \cdot 10^{-20}$	»	$0,9158 \cdot 10^{-20}$
» 336	» 11 :	§ 28, P. II	»	§ 27, P. II

PARTE I

SVILUPPO STORICO E BASI SPERIMENTALI

Questa prima parte serve a dare una veduta d'insieme sull'evoluzione della meccanica atomica, seguendo l'ordine storico e prendendone occasione per descrivere alcune delle più importanti esperienze su cui si fonda la teoria. Essa ha carattere più elementare del resto del libro, che è invece dedicato all'esposizione della meccanica atomica nel suo assetto attuale.

CAPITOLO I.

IL MODELLO DELL'ATOMO

§ 1. - IL CONCETTO DI ATOMO NELLA CHIMICA E NELLA FISICA.

L'ipotesi di una struttura corpuscolare della materia, quantunque sia stata enunciata da tempi antichissimi, si impose come necessaria all'interpretazione dei fatti sperimentali soltanto quando furono scoperte le leggi fondamentali della chimica, cioè quella delle proporzioni fisse e quella delle proporzioni multiple. Queste leggi infatti conducono naturalmente a pensare ogni corpo semplice come formato di particelle tutte uguali tra loro, le quali si possono associare variamente in modo da formare innumerevoli composti, ma intervengono sempre, in tutti i fenomeni chimici, come unità indivisibili, cioè come *atomi* nel senso etimologico della parola. La chimica permette di determinare con grande esattezza i *rapporti* tra i pesi dei vari atomi, ossia di esprimere tutti i pesi atomici con una unità definita mediante uno di essi (p. es., l'atomo di ossigeno) ma non permette di confrontare il peso di un atomo col peso dei corpi ordinari, e quindi di esprimere i pesi atomici in frazioni di grammo. Tanto meno poi la chimica dà indizi sulla grandezza e sulla forma degli atomi: per essa basta che gli atomi siano così piccoli da sfuggire all'osservazione diretta. Perciò, per la comprensione della maggior parte dei fenomeni chimici basta raffigurarsi gli atomi, p. es., come minutissime sferette compatte e indivisibili (il cui diametro non interessa) dotate di una particolare specie di forza attrattiva per cui possono

raggrupparsi in *molecole*, le quali a loro volta, riunite in numero sterminato, formano i corpi che noi vediamo.

Questa concezione semplicistica della struttura della materia non solo è sufficiente a render conto dei fenomeni chimici, ma permette anche di svolgere la teoria di molti fenomeni fisici: ne è un esempio la teoria cinetica dei gas, la quale può raggiungere buona parte dei suoi risultati riguardando le molecole come semplici punti o sferette, o aggregati di sferette. Si osservi inoltre che in molte formule della teoria cinetica dei gas interviene la massa m delle singole molecole, ma nelle formule finali, destinate al confronto con l'esperienza, la m appare *quasi sempre* moltiplicata per il numero totale delle molecole, cosicchè interviene in realtà solo la massa totale del gas. Ciò significa che la maggior parte della teoria è indipendente dalla grandezza che si attribuisce alle molecole e dal loro numero, e quindi non permette di trarre su tali dati alcuna conclusione. Perciò, fino a tempi assai recenti, la teoria atomica si è sviluppata prescindendo dalla conoscenza della grandezza assoluta delle molecole e degli atomi.

Ma si sono poi scoperti alcuni fenomeni nei quali interviene direttamente la massa delle singole molecole o — ciò che è lo stesso — il *numero di Avogadro* N , cioè il numero di molecole contenute in un grammo-molecola ⁽¹⁾; tali sono, p. es., i fenomeni del moto browniano. Questi fenomeni hanno quindi permesso di determinare sperimentalmente questo numero, il quale rappresenta, per così dire, la chiave per il passaggio dal mondo dei corpi ordinari al mondo degli atomi. Anzi, è stato possibile determinare il numero di Avogadro per una quindicina di vie indipendenti, e l'aver trovato risultati concordanti entro i limiti degli errori spe-

⁽¹⁾ Come è noto, un grammo-molecola è un numero di grammi uguale al peso molecolare. Perciò, conoscendo il numero di Avogadro N , si ottiene il peso in g. di una singola molecola dividendo per N il peso molecolare. Similmente il peso in g. di un atomo si ottiene dividendo per N il peso atomico.

rimentali è la prova migliore della realtà dell'ipotesi atomica ⁽¹⁾.

In altri fenomeni (p. es. la viscosità dei gas) intervengono le dimensioni delle molecole o degli atomi, e quindi questi fenomeni hanno permesso di determinare tali dimensioni almeno come ordine di grandezza. Si è trovato che gli **atomi** hanno diametri dell'ordine di 10^{-8} cm., e che di questo ordine sono pure le distanze alle quali si raggruppano nel formare le molecole, cosicchè anche per queste le dimensioni risultano generalmente di qualche unità dell'ordine di 10^{-8} cm. Si vedrà più oltre come queste dimensioni si possano definire più esattamente, e come le teorie odierne ne forniscano i valori, con notevole precisione, per i diversi casi.

Ma la concezione degli atomi quali semplici ed inerti frammenti di materia, è ancora insufficiente a spiegare un **grandissimo** numero di fenomeni fisici. Infatti, l'emissione e l'assorbimento della luce, i fenomeni magnetici, quelli dei **raggi X** e della radioattività, quelli che avvengono nel passaggio dell'elettricità attraverso i gas, l'elettrolisi, l'effetto **fotoceltrico**, quello termoionico, ed altri ancora, tutti **condono a ritenere** che gli atomi siano dei veri meccanismi più o meno complessi, composti di parti più piccole e contenenti **cariche elettriche positive e negative**, in quantità normalmente uguali. Queste cariche elettriche si presentano sempre come multiple di una carica elementare e : si è così stati condotti ad attribuire anche all'elettricità, come già alla materia, una struttura corpuscolare, anzi, a considerare i corpuscoli elettrici come costituenti degli atomi. Questi

⁽¹⁾ Il valore più attendibile (al dicembre 1932) del numero di Avogadro è $N = 6,068 \cdot 10^{23}$, con l'incertezza di qualche unità sull'ultima cifra significativa. Se ne deduce che l'unità adottata per i pesi atomici ($\frac{1}{16}$ dell'atomo di O, anzi, più esattamente, dell'isotopo più abbondante dell'ossigeno: v. § 4) vale grammi $\frac{1}{N} = 1,648 \cdot 10^{-24}$: il peso di un atomo di idrogeno differisce, come è noto, assai poco da questo (g. $1,661 \cdot 10^{-24}$).

hanno quindi perduto il primitivo significato di particelle « indivisibili » (ed infatti, si conoscono oggi diversi fenomeni, anche assai comuni come la ionizzazione, nei quali l'atomo si divide), ma vanno invece pensati come sistemi meccanici, formati di parecchie particelle di diversa specie, delle quali parleremo più oltre (elettroni positivi e negativi, protoni, neutroni), alcune cariche positivamente, altre negativamente, altre neutre, in moto le une rispetto alle altre sotto l'azione delle reciproche forze.

Questa concezione complessa dell'atomo come macchina contenente dell'elettricità in movimento, unita alla concezione elettromagnetica della luce, apre, come è facile comprendere, la possibilità di comprendere i legami tra materia e radiazione rappresentati dai fenomeni di emissione ed assorbimento, diffusione, risonanza, ecc.: è noto infatti dall'elettromagnetismo che il moto di cariche elettriche può generare onde elettromagnetiche, e queste, viceversa, possono mettere in moto le cariche elettriche che investono. Inoltre, poichè cariche elettriche in moto generano un campo magnetico, si intravede la possibilità di spiegare per questa via le proprietà magnetiche della materia; anzi si può osservare che una delle prime concezioni non puramente statiche dell'atomo (o della molecola) fu ideata da AMPÈRE proprio allo scopo di ricondurre i fenomeni magnetici a correnti elettriche circolanti entro i più piccoli costituenti della materia.

Tuttavia, come vedremo tra breve, queste ampie possibilità che si intravedono ad un esame superficiale della questione, si sono realizzate soltanto attraverso non lievi difficoltà, e hanno condotto assai più lontano di quanto non si potesse a prima vista pensare.

§ 2. - L'ELETTRONE.

I corpuscoli elettrici negativi, o *elettroni*, sono noti da molto tempo come costituenti i raggi catodici ed i raggi β delle sostanze radioattive: essi inoltre vengono emessi dai metalli incandescenti (*effetto termoionico*) e da molte sostanze quando sono colpite da luce di piccola lunghezza d'onda

(*effetto fotoelettrico*) o da raggi X. Gli elettroni ottenuti in questi diversi modi sono risultati sempre identici per tutte le loro proprietà: se ne è potuto misurare (mediante la deflessione subita in un campo elettrico e magnetico) il rapporto $\frac{e}{m}$ tra la carica e la massa, cosicchè da una misura indipendente di e si può ricavare m . La carica e dell'elettrone è stata misurata con grande precisione dal WILSON e poi dal MILLIKAN col « metodo della nebbia » e con quello della « goccia equilibrata », il quale ultimo metodo — si noti — permette di determinare la carica di elettroni presi uno per uno e quindi di accertare che le cariche di tutti gli elettroni sono effettivamente uguali tra loro e che il valore trovato non è soltanto un valore medio ⁽¹⁾. I valori più attendibili della carica (in valore assoluto) e della massa dell'elettrone sono (dicembre 1932) ⁽²⁾:

Carica: $e = 4,7668 \cdot 10^{-10}$ u. e. s. = $1,590 \cdot 10^{-20}$ u. e. m.

Massa (per piccole velocità): $m_0 = 9,038 \cdot 10^{-28}$ g.

La massa risulta circa 1837 volte più piccola della massa dell'atomo di idrogeno. Per velocità elevate si è constatato che il rapporto $\frac{e}{m}$ decresce proporzionalmente a $\sqrt{1 - \frac{v^2}{c^2}}$, il che è in accordo con la teoria della Relatività, secondo la quale la massa m deve variare con la legge

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

dove m_0 è il valore della massa per piccole velocità.

È noto che nell'elettrolisi un grammo-molecola di ioni monovalenti trasporta una carica di 96494 coulomb inter-

⁽¹⁾ Per una descrizione di queste esperienze, si vedano i nn. 27, 28 e 31 della Bibliografia.

⁽²⁾ Per una discussione critica dei valori numerici di queste e di altre costanti fisiche, vedasi R. T. BIRGE, *Phys. Rev. Suppl. (Rev. of Mod. Phys.)* 1, 1 (1916); *Phys. Rev.* 43, 211, 1933.

nazionali, pari a $2,8927 \cdot 10^{14}$ u. e. s.: dividendo questa quantità per il numero di Avogadro ($6,068 \cdot 10^{23}$) si ottiene la carica di ciascuno ione monovalente, e si trova che coincide col valore testè riportato della carica dell'elettrone (1). Questa coincidenza induce a ritenere che la carica elettrica degli ioni monovalenti sia dovuta ad un elettrone aggiunto o sottratto all'atomo neutro, e similmente che gli ioni polivalenti non siano altro che atomi aventi un certo numero di elettroni in più o in meno del normale: si identificano insomma gli elettroni osservati come corpuscoli liberi nei raggi catodici, nei raggi β , ecc., con quelli che si manifestano nei fenomeni elettrolitici.

Che poi negli atomi siano sempre contenuti degli elettroni, della stessa specie di quelli che si osservano liberi, è provato da molti fatti che verranno esposti nel seguito. Ci limiteremo qui ad accennare ad una delle prove che storicamente ha avuto la maggiore importanza: l'effetto Zeeman permette di determinare sperimentalmente il rapporto $\frac{e}{m}$ relativo al corpuscolo elettrico al cui movimento si attribuisce l'emissione della luce, e si è trovato un valore coincidente con quello già determinato per gli elettroni col metodo della deflessione. Si riconosceva così la presenza degli elettroni in seno agli atomi, e la loro diretta partecipazione al fenomeno dell'emissione luminosa.

§ 3. - I CORPUSCOLI POSITIVI. L'ATOMO DI RUTHERFORD.

Se i fatti accennati sopra, insieme ad altri, hanno permesso da lungo tempo di mettere fuori discussione la presenza, entro gli atomi, di elettricità negativa in forma di elettroni, restava da conoscere in quale forma si trovasse

(1) Veramente, quando è stata fatta questa osservazione si disponeva solo di un valore approssimativo per N , e quindi la coincidenza era solo approssimativa. Ammettendo, come sembra plausibile, che essa debba verificarsi esattamente, si ha uno dei migliori modi per determinare N .

nell'atomo quella elettricità positiva che è necessaria a dare all'atomo il suo carattere di sistema elettricamente neutro. J. J. THOMSON, nel 1902, emise l'ipotesi che l'atomo fosse costituito da una sfera omogenea di elettricità positiva entro la quale fossero immersi gli elettroni, come granelli di polvere in una goccia d'acqua: per l'azione reciproca tra cariche positive e negative, gli elettroni resterebbero attratti (come è facile dimostrare) verso il centro della sfera da una forza proporzionale alla distanza, e occuperebbero delle posizioni di equilibrio sotto l'azione di questa forza e delle loro mutue repulsioni. Questa ipotesi però ben presto si mostrò inadeguata all'interpretazione dei fatti sperimentali.

La confutazione più diretta dell'ipotesi di THOMSON fu offerta da una celebre esperienza eseguita dal RUTHERFORD nel 1911 sul passaggio delle particelle α attraverso la materia. Tali particelle (che sono come è noto atomi di elio doppiamente ionizzati, cioè carichi positivamente, lanciati dalle sostanze radioattive con velocità dell'ordine di 10^9 cm./sec.) possono attraversare, in virtù della loro piccolezza e della loro alta velocità, delle sottili foglie metalliche: però si osserva che se si lancia contro la foglia un sottile pennello di raggi α , questo non prosegue in linea retta ma viene sparpagliato in varie direzioni, provando così che le particelle α nell'attraversare la materia vengono deviate, alcune più altre meno, e taluna quasi fino a tornare indietro. Queste deviazioni non possono essere dovute agli urti delle particelle α contro gli elettroni contenuti nella sostanza attraversata, perchè questi, avendo una massa varie migliaia di volte minore di quella delle particelle α , non possono farle deviare sensibilmente. Esse debbono essere dunque dovute all'azione della parte rimanente dell'atomo, quella cioè che contiene, oltre la carica positiva, la quasi totalità della massa. Ora, se si ammette il modello di THOMSON, si vede facilmente che l'azione repulsiva esercitata dalla sfera di elettricità positiva sulla particella α , dapprima cresce con l'avvicinarsi di questa, ma poi, quando la particella penetra nell'interno della sfera, la forza *decrece* man mano che

essa si avvicina al centro: la traiettoria risulta allora incurvata come mostra la fig. 1, *a*. Le deviazioni che così si ottengono non risultano molto grandi e sono insufficienti a spiegare i risultati sperimentali. Perciò il RUTHERFORD emise l'ipotesi (oggi universalmente accettata) che la parte positiva dell'atomo consti di un *nucleo*, assai piccolo rispetto alle dimensioni atomiche, sì da poterlo considerare quasi puntiforme, nel quale sia concentrata quasi tutta la massa nonchè la carica positiva. La forza repulsiva da esso esercitata sulla

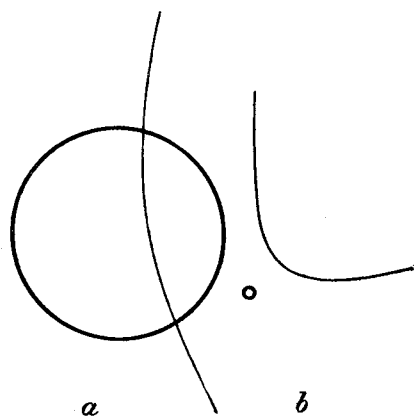


Fig. 1.

particella *a* sarà allora inversamente proporzionale al quadrato della distanza, e quindi crescerà indefinitamente con l'avvicinarsi della particella al nucleo: si capisce dunque che ne debba risultare una deviazione assai più notevole che non nel modello di THOMSON (fig. 1, *b*). E se si calcola come si devono distribuire nei diversi angoli di deviazione le particelle di un fascetto di raggi *a*, si trova una distribuzione in perfetto accordo con quella osservata (eccetto per le particelle che arrivano assai vicine al centro del nucleo, per le quali questo non si può più considerare puntiforme). L'esperienza dunque conferma l'ipotesi di RUTHERFORD del «nucleo quasi puntiforme»: anzi è possibile, dai risultati di tali esperienze, dedurre quale è la carica del nucleo. Si è scoperto così il fatto assai notevole che la carica del nucleo è uguale (salvo il segno) a tante volte la carica elettronica, quanto è il *numero atomico* dell'elemento considerato, cioè il suo numero d'ordine nel sistema periodico di Mendelèiev.

In base a questi risultati il RUTHERFORD propose un modello di atomo, che ha avuto importanza fondamentale nello sviluppo successivo della Fisica Teorica. Egli considerò l'ato-

mo di numero atomico Z come costituito da un nucleo praticamente puntiforme, nel quale è concentrata la quasi totalità della massa, avente una carica positiva Ze (chiamando e la carica dell'elettrone presa in valore assoluto): intorno ad esso si trovano Z elettroni, le cui cariche negative compensano quindi esattamente quella positiva del nucleo. Questi elettroni sono attratti dal nucleo, secondo la legge di Coulomb, con una forza inversamente proporzionale al quadrato della distanza, cioè simile alla forza di gravitazione dei pianeti verso il Sole: essi quindi si muoveranno, secondo le leggi di Keplero, in orbite ellittiche (se si prescinde dalle reciproche repulsioni tra elettroni), mentre il nucleo, per la sua massa preponderante, resterà quasi immobile: tutto l'atomo raffigurerà così un sistema planetario in miniatura. L'atomo di idrogeno, per il quale $Z = 1$, avrà un nucleo di carica e ed un solo elettrone planetario (che in questo caso descriverà esattamente un'orbita ellittica, mancando le perturbazioni degli altri elettroni); l'atomo di elio avrà nucleo di carica $2e$ e due elettroni planetari, e così via, fino all'uranio ($Z = 92$) che avrà carica nucleare $92e$ e 92 elettroni.

Secondo questo modello, quello che si suol chiamare, piuttosto vagamente, «diametro dell'atomo» (e che abbiamo visto essere dell'ordine di 10^{-8} cm.) si deve identificare all'incirca colla massima dimensione delle orbite esterne: si può dire che l'atomo occupa presso a poco una sfera di questo diametro, ma intendendo bene che questa sfera non è riempita di sostanza compatta, anzi è quasi totalmente vuota, ed in essa si aggirano soltanto dei granuli di dimensioni piccolissime rispetto alle loro mutue distanze ⁽¹⁾.

Questo modello che ha dominato la meccanica atomica fino al 1925, secondo le vedute odierne non deve essere inteso come una fedele pittura della realtà intraatomica, ma solo come una sua approssimazione: tuttavia esso conserva

⁽¹⁾ Il raggio del nucleo è dell'ordine di 10^{-12} cm., quello dell'elettrone si ritiene non superiore a 10^{-13} cm.

grande importanza sia come mezzo euristico, sia come mezzo didattico, sia infine perchè fornisce la base di un linguaggio comodo ed espressivo, di cui ci serviremo correntemente.

§ 4. — LA STRUTTURA DEL NUCLEO.

Secondo il modello di Rutherford, la valenza ed in genere le proprietà chimiche di un atomo dipendono esclusivamente dal movimento degli elettroni più esterni: e lo stesso dicasi per le proprietà ottiche come, p. es., lo spettro di emissione o di assorbimento. Gli elettroni più interni intervengono invece nel determinare il comportamento dell'atomo rispetto ai raggi X (spettri Röntgen di emissione e di assorbimento). Infine, negli atomi con proprietà radioattive, si deve pensare che queste abbiano sede nel nucleo, come è confermato dal fatto che tali proprietà sono del tutto insensibili al legame chimico che altera il moto degli elettroni più esterni, ai raggi X che agiscono su quelli più interni, ed in genere a tutti gli agenti fisici che conosciamo, tranne gli urti di particelle o radiazioni estremamente penetranti.

Questo fatto induce a considerare il nucleo non come un corpuscolo elementare al pari dell'elettrone, ma come un sistema in generale complesso, e tale veduta è confermata da diverse altre circostanze. Una di queste è l'esistenza degli isotopi.

Si chiamano « isotopi » gli elementi che hanno le stesse proprietà chimiche e spettroscopiche — quindi vanno collocati nello stesso posto nel quadro di Mendelèiev, donde il nome — ma che differiscono per il peso atomico. L'esistenza di isotopi fu scoperta dapprima fra gli elementi radioattivi, ma più recentemente, per opera specialmente di ASTON, si è riconosciuto che anche molti fra i corpi non radioattivi, ritenuti corpi semplici, sono in realtà miscele di due, tre o più elementi isotopi di peso atomico differente, le quali miscele si comportano come corpi semplici appunto perchè i vari elementi che le compongono hanno proprietà fisiche e chimiche quasi perfettamente identiche. Quindi i pesi ato-

mici determinati dai chimici risultano in realtà da una media dei pesi atomici dei diversi isotopi; questi sono stati determinati da ASTON con grande precisione, con un metodo sul quale non è questo il luogo di soffermarci. Il risultato più importante di queste determinazioni è che i pesi atomici dei singoli isotopi sono sempre (con grande approssimazione) *numeri interi*: le frazioni che figurano nel solito quadro degli elementi chimici derivano dal fatto che quei pesi non si riferiscono a veri corpi semplici, ma, come abbiamo detto, a miscele. Così p. es. il Potassio, di peso atomico 39,10, è in realtà una miscela di due isotopi di pesi atomici 39 e 41. Le sole proprietà fisiche per cui possono sensibilmente differire due isotopi sono — eventualmente — le proprietà radioattive (sia spontanee che provocate), ed è per questo che fra gli elementi spontaneamente radioattivi la scoperta e la separazione degli isotopi è stata assai più facile.

L'interpretazione dell'isotopia nel modello di Rutherford è la seguente. Gli isotopi hanno lo stesso numero atomico, quindi la stessa carica nucleare e lo stesso numero di elettroni planetari, ma differiscono per la massa del nucleo. Poichè il moto degli elettroni è determinato dalle forze elettriche, e la massa del nucleo, che rimane quasi ferma, influisce pochissimo sul moto degli elettroni, questi si muovono in tutti gli isotopi quasi esattamente nello stesso modo: ne segue l'identità delle proprietà fisiche e chimiche. Fanno eccezione le eventuali proprietà radioattive, le quali non dipendono, come si è detto, dagli elettroni planetari, ma solo dalla struttura del nucleo, che è diversa nei diversi isotopi.

Le proprietà radioattive e la loro interpretazione teorica mediante opportune ipotesi sulla struttura dei nuclei costituiscono la « fisica nucleare », che solo assai recentemente (1932-33) sembra aver trovato le sue basi teoriche ed in brevissimo tempo ha fatto progressi enormi anche dal lato sperimentale, con la scoperta di nuove disintegrazioni e della radioattività artificiale. La fisica del nucleo non rientra però nel programma di questo volume, poichè nelle questioni di fisica atomica di cui qui ci occupiamo il nucleo interverrà

sempre come un corpuscolo unitario e di dimensioni trascurabili, caratterizzato solo dal valore della sua massa e della sua carica elettrica positiva (ossia dal numero atomico): ci limitiamo perciò qui solo ad elencare le varie specie di particelle che possono uscire dai nuclei, sia spontaneamente, sia in seguito ad un urto, e che presumibilmente sono i costituenti essenziali di ogni specie di materia:

particelle pesanti	}	<i>protoni</i> : carica $+e$, peso circa una unità atomica. Il nucleo dell'idrogeno è costituito solo da un protone.
		<i>neutroni</i> : carica nulla, peso probabilmente uguale a quello del protone.
particelle leggere	}	<i>elettroni</i> : carica $-e$, peso $1/1823$ unità atomiche. Sono identici a quelli circolanti intorno al nucleo.
		<i>positroni</i> , o <i>elettroni positivi</i> : carica positiva dell'ordine di e , peso dell'ordine di quello dell'elettrone negativo (probabilmente sono esattamente analoghi agli elettroni negativi, salvo il segno della carica).

I neutroni ed i positroni sono stati scoperti solo ultimamente (1932-33) e sono ancora poco conosciuti. Non si sa poi ancora se queste quattro specie di particelle siano da riguardarsi tutte come elementari, cioè indivisibili, o se p. es. il neutrone sia formato da un protone più un elettrone⁽¹⁾, o se il protone sia formato da un neutrone più un positrone.

§ 5. - LEGGE DELLO SPOSTAMENTO RADIOATTIVO. SPIEGAZIONE DELL'ISOTOPIA.

L'identità del numero atomico con la carica nucleare è confermata dalla legge detta dello *spostamento radioattivo*. Si constata infatti che quando una sostanza radioattiva emette raggi β , essa si trasforma in un'altra avente lo stesso

⁽¹⁾ In tal caso essi sarebbero legati assai più intimamente che non nell'atomo di idrogeno.

peso atomico ma numero atomico superiore di un'unità (cioè « si sposta », nel sistema periodico, di una casella in avanti): difatti il nucleo, perdendo un elettrone, non varia sensibilmente di peso mentre la sua carica positiva aumenta di e . Invece, quando una sostanza emette raggi α , essa si trasforma in un'altra avente peso atomico inferiore di 4 unità, e numero atomico inferiore di 2 unità (cioè si sposta di due caselle indietro): difatti una particella α (che come è noto, è identica al nucleo dell'elio) ha peso 4 e carica $2e$, quindi il peso del nucleo che l'ha emessa deve diminuire di 4 unità, e la carica di $2e$.

In particolare, se si susseguono, in un ordine qualunque, due emissioni β ed una α , il peso atomico deve diminuire di 4, ma la carica nucleare non deve variare: si avrà dunque un isotopo della sostanza primitiva. Questo processo avviene effettivamente in certi casi: p. es. il radio D ($A = 210$, $Z = 82$) emettendo una particella β si trasforma in radio E ($A = 210$, $Z = 83$), questo a sua volta emette un'altra particella β diventando polonio ($A = 210$, $Z = 84$), e questo infine, emettendo una particella α , si trasforma in radio G ($A = 206$, $Z = 82$) che ha lo stesso numero atomico, e quindi le stesse proprietà, del radio D , di cui è dunque un isotopo (ed entrambi sono isotopi del piombo).

Nel campo delle sostanze non radioattive, tali trasformazioni non avvengono spontaneamente, ma si può sempre pensare che i nuclei di due isotopi differiscano per un certo numero di elettroni e per altrettanti protoni, oppure (secondo l'ipotesi oggi più generalmente ammessa) per un certo numero di neutroni. Comunque insomma si concepiscano i costituenti elementari dei nuclei, si comprende la possibilità di aggrupparli in modo da ottenere pesi diversi con la stessa carica positiva risultante.

CAPITOLO II.

I QUANTI DI ENERGIA E I QUANTI DI LUCE

§ 6. — INAPPLICABILITÀ DELLE LEGGI CLASSICHE AGLI ATOMI.

Il modello atomico di Rutherford, per quanto seducente per la sua semplicità, porta tuttavia a risultati in pieno disaccordo con l'esperienza se si vuol ammettere che il suo funzionamento sia governato dalle ordinarie leggi della meccanica e dell'elettrologia.

Difatti, secondo le leggi dell'elettromagnetismo, una carica elettrica irradia energia elettromagnetica ogni qual volta sia dotata di una certa accelerazione: quindi un elettrone ruotante attorno al nucleo, avendo costantemente una accelerazione centripeta, dovrebbe irradiare continuamente onde elettromagnetiche. Per conseguenza la sua energia dovrebbe gradualmente diminuire, il che porterebbe ad una graduale diminuzione delle dimensioni dell'orbita, finchè l'elettrone finirebbe per cadere sul nucleo. L'atomo di Rutherford non potrebbe dunque avere carattere permanente, e si può calcolare che la sua vita sarebbe dell'ordine di 10^{-8} secondi.

Inoltre, l'energia sarebbe irradiata sotto forma di radiazioni la cui frequenza fondamentale coinciderebbe con la frequenza del moto orbitale dell'elettrone: ma siccome questa andrebbe continuamente variando a causa dell'impiccolimento dell'orbita, la luce emessa avrebbe frequenza variabile: perciò qualsiasi corpo, che contiene innumerevoli atomi in tutte le possibili fasi della loro vita, dovrebbe emettere radiazioni di tutte le possibili frequenze, ossia uno spettro continuo; è noto invece che i gas emettono spettri di righe, e di frequenza rigorosamente costante.

Quando si presentarono ai fisici queste difficoltà, essi si erano già formata per altra via la convinzione che la meccanica e l'elettromagnetismo classici non si possano applicare nel mondo atomico, così che, anzichè abbandonare il modello di Rutherford, si cercò di determinare delle leggi atte a farlo funzionare in modo da render conto dei risultati sperimentali. Queste leggi furono proposte per la prima volta da BOHR, nel 1913, ed enunciate poi in forma più generale da SOMMERFELD: ma prima di esporle dobbiamo accennare rapidamente agli altri campi nei quali si erano già incontrate contraddizioni con le leggi classiche, ed ai tentativi fatti per sostituirle con leggi nuove. Il primo di questi tentativi, che aperse la strada alla moderna fisica atomica, fu quello fatto dal fisico tedesco MAX PLANCK nel 1900 per dare una teoria dell'emissione del corpo nero: uno successivo, e riguardante la natura della luce anzichè quella della materia, fu fatto da EINSTEIN nel 1905, introducendo il concetto dei « quanti di luce ».

§ 7. - LO SPETTRO DEL CORPO NERO.

È noto che i corpi incandescenti emettono luce: più generalmente, i corpi emettono a qualsiasi temperatura delle radiazioni, che per temperature un po' alte sono visibili al nostro occhio, mentre per temperature inferiori sono invisibili (radiazioni infrarosse). Il tipo di radiazione emessa dipende, entro certi limiti, dalla natura del corpo, ma se questo è un « corpo nero », cioè un corpo capace di assorbire tutte le radiazioni che riceve (come è, con discreta approssimazione, p. es. il nerofumo), si dimostra termodinamicamente, e l'esperienza lo conferma, che la radiazione emessa ad una data temperatura è indipendente dalla natura e dalla forma del corpo. Analizzata allo spettroscopio, tale radiazione mostra uno spettro continuo, del quale si è misurata esattamente l'intensità punto per punto, alle varie temperature, intensità che presenta un massimo per una certa frequenza ν_m e decresce verso zero sia per le alte che per le basse frequenze, come è indicato dal diagramma della fig. 2.

È naturale pensare che la radiazione emessa sia dovuta alla agitazione termica delle cariche elettriche contenute entro la materia: ma se si tenta di precisare quantitativamente questa idea, applicando le ordinarie leggi della meccanica e dell'elettromagnetismo, si è condotti ad un risultato assurdo, cioè si trova, in luogo della curva sperimentale, una curva parabolica ascendente che, presso l'origine, coincide con l'altra, ma poi se ne distacca come è indicato

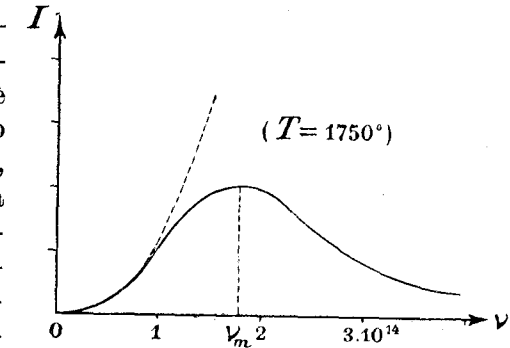


Fig. 2.

dalla fig. 2: essa è rappresentata dalla formula (detta di Rayleigh e Jeans)

$$I = \frac{2\pi}{c^2} kT\nu^2$$

(k = costante di Boltzmann; c = velocità della luce; T = temperatura assoluta; $I(\nu)d\nu$ rappresenta l'energia irradiata dall'unità di superficie nella striscia di spettro tra ν e $\nu + d\nu$). Secondo questa formula (come si vede subito integrando rispetto a ν), l'energia totale irradiata a qualunque temperatura diversa dallo zero assoluto, sarebbe infinita.

Diversi tentativi furono fatti per sfuggire a questa difficoltà, ma essi condussero sempre alla formula di Rayleigh e Jeans, finchè si faceva uso della meccanica e dell'elettromagnetismo classici, pur modificando in vari modi il particolare meccanismo dell'emissione. Nel condurre uno di questi tentativi, verso la fine del 1899, il fisico tedesco PLANCK ⁽¹⁾ scoperse quale era il punto nel quale conveniva staccarsi dalla trattazione classica per giungere ad un risultato pienamente conforme all'esperienza.

⁽¹⁾ *Verh. d. d. Phys. Ges.*, 2, 237 (1900); *Ann. d. Phys.*, 4, 553 (1901).

Egli partì dall'ipotesi che negli atomi fossero contenuti dei minuscoli *oscillatori* ⁽¹⁾ di tutte le possibili frequenze, in perpetua vibrazione (ciascuno col suo periodo proprio), i quali ricevono ed emettono continuamente energia sotto forma di onde elettromagnetiche. Per artificio di calcolo, il PLANCK impostò il problema supponendo che l'energia di ogni oscillatore variasse non con continuità, ma per salti di una certa grandezza ε , che poi si riservava di far tendere a zero. Ma constatò che se si faceva tendere ε verso zero si cadeva nella formula di Rayleigh e Jeans, mentre conservando ad ε un valore finito, e precisamente proporzionale alla frequenza propria ν dell'oscillatore, cioè ponendo

$$(1) \quad \varepsilon = h\nu$$

(con h costante), si arrivava alla formula

$$(2) \quad I = \frac{2\pi\nu^2}{c^3} \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1},$$

la quale, fissato il parametro T , rappresenta una curva del tipo di quella a tratto pieno della fig. 2: anzi, dando ad h un conveniente valore numerico, che secondo le più recenti determinazioni è, in unità *CGS*,

$$h = 6,542 \cdot 10^{-27},$$

la formula (2) rappresenta esattamente la distribuzione dell'energia nello spettro del corpo nero per qualunque valore della temperatura T . In seguito a questo sorprendente risultato, il PLANCK propose l'ardita ipotesi che l'energia avesse, al pari della materia, una struttura atomica, vale a dire si presentasse sempre in quantità multiple di una quan-

⁽¹⁾ Si chiama così un sistema formato da un punto materiale carico di elettricità (p. es. un elettrone) attirato da un centro fisso con una forza proporzionale alla distanza: è noto dalla meccanica che tale punto esegue delle oscillazioni isocrone intorno al centro fisso, con frequenza indipendente dall'ampiezza e caratteristica dell'oscillatore. Però il ragionamento di Planck non richiedeva necessariamente questo modello, ma poteva sussistere anche sotto ipotesi più generali.

tà elementare ε , che egli chiamò *quanto* ⁽¹⁾: questo deve avere, secondo quanto si è detto, valore proporzionale alla frequenza con cui è capace di oscillare il sistema considerato.

Entrava così per la prima volta nella scienza il concetto di una discontinuità nelle leggi fisiche, in contrasto con le ordinarie teorie meccaniche ed elettromagnetiche che — per distinguerle dalle nuove, chiamate *quantistiche* — furono dette *classiche*. In corrispondenza a questo nuovo concetto, entrava nella Fisica una nuova costante universale, h , detta *costante di Planck*, la cui importanza si è poi rivelata sempre maggiore in categorie svariatissime di fenomeni.

L'ipotesi estrema di una struttura atomica dell'energia fu ben presto attenuata dallo stesso PLANCK, il quale mostrò che bastava ammettere la discontinuità nel solo processo di emissione, e non nella natura stessa dell'energia. E per quanto, anche in tale forma, l'ipotesi ripugnasse alla mente di molti scienziati, il successo che ebbero i suoi successivi sviluppi nel render conto dei calori specifici dei solidi (teorie di Einstein e di Debye, 1907 e 1912) confermò in molti il convincimento che si fosse sulla buona strada.

Oggi la formula di Planck si può giustificare per altra via, indipendente dall'ipotesi degli oscillatori, e l'ipotesi dei quanti di energia ci appare come una conseguenza particolare, relativa all'oscillatore, di leggi assai generali fondate su ipotesi che, per esser prive di contraddizioni, sia in sé, sia con le leggi cui soddisfano i corpi ordinari, e per il fatto che rendono conto in modo unitario di un grandissimo numero di fenomeni, appaiono assai più soddisfacenti e plausibili della primitiva ipotesi dei quanti.

§ 8. — L'EFFETTO FOTOELETTRICO E I « QUANTI DI LUCE ».

Mentre nel campo della struttura della materia si veniva facendo strada, con l'ipotesi di Planck, l'idea che le leggi classiche non fossero applicabili all'atomo, una contraddizione

⁽¹⁾ In tedesco « *quantum* », parola latina usata in tedesco nel significato sostantivale di « quantità », « dose ».

non meno grave veniva rilevata da EINSTEIN tra certi fatti sperimentali e la teoria ondulatoria della luce.

Questa contraddizione, sulla quale torneremo più ampiamente in seguito, consiste sostanzialmente in questo. È noto che se si fa cadere della radiazione di frequenza abbastanza elevata (in generale, luce ultravioletta o raggi X o γ) su una superficie metallica, questa emette degli elettroni i quali vengono espulsi con una certa energia, che è senza dubbio comunicata ad essi dalla radiazione incidente. Il fenomeno si chiama *effetto fotoelettrico* nel caso in cui la radiazione eccitatrice è la luce visibile o ultravioletta: noi però useremo questo nome anche nel caso più generale. Ora si è constatato che l'energia che riceve ogni singolo elettrone da parte della radiazione (energia che in parte è consumata nello strappare l'elettrone dal metallo, ed in parte rimane all'elettrone come forza viva) è indipendente dall'intensità della luce, ma dipende solo dalla sua frequenza: l'intensità influisce solo sul numero di elettroni emessi (che è ad essa proporzionale), e non sulla energia della emissione di ciascun elettrone. Così si è stati costretti ad abbandonare l'ipotesi più naturale, che cioè gli elettroni siano costretti ad oscillare sotto l'azione del campo elettrico della luce, e che, quando le loro oscillazioni divengano abbastanza ampie, finiscano per essere divelti dall'atomo cui appartengono ed essere lanciati fuori. Ma vi è una difficoltà più grave, che si presenta indipendentemente da ogni particolare ipotesi sul meccanismo dell'effetto fotoelettrico, ed è la seguente. Diminuendo l'intensità della luce, si può far sì che l'energia che cade su ogni atomo durante tutta l'esperienza (calcolata nell'ipotesi che l'energia cada uniformemente su tutta la superficie) sia assai inferiore alla forza viva con cui viene espulso l'elettrone, eppure anche in queste condizioni gli elettroni vengono emessi con la stessa energia cinetica.

Assai notevole è poi la legge che lega la forza viva w con cui sono emessi gli elettroni alla frequenza ν della radiazione incidente: essa è

$$(3) \quad w = h\nu - w_0,$$

dove h è la costante di Planck, e w_0 è una costante caratteristica del metallo in esame. Per i valori di ν che renderebbero negativo il secondo membro (cioè per $\nu < \nu_0$, dove $\nu_0 = w_0/h$) non vi è emissione, cosicchè l'effetto fotoelettrico presenta la singolare caratteristica di cominciare bruscamente per una certa frequenza ν_0 , che dicesi *soglia fotoelettrica* ed è, al pari di w_0 , caratteristica di ciascun metallo. Anche questa legge sfuggì ostinatamente ad ogni tentativo di giustificazione classica, mentre la presenza in essa della medesima costante h che interviene nella teoria del corpo nero dava l'indizio di un suo legame con qualche profonda legge della natura.

Fu EINSTEIN che, in una celebre memoria ⁽¹⁾ del 1905, mostrò che il singolare comportamento dell'effetto fotoelettrico si poteva inquadrare in un modello corpuscolare della radiazione, modello che egli considerava però solo per il suo valore euristico, senza affermarne il significato concreto. Se infatti si pensa che l'energia di una radiazione monocromatica non sia uniformemente distribuita su tutto il fronte **d'onda**, ma viaggi localizzata in granuli (detti da Einstein **quanti di luce**, ed oggi chiamati anche **fotoni** ⁽²⁾) ognuno dei quali ne contiene una quantità $h\nu$, e viaggia con la velocità della luce, le difficoltà del fenomeno fotoelettrico scompaiono e si arriva immediatamente alla formula (3). Infatti il paradosso relativo all'emissione fotoelettrica con luce debole si spiega immediatamente pensando che, se la luce è debole, i quanti di luce saranno pochi, ma ognuno conterrà sempre l'energia $h\nu$: quindi saranno pochi gli atomi che ricevono un quanto, ma ognuno di quei pochi riceverà sempre l'energia $h\nu$ indipendentemente dall'intensità della luce. La formula (3) poi si giustifica immediatamente se si interpreta w_0 come l'energia che l'elettrone deve spendere per strapparsi

⁽¹⁾ « Ueber einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt ». *Ann. d. Phys.*, 17, 132 (1905).

⁽²⁾ Queste denominazioni non devono far pensare che l'ipotesi si riferisca soltanto alla luce propriamente detta: essa abbraccia ogni tipo di radiazione elettromagnetica.

dal metallo: infatti l'elettrone, che ha ricevuto dal quanto di luce l'energia $h\nu$, e ne spende w_0 per uscire dal metallo, uscirà con una energia residua $h\nu - w_0$ sotto forma di energia cinetica. E se l'energia del quanto di luce è inferiore a w_0 (ossia se $\nu < \nu_0$) l'elettrone non potrà uscire, come si constatò effettivamente.

Si noti che l'ipotesi dei « quanti di luce » non è un puro e semplice ritorno alle vecchie teorie corpuscolari (o balistiche) della luce, poichè l'energia dei quanti di luce non va necessariamente pensata come energia cinetica di un corpuscolo in movimento (ed infatti la loro velocità nel vuoto è sempre c , qualunque sia la loro energia): si è cercato di interpretarla come energia elettromagnetica localizzata in una piccola regione dello spazio, ma non è possibile conciliare tale veduta con le classiche equazioni di Maxwell.

Si osservi pure che l'ipotesi dei « quanti di luce » è distinta da quella dei quanti di Planck: la prima infatti si riferisce al modo col quale l'energia viaggia nello spazio, la seconda si riferisce al modo col quale l'energia viene emessa ed assorbita dalla materia. Tuttavia è evidente che esse si avvalorano mutualmente, e che entrambi i gruppi di fenomeni a cui si riferiscono mettono in evidenza che deve esistere un profondo legame fisico tra la frequenza ν di una radiazione e la quantità $h\nu$ di energia.

Come si dirà più avanti, l'ipotesi dei quanti di luce è stata in seguito confermata da altri fatti sperimentali, e nella teoria attuale il concetto di *quanto di luce* o di *fotone* ha una parte altrettanto importante quanto quelli di elettrone e di protone.

È noto che la teoria elettromagnetica della luce dimostra che all'energia raggianti W deve essere associata una « quantità di moto elettromagnetica » $\frac{W}{c}$, la quale si manifesta, quando viene assorbita dalla materia, col fenomeno della *pressione della luce*: è naturale quindi che ogni teoria della luce per essere in accordo coi fatti deve associare all'energia raggianti W una quantità di moto o *impulso* $\frac{W}{c}$. Perciò ai fotoni si deve attribuire (indipendentemente da ogni concezione concreta corpuscolare) oltre all'energia $h\nu$ un im-

pulso p , di grandezza

$$p = \frac{h\nu}{c} = \frac{h}{\lambda}$$

nella direzione della propagazione della luce. Una riprova sperimentale diretta della esattezza di questa ipotesi è fornita dall'effetto Compton e da talune esperienze ad esso relative, di cui ora parleremo brevemente.

§ 9. - L'EFFETTO COMPTON.

Quando un fascio di raggi X attraversa la materia, una parte di esso viene diffusa in tutte le direzioni come avviene della luce che attraversa un mezzo torbido. Ora, se si analizza, con lo spettrografo a cristallo, questa radiazione diffusa (1) si osserva che essa non ha l'identica costituzione spettrale della radiazione primaria. Se p. es. questa era monocromatica di frequenza ν , lo spettro della radiazione diffusa presenta, oltre alla riga di frequenza ν , una riga di frequenza ν' leggermente inferiore: questa riga ha intensità notevole, anzi prevalente sulla riga « non spostata », se la sostanza diffondente è a numero atomico basso (p. es. carbonio, o alluminio) e se i raggi incidenti sono « duri » (ν elevata) mentre è assai debole o manca del tutto se la sostanza ha peso atomico elevato (p. es. argento) e se i raggi incidenti sono molli. Questo fenomeno di diffusione con lieve aumento di lunghezza d'onda venne scoperto nel 1923 dal fisico americano A. H. COMPTON, e di lui porta il nome.

L'alterazione di lunghezza d'onda è diversa per la radiazione diffusa nelle varie direzioni: nulla nella direzione dei raggi primari, cresce con l'angolo θ fra la direzione di osservazione e quella dei raggi primari, e raggiunge il suo valore massimo (che ammonta a soli 0,048 Å) per $\theta = 180^\circ$, cioè quando si osserva la radiazione diffusa all'indietro; tale variazione è bene rappresentata dalla formula

$$(4) \quad \lambda' = \lambda + 0,024 (1 - \cos \theta)$$

(1) Alla radiazione propriamente diffusa, è frammista la *radiazione secondaria* il cui spettro è quello caratteristico della sostanza diffondente: essa si distingue senza difficoltà da quella diffusa, e non ci interessa ora fissare su di essa l'attenzione.

(dove λ e λ' devono essere espressi in Ångström). Come si vede da questa formula, l'alterazione $\lambda' - \lambda$ nella lunghezza d'onda è indipendente da λ , cosicchè l'effetto *relativo* è notevole per piccoli valori di λ , cioè per raggi « duri », mentre diviene trascurabile quando λ è grande rispetto a 0,048 Å.

Le teorie classiche non danno alcuna spiegazione dell'effetto Compton, poichè secondo tali teorie la diffusione è dovuta al fatto che gli elettroni contenuti nella sostanza diffondente compiono delle oscillazioni forzate sotto l'azione del campo elettrico alternato della radiazione incidente, e quindi divengono a loro volta centri di emissione di onde: ma queste oscillazioni si devono compiere con la frequenza della radiazione incidente, e quindi anche la radiazione diffusa deve avere rigorosamente la stessa frequenza.

§ 10. — TEORIA QUANTISTICA DELL'EFFETTO COMPTON.

Secondo la teoria dei quanti, la diffusione va invece pensata come effetto degli urti tra i fotoni della radiazione incidente e gli elettroni della sostanza diffondente, in seguito ai quali i fotoni vengono deviati dalla loro direzione primitiva, come le particelle α nell'esperienza di Rutherford di cui abbiamo già parlato. Questo urto, qualunque sia il suo meccanismo, è governato dalle due leggi fondamentali della conservazione dell'energia e dell'impulso, come se si trattasse di un urto tra corpi perfettamente elastici. È chiaro quindi che l'elettrone urtato acquista nell'urto una certa velocità, e perciò sottrae una certa quantità di energia al fotone incidente, il quale viene diffuso perciò con energia minore. E siccome l'energia dei fotoni è legata alle loro frequenza dalla relazione $E = h\nu$, il fotone diffuso deve avere una frequenza minore di quello incidente: è così giustificato qualitativamente l'effetto Compton.

Possiamo ora precisare quantitativamente questo ragionamento. Chiamiamo v la velocità acquistata dall'elettrone urtato: la sua forza viva sarà, secondo la meccanica relativista, $mc^2 \left(\frac{1}{\sqrt{1-v^2/c^2}} - 1 \right)$ (dove m è la « massa di quiete »

dell'elettrone) ⁽¹⁾ mentre l'energia del quanto incidente è $h\nu$ e quella del quanto diffuso $h\nu'$: il principio della conservazione dell'energia dà dunque

$$(5) \quad h\nu = h\nu' + mc^2 \left(\frac{1}{\sqrt{1-v^2/c^2}} - 1 \right).$$

Se poi si rappresentano con tre vettori (v. fig. 3) l'impulso del fotone incidente (vettore AO , di lunghezza $\frac{h\nu}{c}$), quello del fotone diffuso (vettore OB , di lunghezza $\frac{h\nu'}{c}$) e quello acquistato dall'elettrone (vettore OD , di lunghezza $\frac{mv}{\sqrt{1-v^2/c^2}}$),

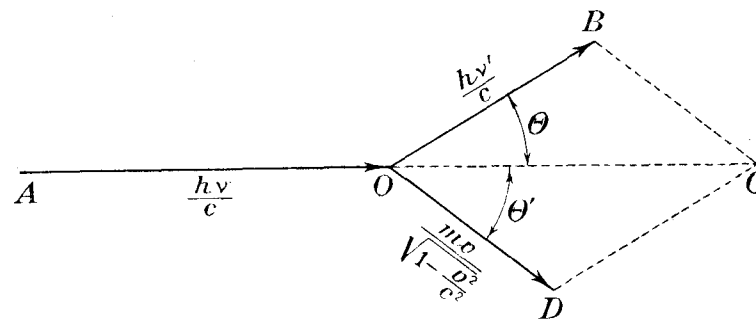


Fig. 3.

e si osserva che il primo vettore deve essere la risultante degli altri due, si ha subito, dal triangolo OBC , per il teorema di Carnot,

$$(6) \quad \frac{m^2 v^2}{1-v^2/c^2} = \frac{h^2 \nu^2}{c^2} + \frac{h^2 \nu'^2}{c^2} - 2 \frac{h\nu}{c} \frac{h\nu'}{c} \cos \theta.$$

Eliminando v tra queste due equazioni si trova

$$(7) \quad \nu - \nu' = \frac{h\nu\nu'}{mc^2} (1 - \cos \theta)$$

ed esprimendo le frequenze mediante le lunghezze d'onda,

(1) Se si usassero le formule classiche per la forza viva e la quantità di moto, si arriverebbe a risultati praticamente non distinguibili da quelli a cui si giunge con la meccanica relativista, ma le formule risulterebbero alquanto più complicate.

cioè ponendo $\nu = \frac{c}{\lambda}$, $\nu' = \frac{c}{\lambda'}$, si trova, dopo facili riduzioni,

$$(8) \quad \lambda' = \lambda + \frac{h}{mc} (1 - \cos \theta).$$

Se per h , m , c si pongono i loro valori numerici, si trova

$$\frac{h}{mc} = 2,4 \cdot 10^{-10} \text{ cm.} = 0,024 \text{ \AA},$$

cosicchè la (8) coincide esattamente con la (4), che descrive, come si è detto, tutti i risultati sperimentali.

In questa teoria si è supposto che gli elettroni diffondenti fossero liberi, mentre nella realtà essi sono più o meno vincolati al nucleo dell'atomo cui appartengono, e quindi all'urto partecipa in certa misura tutto l'atomo (ed eventualmente anche l'intera molecola). Per vedere quale sia l'effetto di questa circostanza, consideriamo il caso estremo che l'elettrone sia rigidamente legato al suo atomo: allora esso nell'urto si comporterà come un corpo non più di massa m , ma di una massa M uguale alla massa dell'atomo, cioè alcune migliaia di volte più grande. Il ragionamento precedente varrà ancora, purchè si sostituisca m con M : quindi nella formula finale (8) la costante $\frac{h}{mc}$, che determina l'entità dell'effetto, diventerà qualche migliaio di volte più piccola, il che significa che l'effetto Compton diventerà inapprezzabile. Questa considerazione giustifica la presenza della « *riga non spostata* »: essa è dovuta agli elettroni più strettamente vincolati al nucleo, mentre la riga spostata è dovuta a quelli periferici che si possono considerare praticamente liberi. Ed è naturale che, per una data frequenza ν , negli elementi di elevato numero atomico, ossia ad elevata carica nucleare, prevalga l'effetto degli elettroni vincolati (che sono più numerosi) mentre negli elementi leggeri quasi tutti gli elettroni si comportino come liberi: è naturale inoltre che, in una data sostanza, gli elettroni si possano considerare tanto più liberi quanto più notevole è l'energia dei fotoni che li urtano, ossia quanto più elevata è la frequenza ν . Tutto questo è, come si è detto, in perfetto accordo con l'esperienza.

§ 11. — ESPERIENZA DI BOTHE E GEIGER.

La teoria precedente dell'effetto Compton (dovuta a COMPTON e a DEBYE) trovò un valido appoggio in una esperienza fatta da BOTHE e GEIGER nel 1925⁽¹⁾, allo scopo di verificare se in realtà, contemporaneamente alla diffusione del fotone, viene lanciato un elettrone (*elettrone di rimbalzo*), come la teoria prevede. Per rivelare sia i fotoni diffusi che gli elettroni di rimbalzo i detti fisici si servirono del *contatore a punta*, strumento già utilizzato da RUTHERFORD e GEIGER per contare le particelle α del Radio, e consistente in una speciale camera di ionizzazione capace di dare un segnale elettrico quando in essa penetra anche un solo elettrone, o un fotone, o un qualunque corpuscolo capace di produrre ionizzazione. Nell'esperienza di Bothe e Geiger, due contatori a punta erano posti ai due lati di un sottile fascetto di raggi X attraversante dell'idrogeno: uno dei due contatori era chiuso da una sottile foglia di platino capace di arrestare gli elettroni, ma non i raggi X: quindi quel contatore registrava solo i fotoni diffusi dall'idrogeno. L'altro invece registrava gli elettroni di rimbalzo, ed oltre a questi, i fotoni diffusi dalla sua parte, e gli elettroni eventualmente prodotti per effetto fotoelettrico. Non era da aspettarsi quindi una coincidenza completa fra le segnalazioni dei due contatori: il secondo contatore dava assai più segnali del primo, ma — e questo è il risultato notevole dell'esperienza — per ogni segnale del primo contatore ve ne era uno *simultaneo* del secondo, il che provava che per ogni fotone diffuso da una parte vi era sempre un elettrone di rimbalzo lanciato dall'altra, con perfetta simultaneità.

§ 12. — ESPERIENZA DI COMPTON E SIMON.

Ma un'esperienza ancora più significativa a conferma della precedente teoria dell'effetto Compton fu eseguita da COMPTON e SIMON nel 1925⁽²⁾, utilizzando il *metodo della nebbia*, di Wilson, che permette, come è noto, di rendere visibili, sotto forma di sottili fili di nebbia, le traiettorie degli elettroni attraverso un gas. In tal modo essi riuscirono non solo a rivelare fotograficamente l'esistenza dell'elettrone di rimbalzo, ma anche a verificare che la direzione in cui esso è lanciato è quella prevedibile con le leggi dell'urto elastico.

⁽¹⁾ W. Bothe e H. Geiger *ZS f. Phys.* **32**, 639 (1925)

⁽²⁾ A. H. Compton e A. W. Simon, *Phys. Rev.* **26**, 289 (1925).

Vediamo anzitutto quale relazione forniscono queste leggi tra la direzione OB di diffusione del fotone e la direzione OD dell'elettrone di rimbalzo, cioè tra gli angoli θ e θ' (fig. 3). Se invece di scrivere la relazione (6) tra le sole grandezze dei tre vettori in questione, utilizziamo in pieno il principio della conservazione dell'impulso scrivendo che si conserva tanto la sua proiezione nella direzione AO , quanto quella in direzione perpendicolare, dobbiamo scrivere invece della (8) le due equazioni

$$\frac{h\nu}{c} = \frac{h\nu'}{c} \cos \theta + \frac{mv}{\sqrt{1-v^2/c^2}} \cos \theta'$$

$$0 = \frac{h\nu'}{c} \sin \theta - \frac{mv}{\sqrt{1-v^2/c^2}} \sin \theta',$$

da cui (lasciando in ciascun secondo membro solo l'ultimo termine, e dividendo membro a membro)

$$\frac{\nu' \sin \theta}{\nu - \nu' \cos \theta} = \operatorname{tg} \theta'.$$

Se ora per ν si sostituisce l'espressione ricavata dalla (7), si ha, con facili trasformazioni

$$(9) \quad \operatorname{tg} \theta' = \frac{1}{1 + \frac{h\nu}{mc^2}} \operatorname{cotg} \frac{\theta}{2}.$$

È questa la relazione cercata. Da essa risulta, in particolare, che, col variare di θ da 0 a 180° , θ' varia da 90° a 0° : quindi, gli elettroni sono sempre diffusi in avanti.

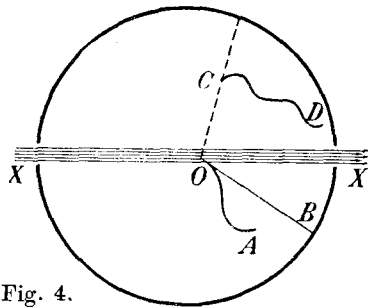


Fig. 4.

Per controllare sperimentalmente questa formula, COMPTON e SIMON fecero un gran numero di fotografie stereoscopiche col metodo della nebbia, in un gas attraversato da un sottile fascio di raggi X. Molte di queste fotografie sorpresero un atto di diffusione, e queste presentavano l'aspetto schematizzato nella figura 4, nella quale XX rappresenta il fascetto di raggi X primari. Da un punto O di questo, parte la traccia OA di un elettrone di rimbalzo. Il fotone diffuso non può lasciar traccia lungo la sua traiettoria, ma viene rivelato appena incontra un atomo C che lo assorba, poichè questo atomo emette allora un elettrone per effetto fotoelettrico, e la traccia CD di questo è visibile nella

fotografia: congiungendo la sua origine C con O si ha la direzione in cui è stato diffuso il fotone (linea punteggiata). La tangente OB alla traccia OA , condotta nella sua origine, dà la direzione iniziale dell'elettrone di rimbalzo. Mediante le due fotografie stereoscopiche si possono così determinare i veri valori (nello spazio) degli angoli θ e θ' , e si trova che soddisfano la (9), entro i limiti (che sono piuttosto ampi) degli errori sperimentali.

§ 13. — IMPOSSIBILITÀ DI UNA TEORIA LETTERALMENTE CORPUSCOLARE DELLA LUCE.

L'ipotesi che la luce abbia natura corpuscolare (che cioè sia composta di corpuscoli nel senso intuitivo della parola) mentre sembra *quasi* imposta dai fenomeni di cui abbiamo parlato nei §§ precedenti, incontra però gravissime difficoltà in un'altra non meno vasta categoria di fenomeni, e cioè in tutti quei fenomeni che hanno costituito, da HUYGHENS in

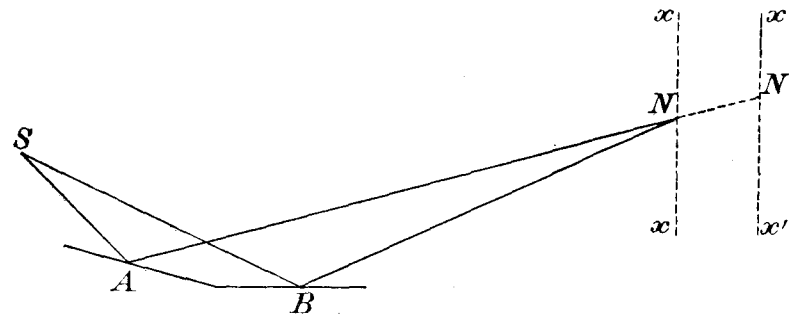


Fig. 5.

poi, la base sperimentale delle teorie ondulatorie della luce: principalissimi fra questi l'interferenza e la diffrazione.

Per rendersi conto della natura di queste difficoltà, si consideri, a titolo d'esempio, un caso particolarmente semplice d'interferenza, come quello degli specchi di Fresnel (fig. 5). Sia S la sorgente, A e B gli specchi, N un punto tale che i cammini ottici SAN e SBN differiscano per un numero dispari di semilunghezze d'onda. Supponiamo dapprima che lo specchio B sia ricoperto da uno schermo: allora in N vi è luce, ossia (usando il linguaggio corpuscolare) vi arrivano i corpuscoli che hanno percorso la linea SAN .

Scopriamo lo specchio B , ed in N si produce, come è noto, una frangia nera, mentre in altri punti (frange chiare) si rinforza l'illuminazione. Potremmo pensare che i corpuscoli che percorrono la linea $SA N$ sono stati influenzati, a distanza, dallo scoprimento dello specchio B e sono stati deviati: ma questa ipotesi, oltre ad essere strana in sè, non regge, perchè se, invece di osservare le frange nel piano xx , le osserviamo in un altro piano $x'x'$, possiamo trovare in N' , cioè sul prolungamento del raggio AN , una frangia chiara. Dovremmo dunque pensare che i corpuscoli provenienti da A sono passati per N , ma in tal punto sono stati resi in qualche modo inefficaci per la presenza simultanea dei corpuscoli provenienti da B . Ma questa ipotesi, oltre a lasciare incomprendibile il fatto che in altri punti dello schermo si ha un rinforzo di illuminazione, è smentita dal fatto, constatato sperimentalmente, che i fenomeni di interferenza si producono anche quando la luce è così debole, che l'apparecchio è attraversato da un corpuscolo per volta: si è quindi costretti a rinunciare ad ammettere che i corpuscoli possano interagire tra loro in modo da rendersi inefficaci in certi punti e rinforzarsi in altri.

Una prova che l'interferenza si produce anche con un solo quanto per volta si ha nel fatto che i cannocchiali astronomici danno l'immagine nitida anche di stelle debolissime, ed è noto che le immagini si formano nei sistemi ottici per un processo di interferenza. Ora, con un cannocchiale dell'apertura di 10 cm. sono perfettamente visibili stelle di 10^3 grandezza, la cui radiazione ha un'intensità di circa $3,5 \cdot 10^{-10}$ erg/sec. cm^2 .

Se si suppone la luce di una lunghezza d'onda media $\lambda = 5500 \text{ \AA}$, e si calcola, con la formula $E = h\nu$, l'energia corrispondente ad ogni quanto, si trova $36 \cdot 10^{-13}$ erg; quindi l'intensità della stella di 10^3 grandezza è di circa 100 quanti per cm^2 e secondo: nel cannocchiale ne entrano allora circa 8000 al secondo. Se si considera che ogni quanto attraversa il cannocchiale in un tempo dell'ordine di 10^{-8} secondi, si vede che i singoli quanti attraversano il cannocchiale isolatamente, anzi a grandi intervalli l'uno dall'altro. La formazione dell'immagine in tali condizioni farebbe pensare che ogni singolo quanto utilizzi per suo conto tutta l'apertura dell'obbiettivo, e condurrebbe quindi ad attribuire ai quanti dimensioni trasversali di parecchi centimetri

o di qualche metro, in contrasto con le esperienze di fotoelettricità, che hanno invece condotto ad attribuir loro dimensioni atomiche.

La interferenza con singoli quanti è stata anche provata con esperienze dirette di G. J. TAYLOR ⁽¹⁾ e di DEMPSTER e BATHO ⁽²⁾, i quali ultimi, mediante un interferometro a gradinata, hanno ottenuto frange di interferenza anche quando la luce era così debole, che certamente ogni quanto era entrato nell'apparecchio dopo che ne era uscito il precedente.

In conclusione, ogni teoria della luce che sia corpuscolare nel senso letterale della parola, cioè che attribuisca ai quanti di luce *determinate traiettorie* come se si trattasse di corpuscoli materiali, si è rivelata assolutamente inconciliabile coi fatti sperimentali, così come, d'altra parte, la teoria ondulatoria di Maxwell si è rivelata in contrasto con l'effetto fotoelettrico e con molti altri fatti sperimentali. Vedremo più innanzi (§ 34) come una profonda revisione dei principi della teoria ondulatoria e di quella corpuscolare sia oggi riuscita a eliminare questo contrasto, fondendole in una e lasciando a ciascuna delle due antiche teorie quella parte di validità che è in accordo coi fatti sperimentali.

⁽¹⁾ *Proc. Camb. Phil. Soc.*, 15, 114 (1909).

⁽²⁾ *Phys. Rev.*, 30, 664 (1927).

CAPITOLO III.

I LIVELLI ENERGETICI

§ 14. — LO SPETTRO DELL'IDROGENO E DEI SISTEMI IDROGENOIDI.

La maggior copia di elementi per la conoscenza degli atomi ci è fornita dalla spettroscopia, la quale, oltre che dal grandissimo numero di dati che fornisce, è caratterizzata dalla estrema precisione che consentono le misure di lunghezza d'onda, le quali si compiono non di rado con errore relativo inferiore ad un milionesimo.

Di tutti gli spettri, il primo ad essere interpretato fu quello dell'idrogeno atomico, che si presenta effettivamente come il più semplice, in accordo col fatto che l'atomo di idrogeno è quello che ha la più semplice struttura.

Se si osserva allo spettroscopio la luce emessa da un tubo Plücker contenente idrogeno, si vede un gran numero di righe che possono facilmente distinguersi in due categorie. Vi è uno sfondo costituito da numerose e fitte righe non molto intense, e su questo spicca un piccolo numero di righe ben visibili, la cui intensità, se si aumenta la corrente che alimenta il tubo, cresce assai più rapidamente che non quella dello sfondo. Si è accertato che le righe numerose e deboli (che costituiscono il cosiddetto *spettro di molte righe*) sono emesse dalle molecole H_2 , mentre quelle più vive (che formano il cosiddetto *spettro balmeriano*) sono dovute agli atomi H , che si producono nel tubo per effetto della dissociazione provocata dal passaggio della corrente. Lo spettro di molte righe, essendo emesso da molecole, è di struttura complicata e solo recentemente è stato interpretato teori-

camente: noi ci occuperemo solo dello spettro balmeriano, e, parlando di spettro dell'idrogeno, intenderemo riferirci esclusivamente a questo.

Esso è costituito di righe situate parte nella regione visibile, parte nell'infrarosso e parte nell'ultravioletto, ed è rap-

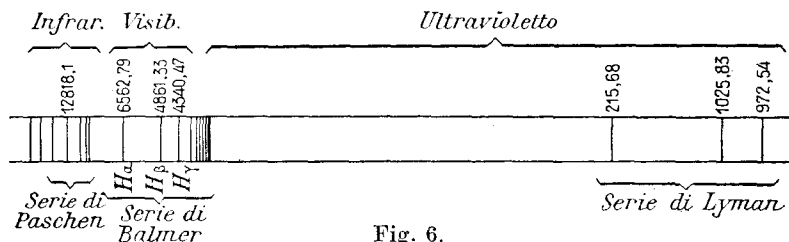


Fig. 6.

presentato schematicamente dalla fig. 6, la quale suppone la dispersione proporzionale alla frequenza. Come è manifesto dalla figura, queste righe si riuniscono in tre gruppi chiamati rispettivamente *serie di Paschen* (infrarossa), *serie di Balmer* (visibile), *serie di Lyman* (ultravioletta).

Il primo passo verso l'interpretazione teorica degli spettri fu compiuta nel 1885 dallo svizzero Balmer, il quale osservò che tutte le righe della serie dell'idrogeno che da lui ha preso poi il nome (e che era la sola allora conosciuta) avevano delle frequenze ⁽¹⁾ rappresentabili con la formula

$$(10) \quad \tilde{\nu} = R \left(\frac{1}{4} - \frac{1}{n^2} \right),$$

nella quale R è una costante, detta oggi *numero di Rydberg*, il cui valore numerico è $R = 109678 \text{ cm.}^{-1}$, ed n un numero

(1) È invalso nella spettroscopia l'uso di chiamare «frequenza» di una radiazione non solo la vera frequenza ν (numero di vibrazioni al secondo, legato alla lunghezza d'onda λ da $\nu = c/\lambda$) ma anche la quantità ad essa proporzionale

$$\tilde{\nu} = \frac{\nu}{c} = \frac{1}{\lambda},$$

che più propriamente si chiama anche «numero d'onde», perchè rappresenta il numero delle lunghezze d'onda contenute in un cm. L'uso del «numero d'onde» è generalmente preferito a quello delle frequenze

intero che può assumere tutti i valori da 3 in su. Come si vede subito, col crescere di n i valori di $\tilde{\nu}$ dati da questa formula tendono verso il limite $\frac{R}{4}$, e difatti le righe della serie di Balmer si addensano verso un limite, tanto che da un certo punto in poi non è più possibile distinguerle tra loro. Tuttavia si è riusciti a distinguerne più di una trentina, e per tutte la frequenza rientra esattamente nella formula di Balmer.

Si riconobbe in seguito che la formula di Balmer non è che un caso particolare di una formula più generale che rappresenta tutte le righe dello spettro dell'idrogeno atomico: tale formula è

$$(11) \quad \tilde{\nu} = R \left(\frac{1}{n'^2} - \frac{1}{n^2} \right),$$

dove R è la costante precedente, ed n' , n sono due numeri interi. Facendo $n' = 1$, ed $n = 2, 3, 4, \dots$ si hanno le frequenze della serie di *Lyman*:

$$\tilde{\nu} = R \left(1 - \frac{1}{n^2} \right);$$

facendo $n' = 2$, ed $n = 3, 4, 5, \dots$ si riottiene la (10) che rappresenta la serie di *Balmer*: e facendo $n' = 3$ ed $n = 4, 5, 6, \dots$ si ottengono le frequenze della serie di *Paschen*:

$$\tilde{\nu} = R \left(\frac{1}{9} - \frac{1}{n^2} \right).$$

La semplicità della formula, il numero notevole di righe che essa rappresenta e l'estrema precisione con cui essa si

propriamente dette, perchè le λ si possono misurare direttamente con grandissima precisione, talora superiore a 1 : 1.000.000, e basta prenderne il reciproco per avere $\tilde{\nu}$, mentre le ν si debbono ricavare da esse con la formula $\nu = c/\lambda$: ora, essendo c nota con precisione di poco superiore ad 1 : 1000, le ν risultano note con assai minor precisione delle $\tilde{\nu}$.

Per evitare equivoci, basta aggiungere alla designazione della «frequenza» quella dell'unità di misura, che per le frequenze propriamente dette ν è il sec^{-1} , per i numeri d'onda $\tilde{\nu}$ è il cm^{-1} . Generalmente, anche il numero d'onde $\tilde{\nu}$ è indicato con la semplice lettera ν .

adatta ai risultati sperimentali rendevano praticamente impossibile che la coincidenza fosse casuale: si trattava però di interpretare la formula con un adatto modello atomico, il che presentò per lungo tempo difficoltà apparentemente insuperabili.

Aggiungeremo che in seguito si è scoperto che degli spettri strettamente analoghi a quello dell'idrogeno sono presentati dagli atomi degli elementi più leggeri, quando sono ionizzati in modo da aver perduto tutti i loro elettroni meno uno, e cioè dagli ioni He^+ , Li^{++} , Be^{+++} , C^{++++} . Tali ioni, essendo costituiti da un nucleo e da un solo elettrone, hanno struttura analoga all'atomo di idrogeno, da cui differiscono solo per la massa del nucleo e per la sua carica, la quale, anzichè essere e come nell'atomo di idrogeno, è Ze nell'elemento di numero atomico Z : essi si designano collettivamente (insieme all'atomo di idrogeno) col nome di *sistemi idrogenoidi*. È prevedibile che a questa analogia di struttura debba corrispondere un'analogia di proprietà spettrali, e precisamente (come diremo) si ha ragione di ritenere che lo spettro dell'atomo di numero atomico Z ionizzato $Z - 1$ volte sia rappresentato dalla formula, analoga alla (11),

$$(11') \quad \tilde{\nu} = Z^2 R \left(\frac{1}{n'^2} - \frac{1}{n^2} \right)$$

e quindi differisca da quello dell'idrogeno solo perchè le frequenze di tutte le righe sono moltiplicate per Z^2 (salvo una lieve correzione dovuta al movimento del nucleo, di cui si parlerà al § 58 p. II), P. es. lo ione He^+ ($Z = 2$) presenta una serie detta di *Pickering*, data da

$$(12) \quad \tilde{\nu} = 4 R \left(\frac{1}{4^2} - \frac{1}{n^2} \right),$$

le cui frequenze sono quadruple di quelle della serie di Lyman dell'idrogeno, e quindi cadono nel campo della luce visibile anzichè nell'ultravioletto: di questa serie sono state osservate varie righe ⁽¹⁾. Sono anche state osservate alcune ri-

(1) Tale serie, osservata dapprima solo nello spettro di una stella, fu attribuita all'idrogeno, a causa del fatto che le righe di posto pari

ghe di Li^{++} di Be^{+++} e di C^{++++} , tutte soddisfacenti la formula (11')

Accenneremo infine al fatto che sia le righe dell'idrogeno, sia quelle degli ioni idrogenoidi, quando sono osservate con mezzi di sufficiente potere risolutivo, si rivelano composte ciascuna di un gruppo di diverse righe vicinissime, il che costituisce la cosiddetta *struttura fina* delle righe: la differenza di lunghezza d'onda delle varie componenti è, nell'idrogeno, al massimo dell'ordine di qualche decimo di Ångström e perciò è rivelabile solo con metodi di alta risoluzione: nell'elio ionizzato invece si giunge a qualche Ångström, e perciò la struttura fina di queste righe si può rivelare anche con un reticolo. Su questo fenomeno torneremo più avanti, per darne la spiegazione teorica.

Molte righe di altri spettri poi presentano una cosiddetta *struttura iperfina*, che richiede sempre mezzi di alta risoluzione e che ha un'origine completamente diversa dalla precedente.

§ 15. - SERIE SPETTRALI. TERMINI SPETTROSCOPICI.

Guidati dalla scoperta delle serie dell'idrogeno, gli spettroscopisti ricercarono delle regolarità analoghe negli spettri di altri elementi, ed infatti in parecchi altri casi riuscirono a scoprire che le righe si raggruppavano in *serie*, essendo caratterizzate le righe di una stessa serie da aspetto generalmente simile per struttura, intensità e nettezza, e da analogo comportamento di fronte ai vari agenti fisici (p. es. campi magnetici) che possono agire sulla sorgente; inoltre, trova-

di detta serie coincidono con le righe della serie di Balmer, come si constatò osservando che la (12) può scriversi anche

$$\tilde{\nu} = R \left(\frac{1}{4} - \frac{1}{(n/2)^2} \right).$$

In seguito, poichè la teoria di Bohr condusse a prevedere che l' He^+ deve emettere una serie siffatta, queste righe furono ricercate, e trovate, nello spettro emesso da un tubo contenente elio, assolutamente privo di idrogeno.

rono per molte serie delle formule empiriche analoghe a quella di Balmer.

In tutte queste formule — come in quella di Balmer — la frequenza di una riga si presenta come differenza di due *termini*, di cui il primo resta costante mentre il secondo assume diversi valori (in corrispondenza ai successivi numeri interi) per le diverse righe di una stessa serie. I termini possono essere assai spesso rappresentati con la formula (detta di Rydberg)

$$(13) \quad \tau_n = \frac{R}{(n - \alpha)^2}$$

dove R è la costante di Rydberg ed α un'altra costante: tenendo fissa α e dando ad n tutti i valori interi da un certo valore in poi, si ottiene una *successione* di termini. In particolare in questa formula rientrano i *termini balmeriani* (per i quali è $\alpha = 0$):

$$(14) \quad \tau_n = \frac{R}{n^2}.$$

Uno stesso elemento presenta in generale diverse successioni di termini, caratterizzate ciascuna da un valore della costante α , e le frequenze di una serie di righe si ottengono in generale combinando un termine fisso di una successione con tutti i termini di un'altra successione: vale a dire, la formula che dà le frequenze di una serie è del tipo

$$\tilde{\nu} = R \left[\frac{1}{(n' - \alpha')^2} - \frac{1}{(n - \alpha)^2} \right],$$

dove n' è fisso ed n assume tutti i valori interi da un certo valore in poi.

Per molti elementi però i termini τ_n hanno una forma più complessa che non quella di Rydberg: un secondo tipo p. es. è rappresentato dalla formula detta di Ritz

$$\tau_n = \frac{R}{(n - \alpha + K\tau_n)^2},$$

dove K è un'altra costante caratteristica di ciascuna successione.

Infine, in molti casi, pur non essendo possibile rappresentare i termini τ_n con delle formule semplici, si possono

tuttavia scrivere le frequenze delle varie righe spettrali sotto la forma

$$(13) \quad \tilde{\nu} = \tau_{n'} - \tau_n,$$

ric conducendo numerose frequenze della stessa sostanza a dipendere da un numero assai più limitato di *termini*. Si può quindi dire che il primo passo nell'interpretazione dello spettro di una sostanza è il ricavare dalla tabella delle frequenze delle sue righe spettrali, la tabella, assai più ridotta, dei «termini» che, combinandosi tra loro, danno luogo ad esse. Come vedremo nel § seguente, i termini di uno spettro hanno un significato fisico assai più importante che non le frequenze delle singole righe.

È evidente che, poichè da un numero relativamente piccolo di termini si ricava, per differenza, un numero assai più rilevante di frequenze, esistono in uno spettro parecchie copie di righe, per le quali la differenza delle frequenze è costante: questo fatto, che fu scoperto empiricamente dal RITZ, prese il nome di *principio di combinazione* e, suggerendo la ricerca dei termini, fornì la chiave dell'interpretazione degli spettri complessi.

§ 16. — CENNO SULLA TEORIA DI BOHR DELL'ATOMO DI IDROGENO.

Il fatto che le frequenze delle righe spettrali si presentino generalmente come differenze di due termini ha ricevuto nel 1913, dal BOHR (1), un'interpretazione che, confermata poi da altri fatti sperimentali, si è mostrata sempre più feconda e costituisce tuttora il fondamento della spettroscopia. Il Bohr propose tale teoria per interpretare la serie di Balmer e le altre affini ed a questo caso ci riferiremo nell'esporgla, ma il suo concetto fondamentale, che è quello di *livello energetico*, si può estendere a tutti gli spettri.

L'atomo di idrogeno, secondo il modello di Rutherford, consta di un nucleo e di un elettrone che descrive intorno a quello un'ellisse kepleriana, o, in particolare, un cerchio: per semplicità riferiamoci a questo caso come fece il Bohr, e,

(1) *Phil. Mag.* 26 (1913), p. 1.

poichè la teoria si applica a tutti i sistemi idrogenoidi, consideriamo addirittura il sistema idrogenoide di numero atomico Z . È noto dalla meccanica che il raggio r dell'orbita è determinato dalle condizioni iniziali del moto, e può essere qualsiasi: basta che la velocità v sia tale che la forza centrifuga eguagli l'attrazione elettrostatica del nucleo, cioè che $mv^2/r = Ze^2/r^2$, cosicchè v è legata ad r da

$$(14) \quad v = \sqrt{\frac{Ze^2}{mr}}.$$

L'energia del sistema (forza viva più energia potenziale) si può quindi esprimere in funzione della sola r , e risulta

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

(Il valore negativo di E significa che per disfare l'atomo allontanando l'elettrone a distanza infinita e riducendolo in quiete, cioè per ionizzarlo senza peraltro imprimere velocità all'elettrone, è necessario spendere lavoro).

Ora Bohr ammise che nella meccanica atomica esistesse una condizione supplementare, per cui è possibile il movimento non su tutti i cerchi ammessi dalla meccanica ordinaria, ma soltanto su alcuni di essi che si chiamano « orbite stabili » o « orbite quantiche », i cui raggi costituiscono una successione infinita ma « discreta » (cioè non continua) r_1, r_2, \dots . Tale condizione supplementare fu dal Bohr formulata nel modo seguente: il *momento angolare* (o momento della quantità di moto) $p = mrv$ deve essere un multiplo intero di $\frac{h}{2\pi}$:

$$(15) \quad p = n \frac{h}{2\pi} \quad (n = 1, 2, 3, \dots).$$

Ora, per la (14)

$$p = \sqrt{mrZe^2},$$

cosicchè, sostituendo nella precedente e risolvendo rispetto ad r , si ricava che il raggio dell'orbita n -esima, che si suol indicare con a_n , è

$$(16) \quad a_n = n^2 \frac{h^2}{4\pi^2 e^2 m Z}.$$

La fig. 7 rappresenta le prime tre orbite quantiche: la più interna ha (nel caso dell'idrogeno, $Z = 1$) il raggio

$$a_1 = \frac{h^2}{4\pi^2 e^2 m} = 0,528 \cdot 10^{-8} \text{ cm.},$$

(che risulta dell'ordine di grandezza richiesto dalla teoria cinetica dei gas per le dimensioni atomiche): le successive hanno raggi rispettivamente quadruplo, nonuplo, ecc. Queste orbite rappresentano, per così dire, dei binari ideali sui quali l'elettrone, secondo la teoria di Bohr, è obbligato a muoversi.

A ciascuna di queste orbite privilegiate corrisponde naturalmente un'energia

$$E_n = -\frac{1}{2} \frac{Ze^2}{a_n}:$$

l'energia dell'atomo può dunque assumere solo certi valori *discreti* E_n ,

che si chiamano *livelli energetici*, ciascuno dei quali corrisponde ad un diverso « stato » dell'atomo (*stato quantico*), caratterizzato, secondo Bohr, dal moto dell'elettrone su una o sull'altra delle orbite quantiche: dall'espressione (16) di a_n si ricava che questi livelli energetici sono dati da

$$(17) \quad E_n = -Z^2 \frac{2\pi^2 e^4 m}{h^2} \frac{1}{n^2}.$$

Il Bohr ammise poi che l'elettrone potesse passare, con un salto brusco (*salto quantico*) da un'orbita n a un'altra n' : sulle modalità di questo passaggio il Bohr non faceva nessuna ipotesi, fuorchè quella, che venisse rispettato il principio di conservazione dell'energia: perciò l'atomo deve assorbire o emettere l'energia corrispondente alla differenza

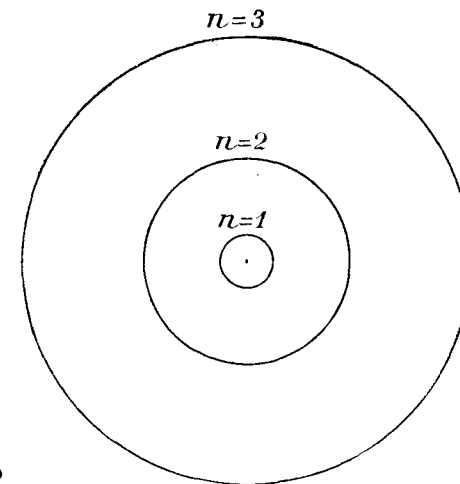


Fig. 7

tra E_n ed $E_{n'}$ (assorbire, se il passaggio avviene da un'orbita più interna ad una più esterna, emettere nel caso contrario): questa energia viene ordinariamente assorbita o emessa sotto forma di radiazione. Il meccanismo dell'emissione e dell'assorbimento, nella teoria di Bohr, non è quindi affatto quello della ordinaria teoria elettromagnetica (fondato sul fatto che, a norma delle equazioni di Maxwell, un elettrone in moto non uniforme e rettilineo irradia): anzi, viene esplicitamente negato che valga tale teoria nel campo atomico, poichè altrimenti l'emissione avrebbe luogo continuamente durante il moto dell'elettrone su un'orbita quantica, e quindi l'energia dell'atomo andrebbe diminuendo gradatamente il che è inconciliabile con l'esistenza dei livelli energetici discreti. L'emissione avviene invece, per così dire, a sprazzi, ogni volta che l'elettrone passa da un'orbita quantica n ad una più interna n' , e sul meccanismo di questa emissione non si fa alcuna ipotesi: soltanto una se ne fa, concernente la frequenza della radiazione emessa. Tale ipotesi è suggerita dal fatto, già rilevato da EINSTEIN (v. § 8) che in vari fenomeni (p. es., effetto fotoelettrico) la radiazione di frequenza ν si presenta raccolta in « quanti » contenenti ciascuno l'energia $h\nu$: identificando ognuno di questi quanti con uno sprazzo di radiazione emesso in un salto quantico, si è condotti ad ammettere che la quantità $E_n - E_{n'}$ di energia emessa, e la frequenza ν (in sec.^{-1}) di essa siano legate dalla relazione

$$(18) \quad \nu = \frac{E_n - E_{n'}}{h},$$

che è tuttora la relazione fondamentale della spettroscopia. Analoga relazione vale per l'assorbimento (salvo lo scambio di E_n con $E_{n'}$).

Dalla (18) si vede subito perchè le frequenze emesse dall'atomo si presentino come differenze di « termini spettrali », ed in pari tempo si riconosce il significato fisico di questi: ciascuno di essi corrisponde ad un diverso livello energetico dell'atomo, e quindi ad un diverso stato quantico. E precisamente, se si divide la (18) per c , per passare dalle ν alle $\bar{\nu}$,

e si pone

$$(19) \quad \tau_n = \frac{-E_n}{hc},$$

(con che, si noti, τ_n risulta positivo), si ottiene

$$(20) \quad \bar{\nu} = \tau_{n'} - \tau_n = Z^2 \frac{2\pi^2 e^4 m}{ch^3} \left(\frac{1}{n'^2} - \frac{1}{n^2} \right),$$

che ha la stessa forma della formula (11) rappresentante la serie di Balmer e le analoghe. Ma ciò che è più notevole è che, inserendo per e , m , h i loro valori numerici tratti da altri fenomeni fisici, il coefficiente $\frac{2\pi^2 e^4 m}{ch^3}$ della (20) diviene numericamente uguale (entro i limiti degli errori sperimentali) alla costante di Rydberg R : difatti (utilizzando i valori di e , e/m , h dati sopra) si trova $R = 109.740 \pm 170 \text{ cm}^{-1}$ mentre la esperienza dà $R = 109.677,76 \text{ cm}^{-1}$. La concordanza diviene poi maggiore se si tien conto della correzione dovuta al movimento del nucleo (v. p. II, § 58) che porta il valore teorico (per l'idrogeno) a 109.680 cm^{-1} . La teoria di Bohr dunque rende conto assai bene di tutte le righe dell'atomo di idrogeno e degli ioni idrogenoidi, cioè complessivamente di più di 50 righe, e fornisce per la costante di Rydberg, la espressione

$$(21) \quad R = \frac{2\pi^2 e^4 m}{h^3 c}.$$

In seguito, la teoria di Bohr è stata ampliata e perfezionata per opera soprattutto di SOMMERFELD il quale, prendendo in considerazione oltre alle orbite circolari anche quelle ellittiche, ed introducendo vari importanti ritocchi, ha potuto non solo rendere conto delle più minute particolarità degli spettri idrogenoidi, ma anche aprire la via allo studio degli atomi più complessi, che si è svolto su queste basi, con grande successo, fino al sorgere della meccanica quantistica. Il SOMMERFELD ha postulato, in luogo della condizione (15) di Bohr, delle condizioni più generali atte a caratterizzare le orbite quantistiche anche in casi più complessi di quello di

Bohr (*condizioni di Sommerfeld*)⁽¹⁾. Queste condizioni, che (al pari di quella di Bohr, che ne è un caso particolare) apparivano allora come postulati, a priori assai strani, giustificabili solo a posteriori attraverso l'esattezza, non sempre completa, delle loro conseguenze, si possono oggi dedurre come risultati di prima approssimazione dalla teoria di Schrödinger. Riservandoci perciò di presentare la teoria di Bohr e Sommerfeld, da questo punto di vista, nella parte II, passiamo ora a esaminare quali sono le idee fondamentali di quella teoria che conservano, anche nella nuova meccanica atomica, tutto il loro valore.

§ 17. - STATI QUANTICI E LIVELLI ENERGETICI.

Sebbene il modello dell'atomo proposto da BOHR sia stato in seguito profondamente modificato, tuttavia il postulato fondamentale di questa teoria, cioè l'esistenza di livelli energetici discreti, è oggi un fatto sperimentalmente accertato, e non solo per l'atomo di idrogeno, ma per tutti gli atomi e le molecole, come si vedrà nei paragrafi seguenti. Si deve cioè ritenere che *un atomo o una molecola sia un sistema meccanico capace di trovarsi in diversi stati (stati quantici) a ciascuno dei quali corrisponde un diverso valore dell'energia*: come rappresentazione intuitiva di questi stati si può pensare (se si vuole) che essi corrispondano al moto degli elettroni atomici su diverse « orbite quantiche » (generalizzazione dei cerchi privilegiati della teoria di Bohr): ma a tale interpretazione non deve oggi essere attribuito un carattere di realtà⁽²⁾. Ciò che diremo nei §§ seguenti è indipendente da tale modello, sebbene si ricorrerà sovente ad esso per aiutare l'intuizione e rendere più espressivo il linguaggio.

⁽¹⁾ Esse furono proposte, indipendentemente e quasi contemporaneamente, da W. Wilson (*Phil. Mag.* 29, 795 (1915)), Ishiwara (*Tokyo Math. Phys. Proc.* 8, 106 (1915)), A. Sommerfeld (*Ann. der Phys.*, 51, 1 (1916)).

⁽²⁾ Si può riguardare tale modello come i notissimi modelli idraulici che aiutano a comprendere le leggi fondamentali dell'elettricità, per cui, p. es., una corrente elettrica è paragonata ad un flusso di liquido in un tubo, ecc.

I diversi livelli energetici di un atomo (o di una molecola) si sogliono rappresentare graficamente mediante dei tratti orizzontali a varia altezza (v. fig. 8, che rappresenta i livelli dell'atomo di idrogeno): poichè l'energia è sempre negativa, essi sono sempre al disotto del livello zero, che corrisponde al limite di ionizzazione, e normalmente si addensano verso tale limite. Il livello più basso ($n = 1$) dicesi « livello fondamentale », perchè corrisponde alle condizioni in cui si trova ordinariamente l'atomo (*stato fondamentale*): gli altri stati si dicono *stati eccitati* (nel modello di Bohr, lo stato fondamentale corrisponde all'orbita più piccola).

Un'altra delle idee fondamentali di Bohr, che ancora conserva tutto il suo valore, è quella per cui un atomo può passare dall'uno all'altro stato assorbendo o emettendo la

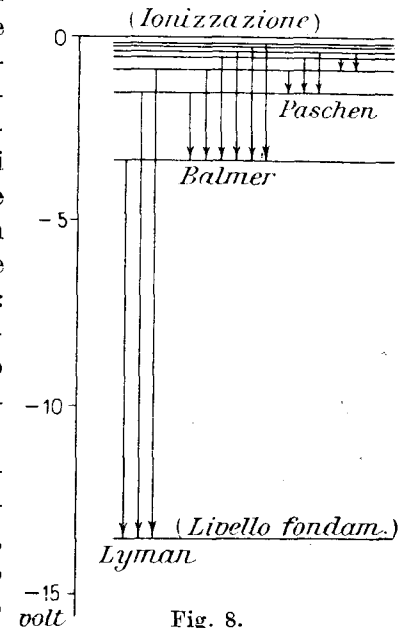


Fig. 8.

differenza di energia, la quale — quando è sotto forma di radiazione — ha la frequenza data dalla legge (18). Così si interpreta, per qualunque atomo o molecola, l'esistenza dei « termini spettrali », i quali corrispondono ai livelli energetici a norma della (19), e permettono quindi di costruire direttamente lo schema dei livelli energetici con soli dati spettroscopici. Nel diagramma si usa indicare con delle frecce verticali i passaggi che danno origine alle righe spettrali osservate: la frequenza di queste è quindi proporzionale alla lunghezza della freccia.

Si può dire che scopo precipuo della meccanica atomica è quello di giungere ad una previsione quantitativa dei livelli energetici degli atomi, il che si fa oggi con criteri diversi da quello di Bohr. Ma noi rimandiamo a più oltre queste con-

siderazioni per accennare ora agli altri fatti sperimentali che confermano l'esistenza di stati quantici discreti e l'esattezza della legge (18) riguardo ai passaggi dall'uno all'altro di essi con emissione o assorbimento di radiazione.

§ 18. — ECCITAZIONE E IONIZZAZIONE PER URTO ELETTRONICO.

Un atomo può acquistare energia non solo per assorbimento di radiazione ma anche se viene urtato da un altro atomo o da un elettrone: ciò significa, nel modello di Bohr, che un elettrone atomico, urtato da un altro corpuscolo, può passare dall'orbita fondamentale ad una più esterna (*eccitazione*), ricavando l'energia necessaria dalla forza viva della particella urtante ⁽¹⁾. E se l'urto è abbastanza violento, l'elettrone può venire addirittura sbalzato via dall'atomo, ossia può aversi la *ionizzazione* di questo: la ionizzazione è dunque un caso limite di eccitazione.

Le esperienze che descriveremo nei §§ successivi riguardano appunto l'eccitazione di atomi urtati da un elettrone: perciò premettiamo alcune considerazioni generali su tale fenomeno.

Osserviamo anzitutto che, data la massa grandissima che ha un atomo in confronto di un elettrone, la forza viva che esso riceve dall'urto di questo è completamente trascurabile, e quindi si può ragionare come se l'atomo rappresentasse per l'elettrone un ostacolo fisso: l'energia comunicatagli dall'urto serve interamente ad aumentare la sua energia interna, cioè a farlo passare dallo stato fondamentale ad uno stato eccitato. Ciò premesso, l'urto può avere diversi effetti secondo l'energia dell'elettrone urtante.

⁽¹⁾ L'eccitazione per urto ha una parte fondamentale nell'ordinaria emissione di radiazione per temperatura. Difatti, questa avviene nel modo seguente: gli urti dovuti all'agitazione termica eccitano alcuni degli atomi o delle molecole (a spese della forza viva, ossia dell'energia termica): questi atomi (o molecole), tornando poi allo stato fondamentale, emettono sotto forma di radiazione l'energia ricevuta.

Innanzitutto, è chiaro che l'urto non può produrre nessun effetto se l'energia dell'elettrone urtante è minore dell'energia che occorre per portare l'atomo nel *primo* degli stati eccitati, vale a dire se è minore di $E_2 - E_1$, il quale limite chiamasi, per ragioni che appariranno dal seguito, « *energia di risonanza* ». In questo caso dunque l'atomo rimane inalterato, e l'elettrone riparte come se avesse urtato un corpo perfettamente elastico: perciò tali urti si chiamano *urti elastici*.

Se invece l'energia cinetica dell'elettrone urtante supera, anche di poco, l'energia di risonanza, allora può avvenire che l'atomo urtato si ecciti, e che quindi l'elettrone riparta con una forza viva minore, e precisamente uguale all'eccesso della forza viva iniziale sulla energia di risonanza $E_2 - E_1$. Sarà quindi come se l'elettrone avesse urtato un corpo molle: tali urti si dicono *anelastici*. Se poi l'energia cinetica iniziale supera anche $E_3 - E_1$, allora potranno prodursi, oltre a questi, anche degli urti anelastici che portano l'atomo nello stato di energia E_3 e quindi sottraggono all'elettrone l'energia $E_3 - E_1$ (*seconda energia di eccitazione* ⁽¹⁾), e così via. **Infine** se l'elettrone urtante ha energia superiore a $-E_1$ (*energia di ionizzazione*), allora si produrranno anche degli **urti (anelastici) aventi per risultato la ionizzazione dell'atomo, e quindi l'apparizione di uno ione positivo oltre ai due elettroni liberi.**

Quando un atomo viene eccitato, esso di regola non rimane in tale stato, ma dopo un tempo brevissimo (che da esperienze di WIEN è risultato dell'ordine di 10^{-8} secondi) ritorna allo stato fondamentale, emettendo l'energia residua sotto forma di radiazione, di frequenza data dalla formula di Bohr (18). Si comprende dunque che quando degli atomi sono bombardati da elettroni di energia sufficiente ad eccitarli, essi debbono emettere della radiazione. E precisamente, se l'energia cinetica degli elettroni è supe-

⁽¹⁾ Ogni atomo ha evidentemente una serie di energie di eccitazione la prima delle quali è quella chiamata « di risonanza »; esse si addensano verso un limite superiore, che è l'energia di ionizzazione.

riore a $E_2 - E_1$ (ma non a $E_3 - E_1$), gli atomi potranno essere eccitati soltanto al livello E_2 , e quindi potranno emettere soltanto la radiazione di frequenza $\frac{1}{h} = (E_2 - E_1)$: la radiazione emessa, analizzata allo spettroscopio, darà dunque una sola riga. Se invece l'energia cinetica degli elettroni supera $E_3 - E_1$, allora alcuni atomi potranno essere eccitati allo stato E_2 , altri allo stato E_3 , e quindi potranno essere emesse le due righe di frequenza $\frac{1}{h} (E_2 - E_1)$, $\frac{1}{h} (E_3 - E_1)$; potrà anche eventualmente succedere che un atomo eccitato allo stato E_3 torni allo stato fondamentale in due tempi, e cioè prima passando allo stato E_2 e poi da questo allo stato E_1 : in tal caso verrà emessa, nel primo di questi passaggi, la riga di frequenza $\frac{1}{h} (E_3 - E_2)$. E così di seguito, aumentando la velocità degli elettroni urtanti, verrà emesso uno spettro sempre più completo. Si ha dunque, con l'osservazione spettroscopica, un altro metodo per controllare i fenomeni della eccitazione per urto.

Nei §§ seguenti mostreremo sommariamente come questi fenomeni si siano potuti verificare sperimentalmente in una numerosissima serie di lavori, tra cui sono rimasti classici quelli eseguiti nel 1913-14 da FRANCK ed HERTZ ⁽¹⁾: il principio fondamentale di tali esperienze risale però ad epoca anteriore, poichè fin dal 1902 LENARD lo aveva applicato in una notevole serie di ricerche sulla ionizzazione per urto.

Prima di descrivere le esperienze vogliamo però fare una osservazione generale per spiegare alcune locuzioni divenute abituali in questo campo di studi. Il dispositivo generalmente usato per lanciare degli elettroni con una data velocità consiste (v. fig. 9) in un filo metallico arroventato mediante il passaggio di una corrente il quale emette, per effetto termoionico, degli elettroni di piccola velocità: di fronte al filo vi è poi una sottile rete metallica, detta *griglia*, mantenuta (mediante una batteria di

⁽¹⁾ *G. Verh. d. Phys. Ges.*, 16, 10 (1914).

pile ed un potenziometro) ad un potenziale *positivo* rispetto al filamento, cosicchè gli elettroni emessi da questo si trovano in un campo elettrico che li accelera verso la griglia: essi passano quindi attraverso le maglie di questa e proseguono al di là per inerzia. Ora, se V è la d. d. p. tra filamento e griglia, e se si suppone trascurabile la velocità con cui gli elettroni sono espulsi dal filamento, si vede, applicando il teorema della forza viva, che essi giungono alla griglia con una energia cinetica uguale ad eV , e quindi con una velocità

$$(22) \quad v = \sqrt{\frac{2eV}{m}}.$$

Variando V mediante il potenziometro, si regola quindi a piacere la velocità v .

Questo modo di operare ha dato luogo all'abitudine di caratterizzare la forza viva e la velocità degli elettroni indicando addirittura il potenziale V , espresso in volt, che occorre a produrla, e di chiamare questo « *energia espressa in volt* ⁽¹⁾ », o, anche più impropriamente, « *velocità espressa in volt* ⁽¹⁾ ». Così p. es. si dirà che un elettrone ha la velocità di 1 volt, quando esso ha quella velocità che acquisterebbe cadendo attraverso alla d. d. p. di 1 volt: cioè la velocità di $5,931 \cdot 10^7$ cm. sec.⁻¹. Si osservi che la *velocità vera* non è proporzionale alla « *velocità espressa in volt* », ma alla sua radice quadrata (v. form. 22).

Per la stessa ragione, è invalso l'uso di caratterizzare l'energia di eccitazione e di ionizzazione degli atomi indicando la differenza di potenziale, in volt, necessaria a conferire agli elettroni tale energia: si parla quindi di « *potenziale di risonanza* », « *potenziale di eccitazione* », « *potenziale di ionizzazione* », ed anche i livelli energetici si esprimono spesso in volt. Dato uno di questi potenziali, V , si passa alla corrispondente energia E con la ovvia formula

$$(22') \quad E = eV,$$

dove e è la carica dell'elettrone in valore assoluto. Di qui si ricava che l'« *energia di 1 volt* » equivale a $1,590 \cdot 10^{-12}$ erg.

Di solito, i potenziali di risonanza, eccitazione e ionizzazione degli atomi sono dell'ordine di qualche volt: raramente superano i 20 volt.

Noteremo infine che ad ogni potenziale di eccitazione o di ionizzazione corrisponde (secondo la relazione di Einstein tra

⁽¹⁾ Talvolta dicesi: in *volt-elettrone*.

energia e frequenza) una frequenza ben determinata, data da

$$(23) \quad \nu = \frac{E}{h} = \frac{eV}{h},$$

la quale rappresenta la frequenza di un quanto di luce avente la stessa energia di un elettrone che è caduto attraverso la d. d. p. V . Se poi ν si esprime in cm^{-1} (e da noi sarà indicata al solito con $\bar{\nu}$) la (23) diviene

$$(23') \quad \bar{\nu} = \frac{eV}{hc}.$$

Molte volte i livelli energetici degli atomi si esprimono appunto in cm^{-1} secondo la (23'): un volt corrisponde a 8107 cm^{-1} .

Se invece si caratterizza la radiazione mediante la lunghezza d'onda λ , allora la relazione (23') va sostituita con la seguente

$$(24) \quad \lambda = \frac{hc}{eV},$$

da cui, sostituendo i valori numerici, si ha tra λ espresso in Å e V espresso in volt la relazione seguente, facile da ricordare:

$$(24') \quad \lambda V = 12335.$$

§ 19. - ESPERIENZE DI FRANCK E HERTZ.

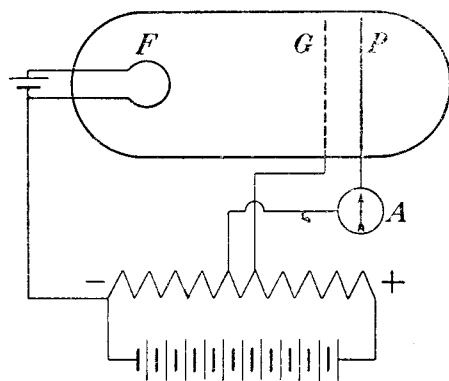


Fig. 9.

Tali esperienze consistono nel provocare l'eccitazione degli atomi per urto elettronico e nel misurare indirettamente la forza viva perduta dagli elettroni nell'urto, ossia l'energia comunicata all'atomo per eccitarlo: si hanno così le altezze dei diversi livelli al di sopra del livello fondamentale, ossia le energie di eccitazione $E_2 - E_1$,

$E_3 - E_1$ ecc.: il confronto dei valori trovati con quelli ricavati dai termini spettrali costituisce una verifica delle ipotesi fondamentali di Bohr.

Il dispositivo di Franck e Hertz è schematicamente il seguente (fig. 9). In un tubo di vetro contenente, allo stato di gas o vapore piuttosto rarefatto, la sostanza che si deve studiare, si trova il sistema per il lancio di elettroni costituito, come si è detto al § precedente, da un filamento incandescente F e da una *griglia* G , mantenuta ad un potenziale positivo, rispetto al filamento F , da una batteria di pile: un potenziometro permette di regolare questa differenza di potenziale, ed un voltmetro (non rappresentato in figura) permette di conoscerla esattamente. Il campo esistente tra F e G accelera gli elettroni, i quali si precipitano sulla rete metallica ed in parte la attraversano. Al di là della griglia, a circa 1 mm. di distanza, vi è una *placca* P , mantenuta ad un potenziale leggermente inferiore a quello della griglia, cosicchè gli elettroni in questo spazio sono lievemente rallentati: giunti sulla placca P , essi tornano al filamento attraverso un galvanometro A .

Se gli elettroni non subiscono urti anelastici, essi si muovono di moto accelerato fino alla griglia, dove raggiungono il massimo della loro velocità: tale massimo corrisponde ad una forza viva uguale ad eV (se V è la differenza di potenziale tra filamento e griglia): tale forza viva è più che sufficiente a far vincere loro il debole controcampo esistente tra griglia e placca, cosicchè quelli che passano tra le maglie della griglia arrivano tutti alla placca e sono registrati dal galvanometro.

Ma supponiamo ora di aumentare il potenziale V , finchè eV sia poco superiore ad $E_2 - E_1$: allora la forza viva degli elettroni, in vicinanza della griglia, sarà sufficiente a produrre degli urti anelastici: alcuni degli elettroni quindi perderanno in un urto quasi tutta la loro forza viva e non saranno più capaci di vincere il controcampo ed arrivare alla placca: dunque il galvanometro, appena eV supera il limite $E_2 - E_1$, segnerà un brusca diminuzione di corrente. Leggendo il potenziale, V' , per cui si ha questa discontinuità, potremo calcolare la prima energia di eccitazione (o di risonanza) con la formula

$$E_2 - E_1 = eV'.$$

Questo V' è quindi quello che abbiamo chiamato *potenziale di risonanza*.

Aumentando ancora il potenziale, la corrente va aumentando, finchè si produce una nuova brusca diminuzione per un valore V'' prossimo al doppio di V' . Infatti con l'aumentare di V succede che gli urti anelastici anzichè prodursi solo in vicinanza della griglia si producono già più indietro (cioè più verso il filamento), e più aumenta V più si arretra

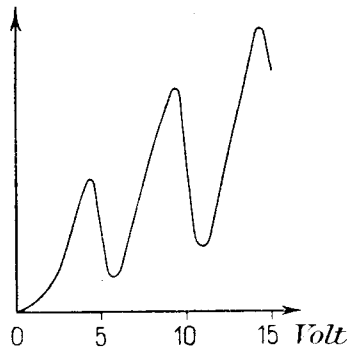


Fig. 10.

la regione dove avvengono tali urti, cosicchè gli elettroni che hanno perduto tutta la loro velocità in un urto possono venire ancora accelerati dal campo prima di arrivare alla griglia: quando V raggiunge $2(E_2 - E_1)$ alcuni di questi elettroni, dopo il primo urto anelastico, riacquistano nel campo un'energia $E_2 - E_1$ e quindi subiscono presso la griglia un secondo urto anelastico, in cui perdono di nuovo la loro forza viva e restano perciò incapaci di arrivare alla placca. Similmente si comprende che vi sarà una nuova diminuzione di corrente per $V = 3(E_2 - E_1)$ e così via. Infatti le curve sperimentali presentano l'andamento rappresentato dalla figura 10.

Di questo fatto si trae profitto per misurare con maggiore esattezza il potenziale di risonanza V' , ricavandolo dalla distanza tra due massimi consecutivi della curva, anzichè dalla posizione del primo massimo: si eliminano così quelle cause d'errore (p. es., differenze di potenziale di contatto) che producono uno spostamento complessivo di tutta la curva, o che perturbano la prima parte di essa.

Abbiamo ragionato finora come se l'atomo potesse assorbire solo l'energia $E_2 - E_1$, prescindendo quindi dai possibili passaggi dal livello fondamentale E_1 ai livelli E_3, E_4 ecc. E ciò è in molti casi lecito, e cioè quando gli urti sono così frequenti che quasi tutti gli elettroni, appena hanno rag-

giunto l'energia sufficiente per portare un atomo al livello E_2 , la perdono immediatamente in un urto e non arrivano così mai a provocare una eccitazione che richieda un'energia maggiore.

Variando le condizioni dell'esperienza si può tuttavia fare in modo che ciò non accada (occorre perciò diminuire la pressione del gas) e allora nella curva si presentano diverse serie di massimi, ogni serie corrispondendo a un livello di eccitazione.

Faremo infine osservare che lo stesso apparecchio permette di determinare direttamente il potenziale di ionizzazione: basta che tra la griglia e la placca, invece di una d. d. p. piccola rispetto al potenziale da determinare, si stabilisca una d. d. p. fissa e notevole, superiore a tutte quelle usate per accelerare gli elettroni (p. es. 40 volt): allora nessuno degli elettroni lanciati dal filamento potrà arrivare alla placca P , perchè saranno tutti ricacciati indietro dal campo esistente tra G e P . Ma se qualcuno degli atomi urtati, in vicinanza della griglia, si ionizza, esso, restando carico positivamente, si precipita sulla placca ed è registrato dal galvanometro. Basterà quindi aumentare gradatamente il potenziale acceleratore degli elettroni finchè il galvanometro comincia ad accusare corrente: questo potenziale V_i sarà quello occorrente a ionizzare il gas, vale a dire sarà tale che

$$eV_i = -E_1.$$

Questa descrizione schematica di tali esperienze sarà sufficiente a far comprendere il loro concetto fondamentale, ed a mostrare la possibilità di determinare per via elettrica i vari livelli energetici. Non possiamo descrivere tutti i particolari tecnici e le molteplici varianti di tali esperienze, che sono state eseguite, su varie sostanze, oltre che da FRANCK e HERTZ, già citati, da molti altri fisici: per essi rimandiamo al n. 28 della bibliografia.

§ 20. - CONTROLLO OTTICO DELLA ECCITAZIONE PER URTO.

Nelle esperienze descritte al § precedente, l'energia assorbita dall'atomo urtato è determinata indirettamente: un me-

todo più diretto di misurarla consiste, come si è accennato, nell'osservare la radiazione emessa dall'atomo tornando al suo stato fondamentale.

L'apparecchio consiste essenzialmente in un tubo di vetro o di quarzo contenente il solito filamento incandescente e la solita griglia, tra i quali si stabilisce un voltaggio acceleratore noto: si proietta poi l'immagine della regione, ove avvengono gli urti, sulla fenditura di uno spettroscopio, e si ricerca per quali valori della tensione appaiono le singole righe dello spettro. Questo metodo fu usato fino dal 1914 da FRANCK e HERTZ col vapore di mercurio: essi trovarono che la riga $\lambda = 2536,6 \text{ \AA}$ (*riga di risonanza*) di questo gas non appariva finchè la tensione era inferiore a 4 volt, ed era invece già fotografabile con la tensione di 5 volt. Misure più precise hanno dato, per il potenziale di risonanza del mercurio, 4,9 volt: dalla lunghezza d'onda della riga di risonanza si ricava, mediante la formula (24'), che il potenziale necessario ad eccitarla deve essere 4,87 volt, in perfetto accordo con quello trovato sperimentalmente.

In seguito esperienze analoghe sono state ripetute da molti altri osservatori e su diverse sostanze sempre con risultati pienamente conformi alla teoria (v. bibl., num. 28).

È da notare che molto spesso le radiazioni ottenute per urto elettronico sono ultraviolette. In tal caso è possibile rivelarle, oltrechè fotograficamente, anche servendosi del loro effetto fotoelettrico. E se si aumenta gradatamente il potenziale acceleratore degli elettroni, seguendo contemporaneamente il variare della corrente fotoelettrica, e si rappresenta graficamente questa dipendenza, si osserva che la curva ha delle inflessioni brusche in corrispondenza dei potenziali di eccitazione, poichè ad ognuno di essi corrisponde l'emissione di nuove righe e quindi un aumento dell'effetto fotoelettrico. È questo il metodo usato da HORTON e DAVIES, FRANCK e KNIPPING, FOOTE e MOHLER e altri.

Tutte queste esperienze, mentre stabiliscono uno stretto legame tra i fenomeni ottici e quelli della collisione elettronica, costituiscono una prova diretta, della realtà fisica dei

livelli energetici, e della validità della formula di Bohr che lega la frequenza emessa alla differenza di energia di due stati quantici.

§ 21. — URTI DI SECONDA SPECIE.

Come abbiamo visto, in un urto tra un elettrone e un atomo può avvenire che l'elettrone ceda parte della sua forza viva all'atomo sotto forma di energia di eccitazione. È stato osservato da KLEIN e ROSSELAND che, oltre a questa specie di urti, detti *di prima specie*, devono essere possibili anche degli urti rappresentanti il fenomeno inverso, vale a dire che un atomo già eccitato, urtato da un elettrone, si scarichi, senza irradiare, ma cedendo invece all'elettrone la sua energia di eccitazione sotto forma di forza viva, e che quindi l'elettrone rimbalzi con una velocità maggiore di quella che aveva prima. La necessità della esistenza di questi urti, chiamati *urti di seconda specie*, è stata dimostrata da KLEIN e ROSSELAND col seguente ragionamento termodinamico.

Supponiamo di racchiudere un gas formato da atomi e da elettroni in una cavità a pareti perfettamente riflettenti sia per gli atomi che per gli elettroni e per la radiazione, così che non siano possibili scambi di energia con l'esterno. In conseguenza degli urti di prima specie, gli elettroni cedono gradatamente la loro energia cinetica agli atomi: se non ci fossero urti di seconda specie, la forza viva media degli elettroni andrebbe continuamente diminuendo rispetto a quella degli atomi, cosicchè si avrebbero due gas mescolati, i quali, anzichè tendere ad un equilibrio di temperatura, se ne allontanerebbero, il che contraddice il secondo principio della termodinamica.

L'esistenza di urti di seconda specie è dunque una conseguenza termodinamica della esistenza, constatata sperimentalmente, degli urti di prima specie.

§ 22. — RISONANZA OTTICA.

Lo schema dei livelli energetici si presta ad interpretare — oltre ai fenomeni di ordinaria emissione ed assorbimento

— anche i fenomeni della risonanza ottica e della fluorescenza: il primo di questi è però interpretabile egualmente bene nello schema classico, come ora diremo.

Il fenomeno detto della risonanza ottica, scoperto da WOOD, consiste in questo: se si illumina un vapore metallico con luce di lunghezza d'onda esattamente corrispondente ad una determinata delle sue righe d'assorbimento, questa luce non viene propriamente « assorbita » dal gas (cioè trasformata in calore) ma viene *diffusa* quasi integralmente in tutte le direzioni. Caratteristica di questo fenomeno è la sorprendente intensità della diffusione anche da parte di un vapore molto rarefatto, e la quasi assoluta monocromaticità della luce diffusa.

L'esperienza, eseguita per la prima volta sul vapore di mercurio, si fa nel modo seguente: in un pallone di quarzo, accuratamente vuotato d'aria, si trova una goccia di mercurio, cosicchè il pallone contiene vapore di mercurio saturo alla temperatura ordinaria, e quindi a densità piccolissima. Se si espone questo pallone ai raggi di una lampada a mercurio (a pareti di quarzo), la quale emette intensamente la riga ultravioletta di lunghezza d'onda $\lambda = 2536,6 \text{ \AA}$, si trova che il pallone riemette intensamente questa luce che si può rivelare fotograficamente. E se si analizza allo spettrografo la luce della lampada dopo che ha attraversato il pallone, si trova che la riga 2536,6 è stata fortemente assorbita. Il fenomeno non avviene se la luce incidente ha una lunghezza d'onda diversa, sia pure pochissimo, da questo valore. Perciò tale riga si chiama *riga di risonanza*.

Il fenomeno si osserva anche con un pallone riempito di vapore di sodio, ed illuminato con la luce della riga gialla D del sodio (o anche di una sola delle sue componenti, $\lambda = 5890$ e 5896 \AA): in tal caso si ha il vantaggio che la luce è visibile senza fotografia, ed il pallone appare all'occhio vivamente splendente: è necessario però riscaldarlo, perchè la densità del vapore di sodio saturo alla temperatura ordinaria è troppo piccola.

L'interpretazione classica della risonanza, dalla quale è derivato il nome del fenomeno, è la seguente. Gli atomi del

vapore conterrebbero degli oscillatori aventi frequenze uguali a quelle che si riscontrano nella luce emessa. Se noi illuminiamo gli atomi con luce di una di queste frequenze, gli oscillatori corrispondenti, trovandosi sollecitati da un campo elettrico alternato con frequenza uguale alla loro propria, entreranno in vibrazione con ampiezza assai grande, a causa del fenomeno, ben noto anche in acustica, della « risonanza meccanica », e quindi emetteranno intensa luce in tutte le direzioni.

Invece la teoria dei quanti interpreta il fenomeno della risonanza nel modo seguente. Gli atomi che si trovano al livello fondamentale possono passare ad un livello superiore non soltanto sotto l'azione dell'urto di un elettrone, ma anche sotto l'azione di radiazione incidente, cioè di fotoni, purchè però ognuno di questi contenga esattamente l'energia necessaria all'eccitazione dell'atomo. Così, se l'atomo è colpito da luce di frequenza

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

esso può passare dallo stato 1 allo stato 2. Esso però normalmente non rimane in questo stato, come abbiamo già detto, ma torna allo stato fondamentale eseguendo il salto inverso, ed emette quindi l'energia che aveva assorbito, precisamente sotto forma di radiazione della stessa frequenza.

Questa teoria permette di spiegare molte particolarità del fenomeno della risonanza, che restano inesplicate nella teoria classica.

§ 23. — FLUORESCENZA.

Fenomeno affine alla risonanza, ma assai più comune e anticamente noto, è quello della fluorescenza, per cui certe sostanze, illuminate, riemettono in tutte le direzioni una luce di composizione spettrale diversa da quella della luce incidente: p. es. una soluzione di fluoresceina, o di estratto di corteccia di ippocastano (esculina) o di solfato di chinina, posta in una vaschetta di vetro ed illuminata con luce bianca, appare (se vista lateralmente) splendente di una luce az-

zurra o verdastra; una soluzione di clorofilla presenta invece una fluorescenza rossa.

Il fenomeno della fluorescenza obbedisce generalmente alla seguente legge, scoperta empiricamente da STOKES, del quale porta il nome: *la luce di fluorescenza non contiene frequenze superiori a quella della luce eccitatrice (che supponiamo monocromatica)*.

Questo fenomeno si spiega quantisticamente nel modo seguente. Gli atomi vengono, come si è detto, eccitati dalla luce incidente: se essi sono portati dal livello fondamentale E_1 a quello immediatamente successivo E_2 si ha la risonanza, di cui si è già parlato: se invece vengono portati in un livello energetico più elevato, p. es. E_3 , allora per tornare allo stato fondamentale essi possono eseguire due (o più) salti quantici successivi, e in ciascuno di questi emettono luce di una frequenza diversa: p. es. dal livello E_3 passeranno al livello E_2 con emissione della riga di frequenza $(E_3 - E_2)/h$, indi passeranno da E_2 ad E_1 con emissione della frequenza $(E_2 - E_1)/h$. Analogamente, se l'eccitazione li avesse portati ad un livello ancora più elevato potrebbero discendere con diversi salti successivi. Si comprende dunque come una sola frequenza eccitatrice possa dar luogo all'emissione di diverse frequenze, e si giustifica anche immediatamente la legge di Stokes, perchè è evidente che tutti i salti quantici eseguiti nella fase di emissione corrispondono a dislivelli più piccoli di quello corrispondente all'eccitazione, e quindi a frequenze minori di quella della luce eccitatrice.

§ 24. - FLUORESCENZA SENSIBILIZZATA.

Una significativa conferma della teoria quantistica della fluorescenza è fornita dalla seguente esperienza eseguita per la prima volta da CARIO ⁽¹⁾, studiata poi da FRANCK e CARIO e da altri. In un tubo di quarzo si trovano mescolati, a temperatura piuttosto elevata, del vapore di mercurio e del vapore di un altro metallo p. es. tallio. Si illumina il tubo

⁽¹⁾ ZS. f. Phys., 10, 185 (1922).

con la luce di lunghezza d'onda 2536,6 Å emessa da un arco a mercurio, e si osserva allo spettroscopio la luce diffusa lateralmente: si trova, oltre alla luce di risonanza del mercurio, un certo numero di righe caratteristiche del tallio. Queste invece non appaiono se si illumina del vapore di tallio non mescolato a quello di mercurio, o se la luce incidente ha frequenza diversa da quella della riga di risonanza del mercurio.

Questo fenomeno chiamato di *fluorescenza sensibilizzata* o *fluorescenza indiretta*, costringe a pensare che lo stato di eccitazione indotto in uno dei gas dalla luce possa trasmettersi all'altro gas, e precisamente si interpreta nel modo seguente. Molti atomi di mercurio assorbono la radiazione incidente e passano ad uno stato eccitato: se, mentre ancora si trovano in questo stato, avviene che essi urtino un atomo di tallio, possono scaricarsi, trasferendo la loro energia di eccitazione all'atomo di tallio e portandolo così in uno stato di eccitazione: l'atomo di tallio a sua volta, tornando in uno o più salti allo stato normale, emette le righe che gli sono caratteristiche.

Naturalmente, l'energia fornita dall'atomo di mercurio eccitato non è uguale a quella occorrente per eccitare l'atomo di tallio: di regola gli è superiore, e l'energia sovrabbondante viene comunicata all'atomo di tallio sotto forma di forza viva.

§ 25. - CENNO SULLO SVILUPPO DELLA SPETTROSCOPIA. L'ELETTRONE ROTANTE.

Non è, in generale, possibile calcolare quantitativamente, con la teoria di Bohr e Sommerfeld, i livelli energetici di un atomo con più elettroni, come si è potuto fare nel caso dell'idrogeno: infatti, già per il caso di due soli elettroni (atomo di elio) il problema presenta gravissime difficoltà matematiche: è stato trattato per approssimazione, e dei risultati di questi laboriosi calcoli parleremo tra poco. Nei casi più complessi le difficoltà matematiche diventano praticamente insormontabili: tuttavia si è potuto affrontare, almeno qualitativamente, il problema, grazie ad un'ipotesi semplifica-

tiva proposta da BOHR che si è mostrata in buon accordo con la realtà. Tale ipotesi consiste nell'ammettere che i salti quantici dai quali ha origine la radiazione interessino in generale un solo elettrone, e precisamente quello più facile da eccitare, ossia quello il cui livello energetico normale è più alto: questo elettrone si chiama « *elettrone ottico* »; gli altri si ammette che continuino a descrivere indisturbati le loro orbite intorno al nucleo, costituendo con questo un complesso invariabile che si suol chiamare *nocciolo* o *resto atomico*. Così il problema si riduce a studiare l'energia degli stati quantici dell'elettrone ottico, e differisce dal problema analogo per l'idrogeno in quanto l'elettrone ottico è soggetto alla forza esercitata da tutto il resto atomico e non solo dal nucleo. Questa forza è naturalmente assai complessa, e variabile col tempo: in generale ci si deve limitare ad una sua valutazione grossolana sostituendola con un campo centrale indipendente dal tempo. Il problema è così ridotto a quello di un moto centrale sotto l'azione di una forza non newtoniana: risolvendolo e applicando poi le tre condizioni di Sommerfeld (essendo il sistema a tre gradi di libertà) si giunge a trovare per l'elettrone ottico una serie di livelli energetici (dipendenti da tre numeri quantici (1)) che sono generalmente in buon accordo qualitativo con quelli osservati: si è ben lontani però dal poter raggiungere risultati quantitativamente esatti.

Tali considerazioni sono servite di base ad un enorme lavoro di interpretazione ed organizzazione di risultati sperimentali nel campo della spettroscopia: lavoro che consiste, in sostanza, nel coordinare ad ogni termine spettrale un gruppo di due numeri quantici n, l . Questa coordinazione conserva il suo valore anche indipendentemente dal modello meccanico di Bohr e Sommerfeld, e costituisce anche oggi la chiave dell'interpretazione degli spettri: difatti vi è una se-

(1) Essi sono detti: *quanto totale* n , *quanto azimutale* l , *quanto magnetico* m . L'ultimo non ha influenza sull'energia, eccetto il caso dell'effetto Zeeman: perciò i termini dipendono normalmente solo dai primi due.

rie di regole, scoperte in parte empiricamente, in parte con considerazioni teoriche (ed oggi giustificate in massima parte dalla meccanica quantistica) che permettono di dedurre dai numeri quantici corrispondenti ai due termini spettrali che danno luogo ad una riga, molte delle caratteristiche di questa (intensità, effetto Zeeman, effetto Stark, ecc.) (1).

Ma da questa ricerca dei numeri quantici corrispondenti a ciascun livello energetico è sorta la necessità di una ipotesi nuova. Si è constatato che tre numeri quantici (anzi due in assenza di campo magnetico) sono insufficienti a definire il livello energetico, e si è raggiunto uno schema soddisfacente solo quando si è introdotto un quarto numero quantico, chiamato « *quanto interno* », il quale però può assumere in ogni caso soltanto due valori. Restava il problema della sua interpretazione meccanica. Dopo una ipotesi di LANDÉ rivelatasi insufficiente allo scopo, si è mostrata pienamente soddisfacente un'ipotesi emessa nel 1925 da UHLENBECK e GOUDSMIT (2), e indipendentemente da BICHOWSKY e UREY (3), la quale ha avuto importanza decisiva nello sviluppo ulteriore della fisica atomica, mostrandosi sempre più feconda di risultati conformi alla esperienza. Tale ipotesi consiste nell'attribuire all'elettrone, oltre alla carica elettrica ed alla massa, anche un momento magnetico di valore uguale a $\frac{eh}{4\pi mc}$, cioè ad un *magnetone di Bohr*) ed un momento angolare (meccanico) avente la stessa direzione, ed avente il valore $\frac{h}{4\pi}$: come se si trattasse di una piccola trottola, il cui asse fosse magnetizzato. Anzi, in un primo tempo si pensò di poter interpretare queste proprietà pensando l'elettrone come una sfera elettrizzata ruotante intorno ad un asse: la rotazione della massa avrebbe prodotto il momento angolare meccanico, mentre la rotazione della carica elettrica, equivalendo ad un sistema di correnti circolari, avrebbe dato luogo al momento magnetico. Questo modello si è ben

(1) Vedasi p. es. il n. 23 della bibl.

(2) *Nature*, 117, 264 (1926); *Physica*, 5, 266 (1925).

(3) *Proc. Nat. Acad. Sc.*, 12, 80 (1926).

presto dovuto abbandonare, ma tuttavia l'ipotesi di Uhlenbeck e Goudsmit ha conservato il nome improprio di « ipotesi dell'elettrone rotante », (*spinning electron*, in inglese) ed il momento angolare di un elettrone si suole anche oggi designare col nome di *spin*. L'elettrone ha dunque perduto la simmetria sferica, per acquistare un asse privilegiato, quello del suo spin.

Applicando al momento angolare dell'elettrone una condizione analoga a quelle di Sommerfeld, si è condotti ad ammettere che, quando l'elettrone si trova in un campo magnetico, esso può assumere due sole orientazioni, e cioè con lo spin parallelo o antiparallelo al campo, a ciascuna delle quali corrisponde un diverso valore dell'energia: i due valori del quanto interno servono appunto a distinguere tra loro i due livelli energetici corrispondenti a queste diverse possibilità di orientazione.

L'ipotesi dell'elettrone rotante si è poi mostrata adattissima anche all'interpretazione delle proprietà magnetiche dei metalli: anzi, alcuni fenomeni di questa categoria (effetti giromagnetici) hanno permesso di misurare il rapporto tra il momento magnetico e quello angolare dell'elettrone, che è risultato conforme ai valori dati sopra.

Accenneremo infine al fatto che i metodi di interpretazione degli spettri atomici sono stati estesi con successo agli spettri emessi dalle molecole (*spettri di bande*), molti dei quali, specialmente di molecole biatomiche, sono oggi perfettamente interpretati mediante gli stessi principi teorici: noi però ci limitiamo alla considerazione della meccanica degli atomi, rimandando per le molecole ad altro volume.

§ 26. — LA CRISI DELLE TEORIE MODELLISTICHE.

La teoria di Bohr e Sommerfeld, sebbene sia servita di guida ai lavori teorici e sperimentali a cui abbiamo accennato nei §§ precedenti, ed a moltissimi altri, poggiava tuttavia su postulati così poco soddisfacenti (soprattutto per l'introduzione arbitraria del discontinuo) che non è mai stata considerata come espressione definitiva di una teoria fisica, ma

piuttosto come una provvisoria codificazione, per così dire, degli strappi da farsi alla meccanica ed all'elettromagnetismo classici per renderli applicabili al mondo atomico. Inoltre, mentre il concetto generale di livello energetico, e la legge di Bohr (18) sulle frequenze, si sono mostrati talmente aderenti, in tutti i casi, ai risultati sperimentali da dare la certezza che rispecchiano una effettiva realtà fisica, assai meno solide erano invece le prove sperimentali della esattezza del metodo di determinazione dei livelli mediante il calcolo delle orbite meccaniche e la successiva applicazione delle condizioni di Sommerfeld: anzi, il solo caso in cui questo metodo conduce a risultati quantitativamente esatti è quello dei sistemi idrogenoidi. Nel caso dell'elio e negli altri pochi casi in cui il metodo si è potuto applicare (superando con procedimenti di approssimazione le gravi difficoltà matematiche) si è giunti a livelli energetici in aperto contrasto con quelli sperimentali. Si è poi constatato che in certi casi (p. es. nella previsione degli effetti Zeeman anomali) le formule dedotte dalla teoria di Bohr e Sommerfeld richiedono dei lievi ritocchi (come la sostituzione di $l(l+1)$ in luogo di l^2 , dove l è un numero quantico) che si sono trovati empiricamente senza che si potesse vederne alcuna giustificazione nell'ambito di quella teoria.

Altre gravi difficoltà incontrava poi la teoria di Bohr e Sommerfeld nell'applicazione alla dispersione ottica, a certe particolarità dei fenomeni di collisione tra elettroni ed atomi (effetto Ramsauer), ed a varie altre questioni.

Mentre queste difficoltà si andavano accumulando nel campo della meccanica atomica, non meno grave appariva la situazione nel campo della teoria della luce, per la quale, come abbiamo già detto, si avevano due modelli, uno ondulatorio ed uno corpuscolare, ciascuno dei quali permetteva di interpretare esattamente una categoria di fenomeni, ma era incompatibile con l'altra categoria.

I diversi tentativi fatti per sostituire al modello atomico di Bohr e Sommerfeld e ai due modelli della luce delle rappresentazioni intuitive più coerenti furono tutti senza successo, cosicchè, verso il 1925, la situazione della fisica teorica era

caratterizzata da un diffuso senso di disagio e dalla sensazione che si rendesse necessaria una revisione profonda dei principi fondamentali della fisica teorica.

Il nuovo assetto logico che dovevano prendere questi principi è apparso chiaro soltanto qualche anno dopo, quando già la nuova meccanica quantistica aveva conseguito, su basi puramente formali, brillanti successi: ma già fin dal 1925 alcuni dei maggiori esponenti della fisica teorica si erano formata la convinzione che questa critica situazione derivasse dal presupposto, insito in tutte le teorie precedenti, che gli atomi, i fotoni e le altre entità del mondo atomico si dovessero concepire alla stregua di corpi o meccanismi immensamente piccoli, sia pure soggetti a leggi diverse da quelle dei corpi ordinari: essi possono invece essere enti essenzialmente diversi, cui non si applicano (e non c'è infatti alcuna ragione perchè debbano applicarsi) i nostri concetti intuitivi di corpo, movimento, ecc.: in altre parole, si considerò la possibilità che il mondo atomico, pur essendo soggetto a leggi esprimibili matematicamente, non fosse rappresentabile con nessun modello intuitivo (v. § 34).

Il nuovo indirizzo che, in conseguenza di queste nuove idee, venne ad assumere la fisica teorica dopo il 1925 si designa abitualmente col nome di *meccanica quantistica* ⁽¹⁾ (*quantenmechanik*): esso si sviluppò in diverse forme, delle quali nel capitolo seguente si darà un cenno puramente storico ed informativo, riservando alle altre parti del volume l'esposizione sistematica delle loro linee essenziali.

(1) Non si confonda questa espressione con quella oggi assai generica di « *Teoria dei quanti* », che abbraccia tutte le teorie nelle quali ha una parte essenziale la costante h di Planck: nella Teoria dei quanti perciò rientrano sia la teoria di Bohr e Sommerfeld (chiamata oggi talvolta « *teoria dei quanti classica* ») sia la « *meccanica quantistica* » nelle sue diverse forme (« *meccanica ondulatoria* », « *metodo delle matrici* », « *metodo degli operatori* »).

CAPITOLO IV.

LA MECCANICA QUANTISTICA

§ 27. — LA MECCANICA QUANTISTICA DI HEISENBERG (METODO DELLE MATRICI).

Il nuovo indirizzo, cui si è accennato alla fine del cap. precedente, fu inaugurato da HEISENBERG con una nota ⁽¹⁾ pubblicata nel luglio 1925. L'idea fondamentale in essa espressa è che alcune delle quantità inerenti al modello atomico in uso nella teoria dei quanti (p. es., le coordinate di un elettrone dell'atomo in un dato istante, la durata di una rivoluzione orbitale, ecc.) non sono mai state direttamente misurate e, dato che i ragionamenti su di esse fondati conducono alle note difficoltà, è lecito dubitare che queste quantità abbiano un effettivo significato fisico e siano in avvenire misurabili. Altre quantità invece (p. es., frequenze emesse, intensità, ecc.) sono direttamente osservabili. Perciò invece di cercare un modello geometrico e meccanico che permetta di ricavare i valori delle quantità osservabili da una struttura non osservabile, è meglio proporsi di collegare direttamente tra loro le quantità osservabili, senza far intervenire nessun modello.

Però le relazioni dirette tra grandezze osservabili non sono in genere esprimibili con i mezzi ordinari dell'algebra, e perciò l'ulteriore sviluppo dell'idea di Heisenberg condusse ad utilizzare un algoritmo matematico che già da tempo era conosciuto, ma che non aveva ancora avuto applicazione nel campo fisico, e cioè l'algebra delle *matrici*.

(1) *ZS. f. Phys.*, **33**, 879 (1925).

Questo metodo fu ampiamente sviluppato soprattutto da HEISENBERG, BORN e JORDAN, e condusse a ritrovare in molti casi i risultati della teoria di Bohr e Sommerfeld, mentre in altri casi condusse a risultati anche più esattamente concordanti con l'esperienza.

Ma non sta soltanto in questo il valore del progresso compiuto da HEISENBERG, bensì nell'aver potuto dedurre tutti i risultati, con unità di metodo, da un medesimo sistema organico di postulati, e quindi nell'aver sostituito ad una teoria fondata su basi parzialmente contraddittorie una teoria perfettamente coerente dal punto di vista logico. Il metodo delle matrici presenta però, accanto ai pregi già rilevati, il grave inconveniente di riuscire piuttosto difficile ad apprendersi e di non soddisfare i bisogni dell'intuizione. Ciò dipende dal fare uso di uno strumento matematico poco usuale ed alquanto complesso, come il calcolo delle matrici, ma soprattutto dalla voluta rinuncia a qualsiasi modello geometrico o meccanico, rinuncia ritenuta necessaria per poter enunciare in forma precisa e coerente le leggi del mondo atomico.

L'intima ragione per cui non è possibile fondare la meccanica atomica su un modello meccanico senza perdere in coerenza logica o in precisione, è stata messa in luce in un successivo lavoro di HEISENBERG, dal titolo « Ueber den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik » (Sul contenuto intuitivo della cinematica e della meccanica quantistica) ⁽¹⁾ nel quale si stabilisce il cosiddetto « principio di indeterminazione », che è, si può dire, la chiave di tutta la meccanica atomica, e che ha permesso di porre la meccanica quantistica nella sua vera luce. Su questo principio, e sulla conseguente impossibilità di un rigoroso modello atomico, torneremo più ampiamente nel seguito.

§ 28. — LA MECCANICA ONDULATORIA

Quasi contemporaneamente al metodo delle matrici sorgeva e si sviluppava un altro metodo di trattazione dei pro-

⁽¹⁾ *ZS. f. Phys.*, **43**, 172 (1927).

blemi atomici: tale teoria, chiamata *meccanica ondulatoria*, fu suggerita primitivamente da L. DE BROGLIE ⁽¹⁾, e poi fu sviluppata e posta in una nuova luce da E. SCHRÖDINGER, che ne espose i fondamenti in una serie di lucide memorie, pubblicate a partire dal febbraio 1926 ⁽²⁾.

Il metodo di Schrödinger prende le mosse dall'osservazione (che risale ad HAMILTON) che le leggi classiche della meccanica del punto si possono mettere in una forma analoga alle leggi dell'ottica geometrica (p. es. il principio della minima azione per il moto di un punto è analogo al principio di Fermat per i raggi luminosi). Alcune di queste analogie anzi si possono vedere anche senza far uso delle formule: p. es., in un mezzo omogeneo la luce si propaga con moto rettilineo uniforme, e così pure si muove un punto in un campo a potenziale uniforme. Se invece l'indice di rifrazione varia da punto a punto, i raggi luminosi sono curvi (come avviene nel miraggio) e la loro velocità non è più uniforme: altrettanto avviene del moto di un punto, se il potenziale è funzione del posto (campo di forza): il potenziale, come si vede, corrisponde, in questa *analogia*, all'indice di rifrazione.

Ora è noto che le leggi dell'ottica geometrica cadono in difetto in tutti quei fenomeni (diffrazione) nei quali intervengono schermi, fessure, fori, ecc. di piccole dimensioni, ossia di dimensioni comparabili con la lunghezza d'onda, e in tali casi si deve impiegare invece l'ottica ondulatoria, di cui l'ottica geometrica rappresenta solo una prima approssimazione. Analogamente — secondo Schrödinger — la meccanica classica del punto rappresenta solo una prima approssimazione di leggi meccaniche più generali: e se si vuol conservare il parallelo con l'ottica, si deve pensare che queste leggi siano analoghe a quelle dell'ottica ondulatoria. Esisterà dunque nei problemi meccanici una grandezza corri-

⁽¹⁾ Tesi presentata all'Università di Parigi, 1924; *Ann. de Phys.* **10**, **3**, 22, (1925).

⁽²⁾ *Ann. d. Phys.*, **79**, 361, 489, 734 (1926). Queste ed altre memorie fondamentali sono raccolte in un volume tradotto anche in francese (v. Bibl. n. 17 e 17^{bis}).

spondente alla lunghezza d'onda, e quando si ha a che fare con sistemi di dimensioni grandi rispetto a questa lunghezza, si potrà usare con approssimazione sufficiente la meccanica classica, mentre nei sistemi di dimensioni confrontabili con la lunghezza d'onda si produrranno dei fenomeni corrispondenti a quelli che in ottica si chiamano di *diffrazione*, e quindi sarà necessario servirsi della meccanica ondulatoria. Se dunque la lunghezza d'onda meccanica è di dimensioni atomiche (cioè dell'ordine di grandezza di un Ångström) si spiega come la meccanica classica si applichi bene ai corpi ordinari ma fallisca nella spiegazione dei fenomeni atomici. L'idea di Schrödinger consisteva dunque nel cercare di costruire una meccanica che stesse a quella classica nello stesso rapporto in cui l'ottica ondulatoria sta all'ottica geometrica.

Questa teoria si sviluppò e si precisò in un modo piuttosto singolare, in quanto che essa ebbe dapprima l'aspetto di un algoritmo puramente matematico operante su una grandezza astratta ψ , soddisfacente l'equazione caratteristica dei fenomeni ondulatori; in seguito SCHRÖDINGER cercò di collegare questa grandezza ad un modello fisico interpretandola come una espressione della densità elettrica: solo più tardi il BORN ⁽¹⁾ ne suggerì l'interpretazione probabilistica che oggi, al lume del principio di indeterminazione, si deve riconoscere come la sola legittima.

La teoria di Schrödinger, anche nella fase preliminare, in cui rappresentava un algoritmo di ignota o dubbia interpretazione, diede tuttavia subito l'impressione di rappresentare un enorme progresso sulla teoria di Bohr e Sommerfeld, e di toccare assai da vicino la profonda natura delle cose. Ciò dipese dal fatto che nella meccanica ondulatoria, invece di postulare l'esistenza di orbite privilegiate e quindi di livelli energetici discreti, si ricavava questa come conseguenza di un postulato assai più conforme alle nostre abitudini mentali: cioè che una certa funzione fosse dappertutto finita e continua.

In questa teoria la discontinuità nasce in modo del tutto

⁽¹⁾ *ZS. f. Phys.*, **38**, 803 (1926), e **40**, 167 (1927).

naturale dal procedimento matematico, in modo abbastanza simile a quello col quale, in acustica, si dimostra che un sistema vibrante può fornire solo delle note discrete.

Su tali principi, SCHRÖDINGER poté calcolare lo spettro dell'idrogeno, gli effetti Zeeman e Stark, l'oscillatore, ecc., e ottenne sempre risultati concordanti con quelli ottenuti col metodo delle matrici di Heisenberg. Questa coincidenza non era casuale, poichè ben presto lo stesso SCHRÖDINGER dimostrò che i due metodi, benchè abbiano avuto origine da concezioni diversissime, sono rigorosamente equivalenti, ossia rappresentano due forme diverse del medesimo procedimento matematico, e quindi debbono in ogni caso condurre ai medesimi risultati. Nella trattazione dei vari problemi si potrà dunque scegliere l'uno o l'altro metodo a seconda della opportunità: per una esposizione generale però il metodo di Schrödinger presenta il vantaggio di appoggiarsi maggiormente all'intuizione e di richiedere un apparato matematico meno eccezionale, e perciò noi ci serviremo di esso come introduzione ai metodi più generali della meccanica quantistica.

§ 29. — LA DIFFRAZIONE DEGLI ELETTRONI E DEGLI ATOMI.

Secondo la meccanica ondulatoria, il comportamento di un fascio di elettroni (p. es. un fascio di raggi catodici) è retto da leggi matematiche assai simili a quelle che governano la propagazione di un fascio di onde, come se gli elettroni fossero, per così dire, *guidati* da queste. Tali onde, che chiameremo « di *De Broglie* », hanno una lunghezza λ che dipende dalla velocità v degli elettroni stessi (o dalla tensione acceleratrice V) e che precisamente, come si vedrà al § 33 è data (per velocità piccole rispetto a c) dalla formula

$$(26) \quad \lambda = \frac{h}{mv} = \frac{h}{\sqrt{2emV}},$$

dove m è la massa di un elettrone. Se ne deduce, in particolare, che se il fascio di elettroni incontra un reticolo, esso deve essere diffratto come un fascio luminoso. Ma poichè

è praticamente impossibile, per ragioni sperimentali, ottenere un fascio ben definito di elettroni molto lenti, la lunghezza d'onda λ non potrà in pratica rendersi maggiore di un certo limite che è dell'ordine di qualche Ångström (corrispondente alla velocità di qualche decina di volt): perciò il parallelo si deve fare, piuttosto che con le onde luminose propriamente dette, con i raggi X, e quindi per la diffrazione si deve normalmente usare non un reticolo ottico ma un cristallo, analogamente alle esperienze di LAUE e di BRAGG.

Questa idea fu suggerita per la prima volta, sembra, da ELSASSER, nel 1925, subito dopo l'apparizione dei primi lavori di meccanica ondulatoria: ma l'esperienza della diffrazione degli elettroni da parte di un cristallo fu eseguita per la prima volta da DAVISSON e GERMER che la resero nota nel 1927 e che riuscirono in tal modo a misurare la « lunghezza d'onda elettronica » in funzione della velocità, trovando pienamente verificata la formula (26). Questa scoperta, che ha profondamente trasformato le abituali concezioni sulla natura dell'elettrone, costituisce la più diretta e mirabile conferma della teoria ondulatoria di Schrödinger. L'esperienza di DAVISSON e GERMER è poi stata modificata e perfezionata da molti altri ricercatori, ed oggi si è costituita una tecnica in un certo senso parallela a quella dell'analisi spettrale dei raggi X: anzi in alcuni casi si è trovato vantaggioso di impiegare la diffrazione di elettroni di nota lunghezza d'onda per lo studio di vari problemi di struttura, così come si fa da tempo, su vastissima scala, con i raggi X. Data la minore penetrazione degli elettroni, il nuovo metodo è particolarmente adatto a studiare la struttura degli strati superficiali del corpo diffrangente.

Aggiungeremo che non solo gli elettroni, ma tutte le particelle materiali lanciate in fascio contro un cristallo, danno luogo a fenomeni di diffrazione: ciò è stato provato sperimentalmente da STERN e suoi collaboratori, da JOHNSON ed altri, mediante i raggi molecolari di idrogeno molecolare, idrogeno atomico ed elio: anche per tali particelle si è confermata la validità della formula (26).

§ 30. — ESPERIENZE ESEGUITE COL METODO DI LAUE.

Descriveremo ora schematicamente le prime memorabili esperienze di DAVISSON e GERMER ⁽¹⁾.

Gli elettroni erano emessi (fig. 11) da un filamento di tungsteno incandescente F , e venivano accelerati dal campo creato tra F ed il diaframma D dalla batteria B : variando la

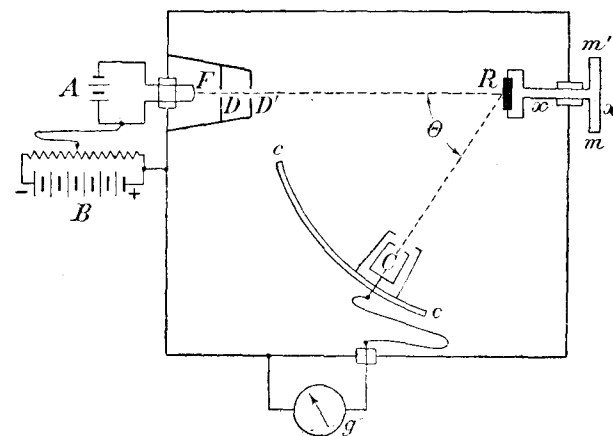


Fig. 11.

tensione di questa si poteva variare la velocità degli elettroni (nelle esperienze descritte la tensione variava da 30 a 370 volt). I diaframmi forati D e D' servivano a limitare un sottile fascio di elettroni, che cadeva normalmente sul cristallo R . Questo era un cristallo di nichel tagliato parallelamente ad una faccia di ottaedro ⁽²⁾. Gli elettroni erano rimandati indietro dal cristallo, e la loro distribuzione nelle varie direzioni si studiava raccogliendoli in un pozzo di Faraday C (protetto da uno schermo contro le azioni elettrostatiche, e munito di una stretta apertura in modo da raccogliere solo gli elettroni diffusi in una determinata dire-

⁽¹⁾ *Phys. Rev.*, **30**, 705 (1927); *J. Chem. Ed.*, **5**, 1041 (1928).

⁽²⁾ Il nichel cristallizza nel sistema monometrico, ed il reticolo cristallino è cubico a facce centrate. Le facce di ottaedro hanno il simbolo (111).

zione), il quale era collegato ad un galvanometro g : spostando il collettore C lungo l'arco graduato cc' ed osservando le indicazioni del galvanometro per le varie posizioni, si poteva studiare la distribuzione degli elettroni diffratti pei vari valori dell'angolo θ . Per studiare poi la distribuzione nei vari azimut si poteva far ruotare il cristallo intorno all'asse xx' , determinandone la orientazione mediante il cerchio graduato mm' . Tutto l'apparecchio era naturalmente racchiuso in un recipiente entro il quale era praticato un vuoto assai spinto.

L'esperienza così disposta è analoga a quella di Laue per i raggi X (con la sola differenza che invece di osservare i fasci diffratti al di là del cristallo si osservano quelli rimandati indietro) e si può farne la teoria sulle stesse basi che servono per l'interpretazione dell'esperienza di Laue. Tuttavia c'è una differenza notevole tra i due casi, ed è che, nel caso degli elettroni, è necessario tener conto del fatto che la lunghezza d'onda nell'interno del cristallo ha un valore λ' diverso da quello λ che ha fuori, e ciò a causa della diversa velocità di propagazione delle onde di De Broglie entro la materia e nel vuoto: tale differenza è invece trascurabile nel caso dei raggi X, pei quali l'indice di rifrazione di tutte le sostanze è praticamente uguale ad 1. La necessità di tener conto di questo fatto è stata messa in rilievo da BETHE, il quale mostrò come le esperienze di Davisson e Germer, correttamente interpretate, fossero in pieno accordo con la teoria di Schrödinger. L'esistenza dell'indice di rifrazione, oltre che prevista per ragioni teoriche, è dunque sperimentalmente dimostrata, e le esperienze di Davisson e Germer permettono anche di calcolarne il valore che chiameremo μ , e che risulta maggiore di 1, e tende ad 1 col crescere di v . Nel fare la teoria di queste esperienze, riterremo perciò che entro il cristallo la lunghezza d'onda sia non λ ma

$$(27) \quad \lambda' = \frac{\lambda}{\mu}.$$

Consideriamo ora (v. fig. 12) un fascio di elettroni che incide normalmente sulla superficie ss del cristallo. Pren-

diamo uno dei piani reticolari che si possono tracciare entro il cristallo, formante p. es. un angolo φ con la superficie: è noto dalla teoria di Bragg che esso, e tutti i piani reticolari ad esso paralleli, si comportano come superficie parzialmente riflettenti, ma in genere le onde piane da essi riflesse interferiscono distruggendosi mutuamente, tranne il caso che la lunghezza d'onda λ' abbia un valore tale che sia verificata la relazione di Bragg

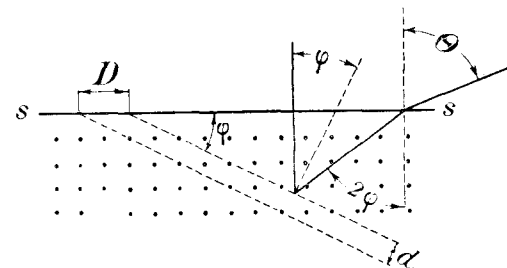


Fig. 12.

$$(28) \quad n\lambda' = 2d \cos \varphi$$

(dove d è la distanza tra due piani reticolari contigui ed n è un numero intero), ossia (v. (27)) per le lunghezze d'onda λ tali che

$$(29) \quad n\lambda = 2d\mu \cos \varphi.$$

Si ha in tal caso una *riflessione selettiva*.

Supposto ora che λ abbia uno di questi valori, per i quali avviene la riflessione, ricerchiamo la direzione in cui si propagheranno all'esterno del cristallo le onde riflesse. Uscendo, esse subiscono una rifrazione, e quindi il raggio emergente (normale alle onde emergenti) forma con la normale alla superficie un angolo θ dato da

$$\frac{\sin \theta}{\sin 2\varphi} = \mu,$$

poichè l'angolo di incidenza nella superficie è, come si vede dalla figura, 2φ .

Si ha quindi

$$\sin \theta = 2\mu \sin \varphi \cos \varphi.$$

Ricavando $\cos \varphi$, e sostituendolo nella (29) si ha

$$(30) \quad n\lambda = d \frac{\sin \theta}{\sin \varphi}.$$

Ora si osservi che, chiamando D la distanza tra le intersezioni dei due piani reticolari con la superficie ss (che è una costante del cristallo, nota dagli studi röntgenografici) si ha, come si vede dalla figura,

$$d = D \sin \varphi$$

e quindi la (30) diviene

$$(31) \quad n\lambda = D \sin \theta.$$

In questa equazione non entra μ : essa è la stessa che vale per i raggi X. È dunque possibile, dalla misura di θ , ricavare λ senza fare nessuna ipotesi sull'indice di rifrazione μ .

Ma se invece si vuole studiare per quali lunghezze d'onda un dato sistema di piani reticolari dà effettivamente luogo alla riflessione selettiva, allora si deve far uso della (29), nella quale interviene μ . Dunque la riflessione selettiva avviene per lunghezze d'onda diverse da quelle per cui avviene nei raggi X, ma, per ciascuna di esse, il fascio riflesso ha la stessa direzione che avrebbe se si trattasse di raggi X della medesima lunghezza d'onda.

In un cristallo esistono, come si sa, numerosissimi piani reticolari, ma di essi sono praticamente efficaci per la diffrazione soltanto quelli nei quali gli atomi sono sufficientemente fitti: a ciascuno di essi corrisponde un valore per il secondo membro della (28) che è una costante caratteristica del sistema di piani reticolari.

Perciò inviando i raggi normalmente al cristallo nel modo descritto, in generale non si avrà nessun fascio diffratto, tranne il caso che λ' abbia un valore tale che, moltiplicato per un intero n , dia un valore coincidente con la costante caratteristica di un piano reticolare realmente efficace. Perciò l'esperienza si dovrà condurre nel modo seguente: fissato un valore della tensione acceleratrice degli elettroni, e quindi della velocità di questi, e quindi di λ' , si esplora, portando il collettore C in diverse posizioni, la distribuzione degli elettroni rinviati: generalmente si trova una distribuzione continua (fenomeno analogo a quello della diffusione ottica, che si sovrappone sempre a quello della diffrazione). Si varia poi la tensione, e quindi λ' ,

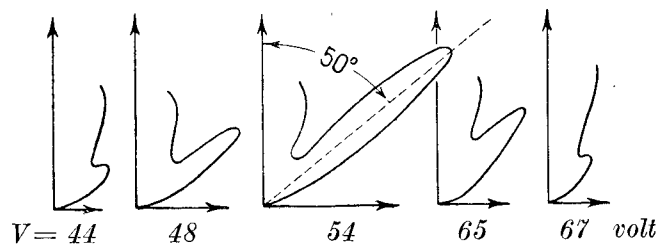


Fig. 13.

e si determina di nuovo la distribuzione: e così via, finché si trova una tensione, per la quale si presenta un massimo in una determinata direzione, ossia al fenomeno della diffusione si

sovrappone quello della diffrazione. La (31) permette allora di conoscere λ : così si ha la lunghezza d'onda corrispondente ad una data tensione e quindi ad una data velocità. Noto λ , mediante la (29) si può poi calcolare l'indice di rifrazione μ corrispondente a quella lunghezza d'onda.

La fig. 13 rappresenta i risultati di una di queste serie di esperienze: ogni diagramma polare è ottenuto riportando in ciascuna direzione un segmento proporzionale al numero di elettroni diffusi in quella direzione ed i diversi diagrammi corrispondono a diverse velocità degli elettroni incidenti: come si vede, per la velocità di 54 volt appare un fascio diffratto ben definito, nella direzione $\theta = 50^\circ$.

§ 31. - ESPERIENZE ESEGUITE COL METODO DI BRAGG.

DAVISSON e GERMER hanno anche eseguito delle misure con un metodo analogo a quello usato da BRAGG per i raggi X.

Si inviano gli elettroni sul cristallo non più normalmente ma con un angolo di incidenza θ (fig. 14). La direzione di propagazione, nell'interno del cristallo, forma con la normale un angolo θ' dato da

$$(32) \quad \sin \theta' = \frac{\sin \theta}{\mu}.$$

Le onde riflesse dai vari piani reticolari paralleli alla superficie ss di regola si distruggono tra loro salvo il caso che sia soddisfatta la relazione di Bragg.

$$(33) \quad n\lambda' = 2d \cos \theta',$$

dove d è la distanza tra i piani reticolari, misurata perpendicolarmente alla superficie, e $\lambda' = \lambda/\mu$.

Ricavando $\cos \theta'$ dalla (32) e sostituendolo nella (33) si ha

$$n\lambda' = \frac{2d}{\mu} \sqrt{\mu^2 - \sin^2 \theta}$$

e quindi

$$(34) \quad n\lambda = 2d \sqrt{\mu^2 - \sin^2 \theta}.$$

Nelle esperienze di Bragg coi raggi X si faceva oscillare il cristallo in modo da far variare θ entro un certo intervallo e si produceva una riflessione regolare ogni volta che θ passava per

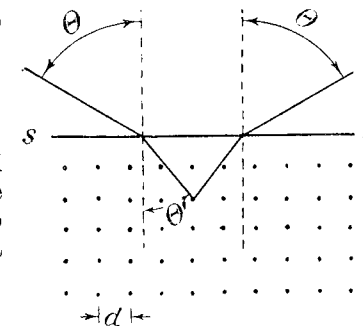


Fig. 14.

un valore soddisfacente la (34): invece nelle esperienze con gli elettroni è più comodo tenere fisso l'angolo di incidenza θ e far variare con continuità λ , variando la tensione acceleratrice V degli elettroni.

DAVISSON e GERMER (servendosi sempre di un cristallo di nichel tagliato secondo un piano 111) disegnarono, per diversi valori di V , il diagramma polare della distribuzione degli elettroni nelle varie direzioni, e trovarono che generalmente si ha una distribuzione pressochè uniforme, salvo per alcuni valori di V per i quali si presenta un massimo assai evidente in una direzione, esattamente corrispondente alla legge della riflessione ottica.

Se l'indice di rifrazione del cristallo per le onde elettroniche fosse 1, la (34) si ridurrebbe alla legge di Bragg, ossia la riflessione regolare si avrebbe per quelle λ per cui $n\lambda = 2d \cos \theta$, ovvero

$$(35) \quad \frac{1}{\lambda} = n \frac{1}{2d \cos \theta} = \text{cost.} \times n.$$

La fig. 15 mostra il diagramma delle intensità rilevate nella direzione corrispondente alla riflessione regolare (con incidenza di 10°) in funzione di $\frac{1}{\lambda}$, essendo λ calcolata dalla tensione V mediante la (26): le

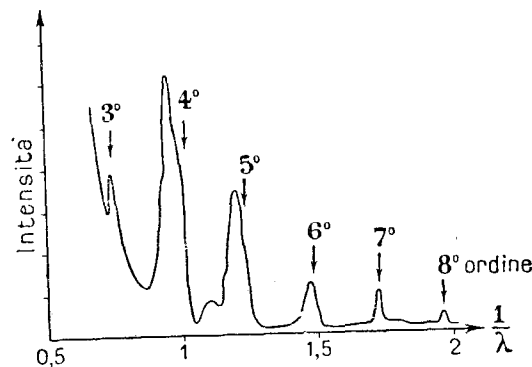


Fig. 15.

freccie rappresentano i punti dove dovrebbero aver luogo i massimi se valesse la (35). Dalla figura si vede bene l'influenza dell'indice di rifrazione, poichè, mentre per piccole lunghezze d'onda la (35) è abbastanza bene verificata, perchè l'indice di rifrazione è in tal caso prossimo ad 1, nelle grandi lunghezze d'onda i massimi sono invece notevolmente spostati, e questo spostamento può servire a calcolare μ in funzione di λ , ossia a stabilire sperimentalmente per le onde di De Broglie una legge analoga a quella della dispersione ottica.

Altre esperienze di diffrazione degli elettroni, sia col metodo analogo a quello di Laue che col metodo analogo a

quello di Bragg, sono state poi eseguite da vari autori con cristalli di diverse sostanze, ottenendo risultati sostanzialmente concordanti con questi.

§ 32. — ESPERIENZE ESEGUITE COL METODO DI DEBYE E SCHERRER.

Il metodo di DEBYE e SCHERRER per l'analisi dei raggi X e per lo studio della struttura dei cristalli consiste, come è noto, nel far cadere un sottile pennello di raggi X su un materiale cristallino ridotto in polvere, e nell'osservare i raggi diffratti: poichè nella polvere si vengono a trovare cristallini orientati in tutte le direzioni, ve ne sono sempre di quelli per cui è soddisfatta la relazione (35) e questi danno luogo a raggi diffratti formanti un cono avente per asse la direzione del fascio incidente. Così su una lastra fotografica posta al di là del tubetto contenente la polvere, si vengono a disegnare degli anelli concentrici dai cui raggi si possono dedurre le lunghezze d'onda presenti nel fascio.

Tale metodo è stato applicato agli elettroni da G. P. THOMSON, da RUPP e da altri, e si è mostrato per molti aspetti più vantaggioso di quello di DAVISSON e GERMER. Con esso THOMSON ha potuto estendere le misure di lunghezza d'onda anche ad elettroni assai più veloci di quelli di DAVISSON e GERMER, giungendo fino a tensioni di 20.000 volt.: RUPP invece ha eseguito le medesime esperienze con elettroni lenti (qualche centinaio di volt) rivelandoli sia fotograficamente, sia col metodo elettrometrico.

§ 33. — DIPENDENZA DELLA LUNGHEZZA D'ONDA DALLA VELOCITÀ.

Considerazioni teoriche suggeriscono, come vedremo al § 26, p. II, che la lunghezza d'onda λ dipenda dalla velocità v degli elettroni secondo una legge del tipo

$$(36) \quad \lambda = \frac{a}{mv},$$

dove a è una costante. Ora, se si osserva che la velocità v è legata alla tensione acceleratrice V dalla legge della forza

viva che è (per velocità piccole rispetto a c)

$$(37) \quad eV = \frac{1}{2} mv^2,$$

si vede che la (36) equivale a

$$(38) \quad \lambda = \frac{a}{\sqrt{2em}} \frac{1}{\sqrt{V}},$$

ovvero, chiamando λ_A la lunghezza d'onda in Ångström e V_v il potenziale in volt,

$$(38') \quad \lambda_A = 10^8 \sqrt{\frac{300}{2em}} \frac{a}{\sqrt{V_v}}.$$

Le misure della lunghezza d'onda λ_A (in Ångström) in funzione della tensione acceleratrice V_v (in volt) fatte con i vari metodi descritti hanno concordemente confermato la validità della formula (1)

$$(39) \quad \lambda_A = \frac{12,25}{\sqrt{V_v}}.$$

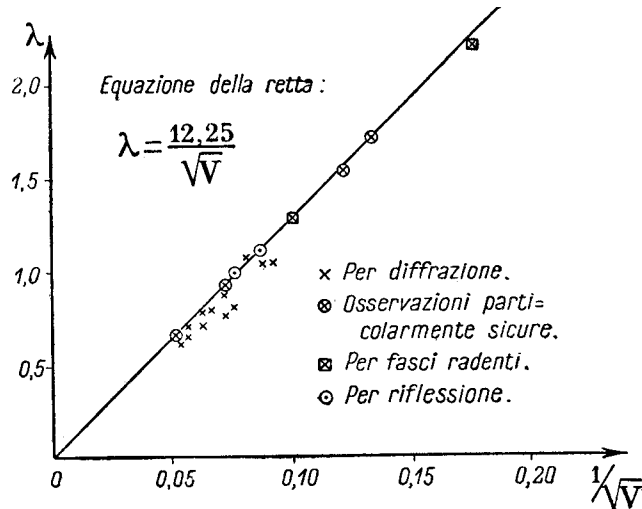


Fig. 16.

L'esattezza con cui questa legge è verificata è mostrata del diagramma della fig. 16, in cui i valori misurati di λ sono

(1) Vedasi p. es. i nn. 30 e 35 della bibl.

portati come ordinate in corrispondenza dei valori di $1/\sqrt{V_v}$ come ascisse: è disegnata la retta di equazione (36), e si vede che i punti sperimentali sono assai prossimi ad essa (entro gli errori sperimentali).

Identificando la (38') con la relazione sperimentale (39), si ricava per la costante a il valore

$$a = \sqrt{\frac{2em}{300}} 12,25 \cdot 10^{-8}$$

e introducendo i valori di e ed m si trova $a = 6,56 \cdot 10^{-27}$. È quindi giustificato identificare a con la costante di Planck h ($= 6,54 \cdot 10^{-27}$), con che la (37) diviene la già citata relazione (26) tra λ e v , che possiamo quindi ritenere fondata su solide basi sperimentali (1).

Convieni aggiungere che le considerazioni precedenti valgono se la velocità v è piccola rispetto a c : altrimenti la legge (37) della forza viva diviene (come si sa dalla teoria della relatività)

$$(40) \quad eV = m_0 c^2 \left(\frac{1}{\sqrt{1 - v^2/c^2}} - 1 \right)$$

e nella (26) la massa m diviene funzione di v secondo la legge

$$m = \frac{m_0}{\sqrt{1 - v^2/c^2}}.$$

Sostituendo nella (26), e ponendovi per v il valore ricavato da (40) si ha (trascurando potenze di $\frac{eV}{m_0 c^2}$ superiori alla prima)

$$(41) \quad \lambda = \frac{h}{\sqrt{2em_0 V}} \left(1 - \frac{1}{4} \frac{eV}{m_0 c^2} \right).$$

Questa formula (che per V piccolo coincide con la (38)) è stata verificata da esperienze eseguite da PONTE su elettroni veloci (fino a 17.250 volt) col metodo di DEBYE e

(1) Indipendentemente da queste, De Broglie aveva dato fin da principio una dimostrazione teorica della (26) basata su considerazioni relativistiche (v. bibl. n. 5).

SCHERRER, spingendo la precisione nella misura di λ fino al 3 per 1000 (la correzione relativistica ammonta a circa 1% per elettroni di 10.000 volt).

§ 34. — ASSETTO LOGICO DELLA MECCANICA ATOMICA. LA DEFINIZIONE OPERATIVA DEI CONCETTI FISICI.

Le esperienze di diffrazione degli elettroni, descritte nei §§ precedenti, hanno dimostrato chiaramente che non solo la radiazione elettromagnetica ma anche i raggi cosiddetti corpuscolari, come i raggi catodici, presentano un comportamento duplice, e cioè in certi fenomeni obbediscono al modello ondulatorio, in altri al modello corpuscolare: si stabilisce così un parallelismo ancora più stretto tra radiazione e materia, ed il problema dei rapporti tra i fotoni e le onde elettromagnetiche (v. § 13) si trasporta senz'altro ai rapporti tra gli elettroni e le onde di De Broglie. Nel caso della radiazione, furono scoperti prima i fenomeni a carattere ondulatorio e poi quelli di tipo corpuscolare; nel caso degli elettroni è avvenuto invece l'opposto: solo per questa ragione storica la nostra mente si è abituata a concepire la radiazione prevalentemente sotto l'aspetto ondulatorio e gli elettroni sotto quello corpuscolare, e il linguaggio usuale favorisce ancora di più questa tendenza, poichè si parla più di onde luminose che di fotoni, più di elettroni che di onde di De Broglie. Ma per essere perfettamente obbiettivi è necessario, sia nel caso della radiazione elettromagnetica che in quello dei raggi catodici, attribuire eguale importanza alle proprietà di tipo ondulatorio ed a quelle di tipo corpuscolare e persuadersi quindi che nè il modello delle onde, nè quello dei corpuscoli, concepiti nel modo usuale (cioè sulla analogia delle onde e dei corpuscoli visibili all'occhio) rappresentano in modo completo la natura di tali enti fisici.

Per parecchio tempo parve di vedere in questo fatto una inestricabile difficoltà, poichè sembrava di trovarsi di fronte ad enti dotati di proprietà contraddittorie: la soluzione della apparente antinomia è venuta solo attraverso ad un processo

di chiarificazione logica dei fondamenti della fisica, che ha accompagnato lo sviluppo formale della meccanica quantistica. Esso è stato guidato soprattutto dal seguente criterio logico fondamentale, che domina, più o meno esplicitamente espresso, tutta quell'opera di revisione dei principi della fisica che si può ritenere iniziata con la teoria della Relatività, e che è continuata sempre più ampia e profonda nella nuova meccanica atomica. La fisica ha per oggetto di eseguire osservazioni ed esperienze e coordinarne i risultati in uno schema quanto più è possibile semplice: quindi, suo oggetto immediato sono i dati di osservazione. È bensì lecito, per interpretare questi, introdurre degli altri enti, ma purchè questi siano definibili e rivelabili mediante esperienze, se non praticamente, almeno *concettualmente possibili*, cioè non vietate da alcuna legge fisica o logica. Enti o relazioni che non si possono rivelare mediante esperienze, anche ideali, ma concettualmente possibili, non possono essere introdotti nei ragionamenti senza correre il rischio di cadere in errore. Così ogni concetto della fisica deve essere suscettibile di una « *definizione operativa* », cioè poter essere definito mediante una serie di operazioni fisiche concettualmente possibili: p. es., la contemporaneità è definita indicando un metodo per decidere se due eventi sono o no contemporanei; l'elettrone è definito indicando un modo per rivelarlo: le coordinate di un elettrone, indicando un procedimento per misurarle, e così via. E le sole questioni fisiche che abbiano un senso sono quelle nelle quali ci si domanda il risultato di una o più esperienze che, almeno concettualmente, si potrebbero eseguire.

È opportuno ora chiarire il concetto di esperienze concettualmente possibili. Si considerano come tali non solo le esperienze effettivamente realizzabili, ma anche i casi limiti di esperienze realmente possibili, purchè ciò che si oppone a raggiungere il limite sia un complesso di difficoltà pratiche e non una legge generale: p. es., da molto tempo nei ragionamenti di termodinamica si fa uso corrente di cicli perfettamente reversibili, e ciò è lecito, poichè — sebbene in pratica essi non si possano ottenere — pure non si conosce

alcuna ragione teorica per cui ciò sia impossibile, e più si perfeziona la tecnica, più ci si avvicina a tali condizioni. Naturalmente, la scoperta di nuove leggi generali potrà anche, in futuro, modificare taluno di questi giudizi, e mostrare che non è concettualmente possibile una operazione prima ritenuta tale, o viceversa: perciò quando si parla di operazioni concettualmente possibili si sottintende « secondo le leggi fisiche attualmente conosciute ». Servendosi soltanto di tali operazioni per la definizione dei concetti, si è almeno certi di non introdurre involontarie contraddizioni con le leggi fisiche conosciute, e quindi di non costruire un edificio logicamente incoerente, come invece può avvenire — ed è avvenuto più volte — se si introducono dei concetti senza preoccuparsi di darne una definizione operativa: esempio ormai celebre è il concetto di tempo *assoluto*, usato dalla fisica pre-relativistica senza accorgersi che la sua definizione operativa richiedeva una trasmissione istantanea di segnali, la quale, nell'ambito dei fenomeni fisici finora noti, è concettualmente impossibile. Il punto di partenza della teoria della Relatività è stato appunto l'analisi del concetto di tempo assoluto, e della sua definizione operativa. Gli stessi metodi critici, applicati ai fenomeni atomici, conducono al principio di indeterminazione ed alla nuova meccanica quantistica.

L'applicazione rigorosa di questo criterio porta a risultati più sorprendenti di quello che a prima vista si potrebbe pensare. Il maggior vantaggio che ne è venuto è stato lo sgombrare il terreno da molte questioni le quali, ad una analisi accurata fatta coi criteri operativi, si sono mostrate prive di senso fisico. Inoltre, si è riconosciuto che la paradossale antinomia tra aspetto corpuscolare ed aspetto ondulatorio della radiazione, e quella analoga sorta più recentemente riguardo alla materia, non contengono una effettiva contraddizione logica, ma derivano solo dalla tendenza della nostra mente ad estendere al mondo atomico quelle rappresentazioni geometriche, cinematiche e causali che si sono formate in essa mediante la continua, inconscia esperienza accumulata nella osservazione del mondo macroscopico, e che chia-

miamo « intuizioni »: siamo così portati ad introdurre nel mondo atomico dei concetti che non sono suscettibili di definizione operativa, e da ciò derivano le accennate apparenti contraddizioni. In altre parole parlando di corpuscoli (elettroni o fotoni) noi siamo condotti a raffigurarci qualcosa di simile ad una pallina materiale straordinariamente impieciolita, e quindi siamo condotti ad attribuire a tali enti tutte le proprietà geometriche e cinematiche che hanno i corpi ordinari (continuità della traiettoria, ecc.) senza accorgerci che in tal modo postuliamo per i corpuscoli delle proprietà che non sono in alcun modo provate sperimentalmente, nè logicamente necessarie: non vi è dunque da meravigliarsi se tali proprietà sono in contrasto con quelle determinate dall'esperienza. Ma se invece al concetto di elettrone o di fotone associamo le sole proprietà determinabili sperimentalmente, possiamo verificare che queste, come è naturale, non sono in contrasto logico tra loro.

Non vi è dunque alcuna ragione perchè debba esistere un *modello* geometrico o meccanico accessibile alla nostra attuale intuizione e capace di interpretare completamente *tutti i fenomeni* del microcosmo. Ciò non toglie che i modelli (come quello di Rutherford-Bohr-Sommerfeld) siano in molti casi di grandissima utilità sia come mezzo euristico, sia come mezzo conciso di espressione, e spesso forniscano interpretazioni esatte o molto approssimate di vaste categorie di fenomeni: però non devono essere intesi alla lettera, bensì devono essere intesi alla stessa stregua delle analogie idrauliche che tanto aiutano a comprendere i fenomeni delle correnti elettriche.

Del resto, la meccanica quantistica fornisce anche la spiegazione del successo che hanno taluni modelli geometrico-meccanici nell'interpretazione di fenomeni atomici. Infatti le leggi della meccanica quantistica conservano taluni dei principi della meccanica ordinaria (p. es. il principio della conservazione dell'energia e della quantità di moto): inoltre vi è un passaggio continuo dalle leggi della micromeccanica a quelle della macromeccanica, per cui in molti casi queste possono servire come espressione approssimata (e in qualche caso esatta) di quelle.

§ 35. — IL METODO DEGLI OPERATORI E LA TEORIA DI DIRAC DELL'ELETTRONE.

A mettere in chiaro l'assetto logico della meccanica atomica hanno contribuito potentemente P. A. M. DIRAC e P. JORDAN, i quali hanno rielaborato l'intera materia da un punto di vista originale fondandosi soprattutto sul concetto fisico di « *stato* » e sul concetto matematico di « *operatore* »: essi così hanno potuto costruire un metodo unitario e generalissimo di trattazione dei problemi fisici (*teoria delle trasformazioni*) del quale il metodo delle matrici e il metodo della meccanica ondulatoria non sono che casi particolari.

Un altro risultato notevole ottenuto dal DIRAC è stato l'aver potuto dimostrare che, nello schema della meccanica quantistica, il momento meccanico e il momento magnetico dell'elettrone sono conseguenza necessaria della sua esistenza e del principio di relatività, cosicchè la cosiddetta « ipotesi dell'elettrone rotante » cessa di essere un postulato a sè, per venire dedotta come conseguenza delle equazioni fondamentali. Questa teoria verrà svolta nel cap. V della parte III.

PARTE II

**MECCANICA ONDULATORIA
DI UNA PARTICELLA**

CAPITOLO I.

INTRODUZIONE MATEMATICA

§ 1. — GENERALITÀ SULLE EQUAZIONI LINEARI OMOGENEE DEL SECONDO ORDINE.

Hanno grande importanza, nella meccanica ondulatoria, le equazioni differenziali (a derivate ordinarie) lineari, omogenee, del secondo ordine, cioè del tipo

$$(1) \quad A(x)y'' + B(x)y' + C(x)y = 0,$$

dove i coefficienti A , B , C sono funzioni analitiche della x , che supporremo reali per x reale. Poichè A è da supporre non identicamente nulla, si può sempre scrivere l'equazione nella forma

$$(1') \quad y'' + P(x)y' + Q(x)y = 0,$$

dove si è posto

$$P = \frac{B}{A} \quad Q = \frac{C}{A}.$$

Ricorderemo anzitutto che qualunque soluzione $y(x)$ è certamente regolare (ossia sviluppabile in serie di potenze intere e positive della x) per tutti i valori di x per cui sono regolari i coefficienti P e Q : solo nei punti dove uno almeno di essi presenta una singolarità può presentarsi una singolarità per la y . Tali punti si dicono *punti singolari*, o *singolarità*, dell'equazione e hanno notevole importanza nello studio delle sue proprietà: per il momento però li escluderemo dalle nostre considerazioni, riservandoci di parlarne in seguito.

Una notissima proprietà fondamentale delle equazioni di cui ci occupiamo è che, trovati due integrali particolari $y_1(x)$, $y_2(x)$, che siano indipendenti (cioè tali che il loro rapporto

non sia una costante), l'integrale generale si ottiene facendone una combinazione lineare mediante due costanti arbitrarie c_1, c_2 : esso è dunque

$$(2) \quad y(x) = c_1 y_1(x) + c_2 y_2(x).$$

Si dice allora che y_1, y_2 sono stati assunti come *integrali fondamentali*: vi è naturalmente larga arbitrarietà nella loro scelta, potendosi ad essi sostituire due loro combinazioni lineari qualunque, purchè indipendenti.

Nei problemi concreti, si deve disporre delle due costanti c_1, c_2 in modo che la y soddisfi a due altre condizioni imposte dal problema: p. es., che in un dato punto $x = a$ la y e la sua derivata y' assumano certi determinati valori (e si vede immediatamente che ciò è sempre possibile, e che la y ne resta univocamente determinata), oppure che la y assuma due dati valori in due dati punti a, b , o che passino due determinate relazioni tra i valori della y e della y' in a ed in b : in molti casi (ma non sempre) siffatte condizioni aggiunte all'equazione determinano univocamente la funzione y . Generalmente i due punti a e b sono sull'asse reale e sono i due estremi dell'intervallo entro il quale interessa la funzione $y(x)$: perciò si parla di *condizioni agli estremi*. A noi interessano soprattutto i due tipi seguenti di condizioni agli estremi:

a) la y si deve annullare ad entrambi gli estremi:

$$(3) \quad y(a) = 0, \quad y(b) = 0;$$

β) la y deve assumere gli stessi valori ai due estremi e così la y' :

$$(4) \quad y(a) = y(b), \quad y'(a) = y'(b).$$

(Entrambi questi tipi rientrano nella categoria delle « condizioni omogenee », alle quali si potrebbero estendere molte delle considerazioni che seguono: esse hanno questo di caratteristico, che se $y(x)$ è un integrale che le soddisfa, anche $cy(x)$, dove entra una costante arbitraria, è un integrale che soddisfa le medesime condizioni). Si osservi poi che siffatte condizioni sono soddisfatte, insieme all'equazione differenziale, se si prende $y = 0$ in tutto l'intervallo, ossia se si attribuisce ad entrambe le costanti della (2) il valore zero:

questa soluzione è evidentemente priva di interesse, cosicchè le nostre considerazioni saranno dirette a ricercare se vi sono anche soluzioni non nulle: acquista cioè interesse il caso (che vedremo essere eccezionale) in cui le condizioni al contorno (α) o le (β) *non* determinano univocamente un integrale dell'equazione data.

Consideriamo p. es. il caso delle condizioni (α). Utilizzando l'espressione (2) dell'integrale generale, si tratta di ricercare due valori, non entrambi nulli, per c_1, c_2 , tali che

$$(5) \quad \begin{cases} c_1 y_1(a) + c_2 y_2(a) = 0 \\ c_1 y_1(b) + c_2 y_2(b) = 0. \end{cases}$$

È noto dall'algebra che questo sistema di equazioni omogenee ammette soluzioni non nulle solo se

$$(6) \quad \begin{vmatrix} y_1(a) & y_2(a) \\ y_1(b) & y_2(b) \end{vmatrix} = 0.$$

Questa equazione vincola i valori, nei punti a e b , dei due integrali fondamentali y_1, y_2 , e naturalmente si ottiene un'equazione della stessa forma se si sostituiscono gli integrali y_1, y_2 con altri due integrali fondamentali.

Se si sanno esprimere esplicitamente $y_1(x), y_2(x)$ in funzione dei coefficienti della equazione, la (6) assume immediatamente l'aspetto di una condizione cui devono soddisfare questi coefficienti affinchè l'equazione ammetta, per le date condizioni agli estremi, una soluzione non nulla (e quindi, infinite soluzioni non nulle, ottenibili da quella con un fattore costante). In generale però non si possiede una tale espressione esplicita di y_1, y_2 : tuttavia la (6) esprime sempre, in forma implicita, una condizione (di tipo funzionale) per i detti coefficienti. Analogamente si può ragionare nel caso delle condizioni (β), e si trova allora in luogo della (6) la condizione:

$$(7) \quad \begin{vmatrix} y_1(a) - y_1(b) & y_2(a) - y_2(b) \\ y_1'(a) - y_1'(b) & y_2'(a) - y_2'(b) \end{vmatrix} = 0.$$

Nel caso delle condizioni (α), la condizione (6) può essere espressa in forma intuitiva nel modo seguente. Ogni integrale dell'equazione data, nel campo reale, si può rappresentare graficamente mediante una curva nel piano xy (v. fig. 17): per indi-

viduarla completamente basta assegnare un suo punto e la tangente ivi. Ora, consideriamo il fascio di tutte le curve integrali uscenti, in varie direzioni, dal punto A ($x = a$): è facile dimostrare che, disegnata una, qualunque altra si ottiene da essa mediante una dilatazione nel senso dell'asse y . Difatti, sia p. es.

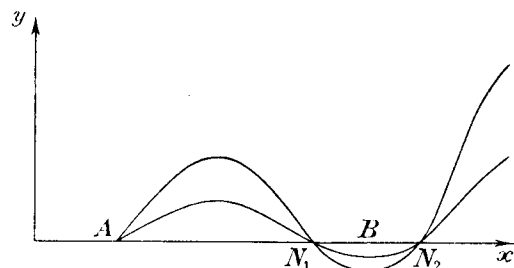


Fig. 17.

$y = f(x)$ l'equazione della curva che esce da A con la inclinazione 1 ($f'(a) = 1$): allora $y = \lambda f(x)$ (dove λ è una costante) sarà un altro integrale, e rappresenterà la curva che esce da A con l'inclinazione λ : dando a λ il valore opportuno si può così

ottenere una qualunque delle curve integrali uscenti da A . Da ciò segue, che tutte le curve integrali uscenti da A tagliano l'asse x negli stessi punti N_1, N_2, \dots (*modi*) corrispondenti alle radici dell'equazione $f(x) = 0$; in altre parole, variando l'inclinazione della tangente in A non si spostano i nodi, la cui posizione, fissato A , è determinata solo dalla natura dell'equazione, cioè dai suoi coefficienti. Se dunque ci si propone di ricercare un integrale soddisfacente alle condizioni (α), cioè tale che la curva che lo rappresenta passi, oltreché per A , per un altro punto prefissato B ($x = b$), si riconosce che *in generale* tale integrale non esiste: esso esiste solo se uno dei nodi N_1, N_2, \dots cade in B (e in tal caso, naturalmente, tutte le infinite curve integrali uscenti da A soddisfano la condizione richiesta). Ora, la condizione che uno dei nodi cada in B , cioè che $f(b) = 0$, equivale proprio alla condizione da noi trovata poc'anzi nella forma (6).

Osservazione. — Nel ragionamento precedente si è supposta la y reale, ma si può osservare che ogni soluzione complessa della (1) che soddisfi agli estremi le (α) o anche le (β) si compone di una parte reale e di una parte immaginaria che separatamente soddisfano l'equazione e altresì le condizioni agli estremi: perciò il ragionamento si può applicare a ciascuna di tali due parti.

§ 2. — EQUAZIONI CON UN PARAMETRO. AUTOVALORI ED AUTOFUNZIONI.

In molti casi (che sono quelli che interessano soprattutto la meccanica ondulatoria, e in genere la teoria delle oscillazioni di qualunque natura) nei coefficienti dell'equazione dif-

ferenziale (1) o (1') è contenuto un *parametro* λ , cioè una costante indeterminata: il caso più interessante è quello in cui λ figura linearmente nel coefficiente C , cosicchè l'equazione può scriversi nella forma seguente

$$(8) \quad A(x)y'' + B(x)y' + [\lambda\alpha(x) + \beta(x)]y = 0.$$

Allora gli integrali fondamentali y_1, y_2 conterranno anch'essi il parametro λ , e quindi lo conterrà anche il primo membro della (6) o della (7) che indicheremo con $\Delta(\lambda)$: si vede allora la possibilità di determinare λ in modo tale che la (6) o la (7) sia soddisfatta, e quindi che esistano soluzioni della proposta equazione differenziale soddisfacenti le volute condizioni agli estremi (α) o (β). A tale scopo bisognerà che λ sia una radice dell'equazione (in generale trascendente)

$$(9) \quad \Delta(\lambda) = 0.$$

Tali valori di λ si chiamano gli *autovalori* dell'equazione differenziale data relativi all'intervallo (a, b) e alle condizioni agli estremi (α) o (β). Trovato un autovalore λ_i , le (5) o le analoghe permettono di ricavare c_1 e c_2 in modo che la y soddisfi le volute condizioni agli estremi (anzi, si noti che le (5), una volta soddisfatta la (6), sono l'una conseguenza dell'altra, cosicchè determinano solo il rapporto c_2/c_1 , e si può scegliere c_1 a piacere). Un siffatto integrale della equazione differenziale si chiama un'*autofunzione* appartenente all'autovalore λ_i .

È da notare però che, salvo casi eccezionali, per questa via non si può praticamente giungere alla determinazione degli autovalori e delle autofunzioni, poichè, per poter impostare l'equazione (9), occorre conoscere l'integrale generale dell'equazione differenziale: le considerazioni precedenti servono solo a chiarire il concetto di autovalori e autofunzioni e a rendere plausibile la loro effettiva esistenza. Il problema della loro determinazione è un problema fondamentale per la teoria delle onde: non si conoscono metodi generali per risolverlo, ma solo degli artifici che in molti casi conducono al successo, alcuni dei quali verranno applicati nel seguito.

§ 3. - FORMA AUTOAGGIUNTA DI UN'EQUAZIONE.

Si dice che un'equazione differenziale del tipo (1) è in forma *autoaggiunta* se fra i coefficienti $A(x)$ e $B(x)$ passa la relazione $B = A'$, cosicchè i primi due termini formano insieme la derivata di Ay' , e l'equazione può scriversi

$$(10) \quad \frac{d}{dx} (Ay') + Cy = 0.$$

Si osservi che qualunque equazione del tipo (1) può mettersi in forma autoaggiunta: infatti, moltiplicando la (1) per il fattore, non nullo,

$$\frac{1}{A} e^{\int \frac{B}{A} dx},$$

i tre coefficienti A , B , C divengono rispettivamente

$$(11) \quad P = e^{\int \frac{B}{A} dx}, \quad Q = \frac{B}{A} e^{\int \frac{B}{A} dx}, \quad R = \frac{C}{A} e^{\int \frac{B}{A} dx},$$

e si verifica immediatamente che $Q = P'$, sicchè l'equazione si può scrivere

$$(12) \quad \frac{d}{dx} (Py') + Ry = 0.$$

Se poi C contiene linearmente un parametro λ , come nella (8), sarà così anche di R , e l'equazione (12) avrà la forma

$$(13) \quad \frac{d}{dx} (Py') + [\lambda e(x) + q(x)] y = 0.$$

Ora, il caso che si presenta più spesso, è quello in cui q è una costante, che si può prendere uguale ad 1, cosicchè la equazione, ridotta in forma autoaggiunta, è allora

$$(14) \quad \frac{d}{dx} (Py') + [\lambda + q(x)] y = 0.$$

(dove P , come risulta dalla (11), è sempre positivo). Nel seguito, ci riferiremo in genere, salvo indicazione contraria, ad equazioni di questo tipo. Si noti tuttavia che la presenza del coefficiente q non porterebbe modificazioni sostanziali (almeno se $q > 0$ in tutto l'intervallo considerato), ma solo una lieve complicazione delle formule.

Si può dimostrare ⁽¹⁾ che un'equazione del tipo (14), con le condizioni agli estremi (α), ammette sempre infiniti autovalori tutti reali, e che tale infinità è *discreta*, cioè numerabile, finchè l'intervallo (a, b) è, come qui si suppone, finito: inoltre (se P è, in tutto l'intervallo, positivo), la successione degli autovalori è limitata inferiormente mentre ha per limite superiore l'infinito. Ad ognuno di questi autovalori corrispondono una, o al più due, autofunzioni linearmente indipendenti (mai più di due, poichè qualsiasi altro integrale si può, come è noto, esprimere come combinazione lineare di quei due).

§ 4. - NORMALIZZAZIONE DELLE AUTOFUNZIONI.

Poichè un'autofunzione qualunque può essere moltiplicata per una costante arbitraria senza cessare di soddisfare le condizioni richieste, si hanno per ognuna di esse altre infinite autofunzioni non indipendenti tra loro, e che quindi non interessa di considerare come soluzioni distinte. Per togliere questa indeterminazione si aggiunge generalmente un'altra condizione, che chiamasi di *normalizzazione* e che è la seguente:

$$\int_a^b |y|^2 dx = 1,$$

ossia (designando, come faremo sempre, con l'asterisco il complesso coniugato)

$$(15) \quad \int_a^b yy^* dx = 1,$$

Tale condizione può essere sempre soddisfatta, perchè, detta $Y(x)$ una autofunzione che non la soddisfi, basta dividere questa per la costante non nulla

$$N = \sqrt{\int_a^b YY^* dx}$$

(fattore di *normalizzazione*), perchè l'autofunzione $y = \frac{1}{N} Y$ così ottenuta soddisfi la (15).

⁽¹⁾ v. Bibl., n. 25.

Si può osservare che la (15) è ancora soddisfatta se si moltiplica la y per una costante complessa di modulo 1 (cioè per un fattore della forma $e^{i\alpha}$): la normalizzazione quindi non determina completamente l'autofunzione. Tuttavia per lo più si prescinde da questa ulteriore arbitrarietà che si presenta nel campo complesso.

§ 5. - ORTOGONALITÀ DELLE AUTOFUNZIONI.

Dimostreremo ora una proprietà fondamentale delle autofunzioni.

Consideriamo due autofunzioni y_n, y_m , della (14), relative alle condizioni (α), ed appartenenti a due *distinti* autovalori λ_n, λ_m : esse soddisferanno identicamente le due relazioni

$$(16) \quad \frac{d}{dx}(Py'_n) + (\lambda_n + q)y_n = 0,$$

$$(16') \quad \frac{d}{dx}(Py'_m) + (\lambda_m + q)y_m = 0.$$

Quest'ultima sarà soddisfatta anche da y_m^* , complessa coniugata di y_m (essendo i coefficienti reali), cioè sarà

$$\frac{d}{dx}(Py_m^{*\prime}) + (\lambda_m + q)y_m^* = 0.$$

Moltiplicando questa per y_n e la (16) per y_m^* , e sottraendo, si ha

$$y_m^* \frac{d}{dx}(Py'_n) - y_n \frac{d}{dx}(Py_m^{*\prime}) + (\lambda_n - \lambda_m)y_n y_m^* = 0.$$

Si verifica immediatamente che l'insieme dei due primi termini è la derivata esatta di $P(y_m^* y'_n - y_n y_m^{*\prime})$, cosicchè, se si moltiplica tutta l'equazione per dx e si integra fra a e b , si ha

$$\left[P(y_m^* y'_n - y_n y_m^{*\prime}) \right]_a^b + (\lambda_n - \lambda_m) \int_a^b y_n y_m^* dx = 0.$$

Poichè agli estremi si annullano tanto y_n che $y_m^{*\prime}$, la prima parte è nulla: siccome poi si è supposto $\lambda_n \neq \lambda_m$, resta

$$(17) \quad \int_a^b y_n y_m^* dx = 0.$$

Tale relazione, di tipo integrale, tra le due autofunzioni y_n, y_m , si chiama (per un motivo che verrà spiegato al cap. I, parte III) relazione di *ortogonalità*: si dirà dunque che: *due autofunzioni della (14) soggette ad annullarsi agli estremi, ed appartenenti ad autovalori diversi, sono ortogonali*.

La stessa proprietà vale evidentemente anche se le condizioni agli estremi sono le (β), purchè il coefficiente P assuma gli stessi valori in a e b .

Si osservi che la (17) porta con sè la relazione coniugata

$$(17') \quad \int_a^b y_n^* y_m dx = 0.$$

Le relazioni di ortogonalità e quelle di normalizzazione si possono esprimere con una formula unica introducendo il simbolo, di uso frequente e assai comodo, δ_{nm} , col significato

$$\delta_{nm} = \begin{cases} 1 & \text{per } n = m \\ 0 & \text{per } m \neq n \end{cases}.$$

Allora la (15) e la (17) si compendiano nella formula

$$(18) \quad \int_a^b y_n y_m^* dx = \delta_{nm}$$

§ 6. - DEGENERAZIONE (AUTOVALORI MULTIPLI).

Ad un dato autovalore possono corrispondere, come si è detto, una o due autofunzioni linearmente indipendenti. Se ve ne corrisponde una sola, ciò significa che tutte le altre soluzioni corrispondenti a quell'autovalore si ottengono da questa moltiplicandola per una costante: se perciò si aggiunge la condizione di normalizzazione si può dire che all'autovalore corrisponde *una sola* autofunzione normalizzata (a meno della costante di modulo 1, di cui al § 4). In tal caso l'autovalore si dice *semplice*, perchè è una radice semplice dell'equazione $\Delta(\lambda) = 0$.

Consideriamo ora il caso ⁽¹⁾ che ad un autovalore corrispondano due autofunzioni linearmente indipendenti, nel

⁽¹⁾ Si riconosce immediatamente che questo caso si può presentare con le condizioni agli estremi (β) ma non con le (α).

qual caso l'autovalore si dirà *doppio* e si dirà, con locuzione divenuta abituale in fisica atomica, che vi è *degenerazione*. Queste due autofunzioni, normalizzate, siano $Y_1(x)$, $Y_2(x)$ (ricordiamo che esse in generale non risultano ortogonali tra loro, poichè il ragionamento del § 5 si applica solo ad autofunzioni appartenenti a due autovalori distinti). È chiaro che potremo da esse ottenere infinite altre autofunzioni, corrispondenti allo stesso autovalore, formandone una combinazione lineare con due coefficienti arbitrari. Dimosteremo ora che è sempre possibile, ed in infiniti modi, costruire due di tali autofunzioni che siano *normalizzate ed ortogonali tra loro*.

Difatti, una prima coppia di autofunzioni ortogonali può essere costituita dalla Y_1 stessa e da una opportuna combinazione lineare $\bar{Y}_2 = \alpha Y_1 + \beta Y_2$: basterà scegliere i coefficienti α , β in modo che risulti

$$\int_a^b (\alpha Y_1 + \beta Y_2) Y_1^* dx = 0,$$

cioè

$$\alpha + \beta \int_a^b Y_2 Y_1^* dx = 0,$$

il che fissa il rapporto tra α e β : si può poi disporre del valore di uno di essi per far sì che anche \bar{Y}_2 risulti normalizzata. Dalla coppia Y_1 , \bar{Y}_2 se ne possono poi ottenere infinite altre mediante le formule

$$\begin{aligned} y_1 &= c_{11} Y_1 + c_{12} \bar{Y}_2, \\ y_2 &= c_{21} Y_1 + c_{22} \bar{Y}_2, \end{aligned}$$

e si trova immediatamente che, affinché y_1 , y_2 risultino ortogonali e normalizzate, i coefficienti devono essere soggetti alle restrizioni

$$(19) \quad \begin{cases} c_{11} c_{11}^* + c_{12} c_{12}^* = 1, \\ c_{21} c_{21}^* + c_{22} c_{22}^* = 1, \\ c_{11} c_{21}^* + c_{12} c_{22}^* = 0. \end{cases}$$

Nel caso di coefficienti reali, queste non sono altro che le relazioni caratteristiche che intercedono, in geometria analitica, tra i coefficienti delle formule per la rotazione degli assi coordinati: tali coefficienti sono, come è noto,

$$(20) \quad \begin{cases} c_{11} = \cos \theta, & c_{12} = \sin \theta, \\ c_{21} = -\sin \theta, & c_{22} = \cos \theta. \end{cases}$$

Per tale ragione, una sostituzione lineare i cui coefficienti soddisfano le (19) si chiama una *sostituzione ortogonale*.

Potremo dunque dire che ad un autovalore doppio corrispondono infinite coppie di autofunzioni normalizzate ed ortogonali tra loro (oltrechè con le altre autofunzioni) le quali coppie si ottengono l'una dall'altra mediante una generica sostituzione ortogonale.

È opportuno spesso considerare un autovalore doppio (radice doppia dell'equazione $\Delta(\lambda) = 0$) come due autovalori E_1 , E_2 , coincidenti, ed a ciascuno di essi far corrispondere, nel modo spiegato sopra, una autofunzione normalizzata ed ortogonale all'altra; con questa convenzione si potrà dire che: ad ogni autovalore corrisponde una autofunzione, e che queste sono tutte ortogonali tra loro.

§ 7. — NODI DELLE AUTOFUNZIONI.

In molte questioni ha interesse la considerazione dei valori di x per i quali una autofunzione si annulla (*nodi* dell'autofunzione): enunceremo perciò qui (senza dimostrarlo) un teorema a questo riguardo (v. bibl. n. 25 o 34).

Consideriamo una qualunque soluzione dell'equazione (14) la quale si annulli in A , e sia rappresentata graficamente da una delle curve della fig. 17: abbiamo già rilevato (§ 1) che la posizione dei nodi dipende solo dai coefficienti dell'equazione. Se ora facciamo variare il parametro λ , in essi contenuto, i nodi si sposteranno e si può dimostrare che *aumentando con continuità λ , i nodi si spostano con continuità verso sinistra*. Ogni volta che uno di tali nodi viene a coincidere col punto B , la curva soddisfa le condizioni (a) e quindi rappresenta una autofunzione, ed il corrispondente valore di λ è un autovalore (relativo alle condizioni a).

Da ciò si comprende, che ogni volta che λ , crescendo, passa per uno degli autovalori, un nuovo nodo *entra* nell'intervallo AB . Indicando con $\lambda_1, \lambda_2, \dots$ gli autovalori (che supponiamo semplici) in ordine crescente, finchè $\lambda < \lambda_1$ non vi sono nodi entro l'intervallo AB , per $\lambda_1 < \lambda < \lambda_2$ vi è un nodo, per $\lambda_2 < \lambda < \lambda_3$ ve ne sono due, e così via. In particolare: l'autofunzione y_n ha $n - 1$ nodi entro l'intervallo AB , senza contare i due agli estremi, o, in forma più espressiva: *i nodi della autofunzione y_n dividono l'intervallo AB in n tratti*.

§ 8. — L'EQUAZIONE DEI MOTI ARMONICI.

Illustreremo le cose dette sulla seguente equazione (ben nota in meccanica, e detta «equazione dei moti armonici») di cui dovremmo occuparci nel seguito:

$$(21) \quad y'' + \lambda y = 0.$$

Essa rientra evidentemente nel tipo (14), ed in questo caso il procedimento del § 1 è praticamente effettuabile, perchè si conosce l'integrale generale della (21), che è notoriamente

$$(22) \quad y = c_1 e^{i\sqrt{\lambda}x} + c_2 e^{-i\sqrt{\lambda}x}.$$

L'intervallo in cui interessa l'integrazione sia $(-l, l)$, e consideriamo separatamente i due tipi, (α) e (β) , di condizioni agli estremi.

Condizioni (α) : debba essere ⁽¹⁾

$$y(-l) = y(l) = 0,$$

cioè

$$(23) \quad \begin{cases} c_1 e^{-i\sqrt{\lambda}l} + c_2 e^{i\sqrt{\lambda}l} = 0, \\ c_1 e^{i\sqrt{\lambda}l} + c_2 e^{-i\sqrt{\lambda}l} = 0. \end{cases}$$

Affinchè questo sistema di equazioni lineari ed omogenee in c_1, c_2 ammetta soluzioni non nulle, deve aversi

$$\begin{vmatrix} e^{-i\sqrt{\lambda}l} & e^{i\sqrt{\lambda}l} \\ e^{i\sqrt{\lambda}l} & e^{-i\sqrt{\lambda}l} \end{vmatrix} = 0.$$

cioè, sviluppando,

$$\sin 2\sqrt{\lambda}l = 0.$$

ossia

$$(24) \quad 2\sqrt{\lambda}l = n\pi \quad (n = 0, 1, 2, \dots).$$

Dunque λ dovrà avere uno dei valori

$$(24') \quad \lambda_n = n^2 \frac{\pi^2}{4l^2}$$

che sono gli autovalori cercati.

⁽¹⁾ Queste condizioni si presentano p. es. nel problema delle onde stazionarie in una corda fissa agli estremi. Un altro esempio si vedrà al § 38.

Determinato λ , si ricava da una qualunque delle (23)

$$(23') \quad c_1 = c_n e^{i\sqrt{\lambda_n}l}, \quad c_2 = -c_n e^{-i\sqrt{\lambda_n}l},$$

dove c_n è una costante arbitraria: sostituendo nella (22) si ha

$$y_n = c_n 2i \sin \sqrt{\lambda_n}(x+l) = C_n \sin n \frac{\pi}{2l}(x+l),$$

dove si è incorporato il fattore $2i$ nella costante arbitraria C_n . La condizione di normalizzazione ci dà poi

$$\int_{-l}^l y_n y_n^* dx = l C_n C_n^*,$$

da cui

$$|C_n| = \frac{1}{\sqrt{l}}.$$

Volendo la y_n reale si dovrà dunque prendere come autofunzione

$$(25) \quad y_n = \sqrt{\frac{1}{l}} \sin n \frac{\pi}{2l}(x+l).$$

Si verifica poi immediatamente il teorema di ortogonalità poichè, per $n \neq m$, si ha

$$\int_{-l}^l \sin n \frac{\pi}{2l}(x+l) \sin m \frac{\pi}{2l}(x+l) dx = 0.$$

L'esistenza di autovalori discreti in questo problema si può intuitivamente spiegare nel modo seguente: le soluzioni della (21) sono rappresentate graficamente da curve esponenziali se $\lambda < 0$, da sinusoidi se $\lambda > 0$: le prime hanno al più un nodo, le altre hanno infiniti nodi distanziati di $\pi/\sqrt{\lambda}$. Se si richiede che agli estremi cadano due nodi, si devono dunque scartare le curve esponenziali e si deve inoltre far sì che l'intervallo $2l$ sia multiplo della distanza tra due nodi $\pi/\sqrt{\lambda}$: questo dà la (24). L' n -esima autofunzione (25) è rappresentata da una curva sinusoidale che ha, oltre ai due nodi in A e B , altri $n-1$ nodi intermedi (v. § 7): la condizione di normalizzazione fissa poi l'ampiezza della sinusoidale.

Condizioni (β) : debba aversi invece

$$\begin{aligned} y(-l) &= y(l), \\ y'(-l) &= y'(l), \end{aligned}$$

cioè

$$(26) \quad \begin{cases} c_1 e^{-i\sqrt{\lambda}l} + c_2 e^{i\sqrt{\lambda}l} = c_1 e^{i\sqrt{\lambda}l} + c_2 e^{-i\sqrt{\lambda}l}, \\ c_1 e^{-i\sqrt{\lambda}l} - c_2 e^{i\sqrt{\lambda}l} = c_1 e^{i\sqrt{\lambda}l} - c_2 e^{-i\sqrt{\lambda}l}. \end{cases}$$

Queste condizioni si possono anche scrivere

$$(27) \quad \begin{cases} (c_1 - c_2) \sin \sqrt{\lambda}l = 0, \\ (c_1 + c_2) \sin \sqrt{\lambda}l = 0. \end{cases}$$

e quindi, se non sono entrambe nulle c_1 e c_2 , dovrà essere

$$\sin \sqrt{\lambda}l = 0,$$

da cui

$$(28) \quad \sqrt{\lambda}l = n\pi \quad (n = 0, 1, \dots).$$

Gli autovalori sono dunque

$$(28') \quad \lambda_n = n^2 \frac{\pi^2}{l^2}.$$

Scelto così λ , le condizioni agli estremi sono entrambe verificate, e quindi c_1 e c_2 restano arbitrarie: la condizione di normalizzazione dà soltanto

$$c_1 c_1^* + c_2 c_2^* = \frac{1}{2l}.$$

Si è dunque in un caso di degenerazione: gli autovalori sono doppi (eccetto l'autovalore 0).

Così una coppia di soluzioni indipendenti, ortogonali e normalizzate si ha (per $n \neq 0$) prendendo nella (22)

$$c_1 = \frac{1}{\sqrt{2l}}, c_2 = 0 \text{ o rispettivamente } c_1 = 0, c_2 = \frac{1}{\sqrt{2l}},$$

ed è

$$(29) \quad y_n^{(1)} = \frac{1}{\sqrt{2l}} e^{in\frac{\pi}{l}x}, \quad y_n^{(2)} = \frac{1}{\sqrt{2l}} e^{-in\frac{\pi}{l}x}.$$

Un'altra si ottiene prendendo

$$c_1 = \frac{1}{2\sqrt{l}}, c_2 = \frac{1}{2\sqrt{l}}, \text{ o rispettivamente } c_1 = \frac{1}{2i\sqrt{l}}, c_2 = \frac{-1}{2i\sqrt{l}},$$

ed è

$$(30) \quad y_n^{(1)} = \frac{1}{\sqrt{l}} \cos n\frac{\pi}{l}x \quad y_n^{(2)} = \frac{1}{\sqrt{l}} \sin n\frac{\pi}{l}x.$$

(Per l'autovalore 0 si ha invece l'unica autofunzione $y_0 = \frac{1}{\sqrt{2l}}$).

Da una di queste coppie se ne ricavano infinite altre con sostituzioni ortogonali: p. es. dalla (29), con la sostituzione (20) si ricava la coppia

$$y_n^{(1)} = \frac{1}{\sqrt{l}} \cos \left(n\frac{\pi}{l}x + \theta \right), \quad y_n^{(2)} = \frac{1}{\sqrt{l}} \sin \left(n\frac{\pi}{l}x + \theta \right).$$

con θ arbitrario.

Questi risultati si potevano prevedere intuitivamente osservando che le condizioni (β) impongono alla sinusoidale di avere la stessa ordinata e la stessa inclinazione in A ed in B (e quindi l'intervallo AB deve essere un multiplo della lunghezza d'onda $2\pi/\sqrt{\lambda}$, donde la (28)) ma resta arbitraria la fase θ della sinusoidale in A .

§ 9. - SVILUPPO DI UNA FUNZIONE IN SERIE DI AUTOFUNZIONI. SERIE DI FOURIER.

Consideriamo la successione delle infinite autofunzioni normalizzate e ortogonali tra loro, y_1, y_2, \dots , corrispondenti agli autovalori dell'equazione (14), con le condizioni agli estremi (α) o (β), e con la convenzione (spiegata al § 6) che gli eventuali autovalori doppi vengano contati per due. Si può dimostrare che, data una qualsiasi funzione $f(x)$ che soddisfi alle stesse condizioni agli estremi e inoltre a certe condizioni qualitative assai poco restrittive, è possibile svilupparla, entro l'intervallo (a, b) , in una serie della forma

$$(31) \quad f(x) = \sum_1^{\infty} f_n y_n(x).$$

la quale risulta uniformemente convergente nell'intervallo (a, b) .

AmMESSO questo teorema (1), si riconosce subito che i

(1) V. bibl. n. 25 o n. 34. Più generalmente vale il seguente teorema: se la funzione f è tale che esista $\int_a^b |f|^2 dx$, la serie (31) è almeno in media convergente, il che significa che

$$(31^*) \quad \lim_{N \rightarrow \infty} \int_a^b \left| f - \sum_1^N f_n y_n \right|^2 dx = 0.$$

Sviluppando il quadrato che figura in questa formula, ed utiliz-

coefficienti f_n risultano espressi da

$$(32) \quad f_n = \int_a^b f(x) y_n^*(x) dx :$$

difatti, se si moltiplicano i due membri della (31) per $y_r^*(x)dx$ (dove r è uno qualunque degli indici $1, 2, \dots$) e si integrano fra a e b , al secondo membro si annullano, in virtù dell'ortogonalità delle y , tutti i termini tranne quello per cui $n = r$, il quale si riduce a f_r : cambiando poi l'indice r in n si ha la (32).

Questo procedimento, come si vede, non è che una generalizzazione del noto procedimento che serve a determinare i coefficienti dello *sviluppo di Fourier*, il quale difatti non è che un caso particolare di sviluppo in serie di funzioni ortogonali, e precisamente nelle autofunzioni della (21) relative alle condizioni agli estremi (β). Difatti, prendendo tali autofunzioni nella forma (30) ed incorporando nei coefficienti f_n (che denoteremo con a_n e b_n) la costante $1/\sqrt{l}$, lo sviluppo (31) si scrive

$$(33) \quad f(x) = \sum_0^{\infty} \left(a_n \cos n \frac{\pi}{l} x + b_n \sin n \frac{\pi}{l} x \right),$$

dove, conformemente alle (32),

$$(34) \quad \begin{cases} a_0 = \frac{1}{2l} \int_{-l}^l f(x) dx \\ a_n = \frac{1}{l} \int_{-l}^l f(x) \cos n \frac{\pi}{l} x dx, \quad b_n = \frac{1}{l} \int_{-l}^l f(x) \sin n \frac{\pi}{l} x dx. \end{cases}$$

È questo lo *sviluppo di Fourier*: esso rappresenta la funzione $f(x)$ entro l'intervallo $(-l, l)$ anche se essa ha in esso dei

zando la (32) e la proprietà di ortogonalità, si trova l'importante formula di Parseval

$$(31^{**}) \quad \int_a^b |f|^2 dx = \sum_1^N |f_n f_n^*|,$$

che esprime la *completezza* del sistema di autofunzioni y_n (difatti, se si considerasse un sistema qualunque di funzioni ortogonali, sia pure infinito, ma non rappresentante la *totalità* delle autofunzioni di un'equazione differenziale, varrebbe nella (31**) il segno $>$ invece che $=$).

punti di discontinuità (nei quali la serie rappresenta la media aritmetica dei due limiti a destra ed a sinistra). Per la validità dello sviluppo, basta che l'intervallo si possa dividere in tratti entro ciascuno dei quali la f è continua e monotona.

Lo stesso sviluppo si può ottenere in una forma più comoda usando le autofunzioni (29), che si possono raccogliere nell'unica formula

$$(35) \quad y_n = \frac{1}{\sqrt{2l}} e^{in \frac{\pi}{l} x},$$

purchè si convenga che n può assumere anche valori negativi: allora ponendo $f_n/\sqrt{2l} = C_n$, la (31) dà (*sviluppo di Fourier in forma esponenziale*)

$$(36) \quad f(x) = \sum_{-\infty}^{\infty} C_n e^{in \frac{\pi}{l} x}$$

e la (32) dà per i coefficienti l'espressione

$$(37) \quad C_n = \frac{1}{2l} \int_{-l}^l f(x) e^{-in \frac{\pi}{l} x} dx.$$

Se si calcola, mediante la (36), l'integrale di ff^* esteso a tutto l'intervallo $(-l, l)$, si trova facilmente

$$(38) \quad \int_{-l}^l |f|^2 dx = 2l \sum_{-\infty}^{\infty} |C_n|^2.$$

che non è altro che la *relazione di completezza* delle autofunzioni in questione.

§ 10. — CASO DI UN INTERVALLO INFINITO. SPETTRO CONTINUO DI AUTOVALORI.

In alcuni problemi di meccanica ondulatoria si deve integrare l'equazione (14) non già in un intervallo finito (a, b) come fin qui abbiamo supposto, ma in un intervallo infinito, da una o da entrambe le parti, e cioè (a, ∞) oppure $(-\infty, \infty)$. Converrà studiare questo caso come limite del precedente, facendo tendere all'infinito uno od entrambi gli estremi: si riconosce allora che il problema presenta, nel caso

dell'intervallo infinito, delle caratteristiche nuove, di cui la più notevole è l'apparizione di uno spettro continuo di autovalori: ciò significa che vi possono essere per λ (oltre, eventualmente, ad una successione discreta di autovalori) degli intervalli entro i quali qualunque valore di λ è un autovalore, ed a questi corrisponde, naturalmente, un insieme continuo di autofunzioni che converrà designare con $y_\lambda(x)$ (dove l'indice λ varia con continuità, cosicchè, in sostanza, y è funzione di due variabili e si potrebbe anche indicare con $y(\lambda, x)$).

Considereremo solo il caso dell'intervallo $(-\infty, +\infty)$, e ci limiteremo ad alcune considerazioni di carattere intuitivo sull'esempio del § 8, caso β . Se si parte da uno degli autovalori (28') e dalle corrispondenti autofunzioni (29), e si fa tendere l ad ∞ , si vede che λ_n tende all'autovalore 0 e le due autofunzioni pure tendono a 0 per qualunque x : non si ottiene dunque nulla di interessante. Bisogna invece, mentre l si fa tendere ad ∞ , considerare autofunzioni di ordine via via più elevato: cioè far tendere anche n ad ∞ , in modo (1) che l'autovalore $\lambda_n = n^2 \frac{\pi^2}{l^2}$ tenda ad un limite prefissato $\lambda (> 0)$: allora le due autofunzioni tendono a

$$(39) \quad y_\lambda^{(1)} = a_\lambda e^{i\sqrt{\lambda}x} \quad y_\lambda^{(2)} = a_\lambda e^{-i\sqrt{\lambda}x}$$

(dove a_λ è una costante rispetto ad x); ovvero, raccogliendo le due formule in una col porre $\omega = +\sqrt{\lambda}$ per la prima ed $\omega = -\sqrt{\lambda}$ per la seconda,

$$(39') \quad y_\omega = a_\omega e^{i\omega x}.$$

(1) Per comprendere intuitivamente il passaggio, si pensi di rappresentare con dei punti su una semiretta (fig. 18) gli autovalori (28'): è

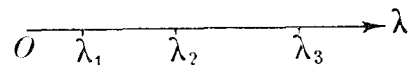


Fig. 18.

chiaro dalla formula che col crescere di l questi punti si spostano indefinitamente verso l'origine, divenendo sempre più fitti: fissato un punto sulla semiretta, per esso passano successivamente gli autovalori di ordine sempre più alto, e cioè un dato valore di λ si identifica con un autovalore di ordine sempre più elevato. È questo il significato del far tendere n ad ∞ insieme a l .

Si può considerare questa formula come la analoga della (35): la variabile continua ω , che può variare da $-\infty$ ad ∞ , tiene in essa il luogo della variabile discontinua n .

L'estensione a queste « autofunzioni di spettro continuo » del teorema di ortogonalità e della normalizzazione richiede alcune avvertenze: è ovvio difatti dalla (39') che $y_\omega y_\omega^*$ non tende a zero per $x \rightarrow \pm \infty$, e quindi l'integrale di normalizzazione non è convergente.

Bisogna invece considerare non una autofunzione, corrispondente ad un valore determinato di λ (ci riferiamo per ora solo alla $y^{(1)}$ od alla $y^{(2)}$), ma la somma di tutte quelle corrispondenti ad un intervallo infinitesimo $(\lambda_0, \lambda_0 + \Delta\lambda)$, ossia l'integrale

$$(40) \quad \Delta y = \int_{\lambda_0}^{\lambda_0 + \Delta\lambda} y_\lambda d\lambda$$

che è infinitesimo dell'ordine di $\Delta\lambda$, e chiamasi *autodifferenziale* (con la notazione Δy , vogliamo sottintendere un passaggio al limite per $\Delta\lambda \rightarrow 0$). Ponendo $\lambda = \lambda_0 + \varepsilon$ e trascurando i quadrati di ε si trova facilmente (scrivendo λ invece di λ_0) (1)

$$(41) \quad \Delta y = a_\lambda e^{i\sqrt{\lambda} + \frac{1}{2} \Delta\lambda} x \frac{\sin \frac{x \Delta\lambda}{4\sqrt{\lambda}}}{\frac{x \Delta\lambda}{4\sqrt{\lambda}}} \Delta\lambda,$$

ovvero, ponendo come sopra $\sqrt{\lambda} = \omega$

$$(41') \quad \Delta y = a_\omega e^{i(\omega + \frac{1}{2} \Delta\omega)x} \frac{\sin \frac{x \Delta\omega}{2}}{\frac{x \Delta\omega}{2}} 2\omega \Delta\omega,$$

formula che si estende anche ai valori negativi di ω .

La funzione $\frac{\Delta y}{\Delta\omega}$ tende a 0, come si vede, per $x \rightarrow \pm \infty$ (il che si può interpretare dicendo che le infinite funzioni sinusoi-

(1) Si potrebbe pensare che, essendo l'integrale (40) esteso ad un intervallo infinitesimo $\Delta\lambda$, esso si riduca ad un solo elemento, e si possa ottenere scrivendo $\Delta\lambda$ in luogo di $d\lambda$ nella espressione da integrare: ciò vale però soltanto per valori finiti di x (come si vede dall'espressione (41)) poichè una variazione, anche infinitesima, di λ , porta una differenza, nel valore della funzione, tanto maggiore quanto più grande è x .

dali, che in essa sono sovrapposte, si elidono all'infinito per mutua interferenza).

Inoltre le funzioni Δy godono una proprietà che è la naturale estensione della ortogonalità di due autofunzioni relative ad autovalori diversi: difatti, se $\Delta_1 y$, $\Delta_2 y$ sono le Δy relative a due intervallini di autovalori $(\lambda_1, \lambda_1 + \Delta_1 \lambda)$ e $(\lambda_2, \lambda_2 + \Delta_2 \lambda)$ che non si sovrappongono (nemmeno in parte) si ha, come può verificarsi con una integrazione alquanto laboriosa,

$$(42) \quad \int_{-\infty}^{\infty} \Delta_1 y \Delta_2 y^* dx = 0$$

Se invece i due intervallini hanno in comune un tratto $\Delta \lambda$, l'integrale risulta uguale a

$$4\pi a_\lambda a_\lambda^* \sqrt{\lambda} \Delta \lambda$$

(designando con λ indifferentemente λ_1 o λ_2 , che differiscono infinitamente poco): ciò suggerisce di porre come condizione di normalizzazione

$$(43) \quad \frac{1}{\Delta \lambda} \int_{-\infty}^{\infty} \Delta_1 y \Delta_2 y^* dx = 1:$$

tenendo conto del valore sopra riportato dell'integrale, si vede che la costante di normalizzazione a_λ va allora presa in modo che sia

$$a_\lambda a_\lambda^* = \frac{1}{4\pi \sqrt{\lambda}}.$$

Si può quindi prendere, come autofunzione normalizzata della (21) nell'intervallo $(0, \infty)$

$$(44) \quad y_\lambda = \frac{1}{2\sqrt{\pi \sqrt{\lambda}}} e^{\pm i\sqrt{\lambda}x},$$

ovvero

$$(44') \quad y_\omega = \frac{1}{2\sqrt{\pi \omega}} e^{i\omega x}.$$

Generalizzando ciò che si è visto su questo esempio particolare, diremo che: l'equazione (14), per un intervallo infinito, può presentare uno spettro continuo di autovalori; le autofunzioni corrispondenti a questi autovalori in generale non tendono a zero all'infinito, ma tendono a zero i loro auto-

differenziali $\Delta \lambda$: precisamente la condizione agli estremi è ⁽¹⁾

$$(45) \quad \lim_{x \rightarrow \pm \infty} \int_{\lambda}^{\lambda + \Delta \lambda} y_\lambda(x) d\lambda = 0$$

e la proprietà di ortogonalità è

$$(46) \quad \int_{-\infty}^{\infty} dx \left[\int_{\lambda_1}^{\lambda_1 + \Delta_1 \lambda} y_\lambda(x) d\lambda \cdot \int_{\lambda_2}^{\lambda_2 + \Delta_2 \lambda} y_\lambda^*(x) d\lambda \right] = 0,$$

per due intervalli $\Delta_1 \lambda$, $\Delta_2 \lambda$ che non si sovrappongono. La condizione di normalizzazione è

$$(46') \quad \frac{1}{\Delta \lambda} \int_{-\infty}^{\infty} dx \left[\int_{\lambda_1}^{\lambda_1 + \Delta_1 \lambda} y_\lambda(x) d\lambda \cdot \int_{\lambda_2}^{\lambda_2 + \Delta_2 \lambda} y_\lambda^*(x) d\lambda \right] = 1,$$

per due intervallini che hanno in comune un tratto $\Delta \lambda$, o eventualmente che coincidono.

Se poi vi sono, oltre agli autovalori continui, anche degli autovalori discreti λ_n , vale anche la seguente proprietà di ortogonalità tra le autofunzioni dello spettro continuo e quelle dello spettro discreto:

$$(47) \quad \int y_n(x) dx \int_{\lambda}^{\lambda + \Delta \lambda} y_\lambda^*(x) d\lambda = 0.$$

Possiamo ora estendere al caso di un intervallo infinito lo sviluppo di una funzione arbitraria in serie di autofunzioni: si avrà però in corrispondenza degli autovalori continui un integrale invece di una serie. Riferendoci, per maggior generalità, al caso che esistano tanto gli autovalori discreti che quelli continui, avremo uno sviluppo della forma ⁽²⁾

$$(48) \quad f(x) = \sum_1^n f_n y_n(x) + \int f_\lambda y_\lambda(x) d\lambda,$$

⁽¹⁾ Il limite $x \rightarrow \pm \infty$ si riferisce al caso che l'intervallo sia infinito da entrambe le parti: altrimenti si legga solo $+\infty$, o $-\infty$, e si sostituisca, per l'altro estremo, l'ordinaria condizione $y = 0$. Le integrazioni rispetto a x si intendono fatte su tutto l'intervallo.

⁽²⁾ Estendendo il concetto di integrale nel senso di Stieltjes, si potrebbe scrivere questa formula, come tutte quelle analoghe, col solo integrale, facendo rientrare in esso anche gli eventuali termini discreti (v. p. es. bibl. n. 14 p. 123).

dove f_λ è una funzione della variabile continua λ , e l'integrazione rispetto a λ si intende fatta su tutto lo spettro continuo di autovalori. La determinazione dei coefficienti f_n si fa come al § 9 (utilizzando, oltre alla mutua ortogonalità delle y_n , anche il teorema (47)), e si trova la stessa espressione (32); la funzione f_λ (che figura sotto il segno di integrale in modo analogo alle f_n sotto il segno di sommatoria) si determina analogamente nel modo seguente. Anzitutto, decomponiamo lo spettro continuo in intervalli $\Delta_1\lambda, \Delta_2\lambda, \dots$ e decomponiamo l'integrale nel modo seguente:

$$f(x) = \sum_n f_n y_n(x) + \sum_i \int_{\lambda_i}^{\lambda_i + \Delta_i \lambda} f_\lambda y_\lambda(x) d\lambda.$$

Ora moltiplichiamo ambo i membri per

$$\frac{dx}{\Delta_r \lambda} \int_{\lambda_r}^{\lambda_r + \Delta_r \lambda} y_\lambda^*(x) d\lambda$$

ed integriamo rispetto ad x (su tutto il campo di variabilità di x): facendo poi tendere a 0 gli intervalli $\Delta_r \lambda$ si riconosce che, in virtù della (47), al secondo membro si annullano tutti i termini della prima sommatoria, ed in virtù della (46) quelli della seconda si annullano tutti tranne l' r -esimo, che per la (46') si riduce a $f\lambda_r$: sopprimendo l'indice r poichè λ varia con continuità si ha:

$$(49) \quad f_\lambda = \lim_{\Delta\lambda \rightarrow 0} \frac{1}{\Delta\lambda} \int_{-\infty}^{\infty} f(x) dx \int_{\lambda}^{\lambda + \Delta\lambda} y_\lambda^*(x) d\lambda$$

e se $f(x)$ all'infinito diviene infinitesima di ordine sufficiente, si può invertire il segno di lim con quello di integrale e si ha

$$(50) \quad f_\lambda = \int_{-\infty}^{\infty} f(x) y_\lambda^*(x) dx.$$

La relazione di completezza (nel caso dello spettro continuo) è:

$$(51) \quad \int_{-\infty}^{\infty} |f|^2 dx = \int |f_\lambda|^2 d\lambda.$$

§ 11. - SVILUPPO IN INTEGRALE DI FOURIER.

La più notevole applicazione delle considerazioni precedenti è lo sviluppo in *integrale di Fourier*, che è la naturale estensione dello sviluppo in serie di Fourier al caso in cui si vuol rappresentare una funzione $f(x)$ (non periodica) data entro un intervallo infinito. Esso si può ottenere prendendo come autofunzioni le (44) ed è quindi:

$$(52) \quad f(x) = \int_0^{\infty} \frac{c_1(\lambda)}{2\sqrt{\pi\lambda}} e^{i\sqrt{\lambda}x} d\lambda + \int_0^{\infty} \frac{c_2(\lambda)}{2\sqrt{\pi\lambda}} e^{-i\sqrt{\lambda}x} d\lambda,$$

ossia:

$$(53) \quad f(x) = \int_{-\infty}^{\infty} C(\omega) e^{i\omega x} d\omega,$$

avendo posto

$$\left\{ \begin{array}{l} C(\omega) = \frac{c_1(\omega^2)}{2\sqrt{\pi\omega}} \text{ per } \omega > 0, \text{ e} \\ C(\omega) = \frac{c_2(\omega^2)}{2\sqrt{\pi\omega}} \text{ per } \omega < 0. \end{array} \right.$$

I coefficienti c_1 e c_2 , sono dati dalla (50) e da essi si trova

$$(54) \quad C(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx.$$

Si osservi che se $f(x)$ è funzione pari, cioè se $f(-x) = f(x)$, tale è anche $C(\omega)$, e se $f(x)$ è dispari, anche $C(\omega)$ è dispari. Nel primo caso la (53) si può anche scrivere:

$$(53') \quad f(x) = 2 \int_0^{\infty} C(\omega) \cos \omega x d\omega.$$

con

$$(54') \quad C(\omega) = \frac{1}{\pi} \int_0^{\infty} f(x) \cos \omega x dx;$$

e nel secondo:

$$(53'') \quad f(x) = 2i \int_0^{\infty} C(\omega) \sin \omega x d\omega,$$

con

$$(54'') \quad C(\omega) = \frac{1}{i\pi} \int_0^{\infty} f(x) \sin \omega x \, dx.$$

Anche nel caso generale la (53) e la (54) si potrebbero facilmente mettere sotto forma trigonometrica.

§ 12. — INTERPRETAZIONE ONDULATORIA DELLO SVILUPPO DI FOURIER.

Lo sviluppo in integrale di Fourier è suscettibile di una espressiva interpretazione fisica che si può riferire a fenomeni ondulatori di qualsiasi natura (onde elastiche, ottiche acustiche, ecc.): noi la esporremo usando il linguaggio dell'ottica elettromagnetica, ma è immediata la traduzione delle stesse considerazioni nel linguaggio corrispondente agli altri tipi di fenomeni ondulatori.

Se una radiazione monocromatica, di lunghezza d'onda λ , si propaga (per onde piane) con velocità V lungo l'asse delle x , una qualsiasi componente f del campo elettrico o magnetico è espressa, in funzione del tempo t e della x , dalla nota *equazione del raggio*

$$(55) \quad f(x, t) = A e^{i \frac{2\pi}{\lambda} (x - Vt)} = A e^{2\pi i (kx - \nu t)},$$

dove si è posto

$$\begin{cases} k = 1/\lambda & (\text{« numero d'onde »}) \\ \nu = V/\lambda & (\text{« frequenza »}) \end{cases}$$

e dove la A è una costante, generalmente complessa, il cui modulo rappresenta l'ampiezza delle onde mentre l'argomento ne rappresenta la fase: la chiameremo « *ampiezza complessa* » (come di consueto, la grandezza fisica f si può intendere rappresentata dalla parte reale, o da quella immaginaria, della precedente espressione).

In un istante qualsiasi, la f è rappresentata graficamente, lungo l'asse x , da una sinusoide di lunghezza d'onda λ , la quale, col tempo, progredisce nel verso positivo dell'asse x (se $V > 0$) con la velocità uniforme V : diremo che essa rap-

presenta un *treno* ⁽¹⁾ *d'onde* monocromatico. Osserviamo poi che un treno d'onde *regressive* si può rappresentare con la stessa espressione, dando però a k un valore negativo.

Una radiazione non monocromatica sarà rappresentata da una somma di termini del tipo (55) (con diverse λ e quindi diverse k e ν) ognuna corrispondente ad una singola componente monocromatica ossia ad una singola *riga spettrale*: (generalmente anche la velocità V sarà diversa per le diverse radiazioni componenti, cioè vi sarà *dispersione*). Potremo scegliere una qualunque di queste quantità λ , k , ν per caratterizzare le singole componenti monocromatiche: prenderemo il « numero d'onde » k , e supporremo nota la funzione $V(k)$ che caratterizza la dispersione del mezzo: la frequenza risulterà data da

$$(56) \quad \nu(k) = \frac{V(k)}{\lambda} = kV(k).$$

Se poi la radiazione, osservata allo spettroscopio, dà uno spettro continuo anziché uno spettro di righe, dovremo rappresentarla con un integrale anziché con una somma, e cioè sarà

$$(57) \quad f(x, t) = \int_{-\infty}^{\infty} A(k) e^{2\pi i (kx - \nu t)} \, dk.$$

La quantità $|A|^2 dk$ rappresenta l'intensità luminosa corrispondente all'intervallo dk dello spettro continuo: quindi la funzione $|A(k)|^2$ rappresenta la distribuzione dell'intensità nello spettro. Per rappresentare poi con una sola formula la sovrapposizione di una radiazione che si propaga in senso progressivo ed una che si propaga in senso opposto, abbiamo esteso, nella (57), l'integrale da $-\infty$ a ∞ , con la convenzione che $A(-k)$ rappresenti ampiezza e fase delle onde regressive di numero d'onde k .

⁽¹⁾ Con la parola « treno » designamo una successione di onde illimitata nello spazio (da $-\infty$ a $+\infty$) e, quindi, nel tempo: useremo invece la parola « gruppo » quando si tratterà di un numero limitato di onde.

Al tempo 0, la distribuzione della f lungo l'asse delle x è data da

$$(58) \quad f(x, 0) = \int_{-\infty}^{\infty} A(k) e^{2\pi i k x} dk,$$

formula che si identifica con la (53), purchè si ponga $\omega = 2\pi k$, $C(\omega) = \frac{1}{2\pi} A\left(\frac{\omega}{2\pi}\right)$. Con queste sostituzioni la (54) ci fornisce allora per A la seguente espressione:

$$(59) \quad A(k) = \int_{-\infty}^{\infty} f(x, 0) e^{-2\pi i k x} dx.$$

e la relazione di completezza (51) si scrive in questo caso

$$(51') \quad \int_{-\infty}^{\infty} |f(x, t)|^2 dx = \int_{-\infty}^{\infty} |A|^2 dk.$$

Così, data la distribuzione iniziale della f , il teorema di Fourier ci insegna a decomporla nella sovrapposizione di infiniti treni d'onde monocromatici, le cui ampiezze complesse sono date dalla (59): questi treni d'onde si spostano poi, ciascuno con una diversa velocità $V(k)$, e la loro sovrapposizione, al tempo t , dà luogo ad una funzione $f(x, t)$ (in generale rappresentata da una curva di forma diversa dalla iniziale) espressa dalla (57).

Nel caso che la $f(x, 0)$ sia una funzione pari (o dispari) si possono adoperare le formule (53'), (54'), (53''), (54''), che divengono:

per f pari:

$$(58') \quad f(x, 0) = 2 \int_0^{\infty} A(k) \cos 2\pi k x dk,$$

$$(59') \quad A(k) = 2 \int_0^{\infty} f(x, 0) \cos 2\pi k x dx;$$

per f dispari:

$$(58'') \quad f(x, 0) = 2i \int_0^{\infty} A(k) \sin 2\pi k x dk,$$

$$(59'') \quad A(k) = \frac{2}{i} \int_0^{\infty} f(x) \sin 2\pi k x dx.$$

Esempio: gruppo d'onde di ampiezza costante. — Applichiamo le considerazioni precedenti al caso che la distribuzione iniziale della f sia quella rappresentata dalla fig. 19, e cioè

$$\begin{cases} f(x, 0) = \sin 2\pi k_0 x & \text{per } -l < x < l \\ f(x, 0) = 0 & \text{per } x < -l, x > l. \end{cases}$$

Diremo che si tratta di un *gruppo d'onde* di lunghezza $2l$. Esso si può pensare approssimativamente realizzato mediante una sorgente, capace di emettere radiazione monocromatica di lunghezza d'onda λ_0 , accesa però solo durante un intervallo fi-



Fig. 19.

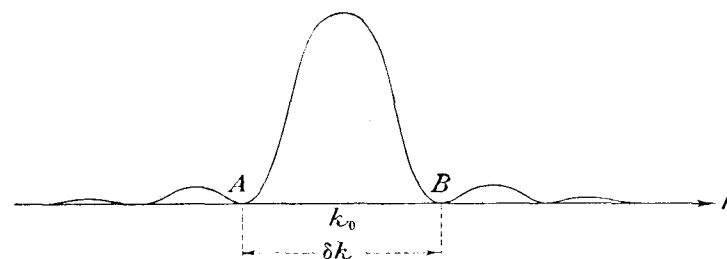


Fig. 20.

nito di tempo, e cioè quello durante il quale la luce percorre uno spazio $2l$: supponiamo di ricevere tale radiazione in uno spettroscopio (di infinito potere risolutivo) e domandiamoci l'aspetto dello spettro.

Trattandosi di funzione dispari, useremo la (58'') e la (59''), che danno

$$f(x, 0) = 2i \int_0^{\infty} A(k) \sin 2\pi k x dk,$$

con

$$\begin{cases} A(k) = \frac{2}{i} \int_0^l \sin 2\pi k_0 x \sin 2\pi k x dx \\ = \frac{l}{i} \left[\frac{\sin 2\pi(k_0 - k)l}{2\pi(k_0 - k)l} - \frac{\sin 2\pi(k_0 + k)l}{2\pi(k_0 + k)l} \right]. \end{cases}$$

Si osservi che la funzione $\left| \frac{\sin u}{u} \right|$ ha un massimo ($= 1$) per $u = 0$, ed altri infiniti massimi a destra e a sinistra di questo, ma rapidamente decrescenti coll'allontanarsi da esso: quindi $A(k)$ sarà rilevante solo per i valori di k prossimi a k_0 e per essi il secondo termine sarà trascurabile rispetto al primo ⁽¹⁾. La distribuzione dell'intensità nello spettro sarà dunque data approssimativamente da:

$$(60) \quad |A(k)|^2 = l^2 \left(\frac{\sin 2\pi(k_0 - k)l}{2(k_0 - k)l} \right)^2,$$

ed è rappresentata graficamente dalla curva della fig. 20. Trascurando le arcate laterali si può dire che lo spettro è composto di una riga corrispondente al numero d'onde k_0 , ma *allargata*: la larghezza della riga (definita provvisoriamente come distanza tra i due minimi A, B che corrispondono a $2\pi(k_0 - k)l = \pm \pi$) è data ovviamente da $\delta k = \frac{1}{l}$. Se la sorgente emettesse luce indefinitamente, cioè se invece che un *gruppo* limitato di onde si avesse un *treno* illimitato ($l = \infty$) l'analisi spettrale darebbe, come è evidente, una riga infinitamente sottile.

Si può dire dunque che: *se si tronca un treno d'onde monocromatiche, esso perde la monocromaticità, e la riga spettrale che gli corrisponde assume una larghezza inversamente proporzionale alla lunghezza totale del gruppo d'onde (o alla durata dell'emissione).*

Dunque, a rigore, la luce non è mai monocromatica se non viene emessa per un intervallo infinito di tempo.

§ 13. — LUNGHEZZA DI UN GRUPPO D'ONDE E LARGHEZZA DELLA RIGA SPETTRALE CORRISPONDENTE.

L'osservazione con cui si chiude il § precedente è di fondamentale importanza, e può venire estesa anche a gruppi d'onde di ampiezza non costante (diremo che si ha un *gruppo d'onde* ogni volta che la f ha valori sensibili solo entro un tratto limitato dell'asse x , ed è nulla o trascurabile nel rimanente). Bisogna però, per questa estensione, definire prima in modo preciso e generale la *larghezza* di una riga spettrale e la *lunghezza* di un gruppo d'onde.

⁽¹⁾ Poichè supponiamo di osservare solo le onde « progressive », consideriamo solo i valori positivi di k .

È naturale intanto definire come *centro* della riga spettrale il « baricentro » dell'intensità, cioè il valore \bar{k} di k definito da

$$(61) \quad \bar{k} = \frac{1}{I} \int_{-\infty}^{\infty} k |A|^2 dk,$$

ove si è posto (intensità totale dello spettro):

$$(62) \quad I = \int_{-\infty}^{\infty} |A|^2 dk.$$

La semilarghezza Δk della riga si definirà allora con la formula seguente (analogo a quella che definisce in meccanica il « raggio d'inerzia », o, nella teoria degli errori, l'errore quadratico medio):

$$(63) \quad (\Delta k)^2 = \frac{1}{I} \int_{-\infty}^{\infty} (k - \bar{k})^2 |A|^2 dk.$$

In modo perfettamente analogo definiremo come *centro* del gruppo d'onde il punto \bar{x} dato da:

$$(64) \quad \bar{x} = \frac{1}{I} \int_{-\infty}^{\infty} x |f|^2 dx,$$

dove l'integrale

$$(62') \quad I = \int_{-\infty}^{\infty} |f|^2 dx,$$

è lo stesso della (62) (in virtù della (51') del § 12), e come *semilunghezza* ⁽¹⁾ del gruppo il Δx dato da

$$(65) \quad (\Delta x)^2 = \frac{1}{I} \int_{-\infty}^{\infty} (x - \bar{x})^2 |f|^2 dx.$$

Ciò premesso, si può dimostrare che più è corto il gruppo d'onde, più larga è la riga spettrale che gli corrisponde, e

⁽¹⁾ Applicandola p. es. al gruppo d'onde della fig. 19, questa definizione darebbe (approssimativamente), $\Delta x = \frac{l}{\sqrt{3}}$.

precisamente che Δx e Δk sono vincolati dalla disuguaglianza (assai importante in meccanica ondulatoria)

$$(66) \quad \Delta k \Delta x \geq \frac{1}{4\pi},$$

la quale deriva dal teorema di Fourier, indipendentemente dal significato fisico delle quantità in questione.

Per dimostrarlo introduciamo la funzione

$$(67) \quad F(x) = f(x,0) e^{-2\pi i \bar{k}x}$$

e partiamo dalla ovvia osservazione, che la quantità

$$D = \left| \frac{x - \bar{x}}{2(\Delta x)^2} F + \frac{\partial F}{\partial x} \right|^2$$

è sempre ≥ 0 . Sviluppando questa espressione si ha

$$\left\{ \begin{aligned} D &= \frac{(x - \bar{x})^2}{4(\Delta x)^4} FF^* + \frac{x - \bar{x}}{2(\Delta x)^2} \left(F \frac{\partial F^*}{\partial x} + F^* \frac{\partial F}{\partial x} \right) + \frac{\partial F}{\partial x} \frac{\partial F^*}{\partial x} = \\ &= \frac{(x - \bar{x})^2}{4(\Delta x)^4} FF^* + \frac{1}{2(\Delta x)^2} \frac{\partial}{\partial x} \left[(x - \bar{x}) FF^* \right] - \frac{FF^*}{2(\Delta x)^2} \\ &\quad + \frac{\partial}{\partial x} \left(F^* \frac{\partial F}{\partial x} \right) - F^* \frac{\partial^2 F}{\partial x^2} \geq 0. \end{aligned} \right.$$

Moltiplichiamo per dx , e integriamo da $-\infty$ a ∞ , osservando che $FF^* = ff^* = |f|^2$, e tenendo conto delle (65), (62'): si ottiene

$$\int_{-\infty}^{\infty} D dx = \frac{I}{4(\Delta x)^2} + \frac{1}{2(\Delta x)^2} \left[(x - \bar{x}) FF^* \right]_{-\infty}^{\infty} - \frac{I}{2(\Delta x)^2} + \\ + \left[F^* \frac{\partial F}{\partial x} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} F^* \frac{\partial^2 F}{\partial x^2} dx \geq 0.$$

e se la F , per $x = \pm \infty$, è infinitesima di ordine sufficiente, scompaiono il secondo ed il quarto termine e resta

$$(68) \quad -\frac{I}{4(\Delta x)^2} - \int_{-\infty}^{\infty} F^* \frac{\partial^2 F}{\partial x^2} dx \geq 0.$$

Questo integrale si può mettere in relazione con Δk nel modo seguente.

Si ha dalla (63), dalla (59) e dalla (67)

$$\left. \begin{aligned} (\Delta k)^2 &= \frac{1}{I} \int_{-\infty}^{\infty} (k - \bar{k})^2 AA^* dk = \\ &= \frac{1}{I} \int_{-\infty}^{\infty} (k - \bar{k})^2 A dk \int_{-\infty}^{\infty} f^*(x,0) e^{2\pi i kx} dx = \\ &= \frac{1}{I} \int_{-\infty}^{\infty} (k - \bar{k})^2 A dk \int_{-\infty}^{\infty} F^*(x) e^{2\pi i (k - \bar{k})x} dx, \end{aligned} \right\}$$

e, invertendo le integrazioni:

$$(69) \quad (\Delta k)^2 = \frac{1}{I} \int_{-\infty}^{\infty} F^*(x) dx \int_{-\infty}^{\infty} (k - \bar{k})^2 A e^{2\pi i (k - \bar{k})x} dk.$$

L'integrale rispetto a k si può ottenere osservando che la (67), per la (58), si può scrivere

$$F(x) = \int_{-\infty}^{\infty} A e^{2\pi i (k - \bar{k})x} dk,$$

da cui, derivando due volte,

$$\frac{\partial^2 F}{\partial x^2} = -4\pi^2 \int_{-\infty}^{\infty} (k - \bar{k})^2 A e^{2\pi i (k - \bar{k})x} dk.$$

Sostituendo nella (69) l'espressione così trovata per l'integrale rispetto a k , essa diviene

$$(\Delta k)^2 = \frac{-1}{4\pi^2 I} \int_{-\infty}^{\infty} F^* \frac{\partial^2 F}{\partial x^2} dx.$$

Così abbiamo ottenuto l'integrale che figura nella (68), la quale perciò diviene

$$-\frac{I}{4(\Delta x)^2} + 4\pi^2 I (\Delta k)^2 \geq 0,$$

da cui, essendo $I > 0$, si trae la (66).

Il prodotto $\Delta k \Delta x$ assume il valore minimo $\frac{1}{4\pi}$ quando $D = 0$, cioè quando

$$\frac{x - \bar{x}}{2(\Delta x)^2} F + \frac{\partial F}{\partial x} = 0,$$

da cui, integrando

$$F = Ce^{-\frac{(x-\bar{x})^2}{4(\Delta x)^2}},$$

ossia,

$$(70) \quad f(x, 0) = Ce^{-\frac{(x-\bar{x})^2}{4(\Delta x)^2} + 2\pi i k \bar{x}}.$$

Questa funzione è rappresentata da una curva di andamento sinusoidale iscritta entro la curva

$$(71) \quad y = Ce^{-\frac{(x-\bar{x})^2}{4(\Delta x)^2}},$$

che è la nota *curva gaussiana degli errori* col centro in $x = \bar{x}$.

Questa speciale forma del gruppo d'onde presenta la particolarità che la $A(k)$ è rappresentata da una formula analoga alla f : si trova difatti usando la (59):

$$(72) \quad A(k) = C'e^{-\frac{(k-\bar{k})^2}{4(\Delta k)^2} - 2\pi i \bar{x} k},$$

dove

$$(72') \quad C' = \frac{C}{\sqrt{8\pi} \Delta k} e^{2\pi i \bar{x} \bar{k}}.$$

Un altro teorema sull'integrale di Fourier, che può essere utile tener presente, riguarda la « lunghezza » del gruppo d'onde, e la « larghezza » della riga, definite però non nel modo precedente, ma come il minimo intervallo, che chiameremo $2\Delta'x$ (o $2\Delta'k$), il quale contiene tutti i valori di x (o rispettivamente di k) pei quali la f (o rispettivamente la A) è diversa da 0. P. es., nel caso rappresentato dalle figg. 19 e 20, sarebbe $\Delta'x = l$, $\Delta'k = \infty$. Il teorema in discorso è il seguente: $\Delta'x$ e $\Delta'k$ non possono mai essere entrambe finite. Ciò implica, in particolare, che se una sorgente luminosa emette luce per un tempo finito, la luce emessa non solo non è monocromatica, ma il suo spettro è (teoricamente) illimitato.

§ 14. - VELOCITÀ DI GRUPPO.

Consideriamo un gruppo d'onde *pressochè* monocromatico, cioè tale che nello sviluppo di Fourier compaiano con

intensità apprezzabile soltanto le frequenze comprese in un piccolo intervallo, ossia che $A(k)$ si possa ritenere diverso da zero solo per k compreso entro un piccolo intervallo $(k_0 - \varepsilon, k_0 + \varepsilon)$. Potremo allora introdurre, entro tale intervallo, una variabile η , che assuma sempre valori piccoli, ponendo $k = k_0 + \eta$, e quindi

$$v = v_0 + v'_0 \eta$$

dove

$$v_0 = v(k_0), \quad v'_0 = \left(\frac{dv}{dk}\right)_{k=k_0}.$$

Con ciò lo sviluppo della f si scrive (v. (57)):

$$(73) \quad f(x, t) = e^{2\pi i(k_0 x - v_0 t)} \int_{-\varepsilon}^{\varepsilon} A(k_0 + \eta) e^{2\pi i(x - v' t)\eta} d\eta.$$

Confrontando questa formula con la (55) che rappresenta delle onde rigorosamente monocromatiche, si vede che l'integrale che figura nella (73) tiene il luogo dell'*ampiezza* A della (55): esso però non è una costante ma una funzione di x

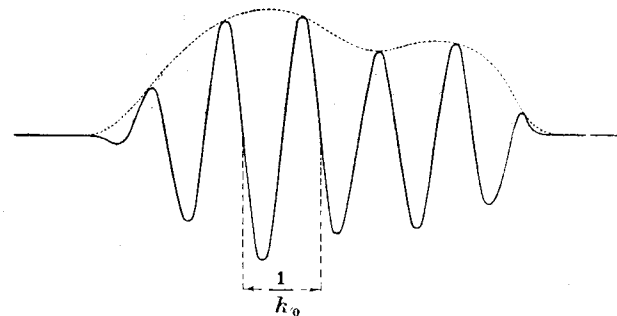


Fig. 21.

e t . Perciò si può dire che la (73) rappresenta delle onde di frequenza v_0 e lunghezza d'onda $1/k_0$, ma di ampiezza (e fase) variabile. Se, per un dato t , rappresentiamo graficamente il modulo dell'integrale in funzione di x (curva tratteggiata della fig. 21, che chiameremo « profilo » del gruppo), la f sarà rappresentata da una specie di sinusoida ad ampiezza variabile; iscritta entro questa curva.

Si osservi ora che l'integrale contiene x e t solo nella combinazione $x - v_0 t$: ciò significa che il profilo del gruppo si sposta senza deformarsi, con velocità v_0 : si può quindi dire che tutto il gruppo d'onde progredisce con questa velocità, la quale perciò si chiama *velocità di gruppo* e verrà indicata con v . Essa sarà dunque (sottintendendo $k = k_0$)

$$(74) \quad v = \frac{dv}{dk},$$

ovvero

$$(74') \quad \frac{1}{v} = \frac{d\left(\frac{1}{\lambda}\right)}{dv}.$$

Si osservi che la velocità di gruppo v è generalmente diversa dalla velocità $V = \frac{v}{k}$ con cui si spostano le singole onde (*velocità di fase*): difatti

$$\frac{dv}{dk} = \frac{d}{dk} [kV(k)] = V(k) + kV'(k)$$

e quindi

$$(75) \quad v = V(k_0) + k_0 V'(k_0).$$

Solo quando $V' = 0$, cioè quando non vi è dispersione, si ha $v = V(k_0)$.

Per comprendere come un gruppo d'onde possa, nel suo insieme, spostarsi con velocità diversa da quella delle singole onde, si osservino le onde prodotte da un sasso alla superficie dell'acqua: si vedrà che mentre il gruppo si propaga, le onde più esterne si smorzano poco a poco, mentre altre ne appaiono all'interno cosicché, mentre il numero delle onde del gruppo rimane costante, ciascuna di esse progredisce gradatamente verso la testa del gruppo: la velocità di ogni onda è quindi maggiore della velocità con cui si sposta il gruppo. Ciò dipende dal fatto che le onde dell'acqua si propagano con velocità tanto maggiore quanto minore è la frequenza ($V'(k) < 0$).

§ 15. — ONDE NELLO SPAZIO. PACCHETTI D'ONDE.

Le cose dette nei §§ precedenti si estendono senza difficoltà alla propagazione di onde in tre dimensioni. Un treno

di onde piane monocromatiche può venir definito dalla lunghezza d'onda λ , dai coseni α, β, γ della direzione di propagazione (normale alle superficie d'onda), e dall'ampiezza complessa A : l'equazione che lo rappresenta, analoga alla (55) (da cui si potrebbe ottenere con un cambiamento di assi) è

$$(76) \quad f(x, y, z, t) = A e^{\frac{2\pi i}{\lambda} (\alpha x + \beta y + \gamma z - vt)}$$

ovvero, introducendo, invece di λ , il numero d'onde $k = 1/\lambda$, e ponendo

$$(77) \quad k_x = k\alpha \quad k_y = k\beta \quad k_z = k\gamma,$$

$$(78) \quad f(x, y, z, t) = A e^{2\pi i (k_x x + k_y y + k_z z - vt)}.$$

Convien considerare k_x, k_y, k_z come componenti di un vettore \mathbf{k} (*vettore di propagazione*) che rappresenta col suo modulo k il numero d'onde $1/\lambda$ e colla sua direzione la direzione di propagazione (1): indicando allora con \mathbf{r} il vettore di componenti x, y, z che individua il punto dove si considera la f , si può scrivere la (78) in forma intrinseca, cioè indipendente dagli assi

$$(78') \quad f(\mathbf{r}, t) = A e^{2\pi i (\mathbf{k} \times \mathbf{r} - vt)}.$$

Sovrapponendo infiniti treni d'onde siffatti di tutti i possibili numeri d'onde e di tutte le direzioni, si ottiene una f rappresentata da

$$(79) \quad f(\mathbf{r}, t) = \int_{-\infty}^{\infty} \int \int A(\mathbf{k}) e^{2\pi i (\mathbf{k} \times \mathbf{r} - vt)} dk_x dk_y dk_z,$$

dove v è funzione in generale di k_x, k_y, k_z , ovvero, se il mezzo è isotropo come supporremo, solo di k .

(1) È utile rilevare fin da ora che nel caso delle onde elettromagnetiche questo vettore è strettamente legato all'impulso \mathbf{p} dei fotoni (v. § 3): questo ha infatti la stessa direzione di \mathbf{k} e la grandezza $p = \frac{h}{\lambda} = hk$: si ha quindi

$$(77') \quad \mathbf{k} = \frac{1}{h} \mathbf{p}.$$

Questa formula, per un dato valore di t (p. es. $t = 0$) si può considerare come lo sviluppo in integrale di Fourier di una assegnata funzione di x, y, z , sviluppo che si può ottenere applicando successivamente tre volte la formula del § 11: si vede allora che il coefficiente A , ampiezza dei singoli treni d'onda ⁽¹⁾, è dato dalla seguente formula, che è la generalizzazione della (59):

$$(80) \quad A(\mathbf{k}) = \iiint_{-\infty}^{\infty} f(\mathbf{r}, 0) e^{-2\pi i \mathbf{k} \times \mathbf{r}} dx dy dz.$$

Quando la f ha valori sensibili solo entro una regione limitata di spazio, ed è nulla o trascurabile al di fuori di questa regione, diremo che essa rappresenta un *pacchetto d'onde* (che è l'analogo tridimensionale del « gruppo d'onde » definito nel § 13). Un particolare pacchetto d'onde si può p. es. realizzare fisicamente nel modo seguente: si consideri un proiettore capace di emettere un fascio cilindrico di luce, e lo si accenda per un brevissimo intervallo di tempo, durante il quale la luce percorra uno spazio $2l$: allora partirà dal proiettore uno « spezzone » cilindrico di luce che costituisce un pacchetto d'onde luminose ⁽²⁾.

Come *centro* del pacchetto si definisce il baricentro di $|f|^2$, cioè il punto le cui coordinate $\bar{x}, \bar{y}, \bar{z}$ sono date da

$$(64') \quad \bar{x} = \frac{1}{I} \iiint_{-\infty}^{\infty} x |f|^2 dx dy dz \quad (\text{con } I = \iiint_{-\infty}^{\infty} |f|^2 dx dy dz)$$

⁽¹⁾ Più esattamente: $A dk_x dk_y dk_z$ rappresenta l'ampiezza (complessa) dei treni d'onda aventi vettori di propagazione compresi fra (k_x, k_y, k_z) , e $(k_x + dk_x, k_y + dk_y, k_z + dk_z)$.

⁽²⁾ È importante notare che il pacchetto d'onde non conserva rigorosamente invariata la sua forma e grandezza, ma ha tendenza a sparpagliarsi man mano che si propaga. Ciò ha origine da fenomeni di diffrazione causati dai diaframmi che limitano il fascio di luce, ed inoltre, se il mezzo di propagazione non è il vuoto, dall'eventuale « dispersione » del mezzo. Però, entro i limiti di approssimazione in cui vale l'ottica geometrica, si può molte volte prescindere da questo sparpagliamento. In tal caso il pacchetto (ove si possa considerare puntiforme) descrive evidentemente il cammino di un « raggio », seguendo le leggi dell'ottica geometrica.

ed altre due formule analoghe: esso si sposta con una velocità uguale alla *velocità di gruppo* già definita nel caso unidimensionale. Similmente si definiscono, con formule analoghe alla (65), le tre semilunghezze $\Delta x, \Delta y, \Delta z$ che danno una idea delle dimensioni del pacchetto lungo i tre assi.

Mediante lo sviluppo di Fourier un pacchetto d'onde si può considerare ottenuto sovrapponendo infiniti treni d'onde monocromatici, di diverso vettore di propagazione, i quali interferiscono dappertutto tranne che nella regione occupata dal pacchetto.

Analogamente alla (61), si può introdurre un *vettore di propagazione* medio $\bar{\mathbf{k}}$ definito da:

$$(81) \quad \bar{k}_x = \frac{1}{I} \iiint |A|^2 k_x dk_x dk_y dk_z$$

e da altre due formule analoghe, ed inoltre le quantità $\Delta k_x, \Delta k_y, \Delta k_z$, definite, analogamente alla (63), da

$$(82) \quad (\Delta k_x)^2 = \frac{1}{I} \iiint (k_x - \bar{k}_x)^2 |A|^2 dk_x dk_y dk_z,$$

ecc. Alla (82) si può applicare, rispetto a ciascuna delle tre variabili, il ragionamento del § 13, e si trovano le disuguaglianze

$$(83) \quad \begin{cases} \Delta x \Delta k_x \geq \frac{1}{4\pi} \\ \Delta y \Delta k_y \geq \frac{1}{4\pi} \\ \Delta z \Delta k_z \geq \frac{1}{4\pi} \end{cases}$$

le quali esprimono che: *quanto meno il pacchetto è esteso nello spazio, tanto più debbono differire tra loro i vettori di propagazione dei treni che lo compongono*. In pratica, si ha che fare assai spesso in ottica con pacchetti aventi dimensioni assai grandi (rispetto alla lunghezza d'onda media) e quindi costruibili con treni d'onde pressochè della stessa frequenza e direzione di propagazione.

In modo analogo, considerando $f(\mathbf{r}, t)$ come funzione solo di t (cioè fissando l'attenzione su un determinato punto dello

spazio) e sviluppandola in integrale di Fourier, si troverebbe la relazione

$$(84) \quad \Delta v \Delta t \geq \frac{1}{4\pi},$$

dove Δv è definita in modo analogo a Δk_x ecc., e Δt in modo analogo a Δx ecc.: se dunque si tratta di un pacchetto di onde limitato, si può dire che Δt misura la durata del suo passaggio per un determinato punto dello spazio, e Δv la larghezza della riga spettrale corrispondente, nella scala delle frequenze: si ritrova così, da un altro punto di vista, che una luce di durata limitata non può essere monocromatica, e che la larghezza della riga spettrale è tanto maggiore quanto più breve è la durata.

§ 16. — CENNO SULLE EQUAZIONI CON SINGOLARITÀ.

Nei §§ precedenti si è sempre supposto che i coefficienti dell'equazione differenziale fossero regolari entro tutto l'intervallo considerato, compresi gli estremi: in molti problemi di meccanica ondulatoria si presentano però equazioni in cui qualcuno dei coefficienti dell'equazione (scritta nella forma (1')) diviene infinito in un punto dell'intervallo, generalmente ad un estremo (punti singolari). Perciò riassumeremo qui brevemente alcune nozioni sulle singolarità delle equazioni differenziali lineari (omogenee, del secondo ordine), i cui coefficienti supporremo siano funzioni monodrome della x ⁽¹⁾.

Si dice che in un punto x_0 l'equazione (1') presenta una singolarità *fuchsiana* (o, come taluni dicono, *non essenziale*), se per $x = x_0$ uno almeno dei coefficienti P , Q diventa infinito, ma tuttavia di ordine non superiore al 1° per P , e al 2° per Q , cosicchè l'equazione si possa scrivere

$$(85) \quad y'' + \frac{\mathfrak{S}(x-x_0)}{x-x_0} y' + \frac{\mathfrak{Q}(x-x_0)}{(x-x_0)^2} y = 0$$

dove $\mathfrak{S}(x-x_0)$ e $\mathfrak{Q}(x-x_0)$ rappresentano delle ordinarie serie di potenze di $x-x_0$ (ad esponenti interi non nega-

(1) Per maggiori particolari vedasi p. es. il n. 34 della bibl.

tivi). Se invece per $x = x_0$ il coefficiente P diventa infinito di ordine superiore al 1°, o Q di ordine superiore al 2°, si dice che la singolarità è *non fuchsiana* ovvero *essenziale*.

La ragione della distinzione sta nel comportamento delle soluzioni nell'intorno del punto $x = x_0$: nel caso della singolarità fuchsiana qualunque integrale dell'equazione o è regolare in $x = x_0$, o al più presenta una singolarità che consente di metterlo nella forma

$$(86) \quad y = (x-x_0)^\alpha \Phi(x-x_0),$$

con α reale (in generale non intero) e Φ funzione regolare (o, in un caso eccezionale cui si accennerà più sotto, con una singolarità logaritmica); invece se i coefficienti in $x = x_0$ non soddisfano la condizione anzidetta, possono esistere degli integrali con singolarità di altro tipo, o che divengano del tutto indeterminati in $x = x_0$.

Escludendo dalle nostre considerazioni le singolarità non fuchsiane, partiamo dall'equazione (85), e cerchiamo di soddisfarla con una soluzione della forma (86), prendendo per Φ una funzione regolare, cioè una serie di potenze intere e positive di $x-x_0$.

Sostituendo l'espressione (86) nell'equazione (85), e uguagliando a zero i coefficienti delle singole potenze di $x-x_0$, si determinano formalmente i coefficienti della serie Φ : la prima delle condizioni così ottenute serve a determinare α ed è, come si trova facilmente,

$$(87) \quad \alpha(\alpha-1) + \alpha \mathfrak{S}(0) + \mathfrak{Q}(0) = 0.$$

È questa un'equazione di secondo grado che fornisce in generale due radici, α_1 , α_2 , a cui corrispondono, *in generale*, due integrali della forma (86) che (si dimostra) sono indipendenti, e possono quindi essere presi come integrali fondamentali: il raggio di convergenza di tali serie si estende fino al più prossimo punto di singolarità. Se però le due radici della (87) differiscono per un numero intero (o, in particolare, coincidono), allora una sola di esse (quella la cui parte reale è più grande) fornisce una soluzione del tipo (86), mentre l'altra dà luogo a difficoltà nella formazione dei suc-

cessivi coefficienti della Φ : in tal caso, invece di una serie di potenze intere e positive, si deve prendere per Φ un'espressione del tipo $\Phi_1 + \Phi_2 \lg(x - x_0)$, con Φ_1 e Φ_2 funzioni regolari: ma non insistiamo su questo caso eccezionale, bastandoci rilevare che in tutti i casi si può trovare, mediante la (87) almeno un'integrale della forma (86). L'equazione (87) si dice « equazione ~~fundamentale~~ » relativa alla singolarità x_0 .

Accenneremo infine al caso in cui la singolarità si trova all'infinito, caso che si riconduce, come è noto, al precedente, con la trasformazione $x = 1/\zeta$: si trova così facilmente che la condizione necessaria e sufficiente perchè al punto all'infinito non vi sia una singolarità non fuchsiana è che per $x \rightarrow \infty$ il coefficiente P sia infinitesimo almeno del 1° ordine, ed il coefficiente Q almeno del 2°, cioè che essi possano scriversi (per x abbastanza grande)

$$(88) \quad P(x) = \frac{1}{x} \mathfrak{S}\left(\frac{1}{x}\right), \quad Q(x) = \frac{1}{x^2} \mathfrak{Q}\left(\frac{1}{x}\right),$$

dove \mathfrak{S} e \mathfrak{Q} sono serie di potenze intere e positive di $\frac{1}{x}$.

§ 17. — CENNO SULLE EQUAZIONI A DERIVATE PARZIALI.

Le equazioni che interessano la meccanica ondulatoria sono, nella maggior parte dei casi, equazioni a derivate parziali, lineari ed omogenee: a queste si possono estendere molte delle considerazioni fatte per il caso di una variabile sola. Accenneremo brevemente a talune di queste estensioni, riferendoci, per comodità di scrittura, al caso di due variabili indipendenti, x ed y (l'estensione a tre o più variabili è immediata).

L'equazione si dirà *autoaggiunta* se ha la forma seguente (analoga alla (12))

$$\frac{\partial}{\partial x} \left(P \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left(P \frac{\partial u}{\partial y} \right) + Ru = 0,$$

dove P ed R sono due funzioni di x ed y (che supporremo analitiche): spesso in R figura un parametro λ (come nella (14)), cioè l'equazione è

$$(89) \quad \frac{\partial}{\partial x} \left(P \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left(P \frac{\partial u}{\partial y} \right) + [\lambda + q(x, y)] u = 0.$$

Caso particolare assai notevole è quello in cui $P = 1$, cioèchè i primi due termini formano il laplaciano Δu .

Il problema delle autofunzioni, enunciato pel caso di due variabili, è il seguente: *data una regione S del piano limitata da un contorno σ , determinare una $u(x, y)$ che entro S soddisfi la (89) e che sul contorno σ si annulli, o abbia la derivata normale nulla.*

Tale problema in generale non ammette soluzioni regolari e non nulle, fuorchè per certi valori λ_i del parametro λ , che si chiamano *autovalori*. Essi formano (se la regione S ha estensione finita) una infinità numerabile. A ciascuno di essi corrispondono una o più *autofunzioni* indipendenti $u_i(x, y)$, (se più di una l'autovalore si dice *multiplo*): due autofunzioni u_n, u_m , corrispondenti a due diversi autovalori λ_n, λ_m , godono la proprietà di *ortogonalità*:

$$(90) \quad \int_S u_n u_m^* dS = 0$$

(dove l'integrale è esteso a tutto il campo S , e $dS = dx dy$), la quale si dimostra in modo perfettamente analogo a quello seguito nel caso di una sola variabile.

Similmente si imporrà alle autofunzioni la condizione di *normalizzazione*:

$$(91) \quad \int_S u_n u_n^* dS = 1.$$

Per risolvere i problemi a derivate parziali non esistono metodi di generale applicazione: quello che riesce più spesso è il metodo della *separazione delle variabili*: esso consiste nel cercare una soluzione $u(x, y)$ che sia il prodotto di una funzione X della sola x per una funzione Y della sola y . Perciò si deve porre nella (89)

$$(92) \quad u = X(x) Y(y),$$

e si deve cercare, se è possibile, di ridurre l'equazione all'eguaglianza di un'espressione contenente sole funzioni di x con una contenente sole funzioni di y : se ciò riesce, ciascuno dei due membri dovrà separatamente essere uguale ad una

costante (perchè x ed y devono poter variare indipendentemente): si hanno così due equazioni *a derivate ordinarie* per le due funzioni X ed Y . Così il problema è ricondotto a quello per una sola variabile.

La separazione delle variabili è effettivamente possibile nella maggior parte dei casi che ci interessano, e ne vedremo degli esempi più oltre. È da notare che la possibilità di separazione delle variabili è relativa al sistema di coordinate scelto: operando un cambiamento di coordinate (p. es., da cartesiane a polari) può così talvolta rendersi possibile la separazione che nel sistema primitivo non era possibile. In generale, il sistema di riferimento che consente la separazione delle variabili ha una certa suggestiva relazione coi dati geometrici e con la simmetria del problema: p. es., nei problemi a simmetria sferica sono le coordinate polari che si separano, ecc. Ma non vi è alcun metodo generale e rigoroso per trovare, in ogni caso, quale è il sistema di riferimento che consente la separazione delle variabili, nè si conosce alcun mezzo per poter dire, data l'equazione, se un tale sistema di riferimento esiste.

CAPITOLO II.

IMPOSTAZIONE PROBABILISTICA DEI PROBLEMI ATOMICI

Mostreremo in questo capitolo come si debbano impostare i problemi fondamentali riguardanti i fenomeni atomici per essere coerenti al principio logico posto da HEISENBERG a fondamento della fisica teorica (v. § 34), che cioè non si debbano mai far intervenire nei ragionamenti in modo essenziale enti che non siano almeno concettualmente osservabili, o relazioni che non siano almeno concettualmente verificabili. Questa impostazione, come vedremo, porta a risolvere le apparenti contraddizioni tra natura ondulatoria e natura corpuscolare sia della materia che della radiazione: siccome però il ragionamento è forse più facile a comprendersi nel caso della radiazione, incominceremo da questo, e poi trasporteremo per analogia le nostre considerazioni al caso della materia.

§ 18. — ANALISI DEL CONCETTO DI FOTONE.

Se si pensa ai diversi mezzi di cui ci serviamo abitualmente per rivelare la radiazione (visione, lastra fotografica, cella fotoelettrica e simili) si riconosce che essi rivelano la radiazione solo per mezzo delle modificazioni che subiscono le particelle materiali che la assorbono, le quali modificazioni consistono essenzialmente nell'acquisto di una determinata quantità di energia (*quanto*) che si manifesta con azioni chimiche (sulla retina o sulla lastra) o termiche, o con emissione di elettroni, ecc.: chiameremo siffatti fenomeni « *atti elementari di assorbimento* ». Un altro mezzo, meno comune,

di rivelare la radiazione, consiste nel servirsi della *pressione* da essa esercitata, o dell'*impulso* impresso da essa ad un elettrone (p. es. nell'effetto Compton): in questi casi si utilizza non l'energia ma la « quantità di moto » (o « impulso ») comunicata dalla radiazione alla materia. Similmente, ci possiamo accorgere che una particella materiale ha compiuto un « *atto elementare di emissione* », o dal fatto che la sua energia è diminuita senza che vi sia stata altra causa, o dal fatto che è variato il suo impulso (rinculo), senza che ciò possa attribuirsi ad altra ragione. In tutti i casi insomma, la radiazione ci si rivela attraverso la sua interazione con le particelle materiali: non è possibile mai sorprenderla nel suo tragitto dall'una all'altra di queste, ma solo, per così dire, alla partenza, o all'arrivo.

Per meglio comprendere la portata di questa impossibilità giova discutere alcune apparenti infrazioni.

Una di queste, assai banale, potrebbe riferirsi al fatto ben noto che quando un fascio di luce attraversa un mezzo torbido il suo cammino diventa lateralmente visibile: ma è facile comprendere che in questo caso è solo la grossolanità dei mezzi di osservazione che dà l'illusione di vedere una striscia luminosa continua: se osserviamo invece con sufficiente accuratezza, vediamo le singole particelle diffondenti sotto forma di punti brillanti staccati (effettivamente visibili ad occhio nudo nel caso del pulviscolo, o al microscopio nel caso di soluzioni torbide: se poi le particelle diffondenti sono singole molecole, esse sono solo praticamente, non concettualmente invisibili). Ma la successione di quei punti brillanti non segna affatto i punti della traiettoria di un medesimo fotone: si tratta invece, come è chiaro, del fatto che *alcuni* dei quanti emessi dalla sorgente vengono diffusi lateralmente, mentre quelli che non lo sono passano inosservati. Chè se la sorgente emettesse un solo quanto, o questo sarebbe diffuso da una particella materiale, e allora l'osservatore vedrebbe un *solo* punto brillante, o non lo sarebbe, e allora l'osservatore, che sta di lato, non vedrebbe nulla.

Esaminiamo ora la validità del ragionamento seguente: se in *A* è una sorgente luminosa che compie un atto elementare di emissione, e in *B* è una particella che compie un atto elementare di assorbimento, sembra di poter affermare che il quanto emesso da *A* ha percorso il segmento di retta *AB* fino a giungere in *B* dove è stato assorbito. Ma se cerchiamo un mezzo qualunque per controllare questa affermazione, ci accorgiamo subito della sua in-

consistenza. Infatti, per dare valore a tale affermazione, bisognerebbe che, ponendo tra *A* e *B* un tramezzo opaco con un foro sottile, si constatasse che la particella *B* riceve il quanto *solo* se il foro si trova sulla retta *AB*: invece è ben noto che un foro sottile dà luogo a fenomeni di diffrazione in virtù dei quali, se esso si trova sulla retta *AB*, può avvenire che la particella *B* non riceva luce, e viceversa può darsi che essa la riceva anche se il forellino si trova fuori della retta *AB*.

Si potrebbe pensare che, posto il forellino del tramezzo in un punto qualunque *C*, e ripetendo molte volte l'esperienza, quando finalmente avviene che la particella *B* riceva l'energia emessa da *A*, sia lecito attribuire al fotone la traiettoria costituita dalla spezzata *ACB*: ma evidentemente si può allora ripetere per ciascuno dei due tratti *AC* e *CB* il ragionamento fatto dianzi, e ci si accorge che è gratuita la costruzione della traiettoria con due segmenti di retta.

Si potrebbe allora pensare di raggiungere la determinazione completa della traiettoria con una esperienza-limite, cioè mediante una successione di infiniti schermi, ognuno con un foro infinitamente sottile, ovvero, ciò che è equivalente, costruendo tra *A* e *B* un tubo a pareti assorbenti, infinitamente stretto, di forma qualunque, rettilinea o no. Allora, qualora avvenisse che la particella *B* ricevesse il quanto, si potrebbe affermare che esso ha percorso la traiettoria costituita dal tubo. Ma tale affermazione sarebbe priva di significato fisico, poichè per controllarla non avremmo altro mezzo che quello di inserire, in un punto intermedio del tubo, una particella materiale capace di assorbire o diffondere la luce, e questa allora non giungerebbe più a *B*: vale a dire, il fenomeno in studio resterebbe alterato in modo essenziale per il fatto stesso di essere osservato.

Da queste considerazioni risulta che è possibile definire operativamente l'emissione o l'assorbimento di un fotone, ma non già la traiettoria seguita da esso: quindi non è lecito raffigurarsi il fenomeno come il movimento di un punto o di un corpuscolo, e non è molto appropriato parlare (come, per comodità di linguaggio, abbiamo fatto poc'anzi) di « partenza » o di « arrivo » di un fotone su un atomo. È vero che l'atto elementare di emissione ha alcune analogie col lancio di un proiettile materiale (rinculo e perdita di energia) e l'atto elementare di assorbimento ha analogia con l'urto di un proiettile con l'atomo assorbente: inoltre (nel caso semplice che non vi siano mezzi interposti) il secondo fenomeno

avviene con un ritardo, rispetto al primo, uguale al tempo che impiegherebbe un mobile animato dalla velocità c . Ma l'analogia non si estende alla fase intermedia del fenomeno, poichè viene a mancare la proprietà fondamentale della cinematica del punto, cioè la continuità della traiettoria. La nostra abitudine mentale ci fa considerare questa proprietà come indissolubilmente legata alle altre, mentre ne è logicamente indipendente: essa ci conduce falsamente a costruire un modello *letteralmente* corpuscolare dei fenomeni ottici, cioè a pensare i fotoni come corpi ordinari straordinariamente impiccioliti, ma dotati delle ordinarie proprietà cinematiche: con ciò si fa lo stesso errore che si commetterebbe se le note analogie tra correnti elettriche e correnti idrauliche si interpretassero pensando che l'elettricità sia veramente un liquido. Non vi è dunque da meravigliarsi se una teoria *letteralmente* corpuscolare porta a risultati falsi, poichè in essa si introduce tacitamente una premessa contraria all'esperienza, cioè quella di una completa analogia tra fotoni e corpi ordinari.

Bisogna dunque attribuire al modello corpuscolare solo il valore di una analogia limitata a certi aspetti dei fenomeni (conservazione dell'energia e dell'impulso, ecc.): tenendo ben presente questo, il modello corpuscolare è però di grandissimo aiuto, poichè fornisce un linguaggio espressivo per la descrizione di molti fenomeni, evitando lunghe perifrasi. P. es., diremo che su un certo elemento $d\sigma$ di superficie « arriva un fotone », invece di dire che: « se quell'elemento di superficie fosse coperto di una sostanza perfettamente assorbente, un atomo di questa compirebbe un atto elementare di assorbimento »; similmente, l'espressione « trovare un fotone in un certo elemento di volume dS » significa « riempire dS di materia completamente assorbente, e constatare che un atomo di questa compie un atto elementare di assorbimento ».

Per vedere ora con quali limitazioni è lecito usare il linguaggio corpuscolare, riprendiamo l'esempio (v. § 15) di un proiettore che, acceso per un breve istante, emetta un cilindro o « pacchetto » di luce (che supporremo pressochè monocromatica): se vogliamo descrivere obbiettivamente il fe-

nomeno senza fare uso di nessun modello, nè ondulatorio nè corpuscolare, dovremo dire che vi è una regione S dello spazio (spostantesi con velocità c) che è caratterizzata dalla proprietà che, quando un atomo viene a trovarsi nel suo interno, esso può compiere un atto elementare di assorbimento, ed in tal caso assorbe una quantità determinata w di energia, caratteristica della radiazione: la probabilità che l'assorbimento avvenga è (per un dato tipo di atomo assorbente) proporzionale alla densità di energia W (che supporremo uniforme). Se vogliamo tradurre questo nel modello corpuscolare, diremo che la regione S è popolata di fotoni, ciascuno dotato di energia w , cosicchè in ogni unità di volume ce ne saranno in media W/w . È chiaro poi che se prendiamo un elemento di volume dS tale che $\frac{W}{w} dS$ sia assai piccolo

rispetto ad 1, questa espressione rappresenta non il numero dei fotoni che si trovano in dS , ma la probabilità di trovarcene uno: perciò W/w potrà anche dirsi « densità della probabilità » di trovare un fotone. Tutto ciò è perfettamente identico a quanto potrebbe dirsi nel caso di uno sciame di pallini da caccia. Vi è però una differenza profonda in questo: nel caso dei pallini da caccia una conoscenza perfetta di tutte le modalità del lancio permetterebbe, almeno in teoria, di calcolare la posizione di ciascun pallino istante per istante e di sapere quindi, con certezza, se nell'elemento dS , in un dato istante, si trovi o no un pallino, cosicchè il concetto di probabilità entra nella questione solo per l'impossibilità pratica di conoscere tutti gli elementi occorrenti al calcolo preciso delle traiettorie: invece, nel caso della luce, non ha senso assegnare a ciascun fotone una posizione ed una velocità entro lo sciame, cosicchè nessun calcolo, nemmeno in teoria, può farci prevedere se nell'elemento dS , in un dato istante, si trovi o no un fotone: tutto ciò che si può dire è che la probabilità di trovarcelo, quando lo si osservi, è $\frac{W}{w} dS$. In questo caso dunque il concetto di probabilità entra nella questione non per la nostra ignoranza di alcuni elementi del problema, ma per la natura stessa del fenomeno. Tuttavia,

una volta rilevata questa importante differenza concettuale tra i due casi, si può talvolta prescindere formalmente, ossia usare a proposito dei fotoni lo stesso linguaggio che si userebbe per dei proiettili materiali, a patto però di considerare *imperfettamente conosciuti* (ossia, conosciuti con un errore di cui è nota la legge probabilistica) gli elementi del loro movimento (che in realtà sono, almeno in parte, *fisicamente indeterminati*, ossia *concettualmente inosservabili*, come si preciserà meglio in seguito).

Bisogna sempre tener presente che il dire che gli elementi del moto di un fotone (posizione, traiettoria, impulso, ecc.) non sono « fisicamente determinati » ha un significato profondamente diverso dal dire che questi elementi, pur essendo fisicamente determinati, non sono « conosciuti », ma si potrebbero conoscere a priori con una osservazione concettualmente effettuabile senza perturbare il fenomeno. Per comprendere la differenza, si pensi all'esempio seguente. Tra una sorgente O , che emette dei fotoni (a intervalli di tempo eventualmente anche grandi) ed uno schermo s su cui essi vengono ricevuti, c'è un tramezzo opaco con due fori A e B : si formano allora sullo schermo delle frange di interferenza, e per ciascun fotone è *indeterminato* se è passato dal foro A o dal foro B : qualunque tentativo di decidere la questione p. es. intercettando uno dei due cammini altererebbe il fenomeno di interferenza. Supponiamo ora invece che vi sia un dispositivo che automaticamente chiuda ora l'uno ora l'altro dei due fori, e questo irregolarmente, in modo che il passaggio di ciascun fotone da A o da B venga a dipendere da una specie di sorteggio automatico: per un osservatore che non veda il meccanismo è *sconosciuto* se un dato fotone è passato da A o da B , ma, contrariamente al caso precedente, qui è concettualmente possibile di decidere la questione senza alterare il fenomeno. E la differenza si rispecchia nel fatto che, in questo caso, sullo schermo s non si formano frange di interferenza: il fenomeno è dunque fisicamente diverso.

§ 19. — SIGNIFICATO PROBABILISTICO DELL'OTTICA ONDULATORIA.

Esaminiamo ora il significato che deve attribuirsi alla teoria ondulatoria della luce, ed i suoi rapporti coi fatti sperimentali. È noto che tutte le esperienze di formazione di immagini, interferenza, diffrazione, ecc., confermano concor-

demente ed esattamente le previsioni della teoria ondulatoria. Tali conferme consistono sostanzialmente in questo: data la forma, la posizione e la natura delle sorgenti e dei vari mezzi (lenti, schermi, specchi), si calcola mediante le leggi dell'ottica (ossia, in definitiva, mediante l'equazione delle onde) quale deve essere l'intensità di illuminazione I nei vari punti di uno schermo, che in genere è la retina o una lastra fotografica: si fa l'esperienza, e si trova che l'intensità misurata corrisponde, punto per punto, alla funzione I risultante dal calcolo, riproducendo quindi la forma prevista delle immagini, delle ombre, delle frange di interferenza, ecc. Data la già rilevata natura discontinua dei fenomeni di assorbimento, l'« intensità di illuminazione misurata », nel caso di luce monocromatica al quale per ora ci limitiamo, divisa per l'energia w di un fotone, rappresenta il numero, per unità di superficie e di tempo, dei fotoni che arrivano sullo schermo: la verifica sta dunque nel constatare

che su un elemento $d\sigma$, nel tempo dt , arrivano proprio $\frac{I}{w} d\sigma dt$ fotoni, dove I è data, in funzione del posto, dall'integrazione dell'equazione delle onde (precisamente, I è la proiezione del vettore di Poynting sulla normale alla superficie).

Si noti però che questo si verifica solo se i fotoni sono molto numerosi (come avviene di solito nelle esperienze di ottica): se invece essi sono pochi, o magari uno solo, l'ottica ondulatoria cade in difetto. Se, p. es., la sorgente emette un solo quanto di energia di fronte ad una lastra fotografica, l'ottica ondulatoria ci fa prevedere che l'energia debba ripartirsi su tutta la lastra con una certa intensità I , mentre in realtà la ritroveremo tutta in un solo punto (sono questi casi appunto che hanno suggerito il modello corpuscolare). Quale significato può attribuirsi allora alla intensità I data dall'ottica ondulatoria? Evidentemente, essa misura la « densità di probabilità » che il fotone arrivi in un punto della lastra piuttosto che in un altro: precisamente, se i fotoni emessi sono n , $\frac{I}{nw} d\sigma dt$ è la probabilità che ciascun fotone emesso ha di arrivare sull'elemento $d\sigma$ nel tempo dt .

avviene con un ritardo, rispetto al primo, uguale al tempo che impiegherebbe un mobile animato dalla velocità c . Ma l'analogia non si estende alla fase intermedia del fenomeno, poichè viene a mancare la proprietà fondamentale della cinematica del punto, cioè la continuità della traiettoria. La nostra abitudine mentale ci fa considerare questa proprietà come indissolubilmente legata alle altre, mentre ne è logicamente indipendente: essa ci conduce falsamente a costruire un modello *letteralmente* corpuscolare dei fenomeni ottici, cioè a pensare i fotoni come corpi ordinari straordinariamente impiccioliti, ma dotati delle ordinarie proprietà cinematiche: con ciò si fa lo stesso errore che si commetterebbe se le note analogie tra correnti elettriche e correnti idrauliche si interpretassero pensando che l'elettricità sia veramente un liquido. Non vi è dunque da meravigliarsi se una teoria *letteralmente* corpuscolare porta a risultati falsi, poichè in essa si introduce tacitamente una premessa contraria all'esperienza, cioè quella di una completa analogia tra fotoni e corpi ordinari.

Bisogna dunque attribuire al modello corpuscolare solo il valore di una analogia limitata a certi aspetti dei fenomeni (conservazione dell'energia e dell'impulso, ecc.): tenendo ben presente questo, il modello corpuscolare è però di grandissimo aiuto, poichè fornisce un linguaggio espressivo per la descrizione di molti fenomeni, evitando lunghe perifrasi. P. es., diremo che su un certo elemento $d\sigma$ di superficie « arriva un fotone », invece di dire che: « se quell'elemento di superficie fosse coperto di una sostanza perfettamente assorbente, un atomo di questa compirebbe un atto elementare di assorbimento »; similmente, l'espressione « trovare un fotone in un certo elemento di volume dS » significa « riempire dS di materia completamente assorbente, e constatare che un atomo di questa compie un atto elementare di assorbimento ».

Per vedere ora con quali limitazioni è lecito usare il linguaggio corpuscolare, riprendiamo l'esempio (v. § 15) di un proiettore che, acceso per un breve istante, emetta un cilindro o « pacchetto » di luce (che supporremo pressochè monocromatica): se vogliamo descrivere obbiettivamente il fe-

nomeno senza fare uso di nessun modello, nè ondulatorio nè corpuscolare, dovremo dire che vi è una regione S dello spazio (spostantesi con velocità c) che è caratterizzata dalla proprietà che, quando un atomo viene a trovarsi nel suo interno, esso può compiere un atto elementare di assorbimento, ed in tal caso assorbe una quantità determinata w di energia, caratteristica della radiazione: la probabilità che l'assorbimento avvenga è (per un dato tipo di atomo assorbente) proporzionale alla densità di energia W (che supporremo uniforme). Se vogliamo tradurre questo nel modello corpuscolare, diremo che la regione S è popolata di fotoni, ciascuno dotato di energia w , cosicchè in ogni unità di volume ce ne saranno in media W/w . È chiaro poi che se prendiamo un elemento di volume dS tale che $\frac{W}{w} dS$ sia assai piccolo

rispetto ad 1, questa espressione rappresenta non il numero dei fotoni che si trovano in dS , ma la probabilità di trovarcene uno: perciò W/w potrà anche dirsi « densità della probabilità » di trovare un fotone. Tutto ciò è perfettamente identico a quanto potrebbe dirsi nel caso di uno sciame di pallini da caccia. Vi è però una differenza profonda in questo: nel caso dei pallini da caccia una conoscenza perfetta di tutte le modalità del lancio permetterebbe, almeno in teoria, di calcolare la posizione di ciascun pallino istante per istante e di sapere quindi, con certezza, se nell'elemento dS , in un dato istante, si trovi o no un pallino, cosicchè il concetto di probabilità entra nella questione solo per l'impossibilità pratica di conoscere tutti gli elementi occorrenti al calcolo preciso delle traiettorie: invece, nel caso della luce, non ha senso assegnare a ciascun fotone una posizione ed una velocità entro lo sciame, cosicchè nessun calcolo, nemmeno in teoria, può farci prevedere se nell'elemento dS , in un dato istante, si trovi o no un fotone: tutto ciò che si può dire è che la probabilità di trovarcelo, quando lo si osservi, è $\frac{W}{w} dS$. In questo caso dunque il concetto di probabilità entra nella questione non per la nostra ignoranza di alcuni elementi del problema, ma per la natura stessa del fenomeno. Tuttavia,

una volta rilevata questa importante differenza concettuale tra i due casi, si può talvolta prescindere formalmente, ossia usare a proposito dei fotoni lo stesso linguaggio che si userebbe per dei proiettili materiali, a patto però di considerare *imperfettamente conosciuti* (ossia, conosciuti con un errore di cui è nota la legge probabilistica) gli elementi del loro movimento (che in realtà sono, almeno in parte, *fisicamente indeterminati*, ossia *concettualmente inosservabili*, come si preciserà meglio in seguito).

Bisogna sempre tener presente che il dire che gli elementi del moto di un fotone (posizione, traiettoria, impulso, ecc.) non sono « fisicamente determinati » ha un significato profondamente diverso dal dire che questi elementi, pur essendo fisicamente determinati, non sono « conosciuti », ma si potrebbero conoscere a priori con una osservazione concettualmente effettuabile senza perturbare il fenomeno. Per comprendere la differenza, si pensi all'esempio seguente. Tra una sorgente O , che emette dei fotoni (a intervalli di tempo eventualmente anche grandi) ed uno schermo s su cui essi vengono ricevuti, c'è un tramezzo opaco con due fori A e B : si formano allora sullo schermo delle frange di interferenza, e per ciascun fotone è *indeterminato* se è passato dal foro A o dal foro B : qualunque tentativo di decidere la questione p. es. intercettando uno dei due cammini altererebbe il fenomeno di interferenza. Supponiamo ora invece che vi sia un dispositivo che automaticamente chiuda ora l'uno ora l'altro dei due fori, e questo irregolarmente, in modo che il passaggio di ciascun fotone da A o da B venga a dipendere da una specie di sorteggio automatico: per un osservatore che non veda il meccanismo è *sconosciuto* se un dato fotone è passato da A o da B , ma, contrariamente al caso precedente, qui è concettualmente possibile di decidere la questione senza alterare il fenomeno. E la differenza si rispecchia nel fatto che, in questo caso, sullo schermo s non si formano frange di interferenza: il fenomeno è dunque fisicamente diverso.

§ 19. — SIGNIFICATO PROBABILISTICO DELL'OTTICA ONDULATORIA.

Esaminiamo ora il significato che deve attribuirsi alla teoria ondulatoria della luce, ed i suoi rapporti coi fatti sperimentali. È noto che tutte le esperienze di formazione di immagini, interferenza, diffrazione, ecc., confermano concor-

demente ed esattamente le previsioni della teoria ondulatoria. Tali conferme consistono sostanzialmente in questo: data la forma, la posizione e la natura delle sorgenti e dei vari mezzi (lenti, schermi, specchi), si calcola mediante le leggi dell'ottica (ossia, in definitiva, mediante l'equazione delle onde) quale deve essere l'intensità di illuminazione I nei vari punti di uno schermo, che in genere è la retina o una lastra fotografica: si fa l'esperienza, e si trova che l'intensità misurata corrisponde, punto per punto, alla funzione I risultante dal calcolo, riproducendo quindi la forma prevista delle immagini, delle ombre, delle frange di interferenza, ecc. Data la già rilevata natura discontinua dei fenomeni di assorbimento, l'« intensità di illuminazione misurata », nel caso di luce monocromatica al quale per ora ci limitiamo, divisa per l'energia w di un fotone, rappresenta il numero, per unità di superficie e di tempo, dei fotoni che arrivano sullo schermo: la verifica sta dunque nel constatare che su un elemento $d\sigma$, nel tempo dt , arrivano proprio $\frac{I}{w} d\sigma dt$ fotoni, dove I è data, in funzione del posto, dall'integrazione dell'equazione delle onde (precisamente, I è la proiezione del vettore di Poynting sulla normale alla superficie).

Si noti però che questo si verifica solo se i fotoni sono molto numerosi (come avviene di solito nelle esperienze di ottica): se invece essi sono pochi, o magari uno solo, l'ottica ondulatoria cade in difetto. Se, p. es., la sorgente emette un solo quanto di energia di fronte ad una lastra fotografica, l'ottica ondulatoria ci fa prevedere che l'energia debba ripartirsi su tutta la lastra con una certa intensità I , mentre in realtà la ritroveremo tutta in un solo punto (sono questi casi appunto che hanno suggerito il modello corpuscolare). Quale significato può attribuirsi allora alla intensità I data dall'ottica ondulatoria? Evidentemente, essa misura la « *densità di probabilità* » che il fotone arrivi in un punto della lastra piuttosto che in un altro: precisamente, se i fotoni emessi sono n , $\frac{I}{nw} d\sigma dt$ è la probabilità che ciascun fotone emesso ha di arrivare sull'elemento $d\sigma$ nel tempo dt .

Questa interpretazione probabilistica della I è valida in ogni caso: se poi i fotoni sono molto numerosi, in virtù della legge dei grandi numeri (o di Bernouilli) si può essere praticamente certi che ne arriverà effettivamente su $d\sigma$, nel tempo dt , un numero eguale a questa probabilità moltiplicata per n , cioè a $\frac{I}{w} d\sigma dt$, portando l'energia $I d\sigma dt$: così si ritrova il significato ordinario di I . Si riconosce così che *l'ottica ondulatoria non è che un metodo per calcolare la densità di probabilità di distribuzione dei fotoni*, e quando poi i fotoni sono molto numerosi, per la legge dei grandi numeri la loro densità effettiva di distribuzione risulta proporzionale a detta densità di probabilità, il che rende ragione dell'accordo della teoria con l'esperienza in tutti gli ordinari fenomeni di ottica.

Troviamo così che mentre, come si è detto, è vano ricercare delle leggi di movimento dei fotoni come si possiedono per i corpuscoli materiali, possediamo però un mezzo per determinare statisticamente la loro distribuzione, e questo mezzo è costituito dalle leggi dell'ottica ondulatoria.

§ 20. — PRINCIPIO DI SOVRAPPOSIZIONE DELL'OTTICA.

Passiamo ora al caso di radiazione non monocromatica, p. es., di due radiazioni di frequenza ν_1 , ν_2 , aventi anche, eventualmente, direzione e intensità differenti. L'esperienza insegna che le manifestazioni di queste due radiazioni sono *additive*, e cioè che i loro effetti si sovrappongono senza perturbarsi reciprocamente: ciò che è rispecchiato, nella teoria ondulatoria, dalla linearità delle equazioni differenziali che la governano. Tale additività si conserva anche nell'interpretazione probabilistica: se le onde di frequenza ν_1 danno su una certa superficie un'intensità di illuminazione I_1 , e quelle di frequenza ν_2 un'intensità I_2 , vi è una probabilità proporzionale a $\frac{I_1}{w_1}$ che l'elemento di superficie $d\sigma$ (supposto perfettamente assorbente) assorba un quanto $w_1 = h\nu_1$ nel tempo dt , e una probabilità (indipendente) proporzionale a $\frac{I_2}{w_2}$ che esso assorba un quanto $w_2 = h\nu_2$. Tale indipen-

denza delle probabilità (estensibile a quante si vogliano componenti monocromatiche della radiazione) prende il nome di *principio di sovrapposizione* dell'ottica.

Aggiungiamo due parole riguardo al caso in cui si osservi l'*impulso* dei fotoni, come nell'effetto Compton. Se l'ottica ondulatoria dà, nel punto considerato, il passaggio di onde piane monocromatiche aventi un certo vettore di propagazione \mathbf{k} , ciò significa (v. nota a pag. 125) che tutti i fotoni osservati avranno l'impulso $\mathbf{p} = h\mathbf{k}$. Se invece nel punto considerato si sovrappongono p. es. due treni d'onde piane monocromatiche, di vettori di propagazione \mathbf{k}_1 e \mathbf{k}_2 , e di intensità I_1 , I_2 , allora il principio di sovrapposizione dice che vi sarà una certa probabilità, proporzionale a $\frac{I_1}{w_1}$, di trovare un fotone con l'impulso $\mathbf{p}_1 = h\mathbf{k}_1$, ed una probabilità proporzionale a $\frac{I_2}{w_2}$ di trovare un fotone con l'impulso $\mathbf{p}_2 = h\mathbf{k}_2$. Ed analogamente per la sovrapposizione di quanti si vogliano treni d'onde, vale a dire per una radiazione qualunque.

Si osservi che le condizioni descritte (sovrapposizione di treni d'onde di diversa frequenza e direzione) possono venir rappresentate intuitivamente nel modello corpuscolare, quando i fotoni sono molto numerosi (basta pensare ad una miscela di corpuscoli di energia ed impulsi diversi) ma, se, si tratta di uno solo (o di pochi), esse costituiscono uno stato di cose non rappresentabile con alcun modello classico. Difatti, finchè non si compie un'esperienza per determinare (alterandoli) la frequenza e l'impulso del fotone, non è concettualmente possibile prevedere che risultato darà tale osservazione, ma si possono solo assegnare le probabilità dei vari possibili risultati: ciò significa che il fotone non possiede, fino al momento dell'osservazione, una energia ed un impulso determinati (si ha cioè per l'energia e per l'impulso una *indeterminazione* perfettamente analoga a quella già considerata poco sopra per la posizione dei singoli fotoni). Può essere *talvolta* comodo ricorrere anche in questo caso al modello corpuscolare, sostituendo questa indeterminazione col supporre che l'energia e l'impulso del corpuscolo siano sconosciuti (o meglio che se ne conoscano solo le leggi di probabilità), ma bisogna tener sempre presente che è questo un linguaggio improprio, come mostra l'esempio del § 18.

§ 21. — PRINCIPIO DI INDETERMINAZIONE PER I FOTONI.

Vogliamo ora precisare i limiti di validità del modello corpuscolare, ossia i limiti entro i quali ha un senso parlare di «posizione» ed «impulso» di un fotone in un istante determinato. Riprendiamo perciò l'esempio di un «pacchetto di luce» S come quello considerato ai §§ 15 e 19: supposto che la sorgente abbia emesso un solo fotone, potremo dire con certezza che esso si trova nell'interno della regione S , costituente il pacchetto (determinabile con le leggi dell'ottica ondulatoria), cioè, pur non avendo senso parlare delle esatte coordinate del fotone, potremo assegnare a queste dei limiti tanto più ristretti quanto più piccolo è il pacchetto.

Occupiamoci ora dell'impulso del fotone. Ricordiamo perciò (v. § 15) che il pacchetto d'onde si può considerare risultante dalla sovrapposizione di infiniti treni d'onde monocromatiche di diverso vettore di propagazione \mathbf{k} , secondo la formula di decomposizione (79) del cap. I: per il principio di sovrapposizione, a ciascuno di questi treni corrisponde una possibilità per l'impulso \mathbf{p} del fotone, e cioè, alle onde di vettore \mathbf{k} , corrisponde la possibilità di trovare nel fotone l'impulso $\mathbf{p} = h\mathbf{k}$. Vi è quindi un'indeterminazione anche riguardo all'impulso, come alla posizione. Precisamente, se ricordiamo che l'ampiezza dei treni d'onde aventi vettore di propagazione compreso tra (k_x, k_y, k_z) e $(k_x + dk_x, k_y + dk_y, k_z + dk_z)$ è $A dk_x dk_y dk_z$, dove $A(\mathbf{k})$ è dato dalla formula (80), riconosciamo facilmente che la probabilità che l'impulso del fotone sia compreso tra (p_x, p_y, p_z) e $(p_x + dp_x, p_y + dp_y, p_z + dp_z)$ è proporzionale a $|A|^2 dp_x dp_y dp_z$; questa probabilità si può esprimere mediante p_x, p_y, p_z anziché k_x, k_y, k_z semplicemente esprimendo \mathbf{p} in funzione di \mathbf{k} nella (80), che diventa

$$(80') \quad A\left(\frac{\mathbf{p}}{h}\right) = \int_{-\infty}^{\infty} \int \int f(\mathbf{r}, 0) e^{\frac{2\pi i}{h} \mathbf{p} \cdot \mathbf{r}} dx dy dz.$$

Per rendere la cosa più espressiva, conviene introdurre uno spazio rappresentativo degli impulsi, nel quale ogni vettore \mathbf{p} è rappresentato da un punto di coordinate p_x, p_y, p_z :

potremo allora dire che, come la posizione del fotone è indeterminata nello spazio x, y, z , e la sua densità di probabilità nei vari punti è proporzionale alla densità di energia $W(x, y, z)$ (e quindi a $|f|^2$), così la densità di probabilità dell'impulso nello spazio p_x, p_y, p_z è rappresentata dalla funzione $|A|^2$, (la quale pure definisce un pacchetto d'onde, nello spazio degli impulsi). Come misura della indeterminazione delle coordinate si possono prendere le semidimensioni $\Delta x, \Delta y, \Delta z$ del pacchetto d'onde, definite come al § 13: similmente, l'indeterminazione delle tre componenti dell'impulso sarà misurata dalle quantità:

$$(93) \quad \Delta p_x = h\Delta k_x, \quad \Delta p_y = h\Delta k_y, \quad \Delta p_z = h\Delta k_z,$$

dove $\Delta k_x, \Delta k_y, \Delta k_z$ sono definite dalla (82) e dalle altre due analoghe.

Ricordiamo ora il teorema dimostrato al § 15, secondo il quale, più il pacchetto d'onde nello spazio x, y, z è ristretto, più devono differire tra loro i vettori di propagazione delle onde che lo compongono, ossia più è ampio il pacchetto nello spazio delle p , e giungeremo alla importante conclusione che *quanto più esattamente è determinata la posizione di un fotone, tanto più indeterminato ne è l'impulso e viceversa*. Più precisamente, tra le indeterminazioni $\Delta x, \Delta y, \Delta z$ delle coordinate e quelle $\Delta p_x, \Delta p_y, \Delta p_z$ dei rispettivi impulsi, sussistono le disuguaglianze seguenti, che si deducono immediatamente dalle (83) moltiplicandole per h ,

$$(94) \quad \Delta x \Delta p_x \geq \frac{h}{4\pi}, \quad \Delta y \Delta p_y \geq \frac{h}{4\pi}, \quad \Delta z \Delta p_z \geq \frac{h}{4\pi},$$

Con ciò resta precisata quella parziale indeterminazione, di cui si è fatto cenno al § 18, che è necessario applicare al modello corpuscolare, per adeguarlo ai fatti sperimentali. Si noti che non è teoricamente impossibile localizzare un fotone, in un dato istante, con precisione tanto grande quanto si vuole: però resta allora completamente indeterminato il suo impulso (quindi la sua energia e la sua direzione di propagazione).

Similmente, dalla (84) moltiplicandola per h e ricordando che l'energia w dei fotoni è data da $h\nu$, si ha

$$(95) \quad \Delta w \Delta t \geq \frac{h}{4\pi},$$

la quale esprime che: *quanto più esattamente è determinato l'istante del passaggio di un fotone per un determinato punto dello spazio, tanto più indeterminata ne è l'energia e viceversa.*

Le (94) e la (95) esprimono il *principio di indeterminazione per i fotoni*: esse sono, come si vede, conseguenze puramente matematiche del fatto, provato dall'esperienza, che la probabilità di distribuzione dei fotoni è calcolabile con la teoria ondulatoria.

In molti casi non è necessario precisare il valore dei secondi membri delle (94), (95), ma basta tener presente che essi sono dell'ordine di grandezza di h , ed assumere per Δx , ecc., Δp_x , ecc., una qualsiasi misura dell'ordine di grandezza dell'indeterminazione: la portata concettuale del principio rimane la stessa. Perciò il principio di indeterminazione spesso si esprime con le formule

$$(94') \quad \Delta x \Delta p_x \geq \sim h, \quad \Delta y \Delta p_y \geq \sim h, \quad \Delta z \Delta p_z \geq \sim h,$$

$$(95') \quad \Delta w \Delta t \geq \sim h,$$

(dove il segno \sim indica che si tratta degli ordini di grandezza).

Per meglio illustrare il principio e per mostrare di che natura sia l'indeterminazione di cui si parla, consideriamo un caso unidimensionale (limitandoci alla valutazione degli ordini di grandezza). Si abbia un gruppo d'onde luminose di lunghezza complessiva $2l$, come quello rappresentato dalla fig. 19. In questo caso, la determinazione dell'impulso \mathbf{p} si riduce alla misura del suo modulo $p = h/\lambda$, cioè alla misura della lunghezza d'onda: pensiamo perciò di ricevere la luce su un reticolo. Questo permetterà di misurare λ con una approssimazione $\delta\lambda$ il cui ordine di grandezza è dato (come è noto dalla teoria del potere risolutivo dei reticoli) da:

$$\frac{\delta\lambda}{\lambda} = \sim \frac{\lambda}{L},$$

dove L è la massima differenza di cammino ottico utilizzata dal reticolo. Ora questa non può evidentemente superare la lunghezza $2l$ del gruppo d'onde, e quindi:

$$\frac{\delta\lambda}{\lambda} \geq \sim \frac{\lambda}{2l};$$

e poichè si ha (precindendo dal segno) $\frac{\delta p}{p} = \frac{\delta\lambda}{\lambda}$, la precedente dà

$$\frac{\delta p}{p} \geq \sim \frac{\lambda}{2l}, \quad \text{ossia} \quad \delta p \geq \sim \frac{h}{2l}.$$

Questa relazione tra gli ordini di grandezza di $2l$ e di δp (che con considerazioni più dettagliate si potrebbe precisare) mostra appunto che la p risulta tanto più indeterminata, quanto più corto è il gruppo d'onde, cioè quanto più esattamente è determinata, in ciascun istante, la posizione dei fotoni. Il ragionamento, fatto riferendosi al reticolo, si può naturalmente ripetere per qualsiasi altro mezzo di analisi spettrale, poichè esso ha il suo fondamento nel fatto, puramente analitico, che il gruppo d'onde equivale a tutti gli effetti (come si è mostrato al § 12) alla sovrapposizione di treni d'onde di diverse frequenze.

Il principio di indeterminazione espresso dalle (94), (95) rappresenta, per così dire, il correttivo da applicarsi al modello corpuscolare dei fotoni per tener conto della sua incompleta aderenza alla realtà. Analogamente si potrebbe mostrare che, quando si usa il modello delle onde elettromagnetiche, si deve introdurre una certa *relazione di indeterminazione* nelle misure dei campi \mathbf{E} ed \mathbf{H} (v. bibl., n° 10 e 10^{bis}), dalla quale dipende, p. es., il paradosso dell'effetto fotoelettrico segnalato al § 8 p. I.

§ 22. — PRINCIPIO DI INDETERMINAZIONE PER LE PARTICELLE MATERIALI.

Il fondamento concettuale della meccanica quantistica sta nel fatto (rivelato per la prima volta da HEISENBERG) che il principio di indeterminazione, illustrato nel § precedente, vale non solo per i fotoni, ma anche per qualunque « particella » materiale, come per es. gli elettroni ⁽¹⁾. Vale

⁽¹⁾ Da ciò segue che è improprio usare la parola « particella » per designare questi enti che (al pari dei fotoni) hanno soltanto alcune delle proprietà che corrispondono al concetto ordinario di particella o corpuscolo materiale. Tuttavia, per comodità, useremo correntemente questa parola per indicare collettivamente gli elettroni, i protoni, i neutroni, i nuclei, etc., avvertendo una volta per tutte che la parola non ha il significato ordinario.

a dire: è concettualmente impossibile determinare la posizione e l'impulso di una particella in un dato istante, con una precisione superiore a quella consentita dalle disuguaglianze (94); così pure è concettualmente impossibile, mediante un'esperienza che duri un tempo Δt , determinare l'energia della particella con un'incertezza inferiore a Δw , legato a Δt dalla (95).

Naturalmente (come per i fotoni), ciò non significa che sia impossibile conoscere la posizione della particella con tutta la precisione che si desidera *oppure* il suo impulso (e quindi la sua velocità) con tutta la precisione voluta, ma il principio afferma che non si possono conoscere simultaneamente entrambi questi dati: più esattamente, esso vieta di conoscere simultaneamente, con precisione assoluta, una coordinata e la corrispondente componente dell'impulso (o della velocità). Questa indeterminazione deriva (come quella per i fotoni) dalla inadeguatezza del modello corpuscolare, e segna i limiti di validità di questo. Tale principio, sebbene valido per qualunque corpo, ha in pratica conseguenze importanti solo nei riguardi delle particelle di dimensioni atomiche o subatomiche, perchè, quando si tratta di corpi ordinari, data la piccolezza della costante h , che figura nei secondi membri delle (94), (95), le incertezze Δx , Δp_x , ecc. volute dal principio di indeterminazione scompaiono di fronte a quelle, ben più notevoli, derivanti da errori accidentali di misura o dalla imperfetta definizione dei punti di riferimento.

Che il principio di indeterminazione debba valere anche per le particelle materiali, lo si può dedurre logicamente dalla validità del principio analogo per i fotoni, e dalle leggi di conservazione dell'energia e dell'impulso che, come è sperimentalmente accertato, sono verificate in tutti i fenomeni di interazione tra particelle materiali e fotoni. Difatti, se non valesse il principio di indeterminazione per le particelle materiali, si potrebbe approfittare di uno di questi fenomeni di interazione (p. es. emissione, assorbimento o effetto Compton) per infrangere il principio di indeterminazione per i fotoni: basterebbe determinare le coordinate della particella materiale durante l'interazione col fotone, e il suo impulso

prima e dopo tale interazione, con precisione superiore a quella corrispondente alle (94), per conoscere con pari precisione le coordinate del fotone, e l'impulso impressogli: p. es., nel caso unidimensionale citato alla fine del § precedente, basterebbe misurare p , anzichè con un reticolo, facendo assorbire il fotone da un atomo (di cui si sia determinata anche la x) e misurando l'impulso acquistato da questo. Similmente, detta Δt la durata dell'interazione, se fosse possibile una misura dell'energia della particella, in un tempo piccolo rispetto a Δt , con una incertezza inferiore a $\frac{h}{4\pi\Delta t}$, si potrebbe,

compiendo questa misura immediatamente prima e dopo l'interazione, determinare l'energia ceduta o tolta al fotone con precisione superiore a quella consentita dalla (95).

Ma per comprendere la natura del principio di indeterminazione, bisogna esaminare (per così dire) il meccanismo mediante il quale esso è realizzato nella natura, che è il seguente. Quando si tenta di determinare, con una approssimazione Δx , la coordinata x di una particella materiale, si è obbligati a comunicare ad essa un impulso p_x il cui ordine di grandezza è $\frac{h}{\Delta x}$, ma il cui valore preciso è concettualmente impossibile conoscere: cosicchè se una misura precedente aveva fatto conoscere il valore di p_x , tale conoscenza va parzialmente perduta per la previsione dei fenomeni che avverranno dopo la misura di x ; similmente, una misura della componente p_x dell'impulso, con una approssimazione Δp_x , implica che si imprima alla particella uno spostamento lungo l'asse x di entità concettualmente indeterminata, e di ordine di grandezza $h/\Delta p_x$ cosicchè, se prima di questa misura si conosceva la x , tale conoscenza diviene dopo quell'istante parzialmente inutilizzabile. Si può dunque dire che per la misura di una coordinata, e per quella del corrispondente impulso (o componente di velocità), si devono adoperare dispositivi che si escludono mutuamente: è per questo che i due concetti di posizione e velocità, sebbene separatamente abbiano significato fisico preciso, non possono venire attribuiti simultaneamente in modo preciso alla stessa particella:

questa particolare relazione logica tra i due concetti è stata designata dal BOHR col nome di *complementarità* (1).

Mostreremo ora su qualche esempio la incompatibilità fisica dei due tipi di misure, rimandando il lettore, per una discussione più esauriente, al libro di HEISENBERG (v. bibl., n. 10 e 10^{bis}).

§ 23. — MISURA DELLE COORDINATE DI UNA PARTICELLA.

1° Metodo. — Per determinare la posizione di una particella si può idealmente usare lo stesso procedimento con cui ordinariamente si esaminano i corpi, cioè illuminarla e raccogliere la radiazione diffusa in uno strumento ottico (microscopio o camera fotografica). Ciò che impedisce praticamente di eseguire l'esperienza è soltanto la scarsissima intensità della luce diffusa, ma questa difficoltà pratica nulla toglie alla legittimità del ragionamento che segue.

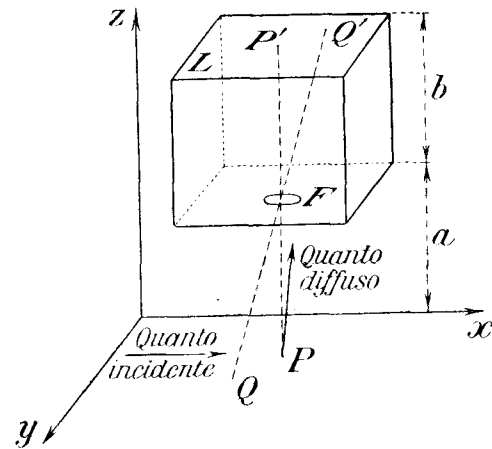


Fig. 22.

Supponiamo dunque che della luce (o, più generalmente, della radiazione) di lunghezza d'onda λ , propagandosi nel senso dell'asse x (fig. 22), arrivi sulla particella della quale vogliamo determinare la posizione: disponiamo poi, nella direzione dell'asse z , uno strumento ottico che, per semplicità, supporremo sia una camera oscura (senza obiettivo) munita di un foro di centro F e raggio R (il ragionamento si potrebbe anche fare sostituendovi una camera con obiettivo o un microscopio). Il dispositivo si presta allora alla determinazione approssimata delle due coordinate x ed y della particella in un certo istante. Difatti, se la particella stesse ferma, p. es. in P , e l'esperienza durasse abbastanza a lungo, sulla lastra L si formerebbe una figura di diffrazione,

(1) V. bibl. n. 21. Si vedrà più oltre che tale relazione si può estendere anche ad altre coppie di grandezze fisiche.

che possiamo schematizzare in un dischetto uniforme di centro P' e raggio

$$r' = \frac{\lambda b}{R},$$

dove b è la distanza del foro F dalla lastra. Ma sappiamo, dalle esperienze sull'effetto Compton, che la particella diffonde la luce per quanti e che ad ogni atto di diffusione riceve un impulso che ne altera la velocità: quindi dovremo limitare l'esperienza ad un solo atto di diffusione, che ci darà sulla lastra non tutta la figura di diffrazione, ma solo un punto Q' di essa: la posizione del punto P' , (dove si formerebbe il centro della figura di diffrazione se la particella continuasse a diffondere quanti sempre dalla stessa posizione) rimane indeterminata, e tutto quello che si può dire di esso, è che la sua distanza da Q' non supera r' . Quindi, se da F il punto Q' si proietta sul piano xy in un punto Q , si può dire che la particella P si trovava, all'atto della diffusione, entro un cerchio di centro Q e raggio

$$r = \frac{a}{b} r' = \frac{\lambda a}{R},$$

dove a è la distanza di F dal piano xy . Così le coordinate x ed y restano determinate con un'incertezza Δx , Δy dell'ordine di grandezza di r :

$$(96) \quad \Delta x \sim \frac{\lambda a}{R}, \quad \Delta y \sim \frac{\lambda a}{R}.$$

(Di qui si vede che, per avere una determinazione assai precisa di x ed y , conviene, a parità di altre condizioni, scegliere radiazioni di lunghezza d'onda assai piccola, p. es. raggi γ anziché luce: ciò del resto corrisponde al fatto ben noto che il potere risolutivo degli strumenti ottici è inversamente proporzionale alla lunghezza d'onda della luce impiegata).

Occupiamoci ora della velocità, o dell'impulso, con cui rimane la particella dopo l'esperienza. Se inizialmente essa aveva un impulso di componenti p_x^0 , p_y^0 , p_z^0 (che, per metterci nelle condizioni più favorevoli, supporremo esattamente note; eventualmente nulle), dopo la misura delle coordinate x , y il suo impulso p_x , p_y , p_z sarà variato per l'impulso ricevuto dal quanto (come sappiamo dalla teoria dell'effetto Compton, confermata dall'esperienza). Ora il quanto, prima della diffusione, aveva un impulso di componenti

$$(97) \quad \frac{h}{\lambda}, \quad 0, \quad 0;$$

dopo la diffusione (se chiamiamo α, β, γ gli angoli formati con gli assi coordinati dalla direzione nella quale il quanto è stato diffuso) l'impulso del quanto avrà le componenti

$$\frac{h}{\lambda} \cos \alpha, \quad \frac{h}{\lambda} \cos \beta, \quad \frac{h}{\lambda} \cos \gamma;$$

quindi l'impulso posseduto dalla particella dopo la diffusione sarà dato da

$$(98) \quad p_x = p_x^0 + \frac{h}{\lambda} (1 - \cos \alpha), \quad p_y = p_y^0 - \frac{h}{\lambda} \cos \beta, \quad p_z = p_z^0 - \frac{h}{\lambda} \cos \gamma.$$

Se fossero noti α, β, γ , sarebbero determinati p_x, p_y, p_z ; bisogna invece osservare che la retta lungo la quale il quanto è stato diffuso non è in alcun modo determinabile, come risulta dalle considerazioni del § 18: si può dire solo che essa passa attraverso il foro della camera oscura, e quindi che $\gamma \leq \gamma_0$, chiamando γ_0 l'angolo sotto cui da P è visto il raggio R del foro, che è dato da $\operatorname{tg} \gamma_0 = R/a$, ovvero, trattandosi di angoli piccoli,

$$(99) \quad \sin \gamma < \frac{R}{a}.$$

Siccome poi, evidentemente,

$$\cos^2 \alpha + \cos^2 \beta = 1 - \cos^2 \gamma = \sin^2 \gamma,$$

si ha:

$$\cos^2 \alpha + \cos^2 \beta < \frac{R^2}{a^2},$$

e questa è la sola limitazione per α e β . Tutto ciò che si può ricavare da essa, riguardo a $\cos \alpha$ e $\cos \beta$ separatamente, è che ciascuno di essi deve esser compreso tra $-R/a$ e $+R/a$. Segue di qui e dalle (98) che p_x può variare entro i limiti

$$p_x^0 + \frac{h}{\lambda} \left(1 \pm \frac{R}{a} \right)$$

e p_y entro i limiti

$$p_y^0 \pm \frac{h}{\lambda} \frac{R}{a};$$

quindi l'incertezza è per entrambi:

$$(100) \quad \Delta p_x = \sim \frac{h}{\lambda} \frac{R}{a}, \quad \Delta p_y = \sim \frac{h}{\lambda} \frac{R}{a}.$$

Da queste e dalle (96) si ha, per moltiplicazione

$$\Delta x \Delta p_x = \sim h, \quad \Delta y \Delta p_y = \sim h,$$

conformemente alle (94'). Similmente si potrebbe ragionare ri-

guardo alla coordinata z ed al rispettivo impulso (disponendo altrimenti la camera oscura).

Lo stesso risultato si troverebbe se — nel caso di una particella luminosa o radioattiva — si utilizzasse la radiazione da essa emessa anziché quella diffusa: difatti l'emissione di un quanto è accompagnata da un rinculo che comunica alla particella un impulso opposto a quello del quanto emesso.

2° Metodo. — Se un fascio parallelo di particelle (p. es. raggi catodici) viene lanciato perpendicolarmente contro uno schermo AB (fig. 23) munito di una fessura di larghezza d , ogni volta che una particella passa attraverso la fessura si può dire che si è determinata la sua coordinata y (supposto che l'asse y sia posto trasversalmente alla fessura) con una incertezza

$$(101) \quad \Delta y = \sim d.$$

Si rammenti ora che risulta dall'esperienza (v. cap. IV, p. I) che un fascio di particelle materiali di impulso p subisce dei fenomeni di diffrazione corrispondenti a quelli di onde di lunghezza $\lambda = h/p$:

quindi al di là della fenditura il fascio non sarà più parallelo, ma avrà (nella direzione y) un'ampiezza angolare $2\alpha_0$ data, secondo le leggi elementari della diffrazione, da

$$\sin \alpha_0 = \sim \frac{\lambda}{d} = \sim \frac{h}{pd}.$$

Ciò significa che la particella, attraversando la fessura, potrà essere deviata dalla direzione primitiva di un angolo α che può andare da $-\alpha_0$ a $+\alpha_0$: perciò la componente $p_y = p \sin \alpha$ del suo impulso che originariamente era zero, resta indeterminata entro i limiti $\pm p \sin \alpha_0$, cioè con un'incertezza

$$(102) \quad \Delta p_y = \sim p \sin \alpha_0 = \sim \frac{h}{d}.$$

Da questa e dalla (101) si ha allora, conformemente a (94')

$$\Delta y \Delta p_y = \sim h,$$

ed analogamente si potrebbe ragionare per x e z .

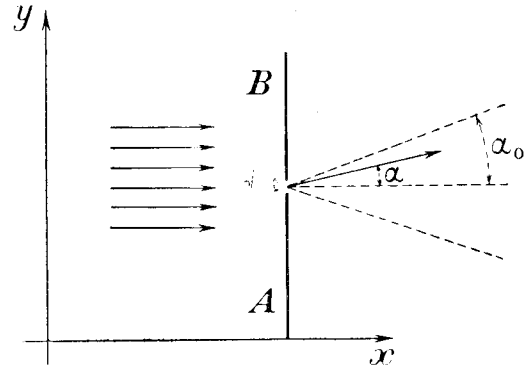


Fig. 23.

3° *Metodo*. — Sappiamo che particelle materiali di grande energia, come le particelle α , possono rendersi isolatamente visibili, sia mediante la scintillazione che producono urtando contro uno schermo fluorescente (spintariscopio), sia facendo attraversare loro un gas soprassaturo di vapore acqueo, il quale si condensa, in forma di nebbia, sugli ioni prodotti dalla particella lungo il suo cammino (camera ad espansione, di Wilson). In entrambi questi casi si localizza, veramente, non la particella in questione, ma l'atomo da essa urtato e, per conseguenza dell'urto, eccitato o ionizzato. Perciò le coordinate della particella al momento dell'urto restano determinate con una incertezza dell'ordine delle dimensioni lineari dell'atomo, che indicheremo con

$$(103) \quad \Delta x, \Delta y, \Delta z.$$

Vediamo ora ciò che può dirsi dell'impulso con cui rimane la particella dopo l'urto: esso è uguale, evidentemente, all'impulso iniziale (che supponiamo noto) meno l'impulso comunicato a quell'elettrone che viene eccitato od espulso. Per determinare quest'ultimo impulso, dobbiamo fare dunque la differenza vettoriale tra l'impulso che ha l'elettrone dell'atomo dopo l'urto e prima dell'urto: il primo può misurarsi con tutta la precisione voluta, con uno dei metodi del § seguente, ma sull'impulso che aveva l'elettrone entro l'atomo prima di venire urtato si può dire soltanto che le sue componenti sono comprese entro certi limiti, tra i quali oscillavano periodicamente per effetto del movimento orbitale. L'ampiezza di queste oscillazioni si valuta facilmente nel caso delle orbite circolari di Sommerfeld: essa è evidentemente (prendendo gli assi x ed y nel piano dell'orbita)

$$\Delta p_x = \Delta p_y = 2mv;$$

e poichè per la condizione di quantizzazione

$$mvr = n \frac{h}{2\pi} \quad (\text{v. § 16 parte I}),$$

sarà

$$\Delta p_x = \Delta p_y = \frac{nh}{\pi r}.$$

D'altra parte l'incertezza su x ed y è data in questo caso dalle dimensioni dell'orbita, cioè

$$\Delta x = \Delta y = 2r,$$

cosicchè, come si vede,

$$\Delta x \Delta p_x = \Delta y \Delta p_y = \frac{2nh}{\pi}.$$

Nel caso più generale di orbite qualunque si troverebbe un risultato dello stesso ordine di grandezza, e cioè in generale

$$\Delta x \Delta p_x \approx \sim nh, \quad \Delta y \Delta p_y \approx \sim nh.$$

Siccome poi n , nel caso più favorevole, ha il valore 1, si ritrovano le relazioni (94').

§ 25. — MISURA DELL'IMPULSO O DELLA VELOCITÀ DI UNA PARTICELLA.

1° *Metodo*. — Il procedimento più naturale per determinare la velocità di una particella consiste nel determinarne la posizione in due istanti t_1, t_2 , separati da un intervallo di tempo noto. Sappiamo però che ogni determinazione di x, y, z introduce un'incertezza su p_x, p_y, p_z data, nelle condizioni più favorevoli, dalle (94): perciò la velocità con cui rimane la particella dopo l'istante t_2 non è uguale a quella che aveva nell'intervallo fra t_1 e t_2 e che sola risulta dalla nostra misura: l'alterazione resta indeterminata, e si sa solo che le sue componenti hanno l'ordine di grandezza dato da

$$(105) \quad \Delta v_x \approx \sim \frac{h}{m \Delta x_2}, \quad \Delta v_y \approx \sim \frac{h}{m \Delta y_2}, \quad \Delta v_z \approx \sim \frac{h}{m \Delta z_2},$$

(chiamando $\Delta x_2, \Delta y_2, \Delta z_2$ le incertezze tollerate nella misura delle coordinate fatta al tempo t_2). Ora, prendendo $t_2 - t_1$ abbastanza grande, si potrà fare in modo che questi errori influiscano tanto

poco quanto si vuole sulla misura dei rapporti $\frac{x_2 - x_1}{t_2 - t_1}$ ecc.,

e quindi questi si potranno considerare esattamente determinabili, ma quella che ci interessa, cioè la velocità dopo t_2 , resterà affetta da una incertezza espressa dalle (105), e che si può rendere piccola soltanto accontentandosi di poca esattezza nella misura della posizione al tempo t_2 . Le (105), che esprimono questa relazione tra le due approssimazioni, non sono altro che le relazioni di indeterminazione (94').

Si osservi che la velocità tra i due istanti t_1 e t_2 (che può essere calcolata, come abbiamo detto, con tutta l'esattezza voluta), è una quantità priva di interesse fisico, poichè la sua definizione stessa presuppone che nell'intervallo considerato la particella non interagisca con il mondo circostante.

2° *Metodo*. — Per misurare la velocità di una particella senza ricorrere a due successive osservazioni di posizione, si può uti-

lizzare l'effetto Doppler. P. es. per misurare la componente v_x si può ricorrere allo schema seguente.

Si manda nella direzione dell'asse x della luce di frequenza nota ν : si raccoglie poi in uno spettroscopio la radiazione diffusa dalla particella nel verso negativo dell'asse x , e si determina la sua frequenza ν' : si ha allora da un ragionamento elementare di ottica:

$$(106) \quad \frac{\nu - \nu'}{\nu} = 2 \frac{v_x^0}{c},$$

dove v_x^0 rappresenta la velocità lungo l'asse x prima della misura. Però va tenuto presente che la particella riceve un impulso $2h\nu/c$ nell'atto della diffusione, e che quindi la velocità dopo la misura (che è quella che ci interessa) è

$$v_x = v_x^0 + 2 \frac{h\nu}{mc}.$$

Poichè ν si suppone noto, la v_x resta determinata con la stessa esattezza con cui si ha la v_x^0 dalla (106), la quale esattezza dipende dalla precisione con cui si misura ν' : si ha cioè:

$$\Delta v_x = \Delta v_x^0 = \frac{c}{2\nu} \Delta \nu'.$$

Il $\Delta \nu'$, come si è visto al § 15, non può essere inferiore a $\sim \frac{1}{\Delta t}$, dove Δt è la durata dell'illuminazione ⁽¹⁾: perciò si ha, nelle condizioni più favorevoli,

$$(107) \quad \Delta v_x = \sim \frac{c}{2\nu \Delta t}.$$

D'altra parte, non si può dire in quale istante dell'intervallo Δt la particella abbia ricevuto l'impulso che ha mutato la v_x^0 in v_x : perciò sulla x della particella resta, dopo la misura, una incertezza uguale a

$$\Delta x = (v_x - v_x^0) \Delta t = \frac{2h\nu}{mc} \Delta t.$$

Da questa e dalla (107) si ricava

$$\Delta x \Delta v_x = \sim \frac{h}{m},$$

⁽¹⁾ O meglio, la durata di un gruppo d'onde *coerenti*: se l'illuminazione durasse più a lungo, ma avvenissero ogni tanto dei bruschi cambiamenti di fase, Δt rappresenterebbe l'intervallo tra due di questi cambiamenti.

ossia

$$\Delta x \Delta p_x = \sim h,$$

che è la prima delle (94').

§ 25. — IMPOSTAZIONE PROBABILISTICA DELLA MECCANICA DI UNA PARTICELLA.

Le considerazioni dei §§ precedenti mostrano che anche per le particelle materiali, come per i fotoni, la rappresentazione letteralmente corpuscolare, che è suggerita dal nome, non corrisponde esattamente alla realtà ma rappresenta solo un modello il cui limite di validità è, per così dire, definito dal principio di indeterminazione. Si deve dunque abbandonare la concezione intuitiva del moto continuo lungo una certa traiettoria, per fissare invece l'attenzione sulla distribuzione spaziale della *densità di probabilità*, la quale sarà definita, come per i fotoni, da una funzione $P(x, y, z, t)$, tale che $P dx dy dz$ rappresenti la probabilità ⁽¹⁾ che, eseguendo un'osservazione al tempo t , si trovi la particella nell'elemento di volume definito da $(x, y, z, x + dx, y + dy, z + dz)$. Analogamente si introdurrà una densità di probabilità $Q(v_x, v_y, v_z, t)$ per le componenti della velocità della particella al tempo t . E la meccanica delle particelle materiali non dovrà quindi servire a determinarne il *moto*, cioè posizione e velocità in funzione del tempo (il che contravverrebbe al principio di Heisenberg, perchè equivarrebbe a postulare l'esistenza di fenomeni atti a definire esattamente posizione e velocità di una particella in un dato istante) ma dovrà

⁽¹⁾ Per precisare il significato che si deve dare in meccanica ondulatoria alla parola « probabilità », si deve pensare di avere un *gran* numero N di sistemi indipendenti identici e sottoposti alle stesse condizioni iniziali, e di eseguire su ciascuno di essi un'osservazione della particella al tempo t : se su N' sistemi l'osservazione dà il risultato considerato, diremo che la « probabilità » di questo è N'/N . (Nel caso ottico, si poteva riferirsi alla presenza *simultanea* di N fotoni: qui non è possibile, perchè N particelle agirebbero tra loro alterando le rispettive distribuzioni di probabilità). Similmente, quando si parla del *valore medio* di una grandezza, si intende che si deve misurare questa negli N sistemi suddetti, e prendere la media.

invece avere un ufficio analogo a quello che ha l'ottica nei riguardi dei fotoni, cioè quello di determinare le densità di probabilità $P(x, y, z, t)$ e $Q(v_x, v_y, v_z, t)$. Se poi si tien presente che i fenomeni di diffrazione, che rappresentano una delle più evidenti prove del carattere ondulatorio delle leggi dell'ottica, si verificano anche per le particelle materiali (come si è visto al cap. IV, p. I), si è indotti naturalmente a ricercare se anche nel caso delle particelle materiali le densità di probabilità non possano ricavarsi da equazioni del tipo di quella delle onde (1). Questa idea si può precisare nel modo seguente (2).

Nel caso dell'ottica, la densità di energia W (e qualcosa di simile potrebbe dirsi per l'intensità I di illuminazione di una superficie) non soddisfa essa stessa ad un'equazione semplice, ma si ricava, con la formula

$$W = \frac{\varepsilon}{8\pi^2} (E_x^2 + E_y^2 + E_z^2) + \frac{\mu}{8\pi^2} (H_x^2 + H_y^2 + H_z^2),$$

dalle sei componenti del campo elettrico \mathbf{E} e di quello magnetico \mathbf{H} , ciascuna delle quali soddisfa l'equazione delle onde, che per E_x , p. es., è:

$$(108) \quad \Delta E_x = \frac{n^2}{c^2} \frac{\partial^2 E_x}{\partial t^2}$$

(con $n = \sqrt{\varepsilon\mu}$, indice di rifrazione). Similmente (3) nel caso della meccanica, anzichè cercare un'equazione differenziale a cui soddisfi P , conviene introdurre un certo numero N (per ora indeterminato) di quantità scalari, generalmente complesse, $\psi_1, \psi_2, \dots, \psi_N$, dalle quali si ricavi P con la formula

$$P = |\psi_1|^2 + |\psi_2|^2 + \dots + |\psi_N|^2,$$

(1) Sui limiti dell'analogia tra ottica e meccanica, v. PAULI, *ZS. f. Phys.*, 80 (1933), p. 573 e segg.

(2) Non occorre dire che il procedimento euristico qui riportato non riproduce affatto lo svolgimento storico della teoria (per il quale rinviamo a quanto si è detto nella parte I) nè pretende di darne una giustificazione rigorosa.

(3) Per rendere l'analogia più evidente anche formalmente, conviene rappresentare il campo elettromagnetico con l'unico vettore com-

ciascuna delle quali soddisfi ad un'equazione differenziale del secondo ordine, del tipo di quella delle onde. DIRAC ha mostrato che il minimo numero di funzioni ψ col quale si può costruire una meccanica ondulatoria adeguata ai fatti sperimentali e soddisfacente al principio di relatività è $N = 4$. Se però ci si contenta di una meccanica ondulatoria valida con l'approssimazione con cui è valida l'ordinaria meccanica non relativistica (ossia per velocità piccole rispetto a c), bastano due funzioni ψ (teoria di Pauli v. cap. V, p. III), se poi si trascurano anche gli effetti derivanti dallo *spin* (v. § 25 p. I), basta allora una sola funzione ψ , ed è questa la meccanica di Schrödinger, della quale soltanto ci occuperemo d'ora in avanti, fino al cap. V, p. III, dove verrà svolta la meccanica ondulatoria relativistica a quattro ψ , di DIRAC. Ci limiteremo anzi per ora allo studio di una sola particella.

Introdurremo dunque una funzione complessa $\psi(x, y, z, t)$, che da SCHRÖDINGER fu chiamata « scalare di campo » e che ora più generalmente si chiama « ampiezza di probabilità » per analogia con l'ampiezza delle onde luminose: essa non ha significato fisico immediato, ma serve a determinare la densità di probabilità P (susceptibile, come si è visto, di controllo sperimentale) mediante la relazione

$$(109) \quad P = |\psi|^2.$$

plesso Ψ , definito da

$$\Psi = \sqrt{\frac{\varepsilon}{8\pi^2}} \mathbf{E} + i \sqrt{\frac{\mu}{8\pi^2}} \mathbf{H}.$$

Si vede subito allora che l'espressione di W diviene

$$W = |\Psi_x|^2 + |\Psi_y|^2 + |\Psi_z|^2 = |\Psi|^2.$$

Il vettore Ψ soddisfa le equazioni

$$\text{rot } \Psi = \frac{i}{c} n \frac{\partial \Psi}{\partial t}, \quad \text{div } \Psi = 0$$

(che compendiano le equazioni di Maxwell e di Laplace per \mathbf{E} ed \mathbf{H}); da queste si ricava subito che ciascuna componente complessa di Ψ soddisfa l'equazione delle onde

$$\Delta \Psi_x = \frac{n^2}{c^2} \frac{\partial^2 \Psi_x}{\partial t^2} \text{ ecc.}$$

Ammetteremo poi che la ψ soddisfi una equazione differenziale analoga alla (106), e cioè ⁽¹⁾

$$(108') \quad \Delta \psi = N^2 \frac{\partial^2 \psi}{\partial t^2}$$

dove N (che rappresenta l'inverso della velocità di fase, ed è analogo ad n/c della (108)) sarà determinato, in modo che vedremo, in funzione di x, y, z . La ψ si propagherà dunque per onde (onde di De Broglie), le quali però non hanno alcuna esistenza materiale, ma rappresentano solo un mezzo analitico per calcolare la densità di probabilità P .

Rileviamo fin d'ora che l'integrale $\int P$ esteso a tutto lo spazio esprime la probabilità totale che la particella venga trovata in un punto qualsiasi, e quindi deve risultare uguale ad 1: perciò la ψ deve soddisfare la condizione

$$(110) \quad \iiint |\psi|^2 dx dy dz = 1.$$

La ψ essendo in generale complessa si può scrivere sotto la forma $\psi = \psi' + i\psi''$, dove ψ' e ψ'' sono due funzioni reali che soddisfano separatamente la (108') (poichè i coefficienti di questa sono reali). L'uso dell'immaginario serve solo a semplificare la scrittura, permettendo di raccogliere in una sola formula le due equazioni cui soddisfano ψ' e ψ'' . Si potrebbe domandarsi se non si possa costruire una meccanica ondulatoria con una sola ψ reale, ma è stato dimostrato ⁽²⁾ che ciò non è possibile, se si vuole che la P risulti, come deve essere, sempre ≥ 0 , che soddisfi la (110) e che si ricavi in termini finiti dalla ψ senza l'intervento di derivate di ψ rispetto a t .

Rimane da vedere come si colleghi questa impostazione probabilistica della meccanica alle ordinarie nozioni di cinematica e dinamica le quali, come è ben noto, si possono in molti casi applicare con successo (almeno entro i limiti degli errori di osservazione) agli elettroni, agli atomi, ecc. Questi

⁽¹⁾ Questa equazione vale, a rigore, solo per onde « monocromatiche », ossia di una sola frequenza. Il principio di sovrapposizione permette però, come si vedrà meglio al § 29, di passare agevolmente al caso più generale.

⁽²⁾ PAULI, *ZS. f. Phys.*, 80 (1933), p. 573 e segg.

casi sono quelli in cui la densità di probabilità P è diversa da zero soltanto in una regione così ristretta da poter essere considerata puntiforme (ossia, quando le onde di De Broglie costituiscono un piccolo « pacchetto », analogo al pacchetto di luce più volte considerato): allora si può praticamente ritenere localizzata la particella nel centro del pacchetto (o in un suo punto qualunque), e si identificherà il suo moto con quello di questo punto. Ogni volta dunque che si ragiona in termini della cinematica classica del moto di una particella, si deve intendere che si parla del moto di un pacchetto d'onde di De Broglie praticamente puntiforme.

§ 26. — MOTO DI UNA PARTICELLA E MOTO DI UN PACCHETTO DI ONDE.

Per dare forma definitiva all'equazione fondamentale della meccanica ondulatoria (108') resta da specificare il coefficiente N , che possiamo chiamare « indice di rifrazione » dello spazio per le onde di De Broglie e che sarà funzione, in generale, oltrechè di x, y, z , anche della frequenza ν di queste onde (come avviene in ottica nei mezzi dispersivi). Per far questo dobbiamo ancora tener conto di due elementi sperimentali e cioè:

I) il fatto che, nei casi in cui (come si è detto alla fine del § precedente) le onde di De Broglie costituiscono un pacchetto pressochè puntiforme, il moto di questo è regolato dalle ordinarie leggi della dinamica del punto;

II) la relazione $\lambda = \frac{h}{p}$ (data dalle esperienze di diffrazione) tra lunghezza d'onda di De Broglie ed impulso delle particelle (v. § 33, p. I).

Per tener conto della condizione I consideriamo una particella di massa m che si muova in un campo di potenziale ⁽¹⁾ $U(x, y, z)$, con un'energia totale E : nella meccanica ondulatoria questa particella (supposto che la sua posizione sia

⁽¹⁾ Chiamiamo « potenziale » l'energia potenziale della particella, mentre di solito in meccanica razionale si chiama « potenziale » questa energia cambiata di segno.

stata determinata con grande precisione) sarà rappresentata da un « pacchetto d'onde di De Broglie » abbastanza ristretto da poterlo considerare puntiforme (v. § precedente) ed il movimento di questo pacchetto, tra due punti qualunque, A e B , della sua traiettoria dovrà (prescindendo dalla diffrazione) identificarsi, sia per la traiettoria che per la velocità, col moto che la meccanica classica assegna ad un punto materiale di massa m ed energia E , nel campo di potenziale U , dal punto A al punto B . Per ottenere questa identificazione abbiamo a disposizione il coefficiente N della (108') ed inoltre la frequenza ν delle onde di De Broglie (o meglio la frequenza media dei treni d'onde che formano il pacchetto).

Procediamo anzitutto all'identificazione delle due traiettorie. Quella del pacchetto d'onde non è altro che un « raggio » ed è quindi determinata dalle leggi della rifrazione, che si possono compendiare in quel principio variazionale che, nel caso delle onde luminose, prende il nome di *principio di Fermat*: fra due punti qualunque A , B il raggio segue la linea per cui:

$$(111) \quad \delta \int_A^B N ds = 0 .$$

Analogamente, secondo la meccanica classica la traiettoria del punto può determinarsi mediante il *principio della minima azione*

$$(112) \quad \delta \int_{t_1}^{t_2} T dt = 0 ,$$

dove T è la forza viva e t_1 , t_2 sono gli istanti nei quali il punto passa per i due punti (fissi) A , B . In questo integrale si può introdurre s anzichè t come variabile d'integrazione, notando che

$$(113) \quad \frac{ds}{dt} = v = \sqrt{\frac{2T}{m}} ,$$

cosicchè si ha

$$\delta \int_{t_1}^{t_2} \sqrt{\frac{m}{2}} \sqrt{T} ds = 0 ,$$

e, tralasciando la costante e notando che

$$(114) \quad T = E - U ,$$

si ha la condizione seguente, puramente geometrica, per determinare la traiettoria:

$$(115) \quad \delta \int_A^B \sqrt{E - U} ds = 0 .$$

Perchè ora la traiettoria del pacchetto d'onde tra A e B coincida con quella che la meccanica classica assegna al punto, bisogna che la (111) e la (115) si equivalgano, per il che basta prendere

$$(116) \quad N = C \sqrt{E - U} ,$$

dove C è una qualsiasi costante (rispetto ad x, y, z). Abbiamo così trovato la distribuzione spaziale dell'indice di rifrazione (che esso fosse funzione di x, y, z soltanto attraverso U , era del resto prevedibile). Osserviamo che, poichè in generale l'indice di rifrazione delle onde di De Broglie dipende dalla loro frequenza, C ed E saranno da considerarsi funzioni di ν .

Identifichiamo ora anche la velocità, punto per punto, dei due movimenti. La velocità v del pacchetto d'onde non è la velocità di fase $1/N$, ma la velocità di gruppo (v. § 14), che è data dalla (74'), che possiamo scrivere (notando che $1/\lambda = N\nu$)

$$\frac{1}{v} = \frac{d(N\nu)}{d\nu} ,$$

ossia, introducendo la (116),

$$\frac{1}{v} = \frac{d}{d\nu} \left(C\nu \sqrt{E - U} \right) = \frac{d(C\nu)}{d\nu} \sqrt{E - U} + \frac{C\nu}{2\sqrt{E - U}} \frac{dE}{d\nu} .$$

D'altra parte, la velocità del punto data dalla meccanica classica è, per la (114),

$$(117) \quad v = \sqrt{\frac{2}{m}} \sqrt{E - U} .$$

Identificando le due espressioni di $1/v$ si ha, con ovvia riduzione,

$$(118) \quad \sqrt{2m} = 2 \frac{d(C\nu)}{d\nu} (E - U) + C\nu \frac{dE}{d\nu} .$$

Si osservi che questa deve essere una identità rispetto ad x, y, z , e che, d'altra parte, x, y, z vi figurano solo attraverso la U : dovrà dunque essere identicamente nullo il coefficiente di U , cioè dovrà aversi

$$\frac{d(Cv)}{dv} \neq 0,$$

ossia

$$(119) \quad Cv = K,$$

con K costante arbitraria. La (118) diviene allora

$$\sqrt{2m} = K \frac{dE}{dv},$$

ossia,

$$\frac{dE}{dv} = a,$$

dove si è posto

$$(120) \quad a = \frac{\sqrt{2m}}{K}.$$

Integrando, si ha dunque per $E(v)$ l'espressione lineare

$$(121) \quad E = av + b$$

dove a e b sono due costanti; l'espressione di $C(v)$ è poi data dalla (119), che si può scrivere, usando la (120),

$$(122) \quad C = \frac{\sqrt{2m}}{av}.$$

In definitiva l'espressione dell'indice di rifrazione che assicura l'identità completa dei due movimenti è

$$(123) \quad N = \frac{\sqrt{2m}}{av} \sqrt{E - U},$$

dove per E si può porre l'espressione (121).

La velocità di fase delle onde di De Broglie è dunque

$$(123') \quad V = \frac{1}{N} = \frac{av}{\sqrt{2m}} \frac{1}{\sqrt{E - U}},$$

e la loro lunghezza d'onda

$$\lambda = \frac{V}{v} = \frac{1}{Nv} = \frac{a}{\sqrt{2m(E - U)}};$$

ovvero, introducendo l'impulso $p = mv$ e badando alla (117)

$$(124) \quad \lambda = \frac{a}{mv} = \frac{a}{p}.$$

§ 27. - L'EQUAZIONE DI SCHRÖDINGER.

L'espressione di N trovata nel § precedente, sostituita nella (108'), fornisce una meccanica ondulatoria che soddisfa alla condizione di identificarsi con la meccanica ordinaria entro i limiti analoghi a quelli dell'ottica geometrica, e ciò qualunque siano le costanti a e b . Ma la condizione II del § precedente obbliga a fissare il valore della costante a : difatti la (124) ha bensì la forma della relazione trovata sperimentalmente tra λ e p , ma perchè si identifichi anche numericamente si deve prendere

$$(125) \quad a = h.$$

La costante b rimane invece indeterminata, poichè essa non figura nell'espressione di λ , ma solo in quella di v , e la frequenza delle onde di De Broglie non è accessibile all'esperienza: ciò rispecchia del resto la circostanza che l'energia E è definita a meno di una costante additiva arbitraria (1). Per semplicità si suole assumere $b = 0$, con che la (121) diviene

$$(121') \quad E = hv$$

e diventa così identica alla formula che vale per i fotoni.

Tenendo conto della (125), la relazione (123') tra la velocità di fase V delle onde di De Broglie di frequenza ν ed il potenziale U diviene

$$(126) \quad \frac{1}{N} = V = \frac{h\nu}{\sqrt{2m}} \frac{1}{\sqrt{E - U}},$$

(1) Ciò vale solo nella meccanica non relativistica: tenendo conto della relatività si ha invece modo di fissare E anche in valore assoluto, e, parallelamente, nella meccanica ondulatoria relativista resta determinata la b (v. cap. V, parte III).

od anche per la (121')

$$(126') \quad \frac{1}{N} = V = \frac{h\nu}{\sqrt{2m}} \frac{1}{\sqrt{h\nu - U}}.$$

Si può dunque dire che per le onde di De Broglie un campo di forza rappresenta quello che per la luce è un mezzo ad indice di rifrazione non uniforme, e che inoltre per esse il mezzo è sempre dispersivo, poichè l'indice di rifrazione dipende da ν .

Se ora introduciamo l'espressione (126) di N nell'equazione generale (108') cui soddisfa la ψ , troviamo l'equazione:

$$(127) \quad \Delta \psi = \frac{2m}{\hbar^2} (E - U) \frac{\partial^2 \psi}{\partial t^2},$$

che esprime la legge di propagazione della ψ ed è l'equazione fondamentale della meccanica ondulatoria: essa presuppone però che le onde siano (per usare il linguaggio ottico) *monocromatiche*, ossia che la ψ sia funzione sinusoidale del tempo di frequenza ν , ciò che significa fisicamente (come mostra la (121')), che l'energia della particella ha un valore E determinato con certezza (*stati semplici*, o *stazionari*, o *quantici*). Per esprimere analiticamente questo fatto, è conveniente (secondo l'uso già invalso in altri rami della fisica, p. es. in elettrotecnica) servirsi delle funzioni esponenziali ad esponente immaginario, e porre ⁽¹⁾

$$(128) \quad \psi = u(x, y, z) e^{-2\pi i \nu t},$$

dove u è una funzione (generalmente complessa) indipendente dal tempo, il cui modulo rappresenta l'ampiezza delle oscillazioni della ψ , e $\nu = E/\hbar$.

Di qui si ha

$$(129) \quad \frac{\partial \psi}{\partial t} = -2\pi i \nu \psi,$$

$$(130) \quad \frac{\partial^2 \psi}{\partial t^2} = -4\pi^2 \nu^2 \psi,$$

(1) Molti autori scrivono questa formula col segno + all'esponente: ciò non porta nessuna differenza sostanziale, salvo alcuni cambiamenti di segno nelle formule che ne derivano (p. es. nel secondo membro della (136)).

e tenendo conto di quest'ultima, la (127) si può scrivere nella forma seguente, che non contiene più derivate rispetto al tempo, e che è quella abitualmente adottata (*equazione di Schrödinger*):

$$(131) \quad \Delta \psi + \frac{8\pi^2 m}{\hbar^2} (E - U) \psi = 0;$$

o anche, ponendo per ψ la (128),

$$(131') \quad \Delta u + \frac{8\pi^2 m}{\hbar^2} (E - U) u = 0,$$

cioè: la parte spaziale, u , della funzione ψ soddisfa la stessa equazione della ψ . Poichè d'altra parte ciò che determina la distribuzione della probabilità è $|\psi|^2$, che è uguale a $|u|^2$, in molti casi è indifferente usare la funzione u invece della ψ (purchè però si tratti, come qui supponiamo, di onde di una sola frequenza).

La funzione coniugata di ψ , cioè

$$\psi^* = u^*(x, y, z) e^{+2\pi i \nu t}$$

soddisfa evidentemente la stessa equazione (131) della ψ , ed ha lo stesso modulo, cosicchè la sua considerazione non ci dà nulla di nuovo.

Resta ancora da osservare, riguardo all'equazione (127) o alla (131), che trattasi di equazioni omogenee, e quindi, trovata una soluzione, se ne ottengono infinite altre moltiplicando quella per una costante arbitraria. Tuttavia questa indeterminazione resta eliminata, se si tien conto del fatto che la ψ deve soddisfare, come si è già osservato al § 25, la condizione (110), che può anche scriversi (ponendo, come d'ordinario, $dS = dx dy dz$):

$$(132) \quad \int \psi \psi^* dS = 1.$$

Questa chiamasi *condizione di normalizzazione*, ed è già stata esaminata nel cap. I per le autofunzioni delle equazioni differenziali: come si è visto al § 4, essa può essere sempre soddisfatta (purchè l'integrale sia convergente, condizione su cui torneremo nel seguito) e determina (a meno di un fattore della forma $e^{i\theta}$) la detta costante moltiplicativa arbi-

traria. L'arbitrarietà che rimane nell'*argomento* della ψ a causa della costante arbitraria θ , non ha influenza nè sul modulo, nè sulla lunghezza d'onda della ψ , che soli hanno significato fisico, e si può quindi prescindere da essa.

§ 28. — I LIVELLI ENERGETICI COME AUTOVALORI DELL'EQUAZIONE DI SCHRÖDINGER.

Per rendere determinato il problema della meccanica ondulatoria di una particella si devono imporre alla ψ , oltre la condizione di soddisfare (nel caso degli stati stazionari) l'equazione (131) con un opportuno valore della costante E (che figura nell'equazione come un parametro a priori indeterminato), anche alcune condizioni di regolarità. E cioè, si richiederà anzitutto che la ψ e le sue derivate prime siano continue e ad un sol valore in tutto lo spazio: inoltre, perchè si possa applicare la condizione di normalizzazione (132), occorrerà che l'integrale di $|\psi|^2$, esteso a tutto lo spazio ⁽¹⁾, sia convergente (il che richiede che la ψ all'infinito sia infinitesima di ordine sufficientemente elevato). Se poi il potenziale $U(x, y, z)$ ha dei punti di singolarità, potranno in quei punti presentarsi delle singolarità anche per ψ (v. § 16): ma richiederemo in tal caso che queste singolarità siano al più poli di ordine inferiore al primo ⁽²⁾.

Con queste condizioni, il problema dell'integrazione dell'equazione di Schrödinger rientra nella categoria di quelli studiati nell'introduzione matematica: come si è visto, esso ammette soluzioni solo se il parametro E (che corrisponde al λ del cap. I) ha uno dei valori che abbiamo detto *autovalori* dell'equazione differenziale. Così le condizioni, assai naturali, di regolarità imposte alla ψ portano automaticamente

⁽¹⁾ In certi casi le condizioni del problema impongono alla particella di restare entro un certo spazio S : allora evidentemente si può integrare l'equazione solo entro questo spazio, con la condizione al contorno $\psi = 0$: v. p. es. § 38.

⁽²⁾ Per una giustificazione di queste condizioni, come pure per la enunciazione delle condizioni cui deve soddisfare la ψ in problemi più generali di questo, v. NEUMANN, *Götting. Nachr.*, (1927), p. 1; o anche PAULI, bibl. n. 14, p. 121.

all'esistenza dei *livelli energetici* (discreti, o talvolta continui) che è, come si è visto, un fatto sperimentale, e la determinazione di questi livelli si riduce al problema matematico della ricerca degli autovalori dell'equazione di Schrödinger. In ciò sta, come si è detto nella parte I, uno dei più brillanti risultati della meccanica ondulatoria.

A ciascun autovalore E_n (dove n può rappresentare in generale un complesso di più indici) corrisponderanno una o più autofunzioni normalizzate ψ_n che dipendono, naturalmente, dal tempo secondo la legge

$$(128') \quad \psi_n = u_n e^{-\frac{2\pi i}{h} E_n t}.$$

Se l'autovalore è semplice, l'autofunzione che gli corrisponde dà senz'altro, col quadrato del suo modulo, la distribuzione della probabilità P : se invece è multiplo, questa distribuzione non resta completamente determinata dal valore dell'energia: è questo il caso detto di *degenerazione*, di cui si vedranno nel seguito importanti esempi.

§ 29. — PRINCIPIO DI SOVRAPPOSIZIONE DELLA MECCANICA ONDULATORIA.

È naturale ammettere che anche per le onde di De Broglie valga un principio di sovrapposizione analogo a quello dell'ottica. Consideriamo perciò due diverse soluzioni normalizzate (indipendenti), ψ_1, ψ_2 , dell'equazione di Schrödinger, corrispondenti in generale ⁽¹⁾ a due diversi valori E_1, E_2 dell'energia ed a due impulsi differenti \mathbf{p}_1 e \mathbf{p}_2 (ci limitiamo, per semplicità di esposizione, al caso di due sole componenti, ma è immediata l'estensione ad un numero qualsiasi): formiamone una combinazione lineare (con due coefficienti costanti c_1, c_2)

$$\psi = c_1 \psi_1 + c_2 \psi_2 = c_1 u_1 e^{-\frac{2\pi i}{h} E_1 t} + c_2 u_2 e^{-\frac{2\pi i}{h} E_2 t},$$

⁽¹⁾ In casi di degenerazione potranno le due energie E_1, E_2 essere uguali: anche in tal caso però richiederemo che ψ_1 e ψ_2 siano ortogonali (v. § 6). Questo caso corrisponde a quello ottico della sovrapposizione di onde della stessa frequenza ma di direzione diversa.

alla quale pure imponremo la condizione di normalizzazione, che si traduce nella condizione $|c_1|^2 + |c_2|^2 = 1$, come si verifica subito ricordando che ψ_1, ψ_2 sono ortogonali e normalizzate. Quale sarà il significato da attribuire a questa ψ ? È evidente che questo caso è analogo a quello ottico (v. § 20) di un fotone di energia ed impulso indeterminati, e che quindi $|c_1 \psi_1|^2 dS$ rappresenta la probabilità che la particella venga trovata nell'elemento di volume dS con l'energia E_1 e l'impulso \mathbf{p}_1 , cioè con le caratteristiche corrispondenti alle onde ψ_1 , mentre $|c_2 \psi_2|^2 dS$ rappresenta la probabilità di trovarla in dS con le caratteristiche delle onde ψ_2 . Si noti che queste due espressioni rappresentano proprio le probabilità e non quantità ad esse proporzionali, come si verifica subito osservando che la probabilità totale di trovare la particella dovunque e con qualunque energia risulta (come deve essere) uguale a 1: infatti

$$\int |c_1 \psi_1|^2 dS + \int |c_2 \psi_2|^2 dS = |c_1|^2 + |c_2|^2 = 1.$$

È evidente poi che $|c_1|^2$ rappresenta la probabilità che la particella abbia l'energia E_1 e l'impulso \mathbf{p}_1 , e $|c_2|^2$ la probabilità dell'energia E_2 e dell'impulso \mathbf{p}_2 .

Consideriamo ora la $\psi(x, y, z, t)$ più generale possibile. Fissato un valore di t , p. es. $t = 0$, la ψ potrà essere sviluppata in serie mediante le autofunzioni dell'equazione di Schrödinger (131') (che formano come si sa, un sistema ortogonale completo): sarà cioè

$$\psi(x, y, z, 0) = \sum_n c_n u_n(x, y, z),$$

con

$$c_n = \int \psi(x, y, z, 0) u_n^* dS.$$

Ora, ciascuna delle u_n rappresenta la distribuzione, per $t = 0$, della ψ in uno « stato semplice »: tale ψ si evolve poi col tempo secondo la legge (128'). Queste varie componenti, come si è detto, si sovrappongono sempre senza alterarsi reciprocamente, e quindi la ψ al tempo generico t sarà data da

$$(133) \quad \psi(x, y, z, t) = \sum_n c_n u_n(x, y, z) e^{-\frac{2\pi i}{h} E_n t}.$$

Questa ψ rappresenta uno *stato* ⁽¹⁾ della particella in cui l'energia e l'impulso non sono determinati: la probabilità che l'energia abbia il valore E_n e l'impulso sia quello corrispondente alle onde ψ_n è, per quanto si è detto sopra, $|c_n|^2$.

La (133) è stata scritta per il caso di autovalori discreti. Se invece gli autovalori della (131') costituiscono uno spettro continuo da $E = E_a$ ad $E = E_b$, la serie sarà sostituita da un integrale, cioè

$$(133') \quad \psi = \int_{E_b}^{E_a} c(E) \psi_E dE,$$

dove

$$\psi_E = u_E(x, y, z) e^{-\frac{2\pi i}{h} E t}$$

e la u_E soddisfa l'equazione di Schrödinger

$$\Delta u_E + \frac{8\pi^2 m}{h^2} (E - U) u_E = 0.$$

In questo caso, $|c(E)|^2$ rappresenta la « densità di probabilità » nello spettro continuo dell'energia, vale a dire, $|c(E)|^2 dE$ è la probabilità che l'energia sia compresa tra E ed $E + dE$.

Il caso più generale è quello in cui vi sono autovalori discreti e autovalori continui, nel qual caso la ψ sarà la somma di una serie e di un integrale. Si può tuttavia considerare la (133') come la forma più generale della ψ , se si intende l'integrale definito al modo di STIELTJES (v. nota al § 10).

§ 30. — EQUAZIONE TEMPORALE DI SCHRÖDINGER.

Va tenuto presente che la ψ nella forma generale (133) o (133') (cioè non « monocromatica ») non soddisfa all'equazione di Schrödinger (131), perchè ciascuno dei suoi termini soddisfa

(1) La forma della funzione ψ che si addice a ciascun caso (ossia lo « stato » del sistema) dipende, come si vedrà meglio nel cap. II della parte III, dalle condizioni iniziali, e in particolare dalle osservazioni a cui è stato inizialmente sottoposto il sistema.

bensi ad un'equazione della forma (131), ma con un diverso valore di E . Si può trovare però facilmente un'equazione indipendente da E cui soddisfano tutte le componenti monocromatiche e perciò anche la ψ che risulta dalla loro sovrapposizione.

Ricordiamo che la componente ψ_n è della forma

$$\psi_n = u_n(x, y, z) e^{-\frac{2\pi i}{h} E_n t}$$

e perciò

$$(134) \quad \frac{\partial \psi_n}{\partial t} = -\frac{2\pi i}{h} E_n \psi_n$$

Essa poi soddisfa l'equazione di Schrödinger

$$(135) \quad \Delta \psi_n + \frac{8\pi^2 m}{h^2} (E_n - U) \psi_n = 0$$

Eliminando E_n tra la (134) e la (135) si ottiene l'equazione

$$\Delta \psi_n - \frac{8\pi^2 m}{h^2} U \psi_n = -\frac{4\pi i m}{h} \frac{\partial \psi_n}{\partial t},$$

in cui i coefficienti sono indipendenti dall'indice n : perciò questa equazione è soddisfatta da tutte le componenti ψ_n , e dunque anche da qualsiasi loro combinazione lineare.

Potremo dire dunque che l'ampiezza di probabilità ψ , anche nel caso in cui non sia determinata l'energia della particella, e quindi le onde non siano « monocromatiche », soddisfa sempre l'equazione

$$(136) \quad \Delta \psi - \frac{8\pi^2 m}{h^2} U \psi = -\frac{4\pi i m}{h} \frac{\partial \psi}{\partial t},$$

che chiameremo *equazione temporale di Schrödinger*.

Si osservi che, essendovi nella (136) un coefficiente immaginario, la coniugata ψ^* della ψ non soddisfa la stessa equazione, ma la seguente

$$(136') \quad \Delta \psi^* - \frac{8\pi^2 m}{h^2} U \psi^* = +\frac{4\pi i m}{h} \frac{\partial \psi^*}{\partial t}.$$

§ 31. - DENSITÀ DI CORRENTE.

Analogamente all'intensità di illuminazione, definita statisticamente al § 19, conviene definire la densità (probabilistica) del flusso di particelle. Mostreremo cioè che esiste un

vettore \mathbf{i} , tale che $\mathbf{i}_n d\sigma dt$ (dove $d\sigma$ è un elemento di superficie, e \mathbf{n} è la sua normale, con un verso positivo determinato) rappresenta la probabilità, algebricamente considerata ⁽¹⁾, che la particella attraversi nel tempo dt l'elemento $d\sigma$.

Supponiamo perciò (come nella nota al § 25) che vi sia non uno ma un gran numero N di sistemi nelle stesse condizioni (e non agenti tra loro): la densità media delle particelle sarà $|\psi|^2$. Perciò in un qualsiasi spazio chiuso S ve ne saranno in media

$$N \int_S \psi \psi^* dS.$$

L'aumento di questo numero per unità di tempo è

$$N \frac{\partial}{\partial t} \int_S \psi \psi^* dS = N \int_S \left(\psi \frac{\partial \psi^*}{\partial t} + \psi^* \frac{\partial \psi}{\partial t} \right) dS,$$

ovvero, ricavando le derivate dalla (136) e dalla (136')

$$N \frac{h}{4\pi i m} \int_S (\psi \Delta \psi^* - \psi^* \Delta \psi) dS.$$

D'altra parte, questa quantità deve essere uguale al numero medio delle particelle che nell'unità di tempo entrano nel volume S attraverso la superficie σ : perciò dovrà aversi

$$N \frac{h}{4\pi i m} \int_S (\psi \Delta \psi^* - \psi^* \Delta \psi) dS = N \int_\sigma \mathbf{i}_n d\sigma$$

(prendendo la normale diretta verso l'interno). Trasformando l'integrale di volume in uno di superficie si ha

$$N \frac{h}{4\pi i m} \int_\sigma \left(\psi^* \frac{\partial \psi}{\partial n} - \psi \frac{\partial \psi^*}{\partial n} \right) d\sigma = N \int_\sigma \mathbf{i}_n d\sigma.$$

Affinchè questa valga per qualunque σ , basta prendere

$$(137) \quad \mathbf{i} = \frac{h}{4\pi i m} (\psi^* \text{grad } \psi - \psi \text{grad } \psi^*).$$

⁽¹⁾ Cioè la differenza tra la probabilità dei passaggi nel verso positivo e quella dei passaggi nel verso negativo.

È questa l'espressione cercata per la *densità di flusso* (probabilistica).

Se si tratta di particelle con una carica elettrica e , il vettore $\mathbf{j} = e\mathbf{i}$ rappresenta ovviamente il valor medio della densità di corrente elettrica.

Nei primi lavori sulla meccanica ondulatoria la $\psi\psi^*$ non veniva interpretata come una densità di probabilità, o una densità media, ma come una densità vera, cosicchè l'elettrone veniva concepito come una distribuzione continua di elettricità, di densità elettrica $\rho = e\psi\psi^*$. Conseguentemente, il vettore $\mathbf{j} = e\mathbf{i}$ definito dalla (137) veniva interpretato come densità vera (non media) della corrente elettrica, ed in base a questi dati si calcolavano, con le ordinarie leggi elettro-magnetiche, il campo elettromagnetico prodotto e la radiazione emessa. Questa ipotesi, sebbene portasse in qualche caso a risultati giusti, era in evidente contrasto con tutti i fenomeni nei quali l'elettrone presenta natura corpuscolare, e fu perciò abbandonata assai presto. Fu il Born che primo propose l'interpretazione probabilistica, che in seguito si inquadra perfettamente nello schema del principio di indeterminazione, come è stato mostrato nei §§ precedenti.

§ 32. - CAMPO ELETTROMAGNETICO E RADIAZIONE.

Quando la particella di cui si parla è carica di elettricità (p. es., se è un elettrone) si presenta il problema di determinare i suoi effetti elettrici e magnetici su altre particelle, e, in particolare, la sua radiazione: a questa ultima questione risponde in modo preciso una teoria svolta da DIRAC, per la quale rimandiamo ad altro volume di questa opera ⁽¹⁾.

Ci limiteremo qui a rilevare che si può avere una valutazione del valore *medio* del campo elettrico, o magnetico (v. nota a pag. 155), applicando le ordinarie leggi dell'elettromagnetismo e prendendo come densità elettrica e come densità di corrente i valori *medi*

$$(138) \quad \rho = e \psi \psi^*$$

$$(139) \quad \mathbf{j} = e\mathbf{i} = \frac{he}{4\pi im} (\psi^* \text{grad } \psi - \psi \text{grad } \psi^*).$$

⁽¹⁾ V. anche la monografia di E. Fermi « Quantum theory of radiation » in *Rev. of Mod. Phys.*, 4 (1932) p. 87-132, oppure bibl. n. 15.

Si osservi che se l'elettrone si trova in uno stato di quelli che al § 29 abbiamo chiamato « semplici », cioè se la sua energia ha un valore ben determinato e quindi la ψ ha la forma (128), le espressioni precedenti divengono

$$(140) \quad \rho = e u u^*,$$

$$(141) \quad \mathbf{j} = \frac{he}{4\pi im} (u^* \text{grad } u - u \text{grad } u^*),$$

ed in esse, come si vede, non figura più il tempo: in tal caso dunque la distribuzione media delle cariche e delle correnti è *stazionaria*. È per questo che gli stati semplici si chiamano anche « *stati stazionari* ». Naturalmente, in tal caso anche i valori medi del campo elettrico e magnetico sono indipendenti dal tempo. Si noti poi che se, in particolare, la u è reale risulta $\mathbf{j} = 0$, cioè la distribuzione media delle cariche è non solo stazionaria ma *statica*. Una tale soluzione esiste sempre, essendo reali i coefficienti della (131'): anzi, nei casi in cui non vi è degenerazione, essa è l'unica (a meno di un fattore inessenziale $e^{i\theta}$ con θ costante).

Quanto alla radiazione emessa, ci limiteremo ad enunciare il risultato essenziale della teoria di Dirac. Se il sistema al tempo 0 si trova in uno stato stazionario di energia E_n (che non sia quello di energia più bassa) osservandolo al tempo t c'è una certa probabilità, crescente con t , di trovarlo in uno stato di energia $E_m < E_n$, essendo stata emessa la differenza di energia sotto forma di radiazione di frequenza

$$(142) \quad \nu_{nm} = \frac{E_n - E_m}{h}.$$

(Questa formula, che nella teoria di Bohr costituiva un postulato a sè, viene invece dedotta, nella teoria di Dirac, dai principi generali della meccanica quantistica). E tale probabilità, in ogni intervallo infinitesimo dt , si accresce di $P_{nm} dt$, dove

$$(143) \quad P_{nm} = \frac{32\pi^4}{3hc^3} \nu_{nm}^3 (X_{nm}^2 + Y_{nm}^2 + Z_{nm}^2).$$

In questa formula, X_{nm} , Y_{nm} , Z_{nm} rappresentano tre integrali in cui entrano le autofunzioni dei due stati stazionari, iniziale e finale, e precisamente

$$(144) \quad X_{nm} = e \int_S X u_n^* u_m dS = X_{mn}^*$$

(e analoghe espressioni per gli altri due). Se si ha inizialmente un gran numero di atomi nello stato n -esimo, l'intensità della radiazione di frequenza ν_{nm} emessa in ogni istante sarà naturalmente proporzionale al loro numero e a P_{nm} : quindi essa è determinata essenzialmente dalle tre quantità X_{nm} , Y_{nm} , Z_{nm} . Se, in particolare, queste risultano tutte e tre nulle, il passaggio dallo stato n -esimo all' m -esimo ha probabilità nulla, cioè è « proibito », e la corrispondente riga spettrale viene a mancare. Si giustificano così le cosiddette « regole di selezione » note da tempo agli spettroscopisti e che nella teoria di Bohr e Sommerfeld venivano giustificate mediante il *principio di corrispondenza* (v. § 64).

Inoltre, si trova che ognuna delle tre quantità X_{nm} , Y_{nm} , Z_{nm} determina una componente dell'ampiezza del campo elettrico nella luce emessa, cosicchè da esse si può anche ricavare lo stato di polarizzazione di questa: dette quantità corrispondono a quello che nella teoria classica sono le componenti dell'ampiezza del momento elettrico del sistema emittente⁽¹⁾. Se p. es. $X_{nm} = Y_{nm} = 0$, $Z_{nm} \neq 0$, la radiazione emessa corrisponde a quella di un oscillatore lineare parallelo all'asse z , e quindi il suo piano di vibrazione (piano del raggio e del vettore elettrico) è parallelo all'asse z .

Considerazioni analoghe si possono fare per l'assorbimento. (v. anche § 43 p. III).

⁽¹⁾ Si vedrà al § 33 p. III che X_{nm} , Y_{nm} , Z_{nm} sono gli elementi delle *matrici* che, nella meccanica quantistica, rappresentano le componenti del momento elettrico.

CAPITOLO III.

PROBLEMI UNIDIMENSIONALI

§ 33. — CARATTERISTICHE DEI PROBLEMI UNIDIMENSIONALI.

Per cominciare a studiare l'equazione di Schrödinger dai casi più semplici, tratteremo in questo capitolo alcuni problemi « unidimensionali », con che intendiamo che nella U , nella ψ ed in tutte le altre quantità che eventualmente interverranno, figura (oltre t) una sola delle coordinate spaziali, p. es. la x .

La caratteristica fisica di questi problemi è che la particella ha eguale probabilità di essere trovata in tutti i punti di ogni piano normale all'asse x , ossia che le sue coordinate y e z sono completamente indeterminate, potendo assumere, con eguale probabilità, qualunque valore: solo della coordinata x si ricerca la distribuzione probabilistica. Si può anche dire che si studia il movimento non della particella P , ma solo della sua proiezione P' sull'asse x : perciò questi problemi corrispondono a quelli della meccanica classica in cui si studia il moto di un punto su una retta: useremo anzi lo stesso linguaggio e parleremo del moto della particella sull'asse x (intendendo riferirci alla sua proiezione P'). Osserveremo incidentalmente che nelle condizioni che ora consideriamo ci si trova necessariamente ogni qualvolta si supponga imposta alla particella la condizione di muoversi *parallelamente* all'asse x (non *sull'asse* x): difatti, per il principio di indeterminazione, questa condizione ($p_x = p_y = 0$) implica la completa indeterminazione delle coordinate y e z .

Ciò premesso, nel nostro caso la (131') diviene l'equazione a derivate ordinarie

$$(145) \quad \frac{d^2u}{dx^2} + \frac{8\pi^2m}{h^2} [E - U(x)]u = 0,$$

che chiameremo *equazione unidimensionale di Schrödinger* (per gli stati stazionari).

Mostreremo ora l'applicazione di questa equazione a qualche problema particolare, con lo scopo soprattutto di illustrare meglio, su esempi analiticamente semplici, lo spirito della meccanica ondulatoria: alcuni poi di questi problemi sono anche suscettibili di dirette applicazioni fisiche.

Sarà utile tener presente, come guida intuitiva alla risoluzione di questi problemi, l'analogia formale di essi con i problemi ottici in cui la luce si propaghi per onde piane lungo l'asse x , e l'indice di rifrazione (indipendente da y e da z) vari con la x proporzionalmente a $\sqrt{E - U}$. In questi problemi unidimensionali poi si può anche invocare l'analogia con le oscillazioni trasversali di una corda, facendo corrispondere u allo spostamento della corda stessa e la densità di essa a $E - U$.

§ 34. - DISCUSSIONE QUALITATIVA DEI PROBLEMI UNIDIMENSIONALI.

La (145) è un'equazione del tipo

$$(146) \quad \frac{d^2u}{dx^2} + f(x)u = 0,$$

con $f(x) = \frac{8\pi^2m}{h^2}(E - U)$. Se si rappresenta con una curva (fig. 24) l'andamento del potenziale U in funzione di x e poi si traccia la retta orizzontale di ordinata E , le intersezioni danno i valori x_1, x_2, \dots di x per cui $f(x) = 0$: tali punti dividono l'asse x in regioni, in alcune delle quali $f > 0$, in altre $f < 0$. Nella meccanica classica, le prime regioni (cioè quelle nelle quali la differenza $E - U$, che deve rappresentare l'energia cinetica, è positiva) sono le sole in cui può muoversi la particella, essendo assegnato il valore E della sua energia totale: le altre regioni invece (in cui $E - U < 0$) sono inaccessibili ad una particella di energia E . Nella meccanica ondulatoria, come vedremo, questo non vale, però anche qui la distinzione delle due regioni ha un significato importante, perchè in esse l'andamento della $u(x)$ ha caratteristiche notevolmente diverse.

Si osservi infatti anzitutto che nelle regioni dove $f > 0$ la u e la $\frac{d^2u}{dx^2}$ hanno segno opposto, e quindi la curva rappresentante la u è in tali regioni sempre *concava* verso l'asse x : viceversa, essa è *convessa* verso l'asse x nelle regioni in cui $f < 0$. È intuitivo allora che nelle prime regioni la curva può attraversare più volte l'asse x con andamento oscillatorio, mentre nelle seconde la curva, se anche attraversa una volta l'asse x , dovendo poi volgergli sempre la convessità non può tornare ad attraversarlo di nuovo (nella stessa regione) e quindi una di tali regioni può contenere al più un solo nodo, e la curva non ha in essa mai carattere oscillatorio. Ciò è reso più chiaro se si pensa al caso particolare di $f = \text{cost.}$: se tale costante è > 0 la curva è una sinusoidale (carattere oscillatorio, concavità verso l'asse x); se < 0 la curva è di tipo esponenziale (carattere non oscillatorio, convessità verso l'asse x).

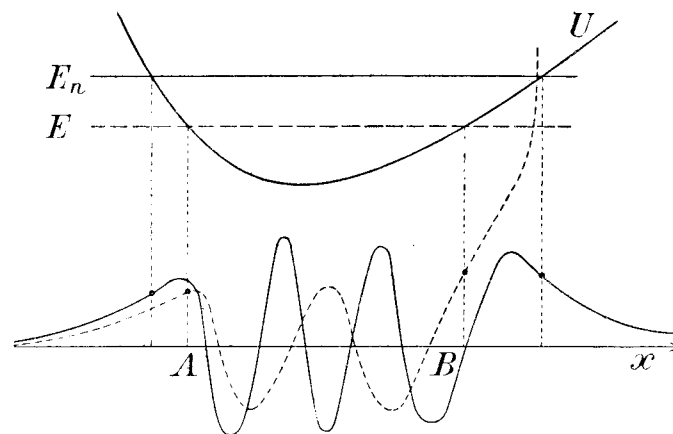


Fig. 24.

Tale considerazione torna assai utile per discutere qualitativamente l'andamento delle autofunzioni corrispondenti ad una assegnata curva del potenziale e ad un dato livello energetico. P. es., consideriamo un potenziale del tipo della fig. 24, cioè dotato di un solo minimo e tendente monotonamente all'infinito sia a destra che a sinistra di questo, con

legge qualunque (in questo caso rientra p. es. l'oscillatore armonico, v. § 39) e, per un valore qualunque di E , consideriamo una curva u (tratteggiata) rappresentante una soluzione che verso sinistra tende asintoticamente all'asse x : essa si scosterà dall'asse x , volgendogli la convessità, fino al punto A , poi inizierà una serie di oscillazioni entro il tratto AB , per riprendere, a destra di B , l'andamento non oscillatorio che la porta in generale a scostarsi indefinitamente dall'asse x . Solo se E ha uno dei valori E_n (autovalori), avviene che la curva torni ad adagiarsi sull'asse x anche verso destra, ed essa rappresenta allora un'autofunzione u_n (curva piena).

Il tratto, entro il quale la curva ha andamento oscillatorio, è evidentemente quello entro cui oscillerebbe la particella, secondo la meccanica classica.

§ 35. — PARTICELLA DI ENERGIA DETERMINATA E NON SOGGETTA A FORZE.

Applichiamo l'equazione unidimensionale di Schrödinger al caso di una particella non soggetta a forze, e libera di muoversi da $x = -\infty$ a $x = +\infty$.

Potremo porre nella (145) $U = 0$, e allora, ponendo per brevità

$$(147) \quad k = \sqrt{\frac{2m}{\hbar^2} E},$$

l'equazione si scrive

$$(148) \quad \frac{d^2 u}{dx^2} + 4\pi^2 k^2 u = 0,$$

con k costante: essa è l'equazione studiata nel § 8 ed ha per integrale generale

$$(149) \quad u = ae^{2\pi i k x} + be^{-2\pi i k x},$$

da cui:

$$(150) \quad \psi = (ae^{2\pi i k x} + be^{-2\pi i k x})e^{-\frac{2\pi i}{\hbar} Et}.$$

Si riconosce subito che la E non può essere negativa (e quindi che k deve essere reale) perchè altrimenti la u diventerebbe

infinita per $x \rightarrow +\infty$ o per $x \rightarrow -\infty$, il che si deve escludere (v. § 28). Del resto, come si è visto, k^2 e quindi E è completamente arbitraria tra 0 e $+\infty$ (spettro continuo di autovalori), ossia la particella può avere qualunque energia (positiva). Supponiamo fissata la E una volta per tutte: ad essa corrispondono due soluzioni linearmente indipendenti rappresentate dai due termini della (149) (siamo dunque in uno dei casi di degenerazione studiati al § 6).

Consideriamo dapprima soltanto il primo termine, cioè poniamo $b = 0$, e prendiamo

$$(151) \quad \psi = ae^{2\pi i (kx - vt)}$$

Questa formula rappresenta (v. § 12) un treno di onde piane *progressive* di lunghezza d'onda

$$(152) \quad \lambda = \frac{1}{k}$$

e di velocità (di fase)

$$(153) \quad V = \frac{v}{k} = \frac{E}{\hbar k}.$$

Si osservi che, secondo la meccanica classica, all'energia E corrispondono una velocità ed un impulso della particella, dati rispettivamente da

$$v = \pm \sqrt{\frac{2E}{m}}, \quad p = \pm \sqrt{2mE} = \pm \hbar k,$$

Per una particella in moto progressivo si ha dunque

$$p = \hbar k$$

ed allora la (152) si identifica con la formula già nota $\lambda = \hbar/p$, mentre la (151) assume la forma seguente:

$$(151') \quad \psi = ae^{\frac{2\pi i}{\hbar} \left(px - \frac{p^2}{2m} t \right)}.$$

La costante a si determina coi criteri di normalizzazione per lo spettro continuo spiegati al § 10. Si osservi che la densità di probabilità della posizione della particella, data da $|\psi|^2$, risulta indipendente da x : ciò è in relazione col fatto

che, essendosi determinato con precisione assoluta l'impulso, resta completamente indeterminata la posizione della particella, conformemente al principio di indeterminazione.

Considerazioni analoghe possono farsi sul secondo termine della (150), che può mettersi ancora nella forma (151'), ma ponendo $p = -hk$: esso rappresenta dunque onde regressive, e corrisponde ad una particella di impulso diretto nel verso negativo. Come si vede, la degenerazione del problema in meccanica ondulatoria ha la sua ragione nel fatto che nella meccanica ordinaria ad un valore di E corrispondono due valori di p , uno positivo ed uno negativo.

Considerando ora la soluzione generale (150), e tenendo presente il principio di sovrapposizione, potremo dire che essa rappresenta il caso in cui è determinata l'energia E della particella ma non il verso del suo impulso, cosicchè vi è una certa probabilità, proporzionale a $|a|^2$, di trovarla in moto progressivo, ed una probabilità, proporzionale a $|b|^2$, di trovarla in moto regressivo.

§ 36. - PARTICELLA DI ENERGIA INDETERMINATA (GRUPPO D'ONDE).

Il caso più interessante nella pratica è quello, più generale del precedente, in cui sono assegnate la densità di probabilità $P_0(x)$ dell'ascissa della particella al tempo 0, e la densità di probabilità $Q_0(p)$ dell'impulso iniziale: queste funzioni saranno rappresentate da due curve la cui forma dipenderà dal tipo di esperienza che è servita a definire lo stato iniziale del sistema.

La ψ sarà in generale, come si è detto al § 29, della forma (133') (somma o integrale di infinite soluzioni semplici corrispondenti ai diversi valori di E): siccome però a ciascun valore di E corrispondono, come si è detto, due soluzioni (distinte per il segno di p), conviene caratterizzare le singole componenti semplici piuttosto dal valore di p che da quello di E , ossia scrivere

$$(154) \quad \psi = a \int_{-\infty}^{\infty} \varphi_0(p) e^{\frac{2\pi i}{h} \left(px - \frac{p^2}{2m} t \right)} dp,$$

dove a è una costante di normalizzazione che verrà fissata

in seguito, e $\varphi_0(p)$ ha un significato analogo al $c(E)$ della (133'), cioè:

$$(155) \quad |\varphi_0(p)|^2 = Q_0(p).$$

Si noti che questa relazione determina solo il modulo di $\varphi_0(p)$, lasciandone arbitrario l'argomento θ : scriveremo dunque:

$$(155') \quad \varphi_0(p) = \sqrt{Q_0(p)} e^{i\theta(p)}$$

e la funzione $\theta(p)$ dovrà essere determinata in modo da soddisfare l'altra condizione iniziale, e cioè che sia

$$(156) \quad |\psi_0|^2 = P_0(x),$$

dove ψ_0 è la ψ per $t = 0$, cioè:

$$(157) \quad \psi_0(x) = a \int_{-\infty}^{\infty} \varphi_0(p) e^{\frac{2\pi i}{h} px} dp,$$

Vedremo tra breve sotto quale condizione questo è possibile.

Si osservi che la (154) si identifica con la (58) del § 12, identificando ψ_0 con la f e ponendo $p = hk$ e

$$(158) \quad A(k) = a\varphi_0(hk) h;$$

la (59), che ne è conseguenza, dà allora

$$(157') \quad \varphi_0(p) = \frac{1}{ah} \int_{-\infty}^{\infty} \psi_0(x) e^{-\frac{2\pi i}{h} px} dx.$$

Osservando poi che, per la (51') del § 12, è

$$\int_{-\infty}^{\infty} \psi_0 \psi_0^* dx = haa^* \int_{-\infty}^{\infty} \varphi_0 \varphi_0^* dp$$

e che, per la condizione di normalizzazione, entrambi gli integrali di questa formula devono valere 1, si trova che si può porre ⁽¹⁾ nelle formule precedenti

$$(159) \quad a = \frac{1}{\sqrt{h}}.$$

⁽¹⁾ La costante arbitraria di modulo 1, per cui potrebbe essere moltiplicata a , e quindi ψ , non influisce sulle probabilità e quindi non ha importanza.

Nei casi più interessanti, la posizione iniziale della particella è definita con una certa approssimazione, cosicchè la curva della $P_0(x)$ presenta l'aspetto « a campana » che ha, p. es., la curva degli errori di Gauss: allora si può dire che la (154) rappresenta un « gruppo d'onde », il quale inizialmente ha il « profilo » (v. § 14) definito dalla curva $P_0(x)$, e poi si sposta (deformandosi) con una velocità di gruppo data dalla (74), che, nel caso attuale, dà immediatamente $v = p/m$: la velocità del gruppo si identifica dunque con la velocità della particella nella meccanica classica, come era da attendersi, in virtù del modo stesso col quale è stata costruita l'equazione di Schrödinger.

L'indeterminazione Δx nell'ascissa della particella in un dato istante è definita dalla formula (analoga alla (65))

$$(160) \quad (\Delta x)^2 = \int (x - \bar{x})^2 \psi \psi^* dx$$

e similmente l'indeterminazione Δp nell'impulso da

$$(161) \quad (\Delta p)^2 = \int (p - \bar{p})^2 \varphi_0 \varphi_0^* dp,$$

formula che, badando alla relazione tra p e k e alla (158), si identifica con la (63). Da ciò si vede che le due indeterminazioni Δx e Δp sono soggette alla limitazione imposta dalla (66), che si scrive ora

$$(162) \quad \Delta p \Delta x \geq \frac{h}{4\pi} :$$

quindi, le due curve $P_0(x)$ ed $Q_0(p)$ non possono essere assegnate ad arbitrio: se esse si assegnano in modo che inizialmente non sia soddisfatta questa condizione, non esiste una $\theta(p)$ che posta nella (155') soddisfi la (156), e non è quindi possibile costruire il gruppo d'onde avente i requisiti imposti. Resta verificato così che lo schema della meccanica ondulatoria rispecchia il principio di indeterminazione per le particelle materiali (come era da attendersi essendo il suo formalismo analogo a quello dell'ottica ondulatoria).

Ha particolare interesse il caso in cui le curve $P_0(x)$ e $Q_0(p)$ sono tali che al tempo 0 sia

$$(163) \quad \Delta x \Delta p = \frac{h}{4\pi}$$

(cioè, che il prodotto delle indeterminazioni sia il minimo possibile). Come si è visto al § 13, ciò richiede che le funzioni $\psi_0(x)$ e $\varphi_0(p)$ siano della forma (70) e (72), che con le notazioni attuali, e ponendo per brevità

$$(164) \quad \alpha = \frac{1}{\sqrt{2}(\Delta x)_0}, \quad \beta = \frac{1}{\sqrt{2}(\Delta p)_0},$$

cioè indicando con α e β le « precisioni » ⁽¹⁾ delle determinazioni iniziali di x e p , divengono

$$(165) \quad \psi_0(x) = C e^{-\frac{\alpha^2}{2}(x - \bar{x}_0)^2 + \frac{2\pi i}{h} \bar{p}_0 x},$$

$$(166) \quad \varphi_0(p) = D e^{-\frac{\beta^2}{2}(p - \bar{p}_0)^2 - \frac{2\pi i}{h} \bar{x}_0 p}.$$

Le curve di probabilità iniziali sono dunque del tipo gaussiano e cioè:

$$(167) \quad P_0(x) = CC^* e^{-\alpha^2(x - \bar{x}_0)^2},$$

$$(168) \quad Q_0(p) = DD^* e^{-\beta^2(p - \bar{p}_0)^2}.$$

La (163) equivale alla seguente relazione tra α e β

$$(163') \quad \alpha\beta = \frac{2\pi}{h}.$$

I moduli delle costanti C e D si determinano con le condizioni di normalizzazione, che danno:

$$(169) \quad CC^* = \frac{\alpha}{\sqrt{\pi}},$$

$$(170) \quad DD^* = \frac{\beta}{\sqrt{\pi}}.$$

⁽¹⁾ Si noti che qui il linguaggio della teoria degli errori viene applicato ad un tipo di indeterminazione di origine del tutto diversa da quella degli errori di osservazione.

Calcoliamo ora la curva di probabilità $P(x)$ della posizione al tempo t .

Dobbiamo anzitutto calcolare la ψ al tempo t , mediante la (154), che, introducendovi l'espressione (166) e ponendo per brevità

$$x - \bar{x}_0 = \xi, \quad p - \bar{p}_0 = \eta,$$

diviene

$$\psi = \frac{D}{\sqrt{h}} e^{\frac{2\pi i}{h} \left[\bar{p}_0 \xi - \frac{\bar{p}_0^2}{2m} t \right]} \int_{-\infty}^{\infty} e^{-\left(\frac{\beta^2}{2} + \frac{\pi i t}{hm} \right) \eta^2 + \frac{2\pi i}{h} \eta \left(\xi - \frac{\bar{p}_0}{m} t \right)} d\eta.$$

L'integrale si riduce facilmente a integrali definiti noti, e risulta uguale a

$$(171) \quad \frac{\sqrt{\pi}}{\sqrt{\frac{\beta^2}{2} + \frac{\pi i t}{hm}}} e^{-\frac{\pi^2}{h^2} \frac{\left(\xi - \frac{\bar{p}_0}{m} t \right)^2}{\frac{\beta^2}{2} + \frac{\pi i t}{hm}}}.$$

Calcolata così la ψ , la $\psi\psi^*$ ci dà la $P(x)$ al tempo t : per scrivere in forma semplice il quadrato del modulo dell'espressione (171), conviene introdurre la notazione

$$(172) \quad \alpha_t = \frac{\pi\beta}{h} \frac{1}{\sqrt{\frac{\beta^4}{4} + \left(\frac{\pi t}{hm} \right)^2}},$$

e così si ottiene

$$P(x) = DD^* \frac{\alpha_t}{\beta} e^{-\alpha_t^2 \left(\xi - \frac{\bar{p}_0}{m} t \right)^2},$$

ovvero, osservando che, per la (169) e la (170),

$$\frac{DD^*}{\beta} = \frac{CC^*}{\alpha}$$

ed introducendo di nuovo la x :

$$(173) \quad P(x) = CC^* \frac{\alpha_t}{\alpha} e^{-\alpha_t^2 \left(x - \bar{x}_0 - \frac{\bar{p}_0}{m} t \right)^2}.$$

Questa formula, confrontata con la (167) mostra che la curva di probabilità al tempo t è ancora una curva gaussiana, ma ha il massimo, invece che in $x = \bar{x}_0$, in

$$x = \bar{x}_0 + \frac{\bar{p}_0}{m} t$$

(cioè il massimo si sposta con la velocità \bar{p}_0/m , come già si sapeva): inoltre, la precisione non è più α ma α_t , data dalla (172), che, per la (163'), si può anche scrivere

$$(172') \quad \alpha_t = \frac{\alpha}{\sqrt{1 + \left(\frac{\alpha t}{\beta m} \right)^2}}.$$

Questa formula dimostra che la precisione diminuisce col tempo, cioè che il gruppo d'onde, nel propagarsi (dopo il tempo 0) si allarga sempre più: ciò corrisponde al fatto (valido anche nella meccanica ordinaria) che, quando la velocità iniziale non è esattamente determinata, l'indeterminazione nella posizione cresce col tempo. Si noti che, se si vuole utilizzare la misura fatta al tempo 0 per calcolare la posizione in un istante *anteriore* ($t < 0$) ci si trova ancora di fronte ad una indeterminazione tanto maggiore, quanto più questo istante è lontano dall'istante della misura.

§ 37. - GRADINO DI POTENZIALE.

Supponiamo che una particella, di energia determinata E , sia soggetta ad un potenziale $U(x)$ avente l'andamento rappresentato dalla fig. 25, e cioè

$$U = 0 \quad \text{per } x < 0 \text{ (regione I)}, \\ U = U_0 \quad \text{per } x > 0 \text{ (regione II)}.$$

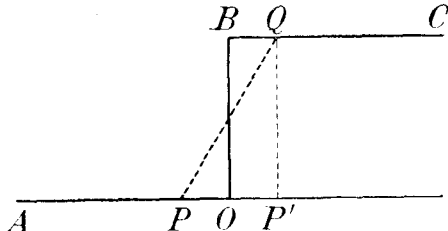
Questa condizione, puramente schematica, si può considerare come caso limite della seguente: supponiamo che in un tratto PP' vi sia un campo di forza che sollecita la particella verso sinistra (come potrebbe realizzarsi, nel caso di un elettrone, con due griglie cariche di segno opposto in P e P'): allora il potenziale avrebbe l'andamento $APQC$.


Fig. 25.

Facendo ora tendere a 0 l'intervallo PP' , ma mantenendo costante la differenza di potenziale tra P e P' (quindi, rendendo infinito il campo) si giunge al caso limite considerato. Questo si può quindi considerare come un « potenziale di doppio strato ».

L'equazione di Schrödinger sarà, nella regione I, ancora la (148), mentre nella regione II avrà la stessa forma salvo che in luogo di k vi figurerà la costante

$$k_0 = \sqrt{\frac{2m}{\hbar^2} (E - U_0)},$$

la quale potrà essere reale o immaginaria, secondochè $E \geq U_0$ oppure $E < U_0$.

La u sarà data dunque da

$$(174) \quad u = A_1 e^{2\pi i k x} + B_1 e^{-2\pi i k x} \text{ per } x < 0,$$

$$(174') \quad u = A_2 e^{2\pi i k_0 x} + B_2 e^{-2\pi i k_0 x} \text{ per } x > 0,$$

e poichè la u deve essere continua, insieme alla sua derivata prima, per $x = 0$, le quattro costanti A_1, A_2, B_1, B_2 dovranno esser legate dalle relazioni

$$(175) \quad \begin{cases} A_1 + B_1 = A_2 + B_2 \\ A_1 - B_1 = \mu(A_2 - B_2) \end{cases},$$

dove si è posto

$$\mu = \frac{k_0}{k} = \sqrt{1 - \frac{U_0}{E}}.$$

Convieni ora distinguere due casi secondo che l'energia E della particella è inferiore o no al dislivello U_0 di potenziale (supporremo in ogni caso che la particella sia lanciata da $-\infty$ verso O).

a) $E \geq U_0$. In questo caso, secondo la meccanica classica, la particella supererebbe il gradino di potenziale, e proseguirebbe il suo moto a destra di O con velocità ridotta nel rapporto $1 : \mu$. Dal punto di vista ondulatorio, dobbiamo invece osservare che ciascuna delle (174), (174') rappresenta la sovrapposizione di onde progressive e regressive: poichè supponiamo la particella proveniente da $-\infty$, e non da $+\infty$, nella regione II non vi dovranno essere onde regressive, e quindi $B_2 = 0$. Le (175) ci danno poi, eliminando A_2 ,

$$B_1 = \frac{1 - \mu}{1 + \mu} A_1.$$

Quindi nella regione I vi saranno in generale, oltre alle onde progressive di ampiezza A_1 , anche le onde regressive, riflesse dallo scalino di potenziale, di ampiezza B_1 . Il significato di queste costanti si enuncia più facilmente se si pensa di lanciare non una sola particella ma un gran numero di esse. Allora $A_1 A_1^*$ è proporzionale al numero di particelle (per unità di lunghezza) che, nella regione a sinistra di O , si muovono in senso progressivo, $B_1 B_1^*$ al numero di quelle che si muovono in senso regressivo: perciò il rapporto

$$R = \frac{B_1 B_1^*}{A_1 A_1^*} = \left(\frac{1 - \mu}{1 + \mu} \right)^2$$

rappresenta la frazione di particelle che vengono respinte indietro o riflesse, ovvero la probabilità che ha una particella di essere riflessa: perciò R può chiamarsi *coefficiente di riflessione* del gradino, mentre

$$\tau = 1 - R$$

rappresenta il *coefficiente di trasmissione*.

Una particella incidente ha dunque, secondo la meccanica ondulatoria, una certa probabilità R di essere riflessa (con la stessa velocità) e una certa probabilità τ di oltrepassare il gradino (con velocità ridotta nel rapporto $1 : \mu$, poichè la lunghezza d'onda da $1/k$ diviene $1/k_0$).

Il problema è analogo alla riflessione della luce contro una superficie semiriflettente: in ottica infatti il coefficiente di riflessione rappresenta la probabilità che ha un fotone incidente di venire riflesso.

b) $E < U_0$. In questo caso, secondo la meccanica classica, la particella verrebbe respinta indietro, senza oltrepassare il gradino.

Nella trattazione ondulatoria, dovremo invece osservare che in questo caso k_0 e μ_0 sono immaginari: perciò porremo

$$k_0 = ik' \quad , \quad \mu_0 = i\mu',$$

(con k' e μ' reali) e scriveremo la (174') nella forma

$$(176) \quad u = A_2 e^{-2\pi k' x} + B_2 e^{2\pi k' x} \quad (x > 0).$$

Osserviamo poi che, perchè la u si conservi finita anche per $x \rightarrow \infty$, dovrà essere $B_2 = 0$: tenuto conto di ciò, le (175) danno, come prima,

$$B_1 = \frac{1 - i\mu'}{1 + i\mu'} A_1.$$

Come si vede, poichè $1 + i\mu'$ e $1 - i\mu'$ hanno lo stesso modulo, le onde riflesse hanno uguale ampiezza di quelle incidenti, si ha cioè

$$R = \frac{B_1 B_1^*}{A_1 A_1^*} = 1,$$

il che significa che tutte le particelle sono riflesse. Però la u è diversa da zero anche a destra di O , dove è data dalla (176), che si riduce a

$$(177) \quad u = A_2 e^{-2\pi k' x}.$$

La parte reale di u (e così la parte immaginaria) hanno dunque l'andamento rappresentato nella fig. 26. Ciò significa che, se si esegue una osservazione di posizione della particella, vi è una certa probabilità di trovarla anche a destra di O , probabilità che è sensibile

fino ad una distanza da O dell'ordine di $1/k'$. In questo risultato la meccanica ondulatoria si differenzia nettamente da quella classica.

Questo risultato può apparire paradossale dal punto di vista classico perchè nella regione a destra di O l'energia potenziale U_0 della particella sarebbe superiore alla sua energia totale E , e quindi la forza viva $E - U_0$ risulterebbe negativa e la velocità immaginaria: è perciò che secondo la meccanica classica tale regione è inaccessibile alla particella. Ma invece dal punto di vista della meccanica ondulatoria il risultato non ha nulla di paradossale. Difatti, non si deve pensare che esso significhi che la particella penetra un poco oltre il gradino prima di essere riflessa (interpretazione contraria allo spirito della meccanica ondulatoria, secondo il quale non si può seguire il moto della par-

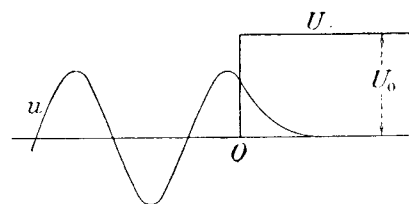


Fig. 26.

ticella) ma significa soltanto questo: che, facendo un'osservazione di posizione, si può trovare la particella anche a destra di O . Ora, bisogna rammentare che un'osservazione diretta a localizzare la particella richiede inevitabilmente che ad essa venga comunicato un impulso (sconosciuto) e quindi che se ne alteri l'energia: perciò, se troveremo la particella a destra di O , non potremo dire di aver colto in fallo il principio della conservazione dell'energia, poichè il fatto si può interpretare, nel linguaggio corpuscolare, dicendo che l'energia necessaria a superare il gradino è stata comunicata alla particella all'atto dell'osservazione (1).

Il fenomeno qui considerato è analogo al noto fenomeno ottico per cui, nella riflessione totale della luce, si ha una perturbazione luminosa anche nel secondo mezzo, nelle immediate vicinanze della superficie riflettente: però nel caso delle onde luminose la riflessione totale si produce solo se l'angolo di incidenza supera l'angolo limite, mentre per le onde di De Broglie il fenomeno si produce, come si è visto, anche con incidenza normale.

§ 38. - PARTICELLA SU UN SEGMENTO.

Mostriamo ora un primo esempio di quantizzazione col metodo di Schrödinger, considerando una particella che possa scorrere (senza forze) su un tratto *limitato* di retta AB , rimbalzando elasticamente agli estremi. Ricerchiamo anche qui, innanzi tutto, le soluzioni semplici (cioè con E determinato).

(1) Questa considerazione può essere precisata mediante un'analisi più dettagliata nella quale non entreremo. Ci limitiamo ad accennare brevemente alla seguente obiezione che potrebbe apparentemente opporsi al ragionamento fatto sopra: non si potrebbe disporre l'osservazione in modo da avere un'indeterminazione Δp nell'impulso inferiore all'impulso che è classicamente necessario a superare il gradino (cioè a $\sqrt{2m}(\sqrt{U_0} - \sqrt{E})$), pur rassegnandosi, per il principio di indeterminazione, ad avere un $\Delta x \geq \frac{h}{4\pi\Delta p}$? In tal caso, se dalla misura risultasse $x > \Delta x$, si sarebbe certi, malgrado l'indeterminazione su x , che la particella è a destra di O . Tale ragionamento è errato perchè Δx e Δp rappresentano i valori medi dello scarto e non quelli massimi i quali, per un teorema enunciato al § 13, non possono essere entrambi finiti; inoltre, poichè la posizione della particella non è *a priori* del tutto indeterminata (essendo praticamente certo che $x < 1/k'$), non è possibile rendere Δp tanto piccolo quanto si vuole.

L'equazione di Schrödinger sarà ancora la (148), ma con la condizione che fuori del segmento AB la ψ si annulli (essendo per ipotesi nulla la probabilità di trovare ivi la particella).

Si possono esprimere le condizioni di questo problema schematico in modo più atto a mostrarne il significato fisico, pensando i due ostacoli A, B , realizzati mediante due *gradini di potenziale* di altezza U_0 (fig. 27) e poi pensando di far tendere U_0 all'infinito.

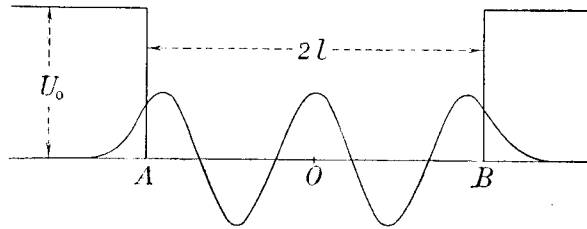


Fig. 27.

nito. Con ciò il k' della (177) tende a ∞ e quindi l'esponentiale della fig. 26 discende a zero immediatamente oltre il gradino: si può quindi fissare l'attenzione solo all'andamento della ψ nel tratto AB , imponendo la condizione che essa debba annullarsi agli estremi.

Poichè l'equazione (148) si identifica con quella già studiata nel § 8 (facendo corrispondere $4\pi^2k^2$ al parametro λ della (21)) il problema attuale è quello già trattato al § 8 sotto le condizioni (a): segue allora dalla (24') che k può assumere solo i valori

$$-k_n = \frac{n}{4l}$$

(cioè, come è intuitivo, quelli per cui la semilunghezza di onda $1/2 k_n$ è sottomultipla della lunghezza $2l$). Se ne ricavano, mediante la (147), i livelli energetici, che sono

$$(178) \quad E_n = \frac{h^2}{32 ml^2} n^2 .$$

A questi corrispondono le autofunzioni (normalizzate) date dalla (25), cioè

$$(179) \quad u_n = \sqrt{\frac{1}{l}} \sin n \frac{\pi}{2l} (x + l) ,$$

e quindi

$$(179') \quad \psi_n = \sqrt{\frac{1}{l}} e^{-\frac{2\pi i}{h} E_n t} \sin n \frac{\pi}{2l} (x + l) .$$

La densità di probabilità di posizione in un istante qualunque è

$$(180) \quad P(x) = \psi \psi^* = uu^* = \frac{1}{l} \sin^2 n \frac{\pi}{2l} (x + l)$$

e presenta quindi dei nodi che dividono AB in n tratti uguali.

La distribuzione della probabilità dell'impulso si ottiene osservando che la (179') si può scrivere

$$\psi_n = c_1 e^{i \frac{n\pi}{2l} x - \frac{2\pi i}{h} E_n t} + c_2 e^{-i \frac{n\pi}{2l} x - \frac{2\pi i}{h} E_n t} ,$$

dove c_1 e c_2 sono date dalle (23') del § 8. Poichè il primo di questi due termini rappresenta delle onde progressive di « numero d'onde » k_n , ed il secondo delle onde identiche, ma regressive, vi sarà in ogni istante una probabilità proporzionale a $|c_1|^2$ di trovare la particella con un impulso $p_n = \frac{h k_n}{h}$

ed una probabilità proporzionale a $|c_2|^2$ di trovarla con impulso $p_n = -\frac{h k_n}{h}$: dalle (23') poi risulta che c_1 e c_2 hanno lo stesso modulo, e quindi le due probabilità sono uguali. La probabilità dell'impulso non è dunque distribuita con continuità, ma è nulla dappertutto tranne pei due valori $\pm k_n/h$, dove vale $1/2$.

Passiamo ora al caso in cui la E non ha un valore determinato (ossia il sistema non è in uno stato stazionario o quantico): in questo caso la ψ è rappresentata da una somma di termini della forma (179'), cioè:

$$(181) \quad \psi = \sqrt{\frac{1}{l}} \sum_1^\infty c_n \sin n\pi \frac{x+l}{2l} e^{-\frac{2\pi i}{h} E_n t}$$

e $|c_n|^2$ rappresenta la probabilità che il sistema sia nello stato n -esimo, cioè che la sua energia sia E_n (le c_n sono, come si

sa, soggette alla restrizione $\sum |c_n|^2 = 1$, che equivale alla normalizzazione di ψ).

In particolare, si possono determinare le c_n in modo da costituire un piccolo gruppo d'onde, e allora si trova che questo si muove di moto alternativo tra A e B in modo simile al movimento previsto dalla meccanica classica, dilatandosi però gradatamente, come il gruppo considerato nel § 36.

§ 39. - OSCILLATORE ARMONICO.

Si chiama *oscillatore armonico* il sistema costituito da un punto materiale mobile su una retta, e attirato verso un punto O di questa da una forza proporzionale alla distanza. Presa la retta come asse x , O come origine, la forza che agisce sul punto sarà dunque $-Kx$, dove K è una costante positiva. È noto che secondo la meccanica classica il punto esegue delle oscillazioni armoniche intorno ad O con ampiezza e fase arbitrarie (dipendenti dalle condizioni iniziali) e frequenza

$$(182) \quad \nu_0 = \frac{1}{2\pi} \sqrt{\frac{K}{m}}.$$

Per studiare il problema corrispondente a questo in meccanica ondulatoria, osserviamo che l'energia potenziale corrispondente alla forza $-Kx$ è

$$(182') \quad U = \frac{K}{2} x^2 = 2\pi^2 m \nu_0^2 x^2$$

e quindi, trasportandola nell'equazione unidimensionale di Schrödinger (146), si ottiene

$$(183) \quad \frac{d^2 u}{dx^2} + \frac{8\pi^2 m}{h^2} (E - 2\pi^2 m \nu_0^2 x^2) u = 0.$$

Si tratta di trovare gli autovalori e le autofunzioni di questa equazione, per l'intervallo da $-\infty$ a $+\infty$.

Riduciamo anzitutto l'equazione ad una forma che ne metta chiaramente in evidenza l'aspetto analitico: per far ciò, basta operare un cambiamento di scala sull'asse x , in-

troducendo la variabile

$$\xi = x \sqrt{\frac{4\pi^2 m \nu_0}{h}},$$

e poi dividere tutta l'equazione per $\frac{4\pi^2 m \nu_0}{h}$, con che essa diviene

$$(183') \quad \frac{d^2 u}{d\xi^2} + (\varepsilon - \xi^2) u = 0,$$

dove si è abbreviata con ε la costante

$$(184) \quad \varepsilon = \frac{2E}{h\nu_0}$$

(che rappresenta l'energia misurata con l'unità $\frac{h\nu_0}{2}$).

Bisogna dunque cercare se la (183') ammette soluzioni finite e continue dovunque, e tendenti a 0 per ξ tendente a $\pm\infty$: si troverà che ciò è possibile solo se ε è uguale ad un numero dispari positivo.

La discussione dell'equazione (183') si può fare nel modo seguente. Si osservi anzitutto che i suoi coefficienti sono finiti per tutti i valori finiti di ξ , quindi eventuali singolarità della u si possono presentare solo per $\xi = \pm\infty$: col criterio del § 16 si riconosce che effettivamente l'equazione ha all'infinito due singolarità non fuchsiane. Un'idea del comportamento asintotico delle soluzioni in questi punti si può avere col seguente metodo euristico: si tenti di soddisfare l'equazione ponendo $u = e^{\lambda\xi^2}$: si troverà che, se si trascurano i termini finiti rispetto a quelli in ξ^2 , l'equazione è asintoticamente soddisfatta prendendo $\lambda = \pm \frac{1}{2}$. Poichè noi cerchiamo soluzioni che non tendono all' ∞ , dobbiamo scartare quelle che hanno per espressione asintotica $e^{\frac{1}{2}\xi^2}$: porremo perciò

$$(185) \quad u = e^{-\frac{1}{2}\xi^2} v$$

e cercheremo di determinare v in modo che l'equazione sia esattamente soddisfatta, e che inoltre la u si conservi finita anche all'infinito, per il che basta che la v non abbia ivi singolarità essenziali.

Sostituendo la (185) nella (183') si trova per v l'equazione

$$(186) \quad v'' - 2\xi v' + (\varepsilon - 1)v = 0.$$

Cerchiamo di integrare questa equazione con una serie della forma

$$(187) \quad v = \sum_0^{\infty} a_r \xi^r$$

(da cui abbiamo escluso le potenze negative di ξ , perchè vogliamo che la soluzione sia finita anche per $\xi = 0$). Sostituendo questa serie nella (186), si trova

$$\sum_0^{\infty} [(r+1)(r+2)a_{r+2} + (\varepsilon - 1 - 2r)a_r] \xi^r = 0.$$

Perchè questa sia identicamente soddisfatta, devono annullarsi tutti i coefficienti, il che dà per le a_r la formula ricorrente

$$(188) \quad a_{r+2} = \frac{2r+1-\varepsilon}{(r+1)(r+2)} a_r.$$

Caratteristica di questa formula è che essa lega ogni coefficiente a quello che lo precede di *due* posti: così, fissato ad arbitrio a_0 , si ricavano da esso mediante la (188) tutti i coefficienti di posto pari, e fissato ad arbitrio a_1 si ricavano tutti quelli di posto dispari. Due soluzioni fondamentali si ottengono p. es. prendendo

$$(189) \quad a_0 \neq 0, \quad a_1 = 0 \text{ (serie di potenze pari),}$$

oppure

$$(189') \quad a_0 = 0, \quad a_1 \neq 0 \text{ (serie di potenze dispari),}$$

e qualunque altra soluzione è una combinazione lineare di queste.

Ora si può dimostrare facilmente che in generale queste due serie hanno per $\xi = \pm \infty$ delle singolarità essenziali: solo nel caso che uno dei coefficienti, p. es. a_{n+r} , si annulli, cioè non avviene, perchè in questo caso, come risulta dalla (188), si annullano anche tutti i coefficienti successivi e la serie si riduce a un polinomio di grado n . La condizione perchè sia $a_{n+2} = 0$ essendo $a_n \neq 0$ è, come si vede dalla (188), che sia

$$(190) \quad \varepsilon = 2n + 1.$$

Dunque: gli autovalori della (183') sono tutti i numeri dispari positivi.

Dalla (184) e dalla (190) risulta che la E deve avere uno dei valori

$$(191) \quad E_n = \left(n + \frac{1}{2}\right) h\nu_0 \quad (n = 0, 1, 2, \dots).$$

Sono questi dunque i livelli energetici dell'oscillatore.

Osserviamo a questo proposito che l'antica teoria di Bohr e Sommerfeld dava (come si vedrà al § 54) in luogo della (191), la formula $E_n = nh\nu_0$ (che si può ricordare enunciandola così: l'energia dell'oscillatore deve essere un multiplo intero dell'energia di un fotone di frequenza uguale alla frequenza propria dell'oscillatore). Le differenze tra due livelli successivi, che intervengono nei fenomeni di emissione e di assorbimento, sono dunque le stesse in entrambe le teorie, ma il livello più basso, che nella teoria di Bohr e Sommerfeld è 0, è qui $\frac{1}{2} h\nu_0$. Numerosi dati sperimentali (p. es., gli spettri di bande) confermano la formula (191).

Passiamo ora ad occuparci delle autofunzioni. Quando ε è dato dalla (190), la formula ricorrente (188) diviene

$$(192) \quad a_{r+2} = -\frac{2(n-r)}{(r+1)(r+2)} a_r.$$

e, se n è pari, si dovrà considerare la soluzione a potenze pari ($a_1 = 0, a_0 \neq 0$, arbitrario), se n è dispari quella a potenze dispari ($a_0 = 0, a_1 \neq 0$, arbitrario). Polinomi siffatti si presentano nello sviluppo delle derivate successive della funzione $e^{-\xi^2}$: infatti si verifica facilmente che

$$(193) \quad \frac{d^n}{d\xi^n} e^{-\xi^2} = (-1)^n H_n(\xi) e^{-\xi^2},$$

dove $H_n(\xi)$ è un polinomio, di grado n , del tipo che stiamo considerando: questi H_n si chiamano *polinomi di Hermite* ⁽¹⁾. I primi quattro di questi, p. es., sono (come si trova facilmente)

$$\begin{array}{ll} H_0 = 1 & H_1 = 2\xi \\ H_2 = -2 + 4\xi^2 & H_3 = -12\xi + 8\xi^3. \end{array}$$

Dunque all'autovalore E_n corrisponde una v data da

$$v_n = \frac{1}{N_n} H_n(\xi)$$

(¹) Vedasi, p. es., bibl. n.° 25 o n.° 34.

(dove N_n è un fattore costante di normalizzazione) e quindi, per la (185), una u data da

$$u_n = \frac{1}{N_n} H_n(\xi) e^{-\frac{1}{2}\xi^2}.$$

La condizione di normalizzazione è

$$\int_{-\infty}^{\infty} u_n^2 dx = \sqrt{\frac{h}{4\pi^2 m v_0}} \int_{-\infty}^{\infty} u_n^2 d\xi = 1,$$

ossia

$$N_n^2 = \sqrt{\frac{h}{4\pi^2 m v_0}} \int_{-\infty}^{\infty} H_n^2(\xi) e^{-\xi^2} d\xi.$$

L'integrale si calcola con successive integrazioni per parti, utilizzando la seguente proprietà dei polinomi di Hermite:

$$(194) \quad \frac{dH_n}{d\xi} = 2nH_{n-1},$$

e si trova

$$N_n = \sqrt{\frac{h}{4\pi^2 m v_0}} \sqrt{2^n n! \sqrt{\pi}}.$$

Con un procedimento simile si verificherebbe che

$$\int_{-\infty}^{\infty} H_n(\xi) H_m(\xi) e^{-\xi^2} d\xi = 0$$

per $n \neq m$, cioè che le funzioni u_n sono ortogonali tra loro.

Le funzioni u_n , per i primi 5 valori di n , sono rappresentate nella fig. 28, prescindendo dal fattore $\sqrt{\frac{h}{4\pi^2 m v_0}}$ (la prima di esse ($n = 0$) non è altro che la nota funzione di Gauss). La funzione u_n^2 , che dà la distribuzione della probabilità di posizione della particella quando si sappia che essa è nello stato quantico n -esimo, è rappresentata, per i primi 5 valori di n , dalle fig. 29: in questi diagrammi il tratto OC rappresenta l'ampiezza delle oscillazioni nel modello classico ($OC = \sqrt{\frac{2E_n}{K}}$). Si osservi a questo proposito che per $x > OC$ risulta $U > E_n$, e quindi la forza viva $E_n - U$, per

una particella a destra di C , risulterebbe negativa ossia la sua velocità immaginaria: perciò tale regione è inaccessibile

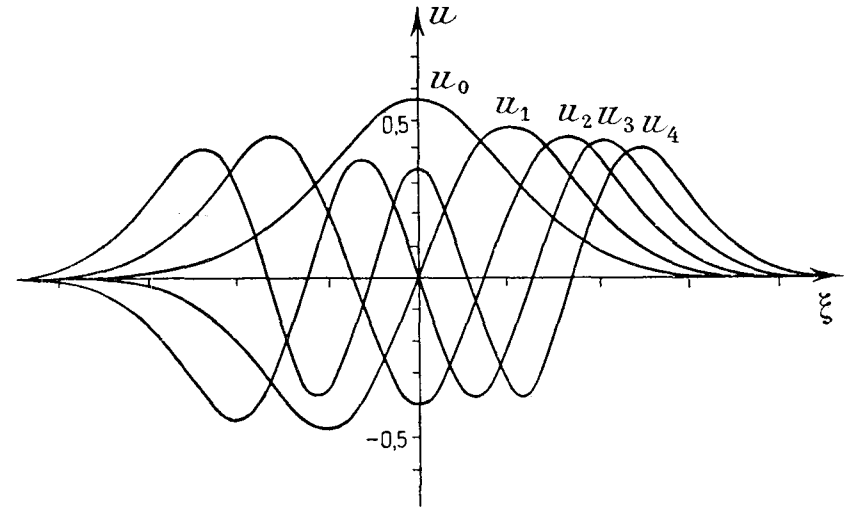


Fig. 28.

alla particella di energia E_n , secondo la meccanica classica. Tuttavia, come risulta dalle curve della fig. 29, vi è la pos-

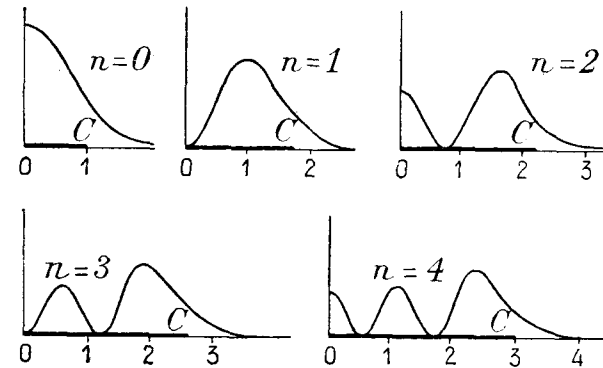


Fig. 29.

sibilità di trovare la particella in tale regione: è questo un apparente paradosso analogo a quello già spiegato nel § 37.

Ricorderemo poi che se, anziché esser determinata l'energia del sistema, sono state determinate la posizione e la ve-

locità iniziale della particella (con l'approssimazione e limitata dal principio di indeterminazione) allora la distribuzione della probabilità, per una successiva determinazione della coordinata x , è data dal quadrato del modulo di una ψ , ottenuta come combinazione lineare di soluzioni corrispondenti a diversi stati quantici e tale da rappresentare un gruppo d'onde tanto più ristretto, quanto più precisa è stata la determinazione iniziale della x : SCHRÖDINGER ha poi dimostrato che questo gruppo d'onde si muove di moto oscillatorio, imitando approssimativamente il moto armonico della meccanica classica. In questo caso anzi, al contrario di ciò che avviene in generale (v. p. es., § 36) il gruppo d'onde conserva indefinitamente le sue dimensioni senza disfarsi.

§ 40. - BARRIERA DI POTENZIALE.

Ha notevole interesse per le applicazioni (ad alcune delle quali si accennerà nel seguito) il problema del moto di una particella lungo una retta x , sotto l'azione di un potenziale

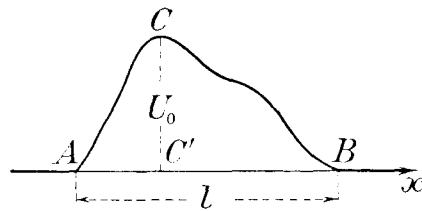


Fig. 30.

nullo dappertutto, tranne che in un certo intervallo AB , entro il quale esso cresce fino ad un massimo U_0 e poi decresce fino a 0: andamento qualitativamente rappresentato dalla fig. 30.

La zona AB costituisce ciò

che chiamasi una *barriera di potenziale*: una particella che provenga p. es. da sinistra trova forza nulla fino ad A , poi incontra una forza ritardatrice da A a C' , indi forza acceleratrice da C' a B , e poi di nuovo forza nulla.

Secondo la meccanica classica, la particella oltrepassa la barriera se la sua forza viva iniziale E è superiore al massimo U_0 del potenziale, altrimenti viene respinta indietro prima di giungere in C' .

La meccanica ondulatoria invece fornisce un risultato diverso: la particella incidente ha in ogni caso una certa probabilità di oltrepassare la barriera (anche se $E < U_0$) e una certa probabilità di rimbalzare indietro (anche se $E > U_0$).

Per trovare questo risultato in modo semplice, schematizziamo ancora il problema dando al potenziale l'andamento della fig. 31, e cioè

$$\begin{aligned} U &= 0 & \text{per } x < 0 \text{ e } x > l & \text{(regioni I e III)} \\ U &= U_0 & \text{per } 0 < x < l & \text{(regione II)} \end{aligned}$$

Nel caso, più generale, della fig. 30, si troverebbe (mediante un procedimento di approssimazioni successive) un risultato qualitativamente analogo a quello che troveremo ora dando alla barriera la forma particolare della fig. 31, costituita di due « gradini di potenziale » in senso opposto.

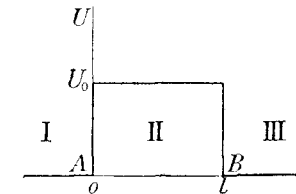


Fig. 31.

Consideriamo dunque separatamente le tre regioni (I, II, III): l'equazione di Schrödinger è, per le regioni I e III, la stessa (148) già studiata nel § 35, il cui integrale generale ha la forma (149), ma le costanti che vi figurano saranno in generale diverse nei due tratti: nella regione II poi, la soluzione avrà la stessa forma salvo la sostituzione di $E - U_0$ in luogo di E . Scriveremo dunque:

$$(195) \quad \left\{ \begin{array}{ll} \text{Regione I} & u = A_1 e^{2\pi i k x} + B_1 e^{-2\pi i k x} \\ \text{» II} & u = A_2 e^{2\pi i k_0 x} + B_2 e^{-2\pi i k_0 x} \\ \text{» III} & u = A_3 e^{2\pi i k x} + B_3 e^{-2\pi i k x} \end{array} \right.$$

dove si è posto

$$k = \sqrt{\frac{2m}{\hbar^2} E}, \quad k_0 = \sqrt{\frac{2m}{\hbar^2} (E - U_0)}.$$

Supponiamo di lanciare contro la barriera da sinistra a destra un gran numero di particelle: allora i quadrati dei moduli delle sei costanti A, B avranno i significati già illustrati nel § 37: dovremo poi porre anche qui $B_3 = 0$ per esprimere il fatto che nessuna particella proviene da destra. Le particelle avranno evidentemente la stessa velocità nelle regioni I e III, ed il *coefficiente di trasmissione* della barriera

(cioè la probabilità che ha ogni particella incidente di oltrepassarla) sarà dato da

$$(196) \quad \tau = \frac{A_3 A_3^*}{A_1 A_1^*}.$$

Osserviamo che le cinque costanti A_1, B_1, A_2, B_2, B_3 devono esser legate tra loro dalla condizione che la u sia continua, insieme alla sua derivata, nei punti A e B . La parte reale di u (come anche la sua parte immaginaria) sarà rappresentata nei tratti I e III da due sinusoidi di lunghezza d'onda $\lambda = 1/k$ (poichè certamente k è reale): nel tratto II,

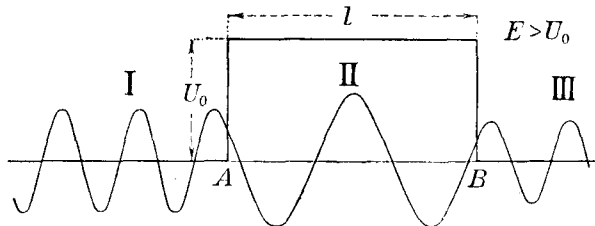


Fig. 32.

invece, la curva può avere due aspetti diversi secondo che la forza viva E supera o no U_0 . Nel primo caso k_0 è reale e quindi la curva è di forma sinusoidale anche nel tratto II, ma con lunghezza d'onda $\lambda_0 = 1/k_0$ maggiore che nei tratti I e III (fig. 32); nel secondo caso k_0 è immaginario e quindi

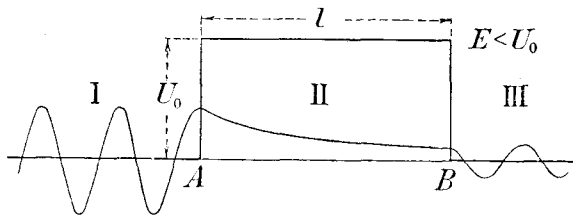


Fig. 33.

la curva nel tratto II è di tipo esponenziale (fig. 33). In entrambi i casi le curve delle tre regioni si raccordano con continuità, come mostrano le figure: questo implica una relazione tra le ampiezze delle due sinusoidi estreme, relazione di cui ora troveremo l'espressione analitica, che ci condurrà al calcolo di τ .

Le condizioni di continuità di u e di $\frac{du}{dx}$ per $x = 0$ danno

$$\begin{aligned} A_1 + B_1 &= A_2 + B_2, \\ A_1 - B_1 &= \mu(A_2 - B_2), \end{aligned}$$

dove si è posto

$$\mu = \frac{k_0}{k} = \sqrt{1 - \frac{U_0}{E}}.$$

Di qui si ricava

$$(197) \quad A_1 = \frac{1}{2} A_2 (1 + \mu) + \frac{1}{2} B_2 (1 - \mu).$$

Analogamente le condizioni di continuità per $x = l$ danno

$$\begin{aligned} A_2 e^{2\pi i k_0 l} + B_2 e^{-2\pi i k_0 l} &= A_3 e^{2\pi i k l}, \\ A_2 e^{2\pi i k_0 l} - B_2 e^{-2\pi i k_0 l} &= \frac{1}{\mu} A_3 e^{2\pi i k l}, \end{aligned}$$

da cui si ricava

$$\begin{aligned} A_2 &= \frac{\mu + 1}{2\mu} A_3 e^{2\pi i (k - k_0) l}, \\ B_2 &= \frac{\mu - 1}{2\mu} A_3 e^{2\pi i (k + k_0) l}. \end{aligned}$$

Sostituendo queste espressioni nella (197) si ha la relazione tra A_1 e A_3 :

$$(198) \quad A_1 = \frac{A_3}{4\mu} e^{+2\pi i k l} \left[(\mu + 1)^2 e^{-2\pi i k_0 l} - (\mu - 1)^2 e^{2\pi i k_0 l} \right].$$

Di qui, mediante la (196), si ricava τ , o, più comodamente, il suo inverso:

$$(199) \quad \frac{1}{\tau} = \frac{|A_1|^2}{|A_3|^2} = \frac{1}{16|\mu|^2} \left| (\mu + 1)^2 e^{-2\pi i k_0 l} - (\mu - 1)^2 e^{2\pi i k_0 l} \right|^2.$$

Per calcolare questa espressione conviene trattare separatamente i due casi di $E > U_0$ o $E < U_0$:

a) Caso di $E > U_0$. In tal caso k_0 e μ sono reali, e perciò il secondo membro della (199) si può scrivere

$$\begin{aligned} \frac{1}{\tau} &= \frac{1}{16\mu^2} \left[(\mu + 1)^2 e^{-2\pi i k_0 l} - (\mu - 1)^2 e^{2\pi i k_0 l} \right] \times \left[(\mu + 1)^2 e^{2\pi i k_0 l} - \right. \\ &\quad \left. - (\mu - 1)^2 e^{-2\pi i k_0 l} \right] = \\ (199') &= \frac{1}{16\mu^2} \left[(\mu + 1)^4 + (\mu - 1)^4 - 2(\mu^2 - 1)^2 \cos 4\pi k_0 l \right]. \end{aligned}$$

Si vede di qui che, fissato μ , il coefficiente di trasmissione τ varia in modo periodico col variare di l (spessore della barriera): per $\cos 4\pi k_0 l = 1$ ossia per

$$(200) \quad l = \frac{n}{2k_0} = n \frac{\lambda_0}{2} \quad (n \text{ intero}),$$

la parentesi quadra acquista il suo massimo valore, uguale a $16\mu^2$, e quindi

$$\tau = \tau_{max} = 1.$$

Per $\cos 4\pi k_0 l = -1$ e quindi per

$$l = \frac{n + 1/2}{2k_0} = \left(n + \frac{1}{2}\right) \frac{\lambda_0}{2},$$

si ha invece

$$\tau = \tau_{min} = \frac{4\mu^2}{(1 + \mu^2)^2}.$$

Tra questi due limiti è sempre compreso τ : come si vede dall'ultima formula, esso non si annulla mai e quindi vi è sempre una certa probabilità che una particella attraversi la barriera. Questa probabilità diviene anzi la certezza, se la larghezza della barriera contiene un numero intero di semilunghezze d'onda, come indica la (200) (si osservi l'analogia di questo fenomeno con la trasmissione della luce attraverso una lamina sottile): ma in ogni altro caso $\tau < 1$, e quindi vi è una certa probabilità $1 - \tau$ che la particella sia respinta indietro, contrariamente a ciò che risulterebbe dalla meccanica classica.

b) Caso di $E < U_0$. In questo caso, che è il più interessante per le applicazioni, k_0 e quindi μ sono immaginari, cosicchè scriveremo

$$k_0 = ik' = i \sqrt{\frac{2m}{\hbar^2} (U_0 - E)}, \quad \mu = i\mu' = i \sqrt{\frac{U_0}{E} - 1},$$

con k' e μ' reali. Perciò la (199) diverrà

$$\frac{1}{\tau} = \frac{1}{16\mu'^2} \left[(1 + i\mu')^2 e^{2\pi k' l} - (1 - i\mu')^2 e^{-2\pi k' l} \right] \times \\ \times \left[(1 - i\mu')^2 e^{2\pi k' l} - (1 + i\mu')^2 e^{-2\pi k' l} \right] =$$

$$(201) \quad = \frac{1}{16\mu'^2} \left[-(1 + i\mu')^4 - (1 - i\mu')^4 + 2(1 + \mu'^2)^2 \cosh 4\pi k' l \right] = \\ = 1 + \frac{(1 + \mu'^2)^2}{8\mu'^2} (\cosh 4\pi k' l - 1).$$

Come si vede, fissato μ' , il secondo membro cresce costantemente col crescere di l , e quindi τ decresce sempre coll'aumentare dello spessore della barriera (anzi, diviene trascurabile appena l è grande rispetto a $1/k'$): però si vede dalla formula che τ non si annulla mai (finchè, beninteso, U_0/E ed l si mantengano finiti) cosicchè vi è sempre una certa probabilità di oltrepassare la barriera, e proseguire indefinitamente al di là, anche per una particella di energia inferiore ad U_0 .

Si presenta qui, apparentemente, la stessa difficoltà rilevata a proposito del gradino di potenziale, inquantochè, se si considerasse la particella come dotata di un effettivo movimento continuo, nel senso ordinario, si sarebbe indotti a ritenere che essa avesse attraversato una regione in cui la sua energia cinetica doveva esser negativa; ma in realtà, come sappiamo, non si può parlare del movimento della particella nell'intervallo tra due successive osservazioni. E se una osservazione di posizione ci facesse trovare la particella nella regione dove $U > E$, ciò non implicherebbe alcun paradosso, poichè (come già si osservò) col solo fatto di osservarne la posizione si altera il valore dell'energia.

L'analogo ottico del fenomeno qui considerato è il seguente. È noto che la speciale perturbazione luminosa che si verifica nel secondo mezzo durante la riflessione totale può venire, per dir così, catturata e trasformata in radiazione ordinaria, ponendo, a brevissima distanza dalla superficie riflettente, la superficie di un secondo mezzo uguale a quello da cui proviene la luce: così se AMN (fig. 34) è un prisma di vetro sulla cui ipotenuusa MN si produce la riflessione totale, accostandovi a brevissima distanza (dell'ordine della lunghezza d'onda) un secondo prisma di vetro $A'M'N'$, si vede che

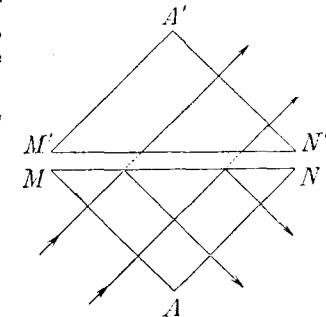


Fig. 34.

una parte della luce incidente passa in questo prisma e ne emerge dalla faccia $A'N'$. E ciò, sebbene nel passaggio dal primo prisma alla lamina d'aria risulti (secondo l'ottica geometrica) un angolo di rifrazione immaginario. La lamina d'aria $MNM'N'$ è perfettamente analoga alla barriera di potenziale dianzi considerata⁽¹⁾.

Quanto alla possibilità di una verifica sperimentale diretta del risultato della meccanica ondulatoria, bisogna pensare che essa sembra praticamente impossibile per il fatto che la distanza l tra le due barriere di potenziale dovrebbe essere dell'ordine di grandezza della lunghezza d'onda di De Broglie, la quale, anche per i raggi catodici più lenti, è sempre brevissima. Tuttavia, come si vedrà nei prossimi §§, vi sono dei fatti sperimentali che si possono interpretare come una indiretta conferma di questo notevole fenomeno.

§ 41. — CENNO SULLA TEORIA DELL'EMISSIONE ELETTRONICA A FREDDO.

È stato constatato sperimentalmente che applicando alla superficie di un metallo, nel vuoto, un forte campo elettrico X diretto verso il metallo, si provoca l'uscita da questo di elettroni, in numero proporzionale alla quantità $e^{-b/X}$; dove b è una costante. Di questo fenomeno, le considerazioni del paragrafo precedente permettono di dare una semplice spiegazione, in parte anche quantitativa.

È noto che per far uscire un elettrone da un metallo si richiede, all'atto dell'attraversamento della superficie, una certa quantità U_0 di energia: si può dunque dire che la superficie del metallo rappresenta per gli elettroni un gradino di potenziale di altezza U_0 , cosicchè, se disponiamo l'asse x normale alla superficie del metallo, il potenziale U delle forze agenti su un elettrone ha l'andamento della linea $AOBC$ nella fig. 35: è lo scalino OB che impedisce agli elettroni (la cui energia è minore di U_0) di uscire dal metallo.

Quando poi all'esterno del conduttore si applica il campo elettrico, il tratto BC del diagramma del potenziale è sostituito dalla retta inclinata BC' (la cui pendenza $\text{tg } \alpha$ misura

⁽¹⁾ Vi è però la differenza, che con le onde di De Broglie il fenomeno si produce anche se l'incidenza è normale, e con la luce no.

la forza eX); viene così a formarsi invece del gradino una «barriera» di potenziale OBD che, secondo la meccanica classica, sarebbe del pari insormontabile agli elettroni. Ma invece, come si è visto nel § precedente, la meccanica ondulatoria consente ad alcuni degli elettroni di sorpassare questa barriera ed uscire dal metallo.

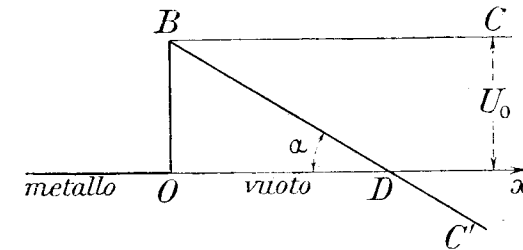


Fig. 35.

FOWLER e NORDHEIM hanno svolto su queste basi una teoria del fenomeno, che ne rende conto in modo soddisfacente anche sotto l'aspetto quantitativo. Noi ci limiteremo a far rilevare che le formule del § precedente lasciano prevedere che l'intensità della corrente elettronica emessa risulti espressa, in funzione del campo X , da una legge del tipo indicato. Difatti, dalla fig. 35 risulta che la larghezza OD della barriera è data da

$$OD = \frac{U_0}{\text{tg } \alpha} = \frac{U_0}{eX}$$

e se (prescindendo dalla differenza di forma della barriera che non modifica l'aspetto qualitativo del ragionamento) si identifica OD con la larghezza l della barriera considerata al § precedente, la quantità $4\pi k'l$ che figura nella formula (201) risulta

$$(202) \quad 4\pi k'l = 4\pi \sqrt{\frac{2m}{h^2} (U_0 - E)} \frac{U_0}{eX} = \frac{b}{X},$$

indicando con b l'insieme delle quantità indipendenti da X . Sostituendo i valori che intervengono nei casi pratici, si trova che $4\pi k'l$ risulta almeno dell'ordine di 100: quindi nell'applicare la formula (201) si può sostituire $\cosh 4\pi k'l$ con $e^{4\pi k'l}$, e di fronte a questo termine si può trascurare l'unità:

la formula si riduce dunque a

$$(203) \quad \frac{1}{\tau} = \frac{(1+\mu'^2)^2}{8\mu'^2} e^{b/x}$$

che dà per il coefficiente di trasmissione τ , e quindi per il numero di elettroni emessi, proprio l'espressione suggerita dai dati sperimentali.

§ 42. — PARTICELLA TRA DUE BARRIERE DI POTENZIALE.

Consideriamo ora il caso in cui il potenziale ha l'andamento della fig. 36, che può considerarsi costituita da due

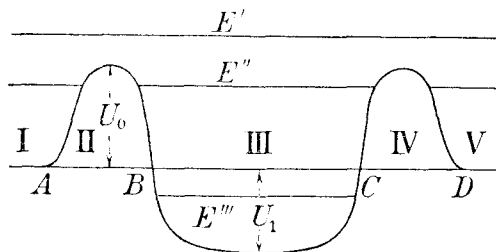


Fig. 36.

barriere di potenziale, simmetriche, di altezza U_0 racchiudenti una « valle » centrale: il fondo di questa può essere anche a un livello più basso del potenziale esterno, come nella figura.

Dal punto di vista della meccanica classica si devono distinguere tre casi. Se $E > U_0$ (p. es. livello E') una particella attraversa tutta la doppia barriera; se $U_0 > E > 0$ (p. es. livello E''), sono possibili tre tipi di movimento nettamente distinti e cioè: o la particella proviene da sinistra, e allora viene respinta dalla prima barriera e torna indietro, o la particella proviene da destra, e compie un simile movimento sulla semiretta positiva, o la particella si trova tra le due barriere, e allora compie delle oscillazioni periodiche. Se poi $0 > E > -U_1$ (livello E''') è possibile solo il moto oscillatorio tra le due barriere.

Per studiare il problema secondo la meccanica ondulatoria, conviene anche qui schematizzarlo dando al potenziale la forma della fig. 37. Si applica poi lo stesso procedimento del § 40, risolvendo l'equazione di Schrödinger separatamente per ciascuno dei cinque tratti in cui i punti A , B , C , D dividono l'asse x , e imponendo poi alla u e alla

$\frac{du}{dx}$ la condizione di essere continue in detti punti, il che stabilisce un legame tra le costanti relative ai diversi tratti. Il risultato che si ottiene si può estendere qualitativamente anche al caso della fig. 36: la sua caratteristica più notevole si riferisce al caso in cui $U_0 > E > 0$ (livelli del tipo E''), ed ora la esporremo sommariamente.

Premettiamo l'ovvia osservazione che, se le due barriere diventassero infinitamente alte, una particella che si trovasse nella regione centrale sarebbe nelle condizioni studiate al § 38 (particella su un segmento) e quindi la sua energia potrebbe assumere solo certi livelli discreti E_1, E_2, \dots :

in una parola, essa sarebbe « quantizzata ». Ma se invece consideriamo finita l'altezza delle barriere, queste non rappresentano più ostacoli insormontabili per la particella, l'energia non è più quantizzata e può assumere qualunque valore; si trova però il risultato seguente. La ψ ha evidentemente carattere oscillatorio sia nella regione centrale (III) che nelle due regioni esterne (I, V) (mentre è di tipo esponenziale nelle regioni II, IV): se l'energia ha un valore qualunque, l'ampiezza della ψ nella regione centrale risulta *in generale* assai minore che nelle regioni esterne (perchè i tratti esponenziali che le raccordano vanno discendendo verso l'interno): ciò significa che la particella ha pochissima probabilità di trovarsi nella regione centrale, e se vi si trova tende fortemente ad uscirne. Solo nel caso che l'energia sia *prossima* ad uno dei valori E_1, E_2, \dots , che nel caso delle barriere infinitamente alte erano i soli valori permessi, avviene che l'ampiezza nella regione centrale sia molto maggiore che nelle regioni esterne, ossia che la particella abbia grande probabilità di esser trovata nella regione centrale. Si può dire quindi che in luogo della quantizzazione rigorosa si ha una *quasi-quantizzazione*,

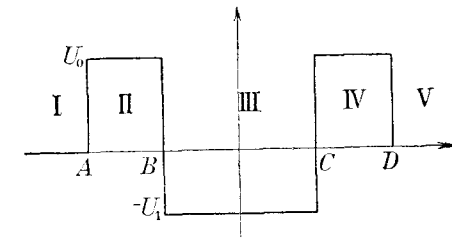


Fig. 37.

nel senso che i livelli discreti $E_1, E_2 \dots$ non sono più rappresentati da rette orizzontali, ma da sottili bande sfumate, e non rappresentano più i soli valori di E permessi, ma bensì i valori di E che conferiscono alla particella una notevole probabilità di esser trovata all'interno della « valle » limitata dalle due barriere.

Per livelli di energia inferiori a 0 (tipo E''') si ha invece una ordinaria quantizzazione, e una ψ limitata praticamente alla regione centrale; invece i livelli più alti delle barriere, cioè del tipo E' , non sono quantizzati in alcun modo, e corrispondono a una ψ di tipo sinusoidale dappertutto.

§ 43. — CENNO SULLA TEORIA DELL'EMISSIONE DEI RAGGI α .

Le considerazioni del § precedente costituiscono il fondamento di una notevole teoria ideata da GAMOW e, indipendentemente, da GURNEY e CONDON, per rendere conto dell'emissione spontanea di particelle α da parte dei nuclei delle sostanze radioattive.

Tale fenomeno dava luogo ad una grave difficoltà teorica. Si è potuto infatti determinare, con esperienze del tipo di quelle che condussero il RUTHERFORD alla scoperta del nucleo atomico (deviazione di particelle α), quale sia la forza che agisce su una particella α nelle vicinanze di un nucleo, e si è trovato che essa è coulombiana e repulsiva finché la distanza dal nucleo non scende al disotto di un certo limite (che è dell'ordine di 10^{-11} o 10^{-12} cm.): in quei casi in cui sperimentalmente si è potuto sorpassare questo limite (e cioè per gli elementi leggeri) si è trovato che, avvicinandosi di più al nucleo, la repulsione cresce meno rapidamente di quello che darebbe la legge di Coulomb, il che lascia presagire che essa in seguito decresca ed infine finisca per diventare attrazione come si richiede che sia nell'interno del nucleo per assicurarne la stabilità. Il potenziale $U(r)$ sarà quindi dato da una curva dall'andamento rappresentato nella fig. 38 con un massimo U_0 corrispondente alla distanza r_0 alla quale la forza diviene attrattiva: le particelle α che fanno parte del nucleo si trovano entro la regione corrispon-

dente alla « valle » di potenziale. Il caso che più interessa è quello dell'uranio I (U I): per questo elemento si è potuta determinare sperimentalmente la curva solo fino alla distanza di circa $3 \cdot 10^{-12}$ cm. (tratto grosso nella figura), ma questo è sufficiente per rilevare il paradosso a cui si è accennato, che è il seguente. Le particelle α che emette l'uranio I hanno una energia di circa 4 milioni di volt, che è molto minore (circa la metà) di quella che corrisponde al più alto punto, M , sperimentalmente accertato, della curva: la loro

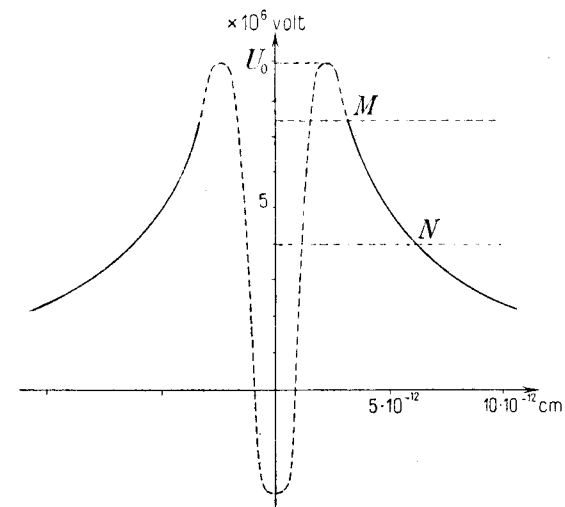


Fig. 38.

energia corrisponde a particelle α che fossero fatte partire, senza velocità iniziale, da una distanza $6 \cdot 10^{-12}$ cm. (punto N della curva). Quindi (anche indipendentemente da ogni ipotesi sulla forma della curva nella parte non sperimentalmente accertata) resta inesplicabile, dal punto di vista classico, come delle particelle abbiano potuto uscire dal nucleo, possedendo un'energia inferiore al massimo, U_0 , della barriera di potenziale che circonda il nucleo.

Dal punto di vista della meccanica ondulatoria invece tale difficoltà non sussiste, perchè, come si è visto, una particella può anche, occasionalmente, attraversare una barriera di potenziale più alta dell'energia che essa possiede.

Le particelle α del nucleo si comporteranno, qualitativamente, come le particelle del problema unidimensionale studiato nel § precedente (da cui questo differisce per avere simmetria sferica, il che non muta nulla di sostanziale) e cioè: staranno *ordinariamente* nell'interno del nucleo, con certi livelli di energia quasi-quantizzati, ma avranno una certa probabilità di uscirne passando (per così dire) attraverso la barriera ma conservando la stessa energia. Ciò è in perfetto accordo col fatto sperimentale che le particelle α emesse da una data sostanza hanno tutte certe energie ben determinate (cioè hanno uno « spettro di righe ») le quali rappresentano appunto i detti livelli energetici (che in pratica risultano esattamente quantizzati, la larghezza delle bande essendo del tutto trascurabile). Una determinazione quantitativa di questi livelli richiederebbe una conoscenza esatta della curva del potenziale, di cui invece si conosce, per la parte interna, solo l'andamento qualitativo: ma fortunatamente questa conoscenza esatta non è necessaria per calcolare il coefficiente di trasmissione della barriera, cioè la probabilità che ha una particella di uscirne, da cui dipende, evidentemente, la vita media dell'elemento radioattivo considerato.

Si è trovata così ⁽¹⁾ una semplice relazione tra la vita media dell'elemento, e la velocità delle particelle α da esso emesse: relazione che era già stata osservata empiricamente da GEIGER e NUTTALL su tutte le sostanze radioattive emettenti particelle α .

CAPITOLO IV.

PROBLEMI A TRE DIMENSIONI

§ 44. — PARTICELLA NON SOGGETTA A FORZE.

In questo capitolo studieremo la meccanica ondulatoria di una particella, togliendo la restrizione del capitolo precedente che tutto dipenda solo da x e da t : interverranno dunque ora tre coordinate spaziali, oltre il tempo.

Il primo problema che tratteremo è quello di una particella libera nello spazio e non soggetta a forze: la sua ψ più generale potrà ottenersi (v. § 29) come somma di « componenti monocromatiche » ognuna corrispondente ad un valore determinato E dell'energia: ricerchiamo dunque queste soluzioni monocromatiche, di frequenza $\nu = E/h$, cioè della forma (128) del § 27.

Riprendiamo l'equazione (131') cui soddisfa la $u(x, y, z)$ e scriviamola esplicitando il Δ e ponendovi $U = 0$.

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + \frac{8\pi^2 m}{h^2} E u = 0 .$$

Questa equazione ben nota si può integrare col metodo della separazione delle variabili, cioè ponendo

$$(204) \quad u(x, y, z) = X(x) \cdot Y(y) \cdot Z(z):$$

difatti l'equazione diviene allora (dividendola tutta per $X Y Z$)

$$\frac{X''}{X} + \frac{Y''}{Y} + \frac{Z''}{Z} + \frac{8\pi^2 m}{h^2} E = 0 .$$

Se si osserva che ciascuno dei tre primi termini dipende da una sola delle coordinate, si riconosce che, affinché l'equa-

⁽¹⁾ v. bibl. n. 29.

zione sia soddisfatta qualunque siano x, y, z , dovrà aversi

$$(205) \quad \frac{X''}{X} = c_1, \quad \frac{Y''}{Y} = c_2, \quad \frac{Z''}{Z} = c_3,$$

con c_1, c_2, c_3 costanti e legate da

$$(206) \quad c_1 + c_2 + c_3 + \frac{8\pi^2 m}{h^2} E = 0.$$

La prima delle (205) ha un integrale generale del tipo

$$X = a_1 e^{\sqrt{c_1} x} + b_1 e^{-\sqrt{c_1} x},$$

il quale, se c_1 fosse positivo, diventerebbe infinito o per $x = +\infty$ o per $x = -\infty$: poichè si richiede invece che la u sia dovunque finita, ne concludiamo che $c_1 \leq 0$, cosicchè nella espressione di X gli esponenti divengono immaginari (o nulli) e la X assume andamento oscillatorio (o costante). Lo stesso potrà ripetersi per Y e Z . Potremo porre dunque (indicando con k_x, k_y, k_z tre nuove costanti),

$$c_1 = -4\pi^2 k_x^2, \quad c_2 = -4\pi^2 k_y^2, \quad c_3 = -4\pi^2 k_z^2,$$

con che la (206) diviene, ponendo $\sqrt{k_x^2 + k_y^2 + k_z^2} = k$,

$$(207) \quad k^2 = \frac{2mE}{h^2}$$

e gli integrali generali delle (205) si scrivono

$$\begin{aligned} X &= a_1 e^{2\pi i k_x x} + b_1 e^{-2\pi i k_x x}, \\ Y &= a_2 e^{2\pi i k_y y} + b_2 e^{-2\pi i k_y y}, \\ Z &= a_3 e^{2\pi i k_z z} + b_3 e^{-2\pi i k_z z}. \end{aligned}$$

Ognuna di queste funzioni è della stessa forma dalla u trovata nel problema unidimensionale (§ 35, form. 149): come si è visto in quel caso, possiamo considerare solo i primi termini (i secondi si ottengono cambiando il segno di k_x, k_y, k_z) ed allora la u assume la forma

$$(208) \quad u = A e^{2\pi i (k_x x + k_y y + k_z z)},$$

e la ψ

$$(209) \quad \psi = A e^{2\pi i (k_x x + k_y y + k_z z - \nu t)},$$

dove (ν , (207))

$$(207') \quad \nu = \frac{1}{h} E = \frac{h}{2m} k^2.$$

Convieni (come al § 15), introdurre il vettore \mathbf{k} di componenti k_x, k_y, k_z (e quindi di modulo k) ed il vettore \mathbf{r} avente l'origine nell'origine degli assi e l'estremo nel punto P dove si calcola ψ , quindi avente le componenti x, y, z : allora la formula precedente si scrive in forma intrinseca, cioè indipendente dagli assi:

$$(210) \quad \psi = A e^{2\pi i (\mathbf{k} \times \mathbf{r} - \nu t)}.$$

Questa formula coincide con la (78') del § 15: come si è visto, essa rappresenta un treno d'onde piane avente per « vettore di propagazione » \mathbf{k} , quindi con $\lambda = 1/k$ e direzione di propagazione data dal vettore \mathbf{k} . Il fenomeno è quindi fisicamente lo stesso studiato al § 35, ma è riferito ora ad assi orientati comunque. Questo treno d'onde rappresenta dunque una particella di impulso \mathbf{p} esattamente determinato ($\mathbf{p} = h\mathbf{k}$), di energia, pure determinata,

$$E = h\nu = \frac{h^2}{2m} k^2,$$

e di posizione totalmente indeterminata. Esprimendo nella (210) \mathbf{k} e ν mediante \mathbf{p} , essa diviene

$$(210') \quad \psi = A e^{\frac{2\pi i}{h} (\mathbf{p} \times \mathbf{r} - \frac{p^2}{2m} t)}.$$

Passiamo ora a considerare le soluzioni più generali, ottenute sovrapponendo infiniti treni d'onde come il precedente con diversi vettori di propagazione, cioè le soluzioni analoghe alla (79) del § 15: esprimendo in questa \mathbf{k} mediante \mathbf{p} e chiamando $\varphi_0(\mathbf{p})$ la funzione $h^{-3/2} A(\mathbf{k})$ (proporzionale all'ampiezza del treno d'onde di impulso \mathbf{p}) si ottiene la seguente espressione per la ψ :

$$(211) \quad \psi = h^{-3/2} \iiint \varphi_0(\mathbf{p}) e^{\frac{2\pi i}{h} (\mathbf{p} \times \mathbf{r} - \frac{p^2}{2m} t)} dp_x dp_y dp_z,$$

ovvero, posto

$$(212) \quad \varphi(\mathbf{p}, t) = \varphi_0(\mathbf{p}) e^{-\frac{2\pi i}{h} \frac{\mathbf{p}^2}{2m} t},$$

$$(213) \quad \psi = h^{-3/2} \iiint \varphi(\mathbf{p}, t) e^{\frac{2\pi i}{h} \mathbf{p} \times \mathbf{r}} dp_x dp_y dp_z.$$

Per $t = 0$ la ψ diventa:

$$(213') \quad \psi_0 = h^{-3/2} \iiint \varphi_0(\mathbf{p}) e^{\frac{2\pi i}{h} \mathbf{p} \times \mathbf{r}} dp_x dp_y dp_z.$$

Ricordando che la A della (79) è data dalla (80), si vede che φ_0 si ricava dalla ψ iniziale ψ_0 con la formula

$$(214) \quad \varphi_0(\mathbf{p}) = h^{-3/2} \iiint \psi_0(\mathbf{r}) e^{-\frac{2\pi i}{h} \mathbf{p} \times \mathbf{r}} dx dy dz,$$

e quindi φ da ψ mediante

$$(215) \quad \varphi(\mathbf{p}, t) = h^{-3/2} \iiint \psi(\mathbf{r}, t) e^{-\frac{2\pi i}{h} \mathbf{p} \times \mathbf{r}} dx dy dz.$$

Si noti l'analogia tra le formule (213) e (215), che si possono considerare inverse l'una dell'altra, e nelle quali le funzioni φ e ψ hanno parti simmetriche.

Ricordando ora il principio di sovrapposizione, possiamo interpretare nel modo seguente la soluzione (213): quando lo stato della particella è rappresentato dalla (213), se si esegue una determinazione dell'impulso vi è probabilità

$$|\varphi_0|^2 dp_x dp_y dp_z = |\varphi|^2 dp_x dp_y dp_z$$

di trovare le componenti dell'impulso comprese tra p_x e $p_x + dp_x$, p_y e $p_y + dp_y$, p_z e $p_z + dp_z$. La funzione $\varphi(\mathbf{p})$ ha dunque, rispetto alle misure di impulso, lo stesso significato che ha la ψ rispetto alle misure di posizione.

È superfluo rilevare che se la ψ_0 data da (213') è diversa da zero solo in una regione limitata dello spazio, essa rappresenta un pacchetto d'onde che si muove approssimativamente di moto rettilineo uniforme, cioè come la particella nel caso classico: però sparpagliandosi gradatamente nelle

tre dimensioni, come si vede generalizzando il calcolo del § 36.

Si può verificare, anche in questo caso, che la meccanica ondulatoria contiene in sé il principio di indeterminazione (generalizzando così quanto si è detto al § 36 per il problema unidimensionale): basta riprendere le (83) del cap. I ed esprimervi \mathbf{k} mediante \mathbf{p} , e si ottiene:

$$(216) \quad \begin{cases} \Delta x \Delta p_x \geq \frac{h}{4\pi}, \\ \Delta y \Delta p_y \geq \frac{h}{4\pi}, \\ \Delta z \Delta p_z \geq \frac{h}{4\pi}. \end{cases}$$

espressione del principio di indeterminazione per una particella nello spazio.

§ 45. - PARTICELLA IN UNA SCATOLA PARALLELEPIPEDA.

Analogamente a quanto abbiamo fatto nel caso unidimensionale (§ 38), possiamo ora considerare brevemente il caso di una particella vincolata a restare entro una scatola parallelepipeda, di lati a , b , c , a pareti perfettamente elastiche: nell'interno della scatola non agiscono forze. Ricercheremo dapprima una soluzione semplice, corrispondente ad un dato valore di E , ossia ad una sola frequenza, lasciandoci guidare dall'analogia col problema delle onde luminose o sonore entro una scatola a pareti riflettenti.

È chiaro che un sistema di onde piane propagantesi nella direzione di coseni α , β , γ , darà luogo, riflettendosi sulle pareti, a sistemi di onde caratterizzati dai coseni $\pm \alpha$, $\pm \beta$, $\pm \gamma$ (dove i segni si possono combinare in tutti gli otto modi possibili): perciò una soluzione come la (209) non potrà sussistere sola, ma dovremo in sua vece prendere una somma di otto termini del tipo:

$$(217) \quad \frac{1}{b c} e^{-\frac{2\pi i}{h} E t} e^{2\pi i(\pm k_x x \pm k_y y \pm k_z z)}.$$

Ora sulle pareti deve essere $u = 0$ per qualunque t : ciò porta ad una limitazione dell'arbitrarietà del vettore di pro-

pagazione k . Difatti, supponiamo che due pareti corrispondano ai piani $x = 0$, $x = a$: siccome la somma anzidetta si può scrivere nella forma

$$u = f_1(y, z, t) e^{2\pi i k_x x} + f_2(y, z, t) e^{-2\pi i k_x x},$$

si vede subito che, affinchè sia $u = 0$ per $x = 0$ e per $x = a$ (qualunque siano y, z, t), deve essere $2\pi k_x a = n_1 \pi$, con n_1 intero; e similmente per k_y e k_z : quindi

$$(218) \quad k = \frac{1}{2} \sqrt{\frac{n_1^2}{a^2} + \frac{n_2^2}{b^2} + \frac{n_3^2}{c^2}},$$

e poichè $k = 1/\lambda$, si conclude che le sole lunghezze d'onda che possono dar luogo ad onde stazionarie sono quelle esprimibili con la formola

$$(219) \quad \lambda_{n_1, n_2, n_3} = \frac{2}{\sqrt{\frac{n_1^2}{a^2} + \frac{n_2^2}{b^2} + \frac{n_3^2}{c^2}}}$$

ed i soli valori possibili per la E sono quindi quelli dati da

$$(220) \quad E_{n_1, n_2, n_3} = \frac{h^2}{8m} \left(\frac{n_1^2}{a^2} + \frac{n_2^2}{b^2} + \frac{n_3^2}{c^2} \right).$$

Il sistema ha quindi livelli energetici discreti, in tripla infinità, caratterizzati dai tre numeri quantici n_1, n_2, n_3 .

Analogamente, in acustica si trova, come è ben noto, che una cavità rettangolare può risuonare solo per certe determinate frequenze (*frequenze proprie*), cui corrispondono lunghezze d'onda date proprio dalla (219).

Nel caso più generale di una particella rinchiusa in una cavità di forma qualunque, il problema è sempre analogo a quello acustico della determinazione delle *frequenze proprie*, o di risonanza, per le quali si possono stabilire onde stazionarie nell'interno della cavità, e conduce sempre a livelli energetici discreti.

§ 46. - FORZE CENTRALI (PARTE GENERALE).

Consideriamo ora una particella soggetta ad una forza centrale: converrà evidentemente servirsi di coordinate po-

lari r, θ, φ , aventi il polo nel centro di forza, ed il potenziale U risulterà funzione della sola r .

L'equazione di Schrödinger (127), esplicitando l'operatore Δ in coordinate polari, si scrive

$$(221) \quad \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial u}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial u}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 u}{\partial \varphi^2} + \frac{8\pi^2 m}{h^2} (E - U) u = 0.$$

Cerchiamo di separare la variabile r da θ e da φ ponendo

$$(222) \quad u = R(r) Y(\theta, \varphi).$$

Sostituendo nell'equazione precedente e moltiplicandola tutta per $\frac{r^2}{R Y}$ si ottiene

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{1}{Y \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{Y \sin^2 \theta} \frac{\partial^2 Y}{\partial \varphi^2} + \frac{8\pi^2 m}{h^2} r^2 (E - U) = 0.$$

In questa equazione il primo e l'ultimo termine dipendono solo da r , gli altri due solo da θ e da φ : quindi l'equazione si spezza nelle due

$$(223) \quad \frac{1}{Y \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{Y \sin^2 \theta} \frac{\partial^2 Y}{\partial \varphi^2} = -C$$

$$(224) \quad \frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{8\pi^2 m}{h^2} r^2 (E - U) = C,$$

dove C è una costante. La prima di queste non contiene la funzione $U(r)$, e quindi è comune a tutti i problemi di forze centrali: nella (222) dunque il fattore $Y(\theta, \varphi)$ è indipendente dalla legge con cui varia la forza in funzione di r , mentre il fattore $R(r)$ dipende da questa legge ed è diverso quindi nei vari problemi. Studieremo in questo § l'equazione (223) che vale per tutti i problemi di forze centrali: essa si può scrivere

$$(223') \quad \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 Y}{\partial \varphi^2} + C Y = 0.$$

e si può dimostrare che ha soluzioni finite, continue e ad un sol valore per ogni direzione, solo se,

$$(225) \quad C = l(l+1),$$

con $l = 0, 1, 2, \dots$ (l'intero l chiamasi «quanto azimutale» perchè corrisponde al quanto azimutale della teoria di Bohr e Sommerfeld). Con questa espressione di C , si riconosce nella (223') l'equazione differenziale delle funzioni sferiche di superficie di ordine l .

A ciò si giunge separando ulteriormente le variabili θ e φ con la posizione

$$(226) \quad Y(\theta, \varphi) = \Theta(\theta) \cdot \Phi(\varphi),$$

con che l'equazione si spezza nelle due

$$(227) \quad \frac{d^2\Phi}{d\varphi^2} + \lambda\Phi = 0,$$

$$(228) \quad \frac{1}{\sin\theta} \frac{d}{d\theta} \left(\sin\theta \frac{d\Theta}{d\theta} \right) + \left(C - \frac{\lambda}{\sin^2\theta} \right) \Theta = 0,$$

dove λ è una costante.

La prima di queste dà

$$(229) \quad \Phi = \cos t e^{\pm i\sqrt{\lambda}\varphi},$$

e poichè la Φ deve essere periodica a periodo 2π nella φ (altrimenti la u non risulterebbe ad un sol valore per ogni punto dello spazio), $\sqrt{\lambda}$ dovrà essere uguale ad un numero intero ⁽¹⁾ m , cioè

$$(230) \quad \lambda = m^2;$$

determinando poi la costante in modo che risulti

$$(231) \quad \int_0^{2\pi} |\Phi|^2 d\varphi = 1,$$

la (229) diviene

$$(229') \quad \Phi = \frac{1}{\sqrt{2\pi}} e^{im\varphi}.$$

⁽¹⁾ Adoperiamo questa lettera per conformarci all'uso ormai universale, sebbene m indichi pure la massa della particella.

dove si è scritto l'esponente senza doppio segno, con l'intesa che m possa essere positivo, nullo o negativo: l'intero m (col segno) si chiama «quanto magnetico».

Per studiare poi la (228) conviene assumere come variabile cos θ , che indicheremo con x : allora l'equazione diviene (tenendo conto della (230)),

$$(232) \quad (1-x^2) \frac{d^2\Theta}{dx^2} - 2x \frac{d\Theta}{dx} + \left(C - \frac{m^2}{1-x^2} \right) \Theta = 0.$$

Gli autovalori e le autofunzioni di questa equazione si studiano con un metodo analogo a quello seguito nel § 39 per l'oscillatore: punti singolari dell'equazione sono $x = \pm 1$ e $x = \infty$, come si riconosce subito dividendola per $1-x^2$ (v. § 16), e si riconosce anche facilmente che si tratta di singolarità fuchsiane. Poichè nel caso nostro x va da -1 a $+1$, ci interessano solo le singolarità agli estremi di questo intervallo: col metodo dell'equazione caratteristica troviamo che in vicinanza di $x = +1$ la Θ può mettersi nella forma

$$(1-x)^{\pm \frac{m}{2}} f_1(1-x),$$

dove f_1 è una funzione olomorfa che non si annulla per $1-x=0$: di queste due forme, quella con esponente negativo ha un polo

per $1-x=0$ e quindi va scartata: perciò $\Theta = (1-x)^{\frac{|m|}{2}} f_1(1-x)$.

Similmente l'altro punto singolare ci dà $\Theta = (1+x)^{\frac{|m|}{2}} f_2(1+x)$, con f_2 olomorfa e non nulla in $1+x=0$. Poichè queste due espressioni devono rappresentare due sviluppi diversi dello stesso ramo della funzione continua Θ , dovrà potersi scrivere

$$(233) \quad \Theta = (1-x^2)^{\frac{|m|}{2}} P(x),$$

con

$$(234) \quad P(x) = \sum_0^{\infty} a_r x^r.$$

Sostituendo la (233) nella (232) si trova per P l'equazione

$$(235) \quad (1-x^2) P'' - 2(|m|+1)xP' + [C - |m|(|m|+1)]P = 0,$$

e cercando di soddisfare questa con la serie (234), si trova per le a_r la formula ricorrente

$$(236) \quad (r+1)(r+2)a_{r+2} = [r(r-1) + 2(|m|+1)r - C + |m|(|m|+1)]a_r.$$

Si può quindi prendere per P una serie di potenze pari o di potenze dispari, col primo coefficiente arbitrario.

La condizione che P sia finita in tutto l'intervallo considerato è poi certo soddisfatta se uno dei coefficienti, p. es. $a_{\gamma+2}$, si annulla (senza che si annulli il precedente), poichè allora si annullano anche tutti i successivi e la serie P si riduce ad un polinomio di grado γ : la condizione perchè $a_{\gamma+2} = 0$ (essendo $a_\gamma \neq 0$) è, come si vede dalla (235), che sia

$$C = \gamma(\gamma - 1) + 2(|m| + 1)\gamma + |m|(|m| + 1) = (\gamma + |m|)(\gamma + |m| + 1),$$

e se si pone

$$(237) \quad \gamma + |m| = l$$

(con che l risulta intero e non negativo), si ha la (225).

Si può poi anche dimostrare che questa condizione è non solo sufficiente ma anche necessaria ⁽¹⁾, cioè che se la P non si riduce a un polinomio essa non può soddisfare le condizioni volute.

Polinomi di Legendre. — Consideriamo l'equazione (235) dapprima nel caso di $m = 0$, nel qual caso la Y non dipende da φ e si identifica (a meno di un fattore costante) con la Θ , e questa con la P . L'equazione si scrive, tenuto conto della (225),

$$(238) \quad (1 - x^2)P'' - 2xP' + l(l + 1)P = 0,$$

e la soluzione cercata è un polinomio $P_l(x)$ di grado l , soddisfacente alla formula ricorrente (236). Tali polinomi sono stati introdotti da tempo nella teoria delle funzioni sferiche, e si definiscono così: si prende $a_1 = 0$ se l è pari e $a_0 = 0$ se l è dispari, in modo che il polinomio risulti nel primo caso di sole potenze pari e nel secondo caso di sole potenze dispari: l'altro coefficiente poi si usa prenderlo in modo che risulti $P_l(1) = 1$. Il polinomio $P_l(x)$ così definito si chiama « *polinomio di Legendre* » di grado l od anche, quando lo si esprime nella forma $P(\cos \theta)$, « *funzione sferica zonale* » (difatti, per ciò che si è detto sopra, esso è un caso particolare di funzione sferica, avente simmetria assiale). Ecco le espressioni dei primi sei polinomi di Legendre:

$$\begin{array}{ll} P_0(x) = 1 & P_1(x) = x \\ P_2(x) = \frac{3}{2}x^2 - \frac{1}{2} & P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x \\ P_4(x) = \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8} & P_5(x) = \frac{63}{8}x^5 - \frac{35}{4}x^3 + \frac{15}{8}x \end{array}$$

⁽¹⁾ V. BECHERT, *Ann. d. Phys.*, **83**, 906 (1927).

Essi costituiscono (essendo autofunzioni della (238)) un sistema di funzioni ortogonali nell'intervallo $(-1, +1)$: non sono però normalizzati, perchè

$$(239) \quad \int_{-1}^{+1} P_l^2(x) dx = \frac{2}{2l+1}.$$

Essi possono anche venir definiti mediante la derivata l -esima dell'espressione $(x^2 - 1)^l$: difatti si ha

$$(240) \quad P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l.$$

Un'altra loro proprietà notevole è espressa dalla formula ricorrente che lega tre polinomi successivi:

$$(241) \quad P_l = \frac{2l-1}{l} x P_{l-1} - \frac{l-1}{l} P_{l-2},$$

la quale permette di calcolare successivamente tutti i polinomi di Legendre a partire da P_0 e da P_1 .

Funzioni associate di Legendre. — Passiamo ora a considerare la (235) senza la restrizione $m = 0$: essa si scrive, tenuto conto della (225),

$$(242) \quad (1 - x^2)P'' - 2(|m| + 1)xP' + [l(l + 1) - |m|(|m| + 1)]P = 0,$$

con $|m| \leq l$.

Le sue autofunzioni si possono ottenere mediante la formula ricorrente (236), ma si può anche osservare che se si deriva l'ultima equazione rispetto ad x si ottiene un'equazione della stessa forma, in cui al posto di P, P', P'' vi sono P', P'', P''' rispettivamente, ed al posto di $|m|$ vi è $|m| + 1$: cosicchè se P_l soddisfa, come si è visto, la (242) per $|m| = 0$, P'_l la soddisfa per $|m| = 1$, P''_l per $|m| = 2$ e così via: perciò la (242) è soddisfatta da $\frac{d^{|m|}}{dx^{|m|}} P_l$ (che è un polinomio di grado $l - |m| = \gamma$) e quindi la (232) è soddisfatta prendendo Θ uguale (o proporzionale) a

$$P_l^m(x) = (1 - x^2)^{\frac{|m|}{2}} \frac{d^{|m|}}{dx^{|m|}} P_l.$$

Queste funzioni si chiamano « *funzioni associate di Legendre* »: esse sono, naturalmente, ortogonali nell'intervallo $(-1, +1)$, ma non sono normalizzate, poichè

$$\int_{-1}^{+1} (P_l^m)^2 dx = \frac{(l+|m|)!}{(l-|m|)!} \frac{2}{2l+1}.$$

Così, affinché Θ sia anche normalizzata (rispetto alla variabile x), basterà prendere per essa l'espressione

$$(243) \quad \Theta_{lm} = \sqrt{\frac{(l-|m|)!}{(l+|m|)!} \frac{2l+1}{2}} P_l^m(x) = \\ = \sqrt{\frac{(l-|m|)!}{(l+|m|)!} \frac{2l+1}{2}} \sin^{|m|} \theta \frac{d^{|m|} P_l(\cos \theta)}{d(\cos \theta)^{|m|}},$$

e risulterà allora

$$(244) \quad \int_0^\pi \Theta_{lm}^2 \sin \theta \, d\theta = 1.$$

Riassumendo, ad ogni autovalore $C = l(l+1)$ della (223') corrispondono $2l+1$ autofunzioni $Y_{lm}(\theta, \varphi)$ (con $m = -l, -l+1, \dots, 0, 1, \dots, l$), date dalle (226), (229'), (243), cioè da

$$(245) \quad Y_{lm} = \frac{1}{N_{lm}} e^{im\varphi} \sin^{|m|} \theta \frac{d^{|m|} P_l(\cos \theta)}{d(\cos \theta)^{|m|}},$$

dove si è messo in evidenza il fattore di normalizzazione

$$(246) \quad \frac{1}{N_{lm}} = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{(l-|m|)!}{(l+|m|)!} \frac{2l+1}{2}}.$$

La Y_{lm} così normalizzata gode la proprietà, conseguenza di (231) e (244), che

$$(246') \quad \iint |Y_{lm}|^2 \sin \theta \, d\theta \, d\varphi = 1,$$

dove l'integrale è esteso a tutta la superficie sferica.

Tali funzioni sono particolari *funzioni sferiche* (di superficie) di ordine l . Di queste, quella corrispondente a $m = 0$ si riduce a

$$(247) \quad Y_{l0} = \frac{1}{N_{l0}} P_l(\cos \theta),$$

e non dipende da φ , ossia ha simmetria assiale.

Si osservi inoltre che all'autovalore nullo ($l = 0$) corrisponde la sola autofunzione

$$(248) \quad Y_{00} = \frac{1}{N_{00}} P_0 = \frac{1}{N_{00}} = \frac{1}{2\sqrt{\pi}},$$

cioè una costante: la u dunque in tal caso dipende solo da r (simmetria sferica).

Diamo qui, per comodità del lettore, le espressioni esplicite delle funzioni sferiche corrispondenti ai primi 4 valori di l , che più spesso intervengono nella meccanica atomica (diamo ad m solo i valori $0, 1, 2, \dots$: per avere le funzioni corrispondenti a valori negativi di m , non c'è che da cambiare il segno all'esponente di e):

$$Y_{00} = \frac{1}{2\sqrt{\pi}},$$

$$Y_{10} = \sqrt{\frac{3}{4\pi}} \cos \theta, \quad Y_{11} = \sqrt{\frac{3}{8\pi}} \sin \theta e^{i\varphi},$$

$$Y_{20} = \sqrt{\frac{5}{4\pi}} \left(\frac{3}{2} \cos^2 \theta - \frac{1}{2} \right), \quad Y_{21} = \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{i\varphi},$$

$$Y_{22} = \frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{2i\varphi},$$

$$Y_{30} = \sqrt{\frac{7}{4\pi}} \left(\frac{5}{2} \cos^3 \theta - \frac{3}{2} \cos \theta \right),$$

$$Y_{31} = \frac{1}{4} \sqrt{\frac{21}{4\pi}} \sin \theta (5 \cos^2 \theta - 1) e^{i\varphi},$$

$$Y_{32} = \frac{1}{4} \sqrt{\frac{105}{2\pi}} \sin^2 \theta \cos \theta e^{2i\varphi}, \quad Y_{33} = \frac{1}{4} \sqrt{\frac{35}{4\pi}} \sin^3 \theta e^{3i\varphi}.$$

Aggiungeremo poi, anticipando un risultato che verrà dimostrato nella parte III, che il quanto azimutale l ed il quanto magnetico m hanno il seguente significato fisico: il modulo p del momento angolare dell'elettrone (momento dell'impulso) rispetto al nucleo è $p = \sqrt{l(l+1)} \frac{h}{2\pi}$, ed il momento dell'impulso dell'elettrone rispetto all'asse z è $m \frac{h}{2\pi}$.

§ 47. - L'EQUAZIONE RADIALE NEL PROBLEMA DELLE FORZE CENTRALI.

Occupiamoci ora del fattore $R(r)$ della (222), che dipende dalla legge della forza. Esso soddisfa l'equazione (224) dove

per C va posto $l(l+1)$: inoltre si può osservare che il primo termine si può scrivere (come è ben noto) in un'altra forma, poichè

$$\frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) = r \frac{d^2}{dr^2} (rR),$$

cosicchè, ponendo

$$(249) \quad y(r) = rR(r),$$

l'equazione si scrive

$$(250) \quad \frac{d^2 y}{dr^2} + \left[\frac{8\pi^2 m}{h^2} (E - U) - \frac{l(l+1)}{r^2} \right] y = 0.$$

Alla y va inoltre imposta la condizione che per $r = 0$ la R si mantenga finita, quindi

$$y(0) = 0,$$

e che per $r \rightarrow \infty$ la R tenda a zero non meno rapidamente di $\frac{1}{r}$, e quindi la y tenda ad un limite finito od a zero.

Queste condizioni conducono (quando sia specificato il potenziale $U(r)$) a determinare per la E una successione di autovalori (in generale, parte discreti e parte continui) che rappresentano i livelli energetici del sistema: nel caso che siano discreti, si potranno caratterizzare con un indice n («quanto totale»), oltre che con l'indice l che già figura nella (250), cosicchè scriveremo E_{nl} , e per le autofunzioni corrispondenti R_{nl} . Per quanto concerne la normalizzazione di queste, si osservi che la $u(r, \theta, \varphi)$ deve essere normalizzata in modo che risulti

$$(251) \quad \int_0^\infty dr \int_0^\pi d\theta \int_0^{2\pi} |u|^2 r^2 \sin \theta d\varphi = 1$$

e, poichè $u = R(r)\Theta(\theta)\Phi(\varphi)$, e i due ultimi fattori si sono già normalizzati conformemente alle (231) e (244), risulta che R dovrà essere normalizzato in modo che sia

$$(252) \quad \int_0^\infty R^2(r) r^2 dr = 1.$$

Si osservi che, sebbene la u venga a dipendere da tre numeri quantici (n, l, m) i livelli energetici dipendono da due soli di essi, poichè il quanto equatoriale m non vi figura: ciascuno di essi dunque è multiplo di ordine $2l+1$ (degenerazione, v. § 6), poichè vi corrispondono altrettante soluzioni indipendenti per la u , corrispondenti ai valori che può assumere m (da $-l$ a $+l$).

Convieni infine rilevare che nella (250) si può conglobare nel potenziale anche il termine $\frac{l(l+1)}{r^2}$, considerando come potenziale la funzione

$$(253) \quad U_1(r) = U(r) + \frac{h^2}{8\pi^2 m} \frac{l(l+1)}{r^2},$$

con che l'equazione assume la forma

$$(254) \quad \frac{d^2 y}{dr^2} + \frac{8\pi^2 m}{h^2} (E - U_1) y = 0,$$

identica a quella del moto unidimensionale sotto l'azione di un potenziale U_1 . Ciò può interpretarsi formalmente dicendo che alla forza derivante dal potenziale U si aggiunge quella derivante dal secondo termine della (253), la quale trova il suo analogo nella forza centrifuga della meccanica classica, e, come quella, dipende dal momento angolare, cioè dal quanto azimutale l :

§ 48. — TEORIA DEI SISTEMI IDROGENOIDI. GLI AUTOVALORI.

Applicheremo ora la teoria delle forze centrali al caso dell'atomo di idrogeno e degli altri sistemi detti «idrogenoidi», formati (v. § 14, p. I) da un nucleo di carica Ze e da un solo elettrone (il caso dell'idrogeno corrisponde a $Z = 1$). Data la massa preponderante del nucleo, possiamo considerare questo come fisso ⁽¹⁾ e perciò considerare l'elettrone soggetto ad una forza centrale attrattiva uguale a $\frac{Ze^2}{r^2}$

⁽¹⁾ La trattazione più rigorosa, che tien conto del movimento del nucleo, sarà fatta nella parte III, § 21.

e, sostituendovi la (263) e uguagliando a zero i coefficienti di ciascuna potenza, si trova per le a_s la formula ricorrente

$$(265) \quad a_{s+1} = a_s \frac{s+l+1-A}{(s+1)(s+2l+2)} .$$

Affinchè la serie si riduca ad un polinomio (di cui indicheremo il grado con n') occorre che sia $a_{n'+1} = 0$, $a_{n'} \neq 0$: quindi che

$$n' + l + 1 - A = 0 ,$$

da cui si ricava che A deve essere uguale al numero intero

$$(266) \quad n = n' + l + 1 ,$$

e ricavando $|E|$ dalla (259), si trova infine che, se E è negativo, esso deve avere uno degli autovalori

$$(267) \quad E_n = - \frac{2\pi^2 m Z^2 e^4}{h^2 n^2} .$$

Sono questi gli stessi livelli energetici dati dalla teoria di Bohr (v. § 16, p. I) in perfetto accordo, come si è visto, con l'esperienza: l'intero n corrisponde al quanto totale, e l'energia dipende solo da esso.

Si osservi che con la (267) l'espressione (256) di r_0 diviene

$$(268) \quad r_0 = n a_1 , \quad \text{dove } a_1 = \frac{h^2}{4\pi^2 m Z e^2} :$$

è utile rilevare che la costante a_1 così definita è uguale al raggio della prima delle orbite circolari fornite dalla teoria di Bohr ($= 0,528 \cdot 10^{-8}$ cm.).

Caso di $E > 0$.

In questo caso l'equazione (258), per x tendente a ∞ , tende alla forma

$$(269) \quad \frac{d^2 y}{dx^2} + \frac{1}{4} y = 0 ,$$

il cui integrale generale è

$$c_1 e^{ix/2} + c_2 e^{-ix/2} ,$$

ed è sempre finito per $x \rightarrow \infty$. Quindi qualunque integrale

della (258) si manterrà finito per $x \rightarrow \infty$: perciò non si è costretti ad imporre alla A alcuna limitazione, ossia, qualunque valore (positivo) di A è un autovalore della (258), e quindi, qualunque valore (positivo) di E è un autovalore della (255) (*spettro continuo di autovalori*). Il fatto che l'energia possa assumere qualunque valore positivo è in perfetta corrispondenza a ciò che si trova nella teoria di Bohr e Sommerfeld: infatti i valori positivi di E corrispondono ad orbite iperboliche, quindi a movimenti non periodici nè moltiplicamente periodici, e perciò non quantizzabili ⁽¹⁾.

Se si opera come nel caso precedente, ponendo

$$(270) \quad y = e^{\pm ix/2} v(x) ,$$

si è condotti per v all'equazione

$$(271) \quad v'' \pm iv' + \frac{A - l(l+1)}{x^2} v = 0 ,$$

e ponendo, come precedentemente,

$$(272) \quad v = x^{l+1} \omega , \quad \omega = \sum_0^\infty a_s x^s ,$$

si trova per le a_s la formula ricorrente,

$$(273) \quad a_{s+1} = a_s \frac{-A \mp i(l+s+1)}{(2l+s+2)(s+1)} .$$

Come si vede, essendo A reale, è impossibile che si annulli il numeratore di questa frazione, e perciò la serie non si riduce mai ad un polinomio: la soluzione generale è una combinazione lineare delle due contenute nella formula (270), e precisamente

$$(274) \quad y = e^{ix/2} x^{l+1} \sum_0^\infty a_s x^s + e^{-ix/2} x^{l+1} \sum_0^\infty a_s^* x^s :$$

(i coefficienti della seconda sommatoria sono i coniugati di quelli della prima, cosicchè la y risulta reale): si verifica facilmente che le serie convergono assolutamente per qualunque valore di x .

⁽¹⁾ Si può anzi dire che tali stati non rappresentano un vero atomo di idrogeno, ma solo l'insieme di un elettrone e di un nucleo che, dopo essersi avvicinati alquanto, si allontanano definitivamente.

§ 49. — LE AUTOFUNZIONI DEI SISTEMI IDROGENOIDI PER $E < 0$.

Occupiamoci ora della forma delle autofunzioni nel caso $E < 0$ (che è il più interessante perchè corrisponde ad un sistema stabile): poichè il fattore dipendente da θ e da φ (cioè la funzione sferica Y_{lm}) è già stato discusso al § 46, essendo comune a tutti i problemi di forze centrali, ci resta da esaminare il fattore $R(r)$. Ricapitolando le successive posizioni (249), (257), (260), (262) avremo

$$R(r) = \frac{y(r)}{r} = \frac{2}{r_0} e^{-x/2} v(x) = \frac{2}{r_0} e^{-x/2} x^l \omega(x),$$

dove $\omega(x)$ è un polinomio di grado n' soddisfacente l'equazione differenziale (264), che scriveremo ora nella forma

$$(264') \quad x\omega'' + (2l + 2 - x)\omega' + n'\omega = 0,$$

e che è determinato (una volta fissato a piacere il primo coefficiente a_0) dalla formula ricorrente (265). Lo studio di tali polinomi è facilitato dal fatto che essi sono legati ad una classe di funzioni studiate da tempo e dotate di notevoli proprietà: i *polinomi di Laguerre*.

Cenno sui polinomi di Laguerre. — Il polinomio di Laguerre di grado K , che si indica con $L_K(x)$, è definito mediante la formula

$$(275) \quad L_K(x) = e^x \frac{d^K}{dx^K} (x^K e^{-x}).$$

Ecco l'espressione esplicita dei primi polinomi di Laguerre:

$$\begin{aligned} L_0 &= 1 \\ L_1 &= 1 - x \\ L_2 &= 2 - 4x + x^2 \\ L_3 &= 6 - 18x + 9x^2 - x^3 \\ &\dots \end{aligned}$$

L'interesse di queste funzioni sta nel fatto che esse sono soluzioni di una notevole equazione differenziale, come può vedersi nel modo seguente. Posto

$$(276) \quad u(x) = x^K e^{-x},$$

talchè

$$(275') \quad L_K = e^x u^{(K)}$$

(dove $u^{(K)}$ denota, come faremo sistematicamente, la derivata K -esima di u), si ha, con una prima derivazione della (276),

$$xu^{(1)} = (K - x)u.$$

Derivando questa relazione ancora $K + 1$ volte (per il che giova osservare che la nota formula di Leibnitz per la derivata n -esima di un prodotto, nel caso in cui questo è della forma $x\varphi(x)$, si riduce a $\frac{d^n}{dx^n} (x\varphi) = x\varphi^{(n)} + n\varphi^{(n-1)}$), si ottiene per la u la seguente equazione differenziale:

$$xu^{(K+2)} + (1+x)u^{(K+1)} + (K+1)u^{(K)} = 0.$$

Sostituendovi la $u^{(K)}$ ricavata da (275'), si ottiene l'equazione caratteristica del K -esimo polinomio di Laguerre:

$$(277) \quad xL''_K + (1-x)L'_K + KL_K = 0.$$

Si osservi che i polinomi di Laguerre non sono autofunzioni di questa equazione, nè sono ortogonali: però godono la proprietà (che si può dimostrare ⁽¹⁾ con successive integrazioni per parti, utilizzando la (275'))

$$(278) \quad \int_0^\infty L_K(x) L_{K'}(x) e^{-x} dx = \delta_{KK'},$$

che si può esprimere così: le funzioni $e^{x/2} L_K$ costituiscono un sistema ortogonale e normalizzato nell'intervallo da 0 a ∞ .

Un'altra proprietà notevole dei polinomi di Laguerre, che ci limitiamo ad enunciare, è quella espressa dalla formula ricorrente

$$(279) \quad L_{K+1} + (2K + 1 - x)L_K + K^2 L_{K-1} = 0.$$

Polinomi generalizzati di Laguerre. — Se si deriva l'equazione (277), si ottiene

$$(280) \quad xL'''_K + (2-x)L''_K + (K-1)L'_K = 0$$

⁽¹⁾ Per questa ed altre proprietà dei polinomi di Laguerre, v. p. es. bibl. n. 25 o n. 34.

che si può considerare come un'equazione di secondo ordine nella funzione L_K , cioè $L_K^{(1)}$: questa dunque soddisfa l'equazione

$$xL_K^{(1)''} + (2-x)L_K^{(1)'} + (K-1)L_K^{(1)} = 0,$$

che è analoga alla (277), salvochè il secondo coefficiente è aumentato di 1, ed il terzo è diminuito di 1.

A questa equazione si può applicare ancora lo stesso procedimento, e così si riconosce, derivando j volte, che la funzione $L_K^{(j)}$, cioè la derivata j -esima di L_K , soddisfa l'equazione

$$(281) \quad xL_K^{(j)''} + (j+1-x)L_K^{(j)'} + (K-j)L_K^{(j)} = 0.$$

Questa funzione è evidentemente un polinomio di grado $K-j$, e si chiama talvolta « *polinomio generalizzato di Laguerre* »: esso è definito da

$$(282) \quad L_K^{(j)} = \frac{d^j}{dx^j} \left[e^x \frac{d^K}{dx^K} (x^K e^{-x}) \right].$$

Si noti che per $j \geq K$ il polinomio $L_K^{(j)}$ è identicamente nullo.

I polinomi generalizzati corrispondenti ad un dato indice superiore j ed a diversi K , se moltiplicati per $x^{j/2} e^{-x/2}$ danno luogo a funzioni ortogonali nell'intervallo da 0 a ∞ : si ha precisamente

$$(282) \quad \int_0^\infty L_K^{(j)}(x) L_{K'}^{(j)}(x) x^j e^{-x} dx = \begin{cases} 0 & \text{per } K \neq K' \\ \frac{(K!)^2}{(K-j)!} & \text{per } K = K'. \end{cases}$$

Si osservi ora che l'equazione (264') cui soddisfa ω si identifica con l'equazione (281) dei polinomi generalizzati di Laguerre, purchè si prenda

$$j+1 = 2l+2, \quad K-j = n',$$

cioè

$$j = 2l+1, \quad K = 2l+1+n' = n+l.$$

D'altra parte anche ω è, al pari di $L_{n+l}^{(2l+1)}$, un polinomio di grado n' : quindi essi possono differire al più per un fattore costante, che indicheremo con N'_{nl} . Scriveremo dunque, mettendo in evidenza che ω e quindi R dipendono dagli indici n ed l ,

$$(283) \quad \omega_{nl} = \frac{1}{N'_{nl}} L_{n+l}^{(2l+1)}.$$

$$(284) \quad R_{nl} = \frac{2}{r_0 N'_{nl}} e^{-x/2} x^l L_{n+l}^{(2l+1)}(x).$$

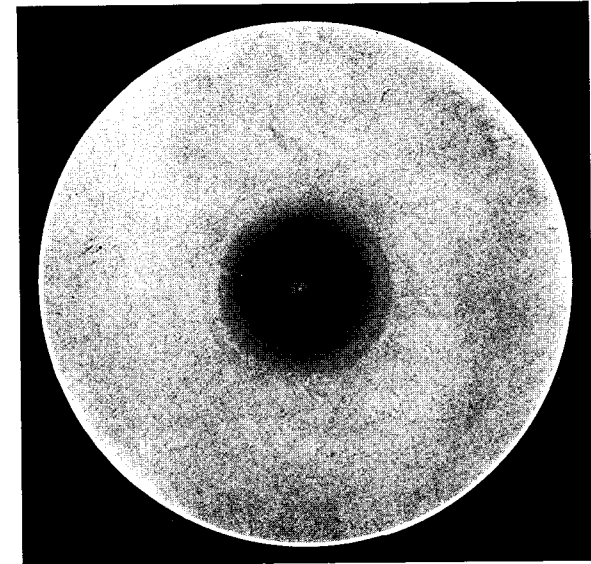


Fig. 39 a.

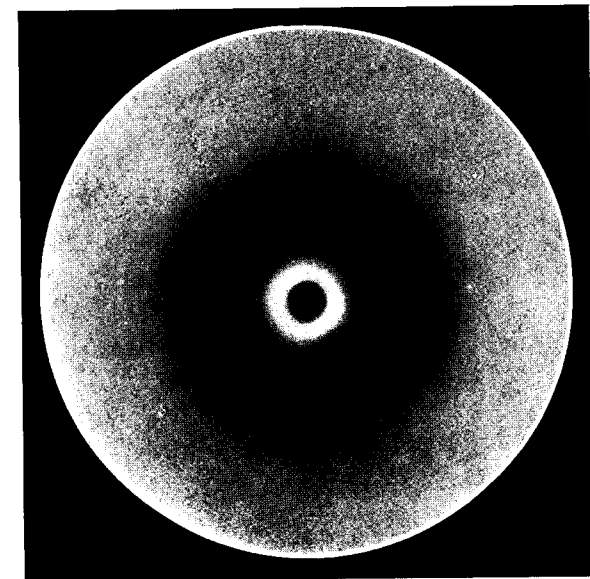


Fig. 39 b.

Il fattore N'_{nl} si determina con la condizione di normalizzazione (252): si trova

$$N'_{nl}{}^2 = r_0 \frac{n [(n+l)!]^3}{(n-l-1)!},$$

cosicchè l'espressione esplicita di R_{nl} in funzione di r è la seguente (dove si è posta per r_0 l'espressione (268), per mettere in evidenza la sua dipendenza da n):

$$(285) \quad R_{nl}(r) = \sqrt{\frac{4(n-l-1)!}{a^3 n^4 [(n+l)!]^3}} e^{-\frac{r}{na}} \left(\frac{2r}{na}\right)^l L_{n+l}^{(2l+1)}\left(\frac{2r}{na}\right).$$

Le espressioni esplicite corrispondenti ai primi valori di n ed l sono, posto $\varrho = r/a$, le seguenti:

$$\begin{aligned} R_{10} &= \frac{2}{a^{3/2}} e^{-\varrho}, \\ R_{20} &= \frac{1}{\sqrt{2} a^{3/2}} e^{-\varrho/2} \left(1 - \frac{1}{2} \varrho\right), \\ R_{21} &= \frac{1}{2\sqrt{6} a^{3/2}} e^{-\varrho/2} \varrho, \\ R_{30} &= \frac{2}{3\sqrt{3} a^{3/2}} e^{-\varrho/3} \left(1 - \frac{2}{3} \varrho + \frac{2}{27} \varrho^2\right), \\ R_{31} &= \frac{8}{27\sqrt{6} a^{3/2}} e^{-\varrho/3} \varrho \left(1 - \frac{1}{6} \varrho\right), \\ R_{32} &= \frac{4}{81\sqrt{30} a^{3/2}} e^{-\varrho/3} \varrho^2. \end{aligned}$$

La (285), insieme con l'espressione già trovata (v. form. 245 e 246) per Y_{lm} , ci permette di scrivere l'espressione completa dell'autofunzione corrispondente ai numeri quantici n, l, m :

$$(286) \quad u_{nlm} = \sqrt{\frac{(l-|m|)! (n-l-1)! (2l+1)}{\pi a^3 n^4 \cdot (l+|m|)! [(n+l)!]^3}} \times \\ \times e^{im\varphi} \sin^{|m|} \theta \frac{d^{|m|} P_l(\cos \theta)}{(d \cos \theta)^{|m|}} e^{-\frac{r}{an}} \left(\frac{2r}{na}\right)^l L_{n+l}^{(2l+1)}\left(\frac{2r}{na}\right).$$

Osserviamo che due autofunzioni corrispondenti a valori di m uguali e di segno contrario differiscono solo per il segno dell'esponente $im\varphi$, e quindi sono complesse coniugate.

È interessante farsi, mediante le formule precedenti, un'idea intuitiva del modo come è distribuita intorno al nucleo la funzione uu^* cioè la « nuvola di probabilità » dell'elettrone.

Ricordiamo anzitutto (v. form. 248) che negli stati in cui $l = 0$ (*stati s*) la u non dipende da θ e da φ , ma solo da r : si ha dunque una nuvola a simmetria sferica. Tale è, in particolare, lo stato corrispondente al livello energetico più basso, o stato fondamentale: infatti $n = 1$ porta necessariamente $l = 0$ (ed $m = 0$).

In tal caso si ha dalla espressione di R_{10} e dalla (248):

$$u = \frac{1}{\sqrt{\pi a^3}} e^{-r/a}$$

e quindi

$$|u|^2 = \frac{1}{\pi a^3} e^{-2r/a}$$

La densità della nuvola sfuma esponenzialmente verso l'infinito: essa è rappresentata dalla fig. 39a, in cui l'annerimento in ciascun punto è presso a poco proporzionale alla densità di probabilità $|u|^2$ di trovare ivi l'elettrone. Si può dire dunque che l'atomo non ha un contorno definito ma, a rigore, si estende a tutto lo spazio, poichè vi è una certa probabilità di trovare l'elettrone a qualunque distanza dal nucleo: tuttavia, come si vede, questa probabilità diviene assai piccola appena r diventa uguale a due o tre volte a (p. es., la probabilità di trovare l'elettrone fuori della sfera di raggio $3a$ risulta circa 0,062). In questo senso, si può dire che le dimensioni dell'atomo nello

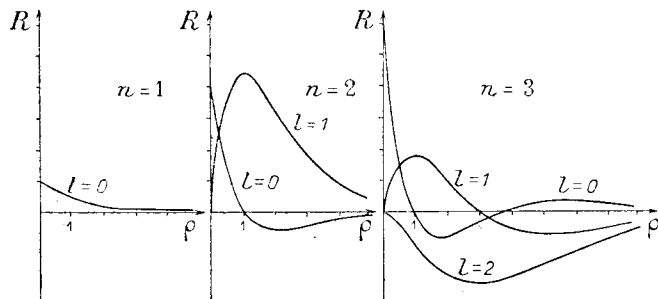


Fig. 40.

stato fondamentale sono dell'ordine di a , cioè dello stesso ordine previsto dalla teoria di Bohr (raggio della prima orbita circolare).

Analoga considerazione può farsi per gli altri stati. Quelli

per cui $l \neq 0$ non presentano simmetria sferica: tuttavia anche per essi ha interesse la considerazione della $R(r)$ e più ancora della funzione $R^2(r)$ poichè, fissata una semiretta uscente dal nucleo (cioè fissati θ e φ), la densità varia lungo tale semiretta proporzionalmente a R^2 : nella fig. 40 è disegnata la curva $R(r)$ per gli stati $n = 1$, $n = 2$, $n = 3$. Ha interesse anche la funzione $z(r) = R^2(r) \cdot r^2$, il cui significato è il seguente: la probabilità che l'elettrone si trovi compreso nello strato sferico infinitesimo di raggio r e spessore dr , è evidentemente

$$dr \iint |u|^2 r^2 \sin\theta d\theta d\varphi$$

e, per la (246'), si riduce a

$$R^2 r^2 dr = z(r) dr.$$

La funzione $z(r)$ è riportata graficamente nella fig. 41 per gli stessi stati della fig. 40.

Vi sono generalmente, come si vede dalle figg. 40 e 41, dei valori di r per cui la $R(r)$ si annulla: essi corrispondono a sfere sulle quali la densità di probabilità è nulla (*sfere nodali*): ne dà un esempio la fig. 39b, che rappresenta la « nuvola di probabilità » per l'idrogeno nello stato $n = 2$, $l = 0$ (stato 2s). Come si vede dall'espressione (285) di R , questa funzione si annulla per $r = 0$ (tranne il caso $l = 0$), per $r = \infty$, ed inoltre in corrispondenza alle radici del polinomio $L_{n+l}^{(2l+1)}$. Come sfere nodali non si contano quella di raggio nullo nè quella di raggio infinito, cosicchè il numero delle sfere nodali è uguale al numero delle radici positive del polinomio $L_{n+l}^{(2l+1)}$ ma si può dimostrare (teorema di Perron) che questo polinomio, di grado n' , ha tutte le sue n' radici positive, cosicchè il numero delle sfere nodali è uguale al quanto radiale $n' = n - l - 1$.

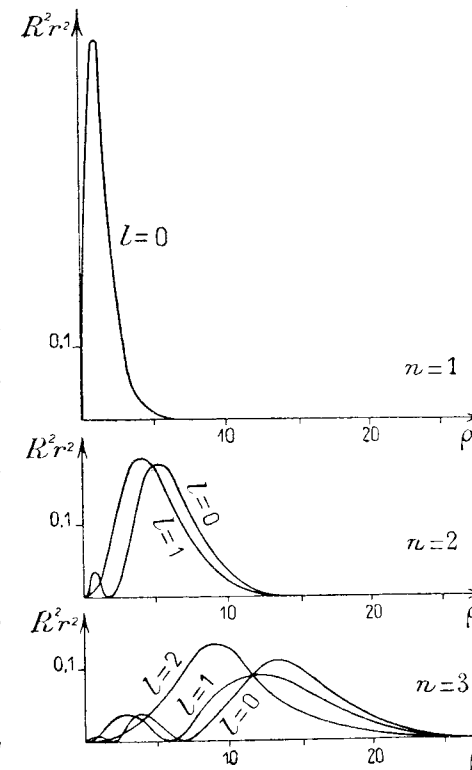


Fig. 41.

§ 50. — REGOLE DI SELEZIONE.

Come si è accennato al § 32, l'intensità e lo stato di polarizzazione della radiazione emessa nel salto quantico da uno stato n ad uno stato m sono determinate dalle quantità x_{nm} , y_{nm} , z_{nm} , definite dalle (144), cioè ottenute con integrazioni operate sul prodotto delle autofunzioni corrispondenti ai due stati in questione ed estese a tutto lo spazio. Accenniamo rapidamente ⁽¹⁾ al calcolo di queste quantità nel caso di un elettrone soggetto a forza centrale (e quindi, in particolare, per i sistemi idrogenoidi).

Poichè ogni stato è individuato da tre numeri quantici, a ciascuno degli indici n , m usati or ora si dovrà sostituire una terna di indici: le denoteremo rispettivamente con n_1 , l_1 , m_1 ed n_2 , l_2 , m_2 (e talvolta solo con 1, 2) cosicchè le quantità da calcolare saranno $x_{n_1 l_1 m_1 n_2 l_2 m_2}$ (che scriveremo, per brevità, x_{12}) e così via.

È poi comodo introdurre, in luogo delle coordinate cartesiane x , y , le loro combinazioni lineari

$$\xi = x + iy \quad , \quad \eta = x - iy \quad ,$$

cosicchè, invece di x_{12} , y_{12} , z_{12} calcoleremo

$$(287) \quad \begin{cases} \xi_{12} = \int_S \xi u_1^* u_2 dS \\ \eta_{12} = \int_S \eta u_1^* u_2 dS \\ z_{12} = \int_S z u_1^* u_2 dS . \end{cases}$$

Evidentemente, da ξ_{12} , η_{12} si passa a x_{12} , y_{12} attraverso le relazioni

$$\xi_{12} = x_{12} + iy_{12} \quad , \quad \eta_{12} = x_{12} - iy_{12} .$$

Introducendo le coordinate polari r , θ , φ si ha evidentemente

$$\xi = r \sin \theta e^{i\varphi} \quad , \quad \eta = r \sin \theta e^{-i\varphi} \quad , \quad z = r \cos \theta \\ dS = r^2 \sin \theta d\theta d\varphi dr .$$

⁽¹⁾ Per trattazione più completa, v. p. es. bibl. n. 1.

Tenendo poi presente che ognuna delle funzioni u_1 , u_2 , si scinde nel prodotto di tre fattori $\Phi(\varphi)$, $\Theta(\theta)$, $R(r)$, si vede che ognuno degli integrali tripli (287) si scinde nel prodotto di tre integrali semplici:

$$(288) \quad \begin{cases} \xi_{12} = \int_0^{2\pi} e^{i\varphi} \Phi_1^* \Phi_2 d\varphi \cdot \int_0^\pi \Theta_1^* \Theta_2 \sin^2 \theta d\theta \cdot \int_0^\infty r^3 R_1^* R_2 dr \\ \eta_{12} = \int_0^{2\pi} e^{-i\varphi} \Phi_1^* \Phi_2 d\varphi \cdot \int_0^\pi \Theta_1^* \Theta_2 \sin^2 \theta d\theta \cdot \int_0^\infty r^3 R_1^* R_2 dr \\ z_{12} = \int_0^{2\pi} \Phi_1^* \Phi_2 d\varphi \cdot \int_0^\pi \Theta_1^* \Theta_2 \cos \theta \sin \theta d\theta \cdot \int_0^\infty r^3 R_1^* R_2 dr . \end{cases}$$

Occupiamoci anzitutto degli integrali rispetto a φ . Il primo di essi, sostituendovi le espressioni di Φ_1 e Φ_2 conformi alla (229') diviene

$$\frac{1}{2\pi} \int_0^{2\pi} e^{i(1-m_1+m_2)\varphi} d\varphi$$

ed è quindi uguale a zero salvo il caso che l'esponente si annulli, cioè che sia $m_2 = m_1 - 1$, nel qual caso l'integrale è uguale ad 1. Similmente, il secondo integrale rispetto a φ si annulla salvo il caso di $m_2 = m_1 + 1$, ed il terzo si annulla salvo il caso di $m_2 = m_1$. Indicando dunque con Δm la differenza $m_2 - m_1$ (salto nel quanto equatoriale) si può dire che, affinché uno almeno dei tre integrali non sia nullo, deve essere

$$(289) \quad \Delta m = \pm 1, 0 .$$

Altrimenti, si annullano tutte e tre le quantità ξ_{12} , η_{12} , z_{12} , e quindi anche x_{12} , y_{12} , z_{12} , il che significa (v. § 32) che la riga spettrale corrispondente ha intensità nulla, ossia che un salto quantico in cui m non varia di ± 1 , o di 0, è « proibito ». Si ha dunque nella (289) la regola di selezione del quanto equatoriale. Si può aggiungere anzi che al salto $\Delta m = 0$ (solo z_{12} diverso da zero) corrisponde l'emissione di radiazione polarizzata rettilineamente come quella che sarebbe emessa da un oscillatore vibrante lungo l'asse z : ai salti $\Delta m = \pm 1$ corrisponde $z_{12} = 0$ e $x_{12} = \mp iy_{12}$: ciò significa che la radiazione corrisponde a quella emessa da un momento elettrico rotante nel piano xy (in senso destrorso

o sinistrorso) e quindi è polarizzata circolarmente per chi l'osserva nella direzione z , rettilineamente per chi l'osserva nel piano xy .

In modo analogo si ricava la regola di selezione pel quanto azimutale l dalla considerazione degli integrali rispetto a θ : il calcolo è però alquanto più lungo, e perciò non lo riportiamo, e ci limitiamo a riferire il risultato essenziale. Si trova che i tre integrali (nei quali, naturalmente, in luogo di m_2 si deve porre rispettivamente $m_1 + 1$, $m_1 - 1$, m_1) si annullano tutti, tranne il caso che sia $l_2 = l_1 \pm 1$, cosicchè pel salto $\Delta l = l_2 - l_1$ si trova la regola di selezione (290)

$$\Delta l = \pm 1 .$$

Infine, la considerazione dell'integrale rispetto ad r (che è lo stesso in tutte e tre le formule) non fornisce nessuna regola di selezione, poichè si trova che esso (quando siano già soddisfatte le regole di selezione per m ed l) non si annulla mai. Cosicchè il salto Δn del quanto totale n può essere qualunque.

Effettivamente l'esperienza dimostra che negli spettri si presentano di regola solo le righe soddisfacenti le regole di selezione ora trovate per m e per l , le quali regole anzi erano state rilevate empiricamente assai prima che sorgesse la meccanica ondulatoria ⁽¹⁾, ed erano state giustificate in vario modo nella teoria di Bohr e Sommerfeld (p. es., mediante il *principio di corrispondenza*, v. § 64).

Il calcolo effettivo degli integrali contenuti nelle (288) permette poi di valutare quantitativamente le probabilità del salto quantico e l'intensità della corrispondente riga spettrale: i casi, non molto numerosi, nei quali è stato possibile eseguire un confronto con l'esperienza hanno confermato i risultati della teoria.

⁽¹⁾ In certi casi eccezionali si presentano anche, con intensità assai ridotta, delle righe che sarebbero « proibite » dalle regole di selezione: ciò dipende da varie cause di cui la teoria, in una ulteriore approssimazione, rende conto perfettamente. Si intende poi che il passaggio tra due stati quantici, anche se proibito da una regola di selezione, può avvenire sempre per via indiretta (cioè attraverso livelli intermedi) o con un processo diverso da quello della radiazione, p. es. per urto.

CAPITOLO V.

LA TEORIA DI BOHR E SOMMERFELD

§ 51. - IL METODO DI WENTZEL E BRILLOUIN.

In questo capitolo tratteremo anzitutto di un metodo assai notevole, dovuto a WENTZEL ⁽¹⁾ e BRILLOUIN ⁽²⁾ e perfezionato da KRAMERS ⁽³⁾ e da vari altri, per trovare delle espressioni approssimate delle autofunzioni e degli autovalori dell'equazione di Schrödinger. Da questo metodo trarremo una *regola di quantizzazione* che sostanzialmente coincide con quella postulata da SOMMERFELD, sulla quale fu fondata la meccanica atomica fino al 1925 (v. parte I). Acquisiti così i fondamenti della teoria di Bohr e Sommerfeld, come prima approssimazione della meccanica ondulatoria, esporremo alcuni dei più notevoli risultati di questa teoria, che sarebbe troppo lungo o malagevole dedurre in modo rigoroso dalla equazione di Schrödinger.

Esporremo ora il metodo di Wentzel e Brillouin, riferendoci al caso unidimensionale, al quale si possono ricondurre anche problemi più generali, mediante la separazione delle variabili. Scriviamo dunque l'equazione unidimensionale di Schrödinger per uno stato di energia E (146, § 34) nella forma

$$(291) \quad u'' + \frac{4\pi^2}{h^2} p^2 u = 0 ,$$

dove si è posto

$$(292) \quad p(x) = \sqrt{2m(E - U)} .$$

⁽¹⁾ *ZS. f. Phys.*, **38** (1926), p. 518.

⁽²⁾ *C. R.*, **183** (1926), p. 24.

⁽³⁾ *ZS. f. Phys.*, **39** (1926), p. 828.

Si osservi che, per tutti i valori di x per cui $U < E$, questa funzione p coincide con l'espressione che avrebbe, nella meccanica classica, l'impulso (preso in valore assoluto) della particella al suo passaggio nel punto x : dove $U > E$ la p

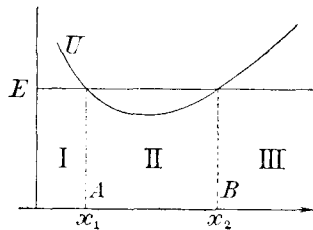


Fig. 42.

risulta immaginaria, e questo indica, secondo la meccanica classica, che tali regioni dell'asse x sono inaccessibili alla particella.

Supporremo che l'energia potenziale U abbia un andamento del tipo della fig. 42, cosicchè vi sia una sola regione AB (regione II, da x_1 a x_2) in cui p è reale, mentre nelle altre due regioni (I e III) è immaginaria a coefficiente positivo: classicamente, la particella eseguirebbe delle oscillazioni tra A e B .

L'equazione lineare del 2° ordine (291) si può trasformare in una del 1° ordine, ma non lineare (del tipo di Riccati) mediante la trasformazione, ben nota nella teoria delle equazioni differenziali,

$$(293) \quad u = e^{\frac{2\pi i}{h} \int y dx}$$

(dove il limite inferiore dell'integrale è un valore qualunque, ma fissato, di x). Si verifica subito infatti, sostituendo nella (291), che la y deve soddisfare l'equazione di Riccati

$$(294) \quad \frac{h}{2\pi i} y' = p^2 - y^2.$$

Considerando che generalmente i risultati della meccanica ondulatoria tendono a quelli della meccanica ordinaria, se nelle formule si trascurano quantità dell'ordine di h , si è indotti a ritenere che si possa ottenere una prima approssimazione trascurando h , una seconda approssimazione trascurando h^2 e le potenze superiori, e così via: ciò suggerisce il tentativo di cercare per y una espressione approssimata Y , della forma

$$(295) \quad Y = Y_0 + \frac{h}{2\pi i} Y_1 + \left(\frac{h}{2\pi i}\right)^2 Y_2 + \dots + \left(\frac{h}{2\pi i}\right)^k Y_k,$$

dove Y_0, Y_1, \dots sono funzioni di x che si determinano formalmente sostituendo la Y nella (294) in luogo di y ed uguagliando nei due membri i coefficienti delle singole potenze di $\frac{h}{2\pi i}$. Si trovano subito così per i primi termini le seguenti

formule ricorrenti:

$$(296) \quad Y_0 = \pm p,$$

$$(296') \quad Y_1 = -\frac{Y_0'}{2Y_0},$$

$$(296'') \quad Y_2 = -\frac{Y_1' + Y_1^2}{2Y_0},$$

e con pari facilità si troverebbero i termini successivi.

Come si vede, si ottengono per Y due espressioni (che indicheremo con Y_a ed Y_b) a seconda che nel primo termine si prende $Y_0 = +p$ o $Y_0 = -p$: una di esse si ottiene dall'altra cambiando il segno a tutti i termini dispari: esse sono (scrivendone solo i primi due termini, e rappresentando coi puntini i successivi fino al k -esimo)

$$(297) \quad Y_a = p - \frac{h}{2\pi i} \frac{p'}{2p} + \dots,$$

$$(297') \quad Y_b = -p - \frac{h}{2\pi i} \frac{p'}{2p} + \dots,$$

Queste due espressioni rappresentano approssimativamente due diversi integrali della (294).

Convieni qui chiarire il significato di queste soluzioni approssimate. Se, anzichè arrestare una delle somme al k -esimo termine, la si prosegue indefinitamente, si ottiene una serie *divergente*, e quindi non si può considerare la y approssimata da una *serie* del tipo (297) o (297'). Tuttavia, si può dimostrare che in generale i primi termini decrescono rapidamente, e che, limitando la somma a questi, Y_a ed Y_b rappresentano con buona approssimazione due integrali particolari della (294), per tutti i punti dell'asse x , *eccetto* quelli prossimi ai due punti critici A e B nei quali $p = 0$: difatti in questi punti Y_a e Y_b divengono infiniti mentre gli integrali della (294) non devono presentare singolarità.

Da ciascuna delle due Y così trovate si otterrà, mediante la (293), un integrale (approssimato) della (291), e quindi

un integrale qualsiasi di questa si potrà approssimare con una espressione del tipo

$$(298) \quad u = c_a e^{\frac{2\pi i}{h} \int Y_a dx} + c_b e^{\frac{2\pi i}{h} \int Y_b dx},$$

con c_a e c_b costanti. Tale approssimazione però cessa di esser valida nelle vicinanze dei due punti critici A e B . Ne segue che per rappresentare un medesimo integrale u nelle tre regioni I, II, III dell'asse reale si devono dare alle costanti c_a e c_b valori diversi in ciascuna di esse: li indicheremo con c_a' , c_b' , c_a'' , c_b'' , c_a''' , c_b''' rispettivamente. Vediamo ora come devono essere legati tra loro questi valori, limitandoci d'ora in poi all'approssimazione data dai primi due termini di Y_a e Y_b , cosicchè la (298) si può anche scrivere (v. (297) e (297')), fissando il limite inferiore in x_1 e convenendo di eseguire l'integrazione lungo l'asse reale:

$$(299) \quad u = \frac{c_a}{\sqrt{p}} e^{\frac{2\pi i}{h} \int_{x_1}^x p dx} + \frac{c_b}{\sqrt{p}} e^{-\frac{2\pi i}{h} \int_{x_1}^x p dx}$$

Si osservi anzitutto che per $x \rightarrow -\infty$, U tende a $+\infty$ e quindi p a $+i\infty$: ne segue che l'esponente del primo termine tende a $+\infty$ e perciò, affinchè la u per $x \rightarrow -\infty$ tenda a zero, come deve, dovrà essere $c_a' = 0$: così nella I regione la u si riduce a

$$(300) \quad u = \frac{c_b'}{\sqrt{p}} e^{-\frac{2\pi i}{h} \int_{x_1}^x p dx}.$$

Nella II regione la (299) si potrà anche scrivere (ponendo $c_a'' = C'' e^{i\delta}$, $c_b'' = C'' e^{-i\delta}$)

$$(301) \quad u = \frac{2C''}{\sqrt{p}} \cos\left(\frac{2\pi}{h} \int_{x_1}^x p dx + \delta\right).$$

Per collegare tra loro queste due espressioni, bisogna trovare un'espressione approssimata di u valida nella zona di confine — cioè in vicinanza del punto A — dove le formule precedenti fanno difetto. Ciò si può ottenere sostituendo ivi la curva del potenziale con un breve tratto rettilineo (cioè riguardando il campo di forza come uniforme nel breve

tratto considerato) ossia, ponendo nell'equazione di Schrödinger

$$\frac{8\pi^2 m}{h^2} (E - U) = K_1 (x - x_1),$$

dove K_1 è una costante (> 0). L'equazione di Schrödinger si può allora integrare rigorosamente, e la u si trova espressa mediante una funzione di Bessel, o, sotto un'altra forma, mediante un integrale definito: questa soluzione (che non presenta singolarità per $x = x_1$) si deve riattaccare con continuità da una parte (per $x < x_1$) alla (300), dall'altra alla (301) il che determina i valori delle costanti c'' e δ della (301). Tralasciamo di riportare i calcoli: si trova ⁽¹⁾ che nella regione II il prolungamento dell'integrale (300) è rappresentato da

$$(301') \quad u = \frac{2c_b''}{\sqrt{i p}} \cos\left(\frac{2\pi}{h} \int_{x_1}^x p dx - \frac{\pi}{4}\right).$$

Passando a considerare la regione III si riconosce, in modo analogo al precedente, che affinchè la u si annulli per $x \rightarrow +\infty$, nella regione III deve mancare il secondo termine della (299), cioè la u deve avere la forma

$$u = \frac{c_a'''}{\sqrt{p}} e^{\frac{2\pi i}{h} \int_{x_2}^x p dx},$$

dove c_a''' è una costante per ora indeterminata, e che non ci interessa. Questa espressione si deve ricollegare ad una della forma (301), valida nella regione II, attraverso il punto critico B : il collegamento può farsi con lo stesso metodo seguito per il punto A e si trova che la u , nella regione II, deve essere rappresentata da

$$(301'') \quad u = \frac{2c_a'''}{\sqrt{i p}} \cos\left(\frac{2\pi}{h} \int_x^{x_2} p dx - \frac{\pi}{4}\right).$$

Ora, affinchè questa espressione si identifichi con la (301'), i due coseni devono risultare uguali, in valore assoluto, per qualunque x : ciò richiede che la somma o la differenza dei

(1) V. KRAMERS, l. cit., o anche bibl. n. 22.

loro argomenti sia un multiplo intero di π : ma poichè la differenza degli argomenti dipende da x , mentre la loro somma, come si vede subito, risulta costante, sarà questa che dovrà uguagliarsi a $n\pi$ (con n intero), il che dà

$$2 \int_{x_1}^{x_2} p dx = \left(n + \frac{1}{2} \right) h .$$

Ricordando che, secondo la meccanica classica, la particella compirebbe delle oscillazioni tra x_1 ed x_2 con impulso $+p$ nel moto da x_1 ad x_2 e $-p$ nella semioscillazione contraria, il primo membro può interpretarsi come l'integrale di $p dx$ (dove p indica ora l'impulso preso col suo segno) esteso ad una oscillazione completa del movimento classico, il che si indica col simbolo \oint ; perciò la formula precedente si può scrivere

$$(303) \quad \oint p dx = \left(n + \frac{1}{2} \right) h .$$

Questa formula non contiene le autofunzioni, ma soltanto la E (contenuta in p): essa può dunque servire a determinare (approssimativamente) i diversi autovalori del problema senza passare attraverso l'equazione di Schrödinger: dando ad n i successivi valori interi si ottengono i successivi autovalori per E . La (303) rappresenta quindi una *regola di quantizzazione* (approssimata) valida per tutti i casi in cui il potenziale ha l'andamento della fig. 42.

È importante rilevare che se n è piuttosto elevato, la (303) può praticamente essere sostituita dalla

$$(303') \quad \oint p dx = nh .$$

Consideriamo ora un altro caso, il cui interesse risulterà meglio in seguito: quello cioè in cui la x , anzichè rappresentare una coordinata cartesiana, rappresenti un angolo che può variare da 0 a 2π (p. es., la θ nelle coordinate polari piane): in tal caso la U (e quindi la p) è una funzione pe-

riodica di x a periodo 2π , e la u , anzichè tendere a zero all'infinito, deve essere anch'essa una funzione periodica di x a periodo 2π . Supponiamo la E abbastanza grande perchè p risulti dappertutto reale, cioè riferiamoci al caso in cui il moto classico sarebbe rotatorio (non oscillatorio). Allora la u avrà dovunque la forma (301) (dove x_1 è un valore qualunque, ma fissato, di x) e sarà periodica a periodo 2π solo se

$$\frac{2\pi}{h} \int_{x_1}^{x_1+2\pi} p dx = 2\pi n ,$$

ossia se

$$\oint p dx = nh ,$$

indicando con \oint l'integrale esteso ad un periodo. La condizione di quantizzazione è dunque, in questo caso, esattamente la (303') anzichè la (303).

§ 52. - LE CONDIZIONI DI SOMMERFELD.

Le considerazioni del § precedente rendono ragione del successo, in molti casi brillante, del metodo di quantizzazione postulato da Sommerfeld assai prima che sorgesse la meccanica ondulatoria. Tale metodo consiste, come si è accennato nella parte I, nel trattare dapprima l'atomo come un sistema di punti materiali soggetto alle ordinarie leggi della meccanica classica (o, in una ulteriore approssimazione, della meccanica relativista), e poi aggiungere a queste leggi delle condizioni restrittive (*condizioni di Sommerfeld*) che rendono possibili solo alcuni tra gli infiniti movimenti consentiti dalla ordinaria meccanica e quindi solo alcuni valori per l'energia. Prima di formulare le condizioni di Sommerfeld, dobbiamo richiamare alcune fondamentali nozioni di meccanica razionale.

Consideriamo un sistema meccanico ad f gradi di libertà, con vincoli indipendenti dal tempo, e soggetto a forze conservative: lo riferiremo ad un sistema di *coordinate lagrangiane* $q_1,$

q_2, \dots, q_f (che, in particolare, possono essere coordinate cartesiane delle singole particelle, se si tratta di un sistema di punti) a cui corrispondono altrettanti momenti ⁽¹⁾ p_1, p_2, \dots, p_f , e supponiamo che il suo movimento sia tale che ognuna delle q_i , e la corrispondente p_i , siano funzioni periodiche ⁽²⁾ del tempo con un periodo T_i . I movimenti che — con una scelta conveniente delle coordinate — presentano questa proprietà si dicono *multiplamente periodici*. Osserviamo incidentalmente che in genere ognuno degli f periodi T_i sarà diverso: se questi periodi sono uguali o commensurabili il moto è «periodico» perchè, trascorso un tempo tale che sia multiplo comune di tutte le T_i , il movimento si ripete identico; ma in generale questo non avviene e perciò i moti periodici sono un caso particolare dei moti multiplamente periodici, caratterizzato dal fatto che si possono trovare f numeri interi m_1, m_2, \dots, m_f tali che

$$\frac{m_1}{T_1} = \frac{m_2}{T_2} = \dots = \frac{m_f}{T_f},$$

ovvero, introducendo le frequenze $\nu_i = 1/T_i$,

$$m_1 \nu_1 = m_2 \nu_2 = \dots = m_f \nu_f.$$

⁽¹⁾ Come si sa dalla meccanica, la forza viva T del sistema è una funzione delle q e delle \dot{q} , e si chiamano *momenti* le quantità

$$p_i = \frac{\partial T(q, \dot{q})}{\partial \dot{q}_i} \quad (i = 1, 2, \dots, f).$$

Essendo la T una funzione quadratica delle \dot{q} , i momenti risultano funzioni lineari delle \dot{q} : è anzi possibile risolverle ed esprimere le \dot{q} come funzioni lineari delle p .

In particolare, se le coordinate q sono le ordinarie coordinate cartesiane x, y, z di un punto, i corrispondenti momenti sono le componenti dell'impulso (o quantità di moto), si ha cioè, essendo $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$,

$$p_x = m\dot{x}, \quad p_y = m\dot{y}, \quad p_z = m\dot{z}.$$

⁽²⁾ Se qualcuna delle q rappresenta un angolo tale che aumentando di 2π si ritrova la stessa configurazione del sistema (p. es., la θ in un sistema di coordinate polari nel piano), allora si considera come periodo T_i relativo a questa coordinata il tempo richiesto perchè essa aumenti di 2π . Si dice allora che questo è un grado di libertà di «rotazione», mentre quelli corrispondenti a coordinate che oscillano periodicamente entro due limiti si chiamano «di librazione» o di «oscillazione».

Noi per ora escluderemo non solo questo caso, ma anche quello più generale che tra le frequenze ν_i passino una o più relazioni del tipo

$$(304) \quad m_1 \nu_1 + m_2 \nu_2 + \dots + m_f \nu_f = 0,$$

con coefficienti interi. Quando vi siano g relazioni di questo tipo, il sistema dicesi *g volte degenerare*: se il moto di un sistema a f gradi di libertà è periodico esso è dunque $f - 1$ volte degenerare. I sistemi degeneri si presentano assai spesso nelle questioni di fisica atomica ma di essi ci occuperemo altrove.

Tra i sistemi multiplamente periodici, ci occuperemo in particolare di quelli che ammettono un sistema di coordinate lagrangiane tale, che ognuno dei momenti p_i sia funzione della sola coordinata q_i e non delle altre. Questa condizione si suole esprimere dicendo che «le variabili sono separabili»: essa infatti equivale ad ammettere che l'equazione di Jacobi ⁽¹⁾ si possa

⁽¹⁾ Ricorderemo dalla meccanica analitica il metodo di Jacobi per integrare le equazioni del moto, quando i vincoli sono indipendenti dal tempo e le forze conservative. Si esprime l'energia totale (somma della forza viva T e dell'energia potenziale U) in funzione delle q e delle p (la funzione $H(q, p)$ così definita si chiama l'*hamiltoniana* del sistema, e contiene in sé tutto ciò che occorre per caratterizzare le proprietà meccaniche di questo), poi si sostituisce, nell'espressione di H , ogni p_i con $\frac{\partial W}{\partial q_i}$ (dove W è una funzione incognita delle q) e si pone l'equazione a derivate parziali (*equaz. di Jacobi*):

$$(305) \quad H\left(q, \frac{\partial W}{\partial q}\right) = \alpha_1$$

dove α_1 è una costante arbitraria. Si tratta ora di trovare una soluzione $W(q, \alpha_1, \alpha_2, \dots, \alpha_f)$ di questa equazione, che contenga, oltre ad α_1 , altre $f - 1$ costanti arbitrarie $\alpha_2, \alpha_3, \dots, \alpha_f$. Trovata questa, le equazioni del moto si hanno scrivendo le $2f$ relazioni seguenti tra le q , le p e t (da cui si potrebbero ricavare esplicitamente le q e le p in funzione di t):

$$(306) \quad \frac{\partial W}{\partial \alpha_1} = t + \beta_1,$$

$$(306') \quad \frac{\partial W}{\partial \alpha_j} = \beta_j \quad (j = 2, 3, \dots, f),$$

$$(307) \quad p_i = \frac{\partial W}{\partial q_i} \quad (i = 1, 2, 3, \dots, f),$$

dove le β sono altre f costanti arbitrarie.

Di queste equazioni, le (306'), che non contengono t , determinano la forma della traiettoria di ciascun punto, la (306) determina la legge

soddisfare con una funzione della forma

$$(308) \quad W = \sum_i f_i(q_i)$$

anzichè della forma più generale $W = f(q_1 \dots q_f)$. Tutte queste condizioni sembrano molto restrittive, ma in pratica la maggior parte dei sistemi che si presentano nello studio della meccanica atomica le soddisfano, e quindi possono ad essi applicarsi le condizioni di Sommerfeld che ora enunceremo.

Il movimento del sistema ad f gradi di libertà dipende, come è noto, da $2f$ costanti, $\alpha_1, \alpha_2 \dots \alpha_f, \beta_1, \beta_2 \dots \beta_f$ definibili dalle condizioni iniziali: ora SOMMERFELD impose una restrizione a queste costanti, obbligandole a soddisfare alle seguenti f condizioni: per ciascuna coordinata, si calcoli l'integrale (detto *integrale di fase*)

$$(309) \quad J_i = \oint p_i dq_i,$$

esteso ad un intero periodo della coordinata stessa: poichè la p_i dipende solo dalla q_i (per ipotesi) e dalle f costanti α questo integrale sarà una funzione $J_i(\alpha_1, \dots, \alpha_f)$ delle sole costanti α : anzi, tali funzioni si potrebbero invertire e se ne potrebbero ricavare le α in funzione delle f costanti J . Le condizioni di Sommerfeld consistono nel porre ciascuno degli integrali J uguale ad un multiplo intero della costante di Planck h , cioè nello scrivere

$$(310) \quad \oint p_i dq_i = n_i h \quad (i = 1, 2, \dots, f),$$

con cui essa viene percorsa, e le (307) determinano i momenti, e quindi le velocità (in funzione delle q e delle α). La costante α_1 ha il significato fisico di *energia totale*; le costanti β , che non intervengono nelle (307), hanno il significato di *fasi*.

Molte volte l'equazione (305) si può integrare col metodo della separazione delle variabili, cioè prendendo W della forma (308) (dove, beninteso, ognuna delle f_i dipende anche dalle f costanti α); allora evidentemente le (307) assumono la forma $p_i = f_i'(q_i, \alpha)$ ossia ogni p_i dipende solo dalla q_i ad essa coniugata, e dalle α . È questo il caso al quale si applicano le condizioni di Sommerfeld.

dove n_i è un numero intero non negativo ⁽¹⁾. Si hanno dunque tante condizioni quanti sono i gradi di libertà, e si introducono altrettanti *numeri quantici* n_i i quali possono sostituire le f costanti α : così in luogo di f costanti suscettibili di valori continui se ne introducono altrettante, ma suscettibili di valori discreti.

E poichè, come si è visto, una delle α si identifica con l'energia E del sistema, si può riguardare questa come una funzione delle f costanti J , e quindi delle n_i : si conclude che l'energia può assumere solo valori discreti, dipendenti dagli f numeri interi n_1, n_2, \dots, n_f e perciò denotabili con $E_{n_1 n_2 \dots n_f}$.

Si trovano così i livelli energetici, in accordo, come si vedrà, generalmente buono e talvolta ottimo, con quelli sperimentali.

Nel caso particolare di un solo grado di libertà, la condizione di Sommerfeld coincide con quella (303') che abbiamo dedotto in via approssimata dalla meccanica ondulatoria: si noti però che nel caso di un grado di libertà oscillatorio abbiamo trovato che la migliore approssimazione si ottiene eguagliando l'integrale a $(n + \frac{1}{2})h$, mentre nel metodo di Sommerfeld lo si eguaglia a nh come per i gradi di libertà rotatori: effettivamente in questi casi l'introduzione di numeri « semi-interi » (cioè del tipo $n + \frac{1}{2}$) al posto dei numeri quantici interi migliora generalmente (come del resto si rilevò empiricamente anche prima che sorgesse la meccanica ondulatoria) l'approssimazione data da Sommerfeld, ed in certi casi (p. es. nel caso dell'oscillatore) dà senz'altro il risultato esatto. Invece, pei gradi di libertà rotatori, i numeri quantici, come si è visto, devono essere interi.

Nel caso di un sistema a più gradi di libertà, a variabili separabili, le condizioni di Sommerfeld si potrebbero ritrovare in modo analogo (estendendo le considerazioni del § 51

⁽¹⁾ Si può infatti dimostrare facilmente che l'integrale a primo membro non è mai negativo.

mediante la separazione delle variabili) come conseguenze di prima approssimazione dell'equazione di Schrödinger e delle condizioni di regolarità imposte alle autofunzioni.

§ 53. — CENNO SULLA SCELTA DEL SISTEMA DI COORDINATE.

Poichè un sistema meccanico può riferirsi a infiniti sistemi di coordinate lagrangiane, sorge la questione: se invece del sistema delle q , si adotta un altro sistema di coordinate q'_1, q'_2, \dots, q'_j (di cui diremo p'_1, p'_2, \dots, p'_j i momenti coniugati), e si applicano a queste le condizioni di Sommerfeld

$$(311) \quad \oint p'_i dq'_i = n_i h,$$

si ottengono le stesse orbite quantiche o no? Bisogna osservare anzitutto che, per poter applicare le condizioni (311), è necessario che anche nel sistema delle q' si possa eseguire la separazione delle variabili, il che implica (se il sistema non è degenere) che la trasformazione che conduce dalle q alle q' deve appartenere ad un tipo particolarmente ristretto, che non altera le linee coordinate nè modifica gli integrali J (o, al più, opera su di essi delle sostituzioni lineari, a coefficienti interi ed a determinante unitario): le (311) determinano in questo caso le stesse traiettorie delle (310), generalmente dando alle n' valori diversi nei due casi.

Ma se il sistema è degenere, non è più così: allora la separazione delle variabili è possibile anche in sistemi di coordinate essenzialmente differenti, che conducono a valori degli J non riducibili ai precedenti con sostituzioni lineari a coefficienti interi e determinante unitario, e quindi si ottengono traiettorie quantiche diverse a seconda del sistema di riferimento scelto. Tuttavia, avviene anche in questi casi che i livelli energetici risultino indipendenti dalla scelta del sistema di riferimento. Ciò dipende dal fatto che i livelli energetici hanno un significato fisico, mentre le orbite della teoria di Bohr e Sommerfeld non sono che una finzione analitica (v. § 57).

Nella trattazione dei sistemi degeneri vi è dunque una certa arbitrarietà nella scelta del sistema di coordinate: ci si lascia perciò guidare, generalmente, dal criterio seguente. Osserviamo che se si *perturba*, anche lievemente, il sistema (p. es., variando leggermente le forze che su di esso agiscono) in generale esso cessa di essere degenere, perchè i suoi periodi T_i variano leggermente

e cessano di soddisfare le relazioni (304). Ora, molte volte un sistema si presenta come degenere solo perchè, per semplicità, si trascurano alcune delle leggere perturbazioni che in realtà non mancano mai: p. es. un atomo si considera in genere immerso a tutte le influenze esterne, ma in realtà è sempre immerso in un campo magnetico, sia pur debole, dovuto agli atomi vicini, al magnetismo terrestre o ad altre cause. Basta perciò prendere in considerazione queste perturbazioni e ricercare un sistema di riferimento nel quale le variabili siano separabili: questo stesso sistema di coordinate poi si adotta per trattare il problema anche trascurando le perturbazioni. Difatti, è intuitivo che se le forze perturbatrici si fanno tendere gradualmente a zero, il loro effetto sul movimento tende pure a zero, mentre il loro effetto nel fissare il sistema di coordinate cui si applicano le condizioni di Sommerfeld è indipendente dalla loro entità.

§ 54. — OSCILLATORE ARMONICO.

Riprendiamo il problema studiato al § 39 per trattarlo col metodo di Sommerfeld. Si ha anzitutto dalla meccanica che la particella eseguirebbe delle oscillazioni secondo la legge

$$(312) \quad x = A \sin(2\pi\nu_0 t - \varphi),$$

dove A e φ sono due costanti arbitrarie: il momento coniugato alla x è

$$p = m\dot{x} = 2\pi\nu_0 m A \cos(2\pi\nu_0 t - \varphi).$$

L'energia totale è (v. (183), § 39)

$$E = U + \frac{1}{2} m\dot{x}^2 = 2\pi^2 m\nu_0^2 A^2 x^2 + \frac{1}{2} m\dot{x}^2 = 2\pi^2 \nu_0^2 A^2 m$$

(e contiene, come si vede, una sola delle due costanti di integrazione). Calcoliamo ora l'integrale di fase

$$J = \oint p dx = \oint m\dot{x} dx = m \int_0^{1/\nu_0} \dot{x}^2 dt = m \cdot 4\pi^2 \nu_0^2 A^2 \int_0^{1/\nu_0} \cos^2(2\pi\nu_0 t - \varphi) dt.$$

e poichè l'ultimo integrale vale $\frac{1}{2\nu_0}$, risulta

$$J = 2\pi^2 \nu_0 A^2 m = \frac{E}{\nu_0},$$

e quindi la condizione di Sommerfeld $J = nh$ dà per E i valori

$$(313) \quad E_n = nh\nu_0 \quad (n = 0, 1, 2, \dots).$$

Si trova così la legge (già postulata da PLANCK nella teoria del corpo nero) che l'energia dell'oscillatore è sempre un multiplo intero del « quanto » $h\nu_0$: se invece si usa la formula (303), (che trattandosi di un moto di tipo oscillatorio rappresenta una migliore approssimazione) si trova

$$(313') \quad E_n = \left(n + \frac{1}{2}\right) h\nu_0,$$

e cioè si aggiunge, agli n quanti $h\nu_0$, una quantità fissa di energia $\frac{1}{2} h\nu_0$ della quale l'oscillatore non può mai privarsi (« energia dello stato zero »). Lo stesso risultato, come si è visto, si ottiene col metodo di Schrödinger, ed è confermato in vari modi dall'esperienza: cosicchè in questo caso particolare il metodo di Wentzel e Brillouin fornisce (casualmente) il risultato esatto, nei riguardi dell'energia.

§ 55. — ROTATORE E ATOMO DI BOHR.

Si chiama « rotatore » un sistema costituito da un corpo rigido girevole intorno ad un asse fisso, non soggetto a forze (o soggetto a forze di momento nullo rispetto all'asse). Il sistema è a un solo grado di libertà, e la sua posizione in ogni istante può essere individuata mediante una coordinata angolare θ , esprime l'angolo fra due piani passanti per l'asse di rotazione, uno fisso e l'altro solidale al corpo. L'energia del sistema, che si riduce alla sola forza viva, è data (indicando con J il momento di inerzia) da

$$E = \frac{1}{2} J\dot{\theta}^2.$$

Il momento coniugato alla coordinata θ è quindi (v. nota al § 52)

$$p = J\dot{\theta},$$

cioè rappresenta il momento dell'impulso, o *momento angolare*. Il moto si svolge, come è ben noto, con la legge $p = \text{cost.}$, cosicchè

$$\oint p d\theta = p \oint d\theta = 2\pi p.$$

La condizione di Sommerfeld dà dunque

$$(314) \quad p = n \frac{h}{2\pi}.$$

e si può esprimere dicendo che *il momento angolare deve essere multiplo di $\frac{h}{2\pi}$* ⁽¹⁾. Quindi la velocità angolare dovrà avere

uno dei valori discreti $n \frac{h}{2\pi J}$, e, sostituendo nella espressione di E , si trovano per l'energia i valori

$$(315) \quad E_n = n^2 \frac{h^2}{8\pi^2 J}.$$

Anche questo risultato coincide (casualmente) in modo perfetto con quello fornito dall'integrazione rigorosa dell'equazione di Schrödinger (convenientemente estesa per abbracciare un sistema siffatto).

Si osservi che l'atomo di idrogeno concepito da BOHR (v. § 16, p. I) in cui l'elettrone è vincolato ad un'orbita circolare, si può meccanicamente identificare con un rotatore, e quindi anche per esso deve valere la (314): così si vede che la condizione postulata da BOHR per caratterizzare le orbite quantiche rientra come caso particolare nelle condizioni di Sommerfeld.

⁽¹⁾ Poichè in molti altri casi i momenti angolari risultano multipli di $\frac{h}{2\pi}$, o almeno in rapporto semplice con questa quantità, si usa spesso assumere $\frac{h}{2\pi}$ come unità di misura del momento angolare (*unità quantistica*). P. es. il risultato espresso dalla (314) si può enunciare dicendo che « il momento angolare del rotatore è espresso da un numero intero n » (sottintendendo l'unità quantistica).

§ 56. — SISTEMI IDROGENOIDI.

Applichiamo ora il metodo di Sommerfeld all'atomo di idrogeno e, in genere, ai sistemi idrogenoidi, senza la restrizione puramente artificiale delle orbite circolari, adottata da BOHR nella prima teoria quantistica di questi sistemi esposta al § 16, p. I: si supporrà ancora però, provvisoriamente, che il nucleo sia fisso. Il sistema è dunque a tre gradi di libertà.

a) *Parte meccanica.* — Si sa dalla meccanica razionale che il movimento dell'elettrone sotto l'azione di una forza centrale attrattiva di intensità Ze^2/r^2 (cioè analoga all'attrazione newtoniana) avviene in orbite ellittiche secondo le leggi di Keplero, essendo da escludere le orbite paraboliche ed iperboliche che corrispondono a stati in cui l'elettrone non è vincolato al nucleo. Introducendo le coordinate polari, nel piano dell'orbita, r , ω (col polo nel nucleo e l'asse polare diretto verso il perielio) l'equazione dell'ellisse si scrive

$$(316) \quad \frac{1}{r} = \frac{Ze^2m}{p^2} (1 + \varepsilon \cos \omega) .$$

Le due costanti p ed ε sono determinate dalle condizioni iniziali e rappresentano: p , il momento angolare ⁽¹⁾ ($p = mr^2\dot{\omega}$), ed ε l'eccentricità (trattandosi di ellisse, $\varepsilon < 1$). Per trovare i semiassi, osserviamo che i valori massimo e minimo di r , cioè la distanza afelica e quella perielica (corrispondenti a $\omega = 180^\circ$, $\omega = 0$), sono

$$(317) \quad r_{max} = \frac{p^2}{Ze^2m} \frac{1}{1 - \varepsilon} , \quad r_{min} = \frac{p^2}{Ze^2m} \frac{1}{1 + \varepsilon} ,$$

e poichè il semiasse maggiore a è evidentemente dato da $\frac{1}{2}(r_{max} + r_{min})$, sarà

$$(318) \quad a = \frac{p^2}{Ze^2m} \frac{1}{1 - \varepsilon^2} ,$$

⁽¹⁾ Sceglieremo il verso positivo di ω coincidente col verso in cui è percorsa l'ellisse, cosicchè p non sia negativo.

mentre, come è noto dalla geometria,

$$(319) \quad b = a \sqrt{1 - \varepsilon^2} .$$

L'energia cinetica è

$$(320) \quad T = \frac{m}{2} (\dot{r}^2 + r^2\dot{\omega}^2)$$

e quella potenziale è $-Ze^2/r$. Per calcolare l'energia totale $E = T + U$ conviene (poichè essa è costante) riferirsi ad un istante particolare del moto, scelto in modo da semplificare il calcolo: p. es. per r_{min} , cui corrisponde $\dot{r} = 0$; si trova allora

$$(321) \quad E = -\frac{1}{2} \frac{Ze^2}{a}$$

e si vede che l'energia dipende solo dal semiasse maggiore, non dal minore.

b) *Condizioni di Sommerfeld.* — Osserviamo che il sistema è doppiamente degenero (poichè le tre coordinate variano tutte con lo stesso periodo). Per la scelta del sistema di coordinate (v. § 53) bisogna dunque tenere conto di qualche perturbazione: una di queste è la correzione relativistica alle leggi della meccanica e l'altra può essere un lieve campo magnetico esterno. Si trova che il sistema così perturbato non è più degenero e che le variabili si separano assumendo coordinate polari (nello spazio) col polo nel nucleo e l'asse polare diretto secondo il campo: adotteremo dunque queste coordinate anche trascurando le dette perturbazioni. Siano perciò r il raggio vettore, θ la colatitudine (angolo del raggio vettore con l'asse polare) e φ la longitudine: la forza viva ha l'espressione nota

$$(322) \quad T = \frac{m}{2} (\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta \dot{\varphi}^2)$$

e quindi i momenti coniugati a r , θ , φ sono rispettivamente

$$\begin{aligned} p_r &= m\dot{r} , \\ p_\theta &= mr^2\dot{\theta} , \\ p_\varphi &= mr^2 \sin^2\theta \dot{\varphi} . \end{aligned}$$

Quest'ultimo ha il significato meccanico di « momento dell'impulso rispetto all'asse polare », ossia *proiezione sull'asse polare del momento angolare p*: esso è quindi costante, mancando le forze esterne.

Ora possiamo scrivere le tre condizioni di Sommerfeld, che sono

$$(323) \quad \oint p_r dr = m \oint \dot{r} dr = n'h ,$$

$$(324) \quad \oint p_\theta d\theta = m \oint r^2 \dot{\theta} d\theta = sh ,$$

$$(325) \quad \oint p_\varphi d\varphi = m \oint r^2 \sin^2\theta \dot{\varphi} d\varphi = m_\varphi h ,$$

dove n' , s , m_φ sono tre numeri interi, non negativi.

L'ultima, essendo p_φ costante, ci dà subito

$$\pm 2\pi p_\varphi = m_\varphi h$$

(dove vale il segno $+$ o $-$ secondo che il moto avviene nel senso delle φ crescenti o decrescenti). Convenendo di indicare ⁽¹⁾ con m^* il numero $\pm m_\varphi$ (dove il segno è scelto col criterio precedente) si avrà

$$(325') \quad p_\varphi = m^* \frac{h}{2\pi} ,$$

vale a dire: il momento angolare rispetto all'asse polare è espresso (in unità $\frac{h}{2\pi}$) da un numero intero m^* ($\begin{matrix} \geq \\ \leq \end{matrix} 0$) che dicesi « *quanto magnetico* » (e che corrisponde al quanto magnetico m della teoria ondulatoria introdotto al § 46).

Una analoga quantizzazione per il momento angolare (totale) $p = |\mathbf{p}|$ può ricavarsi dalla (324) e dalla (325). Si noti però prima che, identificando le due espressioni (320)

⁽¹⁾ Adottiamo provvisoriamente per il quanto magnetico la notazione m^* per evitare confusione con la massa elettronica m : in seguito, quando non vi sia ragione di equivoco, scriveremo m in luogo di m^* , come è uso generale.

e (322) della forza viva, e moltiplicandole per dt , si ha l'identità

$$r^2 \dot{\theta} d\theta + r^2 \sin^2\theta \dot{\varphi} d\varphi = r^2 \dot{\omega} d\omega .$$

Ora, sommando le due ultime condizioni di Sommerfeld (324), (325), e tenendo conto di questa identità, si ottiene

$$m \oint r^2 \dot{\omega} d\omega = (s + m_\varphi) h .$$

Si ricordi poi che il momento angolare è dato da

$$(326) \quad p = mr^2 \dot{\omega} ,$$

perciò, se si definisce un nuovo numero quantico (non negativo) k , ponendo

$$(327) \quad k = s + m_\varphi = s + |m^*| ,$$

si può scrivere ⁽¹⁾

$$(328) \quad \oint p d\omega = kh .$$

Siccome poi p è costante, si ricava di qui

$$(329) \quad p = k \frac{h}{2\pi} ,$$

ossia: *il momento angolare totale è multiplo intero di $\frac{h}{2\pi}$* . Il numero quantico k che misura questo momento angolare in unità $\frac{h}{2\pi}$, nella antica teoria di Bohr-Sommerfeld era chiamato « *quanto azimutale* » (oggi si dà invece questo nome al numero l della teoria ondulatoria, che, come vedremo più oltre, si può identificare con $k - 1$).

⁽¹⁾ Si osservi che, essendo p il momento coniugato ad ω nel sistema di coordinate polari piane (nel piano dell'orbita), la (328) può interpretarsi come una delle due condizioni di Sommerfeld che si troverebbero trattando il problema in due dimensioni come se l'elettrone dovesse muoversi in un piano dato: l'altra è identica alla (323). Questo modo di trattare il problema, cioè con la restrizione alquanto artificiosa che il moto avvenga in un piano dato, è quello abitualmente seguito nelle esposizioni elementari della teoria di Sommerfeld.

c) *Quantizzazione spaziale.* — Le due ultime condizioni di Sommerfeld determinano l'inclinazione del piano dell'orbita rispetto all'asse polare, ossia rispetto al campo magnetico esterno (sia pur debolissimo). Difatti, detto α l'angolo che il vettore \mathbf{p} (che è normale al piano dell'orbita) forma con questo asse, si ha

$$p_{\varphi} = p \cos \alpha,$$

e quindi, sostituendo le espressioni (325') e (329) di p_{φ} e p , si ricava (notando che, come si dirà più avanti, $k \neq 0$)

$$(330) \quad \cos \alpha = \frac{m^*}{k}.$$

Se ne deduce che, fissato k , $\cos \alpha$ può assumere solo la serie di valori discreti $0, \pm \frac{1}{k}, \pm \frac{2}{k}, \dots, \pm 1$: tuttavia, per essere

d'accordo coi fatti sperimentali si deve escludere il caso $m^* = \pm k$, cioè $\cos \alpha = \pm 1$, nel quale il piano dell'orbita risulterebbe normale al campo magnetico. Questa esclusione (della quale l'antica teoria di Sommerfeld non dà nessuna soddisfacente giustificazione ⁽¹⁾) si può giustificare oggi mediante il confronto coi risultati della meccanica ondulatoria, di cui questa teoria deve rappresentare una approssimazione (vedi sez. f). Ammettendola per il momento, si conclude che il quanto magnetico m^* assume i $2k - 1$ valori

$$(331) \quad -(k-1), -(k-2), \dots, -1, 0, 1, \dots, (k-1),$$

cui corrispondono altrettante inclinazioni del piano dell'orbita.

L'esistenza di queste inclinazioni discrete si designa spesso con l'espressione « quantizzazione spaziale ».

d) *Forma delle orbite.* — Dobbiamo ora tener conto della rimanente condizione di Sommerfeld (323), dove n' ($= 0, 1, 2, \dots$) chiamasi *quanto radiale*.

L'integrale si calcola prendendo come variabile d'integrazione ω ed osservando che

$$\dot{r} = \frac{dr}{d\omega} \dot{\omega} = \frac{p}{mr^2} \frac{dr}{d\omega}.$$

⁽¹⁾ Nella antica teoria si usava invece escludere il valore $m^* = 0$, per motivi analoghi a quelli che facevano escludere $k=0$. Il risultato non era però del tutto soddisfacente (v. Geiger e Scheel, *Hdb. der Phys.*, Bd. XIII, 1^a ed., p. 145 e 164).

cosicchè si può scrivere

$$\oint m r \dot{r} dr = p \int_0^{2\pi} \frac{1}{r^2} \left(\frac{dr}{d\omega} \right)^2 d\omega = p \int_0^{2\pi} r^2 \left(\frac{d \frac{1}{r}}{d\omega} \right)^2 d\omega.$$

e, ponendo per $1/r$ l'espressione (316), si ottiene:

$$p \varepsilon^2 \int_0^{2\pi} \frac{\sin^2 \omega d\omega}{(1 + \varepsilon \cos \omega)^2}.$$

La quadratura si può effettuare operando prima per parti e poi con la trasformazione $\cos \omega = \frac{1 - x^2}{1 + x^2}$ (oppure con un metodo più rapido ed elegante dovuto al Sommerfeld ⁽¹⁾) e si trova

$$\oint m r \dot{r} dr = 2\pi p \left(\frac{1}{\sqrt{1 - \varepsilon^2}} - 1 \right).$$

Quindi la (323) dà, tenuto conto anche della (329), la condizione seguente per ε :

$$\frac{1}{\sqrt{1 - \varepsilon^2}} - 1 = \frac{n'}{k},$$

ossia, introducendo un nuovo numero intero (che si identificherà col « quanto totale » del § 47)

$$(332) \quad n = k + n',$$

$$(333) \quad \sqrt{1 - \varepsilon^2} = \frac{k}{n}.$$

Sostituendo nella (318) questa espressione di $1 - \varepsilon^2$, e la (329) per p , si ottiene per a l'espressione (dipendente solo da n)

$$(334) \quad a_n = \frac{h^2}{4\pi^2 Z e^2 m} n^2.$$

Similmente, dalla (319), si vede che il rapporto dei semiassi è

$$(335) \quad \frac{b}{a} = \frac{k}{n}.$$

⁽¹⁾ V. p. es. bibl. n. 18, p. 655.

Di solito si usano i due interi k ed n (anzichè k ed n') per caratterizzare l'orbita.

Si osservi che k è sempre minore od al più uguale ad n , poichè n' non può essere negativo: si ha $k = n$, cioè $n' = 0$, nel caso delle orbite circolari (perciò l' n introdotto nella teoria di Bohr, ristretta alle orbite circolari, si può considerare sia come quanto azimutale che come quanto totale). Inoltre si deve escludere il caso di $k = 0$: anche questa è un'esclusione che si giustifica rigorosamente solo mediante il confronto con la teoria di Schrödinger (v. § 57); nella vecchia teoria si giustificava con la considerazione che il caso di $k = 0$ corrisponderebbe, come mostra la (335), a $b = 0$, cioè l'ellisse degenererebbe in un segmento di retta, e l'elettrone andrebbe ad urtare nel nucleo.

Riassumendo, fissato n , il quanto azimutale k può assumere solo gli n valori

$$(336) \quad k = 1, 2, \dots, n,$$

ed a questi corrispondono altrettante ellissi, tutte con lo stesso semiasse maggiore, ma con diverso semiasse minore:

l'ultimo è il cerchio di raggio a_n : naturalmente, tutte queste ellissi hanno un fuoco nel nucleo (v. fig. 43).

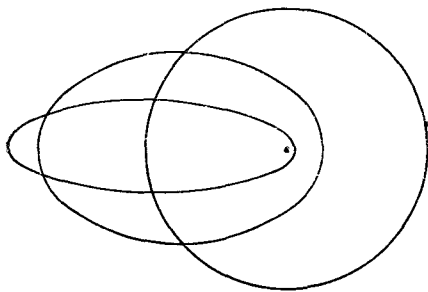


Fig. 43.

e) *Livelli energetici.* — Anzitutto osserviamo che tutte le ellissi corrispondenti allo stesso n avendo lo stesso a hanno la stessa energia: questa dunque

dipende solo dal quanto totale n e non da k . Sostituendo la (334) nella (321) si ha per l'energia di una qualunque delle orbite di quanto totale n

$$(337) \quad E_n = -\frac{2\pi^2 Z^2 e^4 m}{h^2} \frac{1}{n^2},$$

espressione coincidente con quella già trovata per l'orbita circolare n -esima nella teoria di Bohr (v. § 16, p. I).

Dunque la considerazione delle orbite ellittiche, nell'idrogeno e nei sistemi idrogenoidi, non aggiunge nuovi livelli energetici: è per questo che la teoria di Bohr, benchè limitata alle orbite circolari, rendeva già conto di tutte le righe spettrali osservate. La teoria più completa ora svolta ci avverte però che ad ogni livello energetico corrisponde in generale più di un tipo di movimento a causa della già rilevata degenerazione del sistema. Difatti, fissato n , (ossia l'asse maggiore) il quanto azimutale k può assumere i valori (336) e si hanno quindi diverse « forme » dell'orbita: a ciascun k poi corrispondono $2k - 1$ valori di m^* e quindi ciascuna orbita può assumere altrettante inclinazioni: perciò il numero totale di orbite corrispondenti ad un dato quanto totale n e quindi ad un dato livello energetico (che chiamasi anche « peso statistico » o « molteplicità » del livello) è

$$\sum_{k=1}^n (2k - 1) = n^2.$$

Dunque al primo livello corrisponde una orbita, al secondo quattro e così via. Ogni riga spettrale risulta in genere da parecchie specie di salti quantici. P. es. la prima riga della serie di Balmer viene emessa da tutti gli atomi in cui l'elettrone salta da una qualunque delle orbite $n = 3$, $k = 1, 2, 3$, ad una qualunque ⁽¹⁾ delle orbite $n = 2$, $k = 1, 2$.

Si noti che basta una lieve perturbazione (p. es. un campo elettrico o magnetico) per togliere, in tutto od in parte, la perfetta coincidenza tra i valori dell'energia corrispondenti alle n^2 orbite dello stesso quanto totale n : allora ogni livello energetico si scinderà in un gruppo di livelli vicinissimi, corrispondenti ai diversi valori di k e di m^* .

§ 57. — CONFRONTO CON LA TEORIA DI SCHRÖDINGER.

La teoria dell'atomo di idrogeno svolta in questo § deve essere considerata, come sappiamo, una prima approssima-

⁽¹⁾ Si vedrà in seguito che qualcuno di questi salti quantici in realtà non può avvenire, essendo escluso dal principio di selezione.

zione della teoria di Schrödinger, nel senso che i valori numerici che da essa si ricavano per le varie quantità aventi significato fisico (energia, momento angolare, ecc.) sono approssimativamente uguali a quelli deducibili dalla teoria rigorosa svolta al cap. IV. Anzi, i livelli energetici risultano esattamente gli stessi. Anche i valori del momento angolare rispetto all'asse, cioè di p_φ , risultano esattamente gli stessi, purchè al quanto magnetico m^* si attribuiscono i soli valori (331) (escludendo i valori $\pm k$), e purchè si identifichi k con $l + 1$ della teoria ondulatoria (dal che segue l'esclusione del valore $k = 0$): così anche le molteplicità di ciascun livello risultano le stesse nelle due teorie.

È invece soltanto approssimata l'espressione (329) del momento angolare: abbiamo già detto infatti che nella teoria rigorosa questo risulta espresso (in unità $\frac{h}{2\pi}$) non da k ma da $\sqrt{l(l+1)}$ ossia $\sqrt{k(k-1)}$ ovvero $\sqrt{k}l$. In particolare, per $k = 1$ dovrebbe risultare $p = 0$ mentre la (329) dà $p = h/2\pi$.

Talvolta può convenire sostituire la (329) con la formula

$$(329') \quad p = (k-1) \frac{h}{2\pi} = l \frac{h}{2\pi},$$

cioè ritenere p misurato (in unità quantistiche) da l anzichè da k : la (329') dà un'approssimazione un po' migliore della (329), e per $k = 1$ dà addirittura il risultato giusto.

Esaminiamo ora il significato fisico che si può attribuire alle orbite della teoria di Bohr e Sommerfeld ed al moto dell'elettrone su di esse. Per definire fisicamente la traiettoria dell'elettrone, ed il suo moto, bisognerebbe potere (almeno concettualmente) determinare in tempi successivi la posizione di esso, con incertezza piccola rispetto alle dimensioni dell'orbita, e ripetere molte volte l'osservazione sullo stesso atomo. Ma per eseguire una osservazione di posizione con incertezza piccola rispetto al raggio a dell'orbita ⁽¹⁾ bisogna,

(1) Ci riferiamo per semplicità alle orbite circolari, ma il ragionamento si estende subito alle orbite ellittiche, sostituendo a e P con i loro valori medi.

per il principio di indeterminazione, comunicare all'elettrone una variazione di impulso indeterminata di ordine di grandezza $\Delta P \geq \sim h/a$: quindi nella osservazione successiva il moto è già perturbato. Si vede così che la determinazione fisica dell'orbita è concettualmente possibile (approssimativamente) solo se la perturbazione ΔP è trascurabile rispetto all'impulso P posseduto dall'elettrone nel suo moto orbitale.

Ora, $P = \frac{p}{a} = \frac{h}{2\pi a} n$: quindi la condizione è

$$\frac{h}{a} \ll \frac{h}{2\pi a} n \text{ ossia } n \gg 2\pi.$$

Dunque: un'orbita è fisicamente determinabile tanto più esattamente quanto più grande è n . Alle prime orbite (p. es. $n = 1, 2, \dots$) non si può attribuire nessun significato fisico.

Per chiarire meglio quanto precede, e mettere in luce i rapporti del modello di Bohr e Sommerfeld con la meccanica ondulatoria, si supponga di voler determinare le successive posizioni dell'elettrone entro l'atomo illuminandolo con lunghezza d'onda piccola rispetto alle sue dimensioni, ed osservando la luce diffusa, come al § 23 (si tratta, non occorre dirlo, di esperienze ideali). Se l'atomo è in uno dei primi stati quantici, le sue dimensioni sono dell'ordine di 10^{-8} cm., e si dovranno quindi usare raggi X: ma è ben noto che basta un fotone di questi per ionizzare l'atomo e quindi già dalla prima osservazione l'elettrone è sbalzato via e non si può più ripetere l'osservazione sullo stesso atomo. Non è dunque possibile rivelare l'orbita ma solo una posizione dell'elettrone rispetto al nucleo, e questo si può fare con precisione tanto grande quanto si vuole. Se poi si ripete questa osservazione su molti atomi (tutti nello stesso stato) si potrà fare una statistica dei risultati ottenuti, ossia attribuire a ciascun punto dello spazio intorno al nucleo una « densità di probabilità » di trovare ivi l'elettrone: è appunto questa la densità di probabilità che insegna a calcolare la meccanica ondulatoria. Così, p. es., se gli atomi in questione sono tutti nello stato fondamentale ($n = 1$), i risultati delle osservazioni di posizione dell'elettrone si ripartiscono simmetricamente intorno al nucleo, addensandosi intorno ad esso come è mostrato dalla fig. 39a.

Se invece consideriamo atomi in uno stato quantico elevato e perciò di dimensioni maggiori, potremo contentarci di osservare la posizione dell'elettrone con precisione minore e quindi con luce di frequenza minore, incapace di produrre la ionizza-

zione; potremo allora pensare di ripetere l'osservazione molte volte successive sullo stesso atomo, e rilevare così approssimativamente l'orbita dell'elettrone e la legge del suo moto. Ma poichè già con la prima osservazione perturbiamo in modo incontrollabile l'impulso (e l'energia) dell'elettrone, non potremo assegnare questa orbita ad uno stato quantico determinato, e l'indeterminazione dello stato (cioè di n , k , m) sarà tanto maggiore quanto più precise sono state le misure di posizione. Tutto ciò è fedelmente rispecchiato nello schema della meccanica ondulatoria: difatti, sovrapponendo un certo numero di autofunzioni corrispondenti a diversi valori di n , k , m , si può costruire un pacchetto d'onde ristretto quanto si vuole, il quale (v. § 26) si muove intorno al nucleo imitando approssimativamente il moto dell'elettrone nella teoria di Bohr e Sommerfeld: esso, se opportunamente costruito, rappresenta la densità di probabilità in ciascun istante, e quindi permette di prevedere statisticamente i risultati delle successive osservazioni di posizione: se queste sono eseguite con grande precisione (cioè se si osserva con radiazioni di lunghezza d'onda molto piccola, compatibilmente con la necessità di evitare la ionizzazione) dovremo costruire un pacchetto d'onde molto ristretto, per il che sarà necessario sovrapporre autofunzioni corrispondenti ad un ampio intervallo di valori di n , k , m , cioè lasciare ampiamente indeterminato lo stato quantico. Si vede così che il concetto di « orbita » e quello di « stato quantico » sono *complementari* nel senso di Bohr (v. § 22) come quelli di posizione e velocità, ecc.

Risulta poi chiaro che non è possibile costruire un pacchetto d'onde sufficientemente piccolo, se si vuole far uso p. es. solo delle prime tre o quattro autofunzioni: ciò corrisponde al fatto, già rilevato sopra, che le orbite corrispondenti ai primi valori di n non hanno nessun significato fisico.

§ 58. — MOVIMENTO DEL NUCLEO.

Un importante ritocco da farsi alla precedente teoria dei sistemi idrogenoidi consiste nel tener conto del fatto che il nucleo non è rigorosamente fisso, come finora abbiamo supposto, ma descrive una piccola orbita intorno al centro di gravità del sistema. Come si sa dalla meccanica, il problema dei due corpi si riconduce a quello di un solo corpo attirato da un centro fisso, purchè si modifichi lievemente la massa del mobile: il moto di un elettrone di massa m rispetto al nucleo di massa M , è retto dalle stesse equazioni del moto,

rispetto ad un nucleo fisso, di un corpuscolo di massa (*massa ridotta*)

$$(338) \quad m' = \frac{m}{1 + \frac{m}{M}}.$$

Anche le condizioni di Sommerfeld si ottengono da quelle relative al nucleo fisso semplicemente sostituendovi m con m' : ciò si riconosce nel modo più semplice adottando come coordinate lagrangiane del sistema le coordinate polari del nucleo rispetto al baricentro (r_0 , θ_0 , φ_0), e le coordinate polari dell'elettrone rispetto al nucleo (r , θ , φ): si trova allora che i momenti coniugati alle tre prime sono identicamente nulli e quelli coniugati alle altre tre sono

$$p_r = m'\dot{r}, \quad p_\theta = m'r^2\dot{\theta}, \quad p_\varphi = m'r^2\sin^2\theta\dot{\varphi},$$

quindi delle sei condizioni di Sommerfeld le prime tre sono identicamente soddisfatte, e le altre tre si identificano con le (323), (324), (325) salvo la sostituzione di m con m' (1). I livelli energetici sono dunque dati da

$$E_n = -\frac{2\pi^2 Z^2 e^4 m'}{h^2} \frac{1}{n^2} = -Z^2 \frac{R_M h c}{n^2}.$$

dove si è posto (in analogia con la espressione della costante di Rydberg data al § 16, p. I):

$$(339) \quad R_M = \frac{2\pi^2 e^4 m'}{c h^3} = \frac{R}{1 + \frac{m}{M}},$$

ed i termini spettrali risultano

$$\tau_n = Z^2 \frac{R_M}{n^2}.$$

L'effetto del movimento del nucleo è quindi quello di sostituire la costante di Rydberg R con quella, lievemente minore, R_M . Questa correzione, dipendendo da M , è legger-

(1) Si osservi che p_θ ha ancora il significato di momento angolare totale del sistema: perciò il quanto azimutale k conserva il suo significato.

mente diversa per i diversi sistemi idrogenoidi: si trova p. es.

$$\text{per } H : R_H = 109.677,76$$

$$\text{per } He^+ \quad R_{He} = 109.722,40$$

Ciò significa che le frequenze di una stessa riga nei vari spettri idrogenoidi non devono essere esattamente proporzionali a Z^2 : p. es. le righe pari della serie di Pickering non devono coincidere esattamente con quelle della serie di Balmer (come risulta dalla prima approssimazione del § 16, p. I) ma essere leggermente spostate verso il violetto: così è difatti, e la differenza (che è dell'ordine di un Ångström) è perfettamente osservabile nello spettro dato da una miscela di idrogeno ed elio. La misura di queste differenze costituisce anzi uno dei metodi più esatti per determinare il rapporto tra la massa dell'elettrone e quella del nucleo. Il valore di R dato dalla (21) del § 16 è quello che corrisponderebbe ad un nucleo di massa infinita (perciò talvolta è indicato con R_∞).

Una correzione del tutto analoga si può fare nella teoria di Schrödinger e conduce allo stesso risultato (v. § 21, p. III),

§ 59. — ATOMI CON PIÙ ELETTRONI (1).

Negli atomi (o ioni) con più di un elettrone il metodo di Bohr e Sommerfeld condurrebbe a risolvere anzitutto il problema meccanico del moto degli elettroni sotto l'azione dell'attrazione nucleare e della loro mutua repulsione (che non è affatto trascurabile, come è, in prima approssimazione, l'attrazione dei pianeti nel problema astronomico): tale problema è in genere di risoluzione praticamente impossibile. Ma anche nel caso in cui se ne è trovato una soluzione approssimata (quello dell'atomo di elio, con due elettroni) si è riconosciuto che, quantizzando questo movimento con le condizioni di Sommerfeld, si giungeva a livelli energetici grossolanamente errati: ciò significa che in questo caso l'approssimazione data dai primi due termini della (295) è del

(1) V. il volume sulla spettroscopia del presente trattato, oppure il n. 23 o il n. 27 della bibl.

tutto insufficiente. Tuttavia una serie di risultati qualitativi della più alta importanza si è ottenuta con le considerazioni seguenti.

Poichè i termini spettrali di molti atomi e ioni si lasciano rappresentare nella forma di Rydberg od in quella di Ritz (v. § 15, p. I), che sono analoghe a quella di Balmer, il BOHR fu condotto a pensare che l'emissione di tali spettri fosse provocata dal salto quantico di uno solo degli elettroni, gli altri conservando presso a poco immutate le loro orbite. Si vede così l'opportunità di considerare (artificiosamente) l'atomo o lo ione come composto di due parti: una, costituita dal nucleo e da tutti gli elettroni meno uno (questa parte si chiama *nocciolo* o *resto atomico*, e si considera, in prima approssimazione, di struttura invariabile), l'altra costituita dal rimanente elettrone, che chiamasi *elettrone ottico*, e che può percorrere diverse orbite quantiche sotto l'azione delle forze emananti dal nocciolo. Tale elettrone è di regola uno dei cosiddetti *elettroni di valenza*, cioè di quelli che, per essere meno strettamente legati all'atomo, ne vengono più facilmente allontanati quando l'atomo si trasforma in ione positivo.

Le forze con cui il nocciolo agisce sull'elettrone ottico sono naturalmente assai complesse, ma vengono schematizzate considerando tutti gli elettroni del nocciolo come formanti intorno al nucleo una distribuzione continua o « nuvola » di elettricità negativa, statica ed a simmetria sferica, così da generare un campo di forza centrale che si sovrappone a quello dovuto al nucleo (1). Tale campo, come è noto, è, per i punti esterni alla « nuvola », lo stesso come se questa fosse concentrata nel centro, ossia come se la carica positiva del nucleo fosse diminuita di tanto, quanto è la carica complessiva degli elettroni che costituiscono la nuvola stessa: se questi sono in numero di a la carica del nucleo è virtual-

(1) Tale schematizzazione appare oggi, secondo la meccanica ondulatoria, meno arbitraria di quanto poteva sembrare: difatti molto spesso la funzione ψ di Schrödinger corrispondente agli elettroni del nocciolo è tale che $|\psi|^2$, cioè la densità *media* di carica elettrica, è indipendente dal tempo ed a simmetria sferica.

mente ridotta da Ze a $(Z-a)e$ (*carica efficace*): quindi l'elettrone ottico è attratto verso il nocciolo (finchè si trova al difuori di questo) da una forza $(Z-a)e/r^2$ invece che Ze/r^2 . Si dice perciò, alquanto impropriamente, che gli elettroni del nocciolo esercitano un'azione di *schermo* sull'attrazione del nucleo, ed il numero a si chiama *numero di schermo*. Se p. es. si tratta di un atomo neutro, sarà $a = Z - 1$ e quindi la carica efficace sarà e : se si tratta di un atomo ionizzato z volte, sarà $a = Z - 1 - z$ e quindi la carica efficace sarà $(z + 1)e$.

Se però l'elettrone ottico penetra entro il nocciolo, l'azione di schermo diminuisce, perchè, come si sa, il campo generato da una distribuzione sferica di elettricità in un punto interno ad essa, a distanza r dal centro, è lo stesso che sarebbe prodotto da quella parte della carica elettrica che è contenuta entro la sfera di raggio r , mentre le cariche più esterne restano inefficaci. Perciò per un elettrone che penetri entro il nocciolo la carica efficace del nucleo va gradatamente crescendo col decrescere di r , e tende alla carica vera Ze col tendere di r a 0. In tal caso dunque si può dire che l'elettrone ottico si muove in un campo che è newtoniano all'esterno di una certa sfera (limite del nocciolo) e non newtoniano all'interno: la legge con cui varia il campo entro il nocciolo dipende dalla legge con cui si suppone distribuita la densità elettrica in questo, e può venire schematizzata in vario modo, secondo il grado di approssimazione richiesto.

Il calcolo dei livelli energetici di un atomo con più elettroni si riduce dunque, in questa approssimazione, a quello dei livelli energetici di un solo elettrone in un campo non newtoniano, e tale problema si imposterà in modo analogo a quello dei sistemi idrogenoidi: l'ordinaria meccanica dei moti centrali permetterà di determinare una infinità continua di movimenti (e risulta che, in prima approssimazione, questi si possono considerare come movimenti kepleriani cui è sovrapposta una lenta rotazione uniforme, detta *precessione*, cosicchè la traiettoria ha la forma della fig. 44 e chiamasi *rosetta*): imponendo poi le tre condizioni di Sommerfeld si caratterizzano le orbite quantiche, che vengono anche in

questo caso a dipendere da tre numeri interi: un *quanto azimutale* k ($= 1, 2, \dots$) oppure l ($= 0, 1, 2, \dots$) che misura il momento angolare in unità $h/2\pi$, un *quanto magnetico* m ($= 0, \pm 1, \dots, \pm l$) esprime la proiezione di quel momento su un asse, ed un *quanto totale* n . Non si presenterà però, in generale, la circostanza che l'energia dipenda solo da n , come nei sistemi idrogenoidi, bensì essa dipenderà in generale da n e da l (non da m , a cagione della simmetria sferica

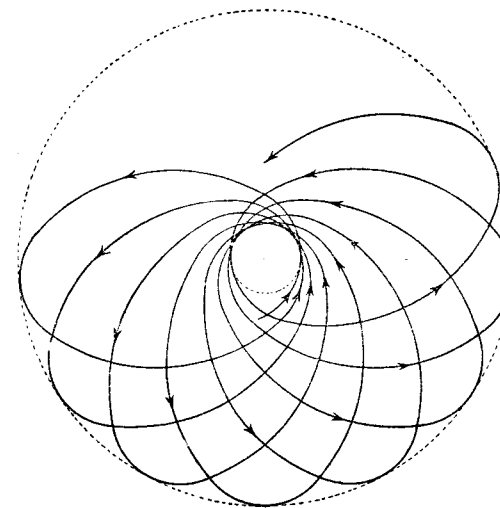


Fig. 44.

del campo ⁽¹⁾), e perciò i livelli energetici, e quindi i termini spettroscopici, costituiranno una serie a *due indici* anzichè ad uno solo: si scriverà cioè E_{nl}, τ_{nl} . È questo il risultato più importante di queste considerazioni (il quale conserva il suo valore anche nella meccanica quantistica): aggiungeremo che, precisando opportunamente il campo creato dal nocciolo, il Sommerfeld è giunto a trovare, per i termini spettrali, la forma di Rydberg e, in ulteriore approssimazione, quella di Ritz ⁽²⁾. È poi evidente che se l'orbita dell'elettrone ottico

⁽¹⁾ L'energia viene a dipendere da m quando l'atomo si trova in un campo magnetico di intensità sufficiente a perturbare il moto: si produce allora l'effetto Zeeman.

⁽²⁾ V. bibl. n. 18.

decorre tutta all'esterno del nocciolo, cioè dove il campo è newtoniano (*orbita non penetrante*) si ritroveranno i livelli energetici dell'idrogeno e quindi i termini balmeriani. Così si giustifica il fatto sperimentale che i termini di ogni atomo o ione in genere si approssimano tanto più alla forma balmeriana, quanto più elevati sono i loro numeri quantici.

È invalso nella spettroscopia l'uso di indicare i termini spettrali, corrispondenti ai livelli energetici di un elettrone, anziché con la notazione τ_{nl} , mediante una lettera minuscola designante il valore del quanto azimutale l preceduto da un numero indicante il quanto totale: il quanto azimutale è indicato mediante le lettere seguenti, derivanti da denominazioni della antica spettroscopia (diamo anche il valore di k , quanto azimutale nella teoria di Sommerfeld):

$l =$	0	1	2	3	4	5...
$k =$	1	2	3	4	5	6...
Lettera usata:	s	p	d	f	g	$h...$

(termini con l più elevato raramente intervengono). Così p. es. il termine per cui $n = 3$ e $l = 0$ si indica con $3s$ anziché con τ_{31} , e si parla di *termini della serie s, della serie p, ecc.*, od anche di *termini s, p, ecc.* Si osservi che, essendo sempre $n \geq l + 1$ la serie s incomincia col termine $1s$, la serie p , con $2p$, la serie d con $3d$, ecc. Spesso però, invece del vero quanto totale n , si usa una numerazione convenzionale, incominciando tutte le serie col numero 1: così, in questa notazione, il termine $1p$ ha il quanto totale $n = 2$, il termine $2p$ ha $n = 3$, ecc.

Graficamente, si usa rappresentare i livelli corrispondenti ai termini su diverse colonne, una per la serie s , una per la p , ecc., come si vede nella fig. 45, a destra.

Nel caso dei sistemi idrogenoidi, i livelli delle varie colonne coinciderebbero tutti (nella nostra approssimazione) e perciò si rappresentano in una colonna unica, come nella fig. 8, o nella fig. 45, a sinistra.

Negli atomi e ioni non idrogenoidi, gli atomi che saltano da uno stato di quanto totale n ad uno di quanto totale n' non emettono tutti la medesima riga spettrale ma righe diverse

secondo i valori iniziale e finale del quanto azimutale k : nel caso dei sistemi idrogenoidi queste righe vengono a sovrapporsi, e da ciò deriva la particolare semplicità di tali spettri. Tuttavia la sovrapposizione non è perfetta, come si dirà più avanti.

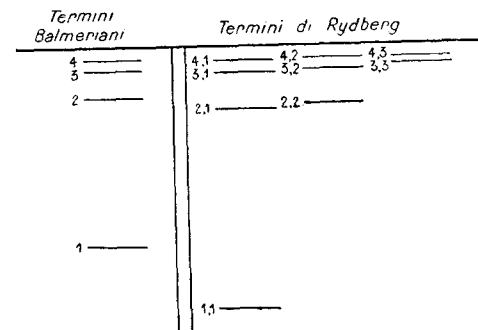


Fig. 45.

§ 60. — CENNO SULLA TEORIA RELATIVISTICA DEI SISTEMI IDROGENOIDI.

Il moto dell'elettrone ottico avviene, come si calcola facilmente, con velocità dell'ordine di 10^8 cm/sec., ossia dell'ordine di qualche millesimo di quella della luce: quindi la meccanica ordinaria è applicabile ad esso solo approssimativamente, ed è da attendersi che una trattazione più rigorosa, fatta utilizzando la dinamica relativista, porti ad applicare ai risultati trovati precedentemente delle correzioni non del tutto trascurabili.

Se, per i sistemi idrogenoidi, si ripete il calcolo del § 56 ma servendosi della meccanica relativista invece di quella ordinaria⁽¹⁾, si trova anzitutto che l'elettrone descrive un'orbita a rosetta, cioè descrive un'ellisse con un fuoco nel nucleo e con legge kepleriana, ma questa ellisse, anziché essere fissa, ruota lentamente intorno al nucleo (l'effetto della correzione relativista, nei riguardi dell'orbita, è dunque analogo a quello che negli atomi non idrogenoidi è prodotto dal fatto che il campo non è esattamente newtoniano). L'orbita

(1) v. p. e. bibl. n° 18.

dunque non è più una linea chiusa ed il moto non è più periodico: la perturbazione relativistica toglie dunque, in parte, la degenerazione del sistema, come si è già accennato al § 56. Quanto ai livelli energetici si trova per essi la seguente espressione approssimata

$$E_{nk} = -\frac{2\pi^2 Z^2 e^4 m}{h^2} \frac{1}{n^2} - \frac{\pi^4 Z^4 e^8 m}{c^2 h^4} \frac{1}{n^3} \left(\frac{8}{k} - \frac{6}{n} \right),$$

ovvero, ricordando l'espressione trovata al § 16, p. I per la costante di Rydberg, ed introducendo la costante⁽²⁾ (detta « costante della struttura fina »)

$$(340) \quad \alpha = \frac{2\pi e^2}{hc} = 7,280 \cdot 10^{-3} = 1/137,37$$

l'espressione precedente diviene

$$(341) \quad E_{nk} = -Z^2 \frac{Rhc}{n^2} \left[1 - \frac{Z^2 \alpha^2}{n^2} \left(\frac{3}{4} - \frac{n}{k} \right) \right].$$

Trascurando il termine in α^2 si ha l'espressione già trovata nella teoria non relativistica: il detto termine rappresenta una correzione il cui rapporto al termine principale è dell'ordine di α^2 , cioè di qualche centomillesimo. Questa correzione, come si vede, dipende non solo da n ma anche da k : perciò tutte le orbite aventi lo stesso quanto totale n e quanto azimutale $k = 1, 2, \dots, n$ (cioè le ellissi dello stesso asse maggiore e di diversa eccentricità) non corrispondono esattamente allo stesso livello energetico, come si era trovato in prima approssimazione, ma a livelli diversi, per quanto vicinissimi tra loro: si parla perciò di *livelli multipli*. Naturalmente, a questa molteplicità dei livelli corrisponde una molteplicità delle righe spettrali: gli atomi che saltano da uno stato di quanto totale n , con diversi valori di k , ad uno di quanto n' , con diversi valori di k' , emettono righe che non si sovrappongono esattamente. Avviene dunque, a rigore, anche nei sistemi idrogenoidi ciò

⁽²⁾ Questa quantità, che interviene anche in relazione all'elettrone rotante, ha le dimensioni di un numero puro, ed è uguale, come si vedrebbe facilmente, al rapporto tra la velocità dell'elettrone nella prima orbita circolare, e la velocità della luce.

che si è detto al § precedente per gli atomi più complessi, sebbene per un motivo diverso: le differenze di frequenza in questione sono però qui così piccole, che si rivelano solo con mezzi di forte potere risolutivo e solo se le righe sono molto sottili: perciò il gruppo di righe vicinissime si suol considerare come una « riga » composta di parecchie « componenti » (*struttura fina*).

La formula (341) permette di calcolare la frequenza delle varie componenti di ciascuna riga: noi ci limiteremo ad osservare che le differenze di frequenza risultano dell'ordine di $\alpha^2 Z^4 R$ e quindi devono essere nello spettro dell'elio ionizzato ($Z = 2$) 16 volte maggiori che nelle corrispondenti righe dell'idrogeno: questo fatto, unito alla circostanza che è più facile ottenere sottili le righe dell'elio che quelle dell'idrogeno, fa sì che la struttura fina si osservi più facilmente nell'elio.

Le previsioni ricavabili dalla formula (341) (integrata da una *regola di selezione* che verrà enunciata al § 64) concordano quasi esattamente coi risultati delle osservazioni eseguite sull'elio e sull'idrogeno: tuttavia questa coincidenza si deve ritenere, in certo modo, fortuita.

Difatti, la teoria precedente trascura due ordini di fatti: anzitutto, che il modello di Bohr e Sommerfeld non rappresenta che una prima approssimazione, in secondo luogo trascura l'influenza dello spin dell'elettrone. Ora si vedrà al cap. V p. III che quando si eliminano questi due errori, usando la meccanica ondulatoria relativistica di Dirac (che tiene conto implicitamente anche dello spin), si giunge ad un risultato assai vicino a quello contenuto nella formula (341) (e confermato esattamente dall'esperienza) cosicchè il parziale successo di questa è dovuto al fatto che gli errori dipendenti dalle due cause anzidette si elidono l'un l'altro quasi esattamente.

§ 61. - IL MAGNETONE DI BOHR

Una conseguenza importante del modello atomico di Rutherford è che un atomo deve possedere in genere un momento magnetico a causa del moto orbitale degli elettroni, le cui

traiettorie equivalgono a tanti circuiti elettrici. Calcoliamo, sulla base di questo modello, il momento magnetico prodotto dal moto dell'elettrone nel caso dei sistemi idrogenoidi.

Osserviamo perciò che, detto τ il periodo del moto kepleriano, per un punto qualsiasi della traiettoria passa, $1/\tau$ volte al secondo, la carica e : perciò è come se l'ellisse fosse percorsa da una corrente di intensità e/τ . Questa ellisse, per un noto principio di elettromagnetismo, equivale dunque ad una lamina magnetica la cui potenza (momento per unità di area) è $\frac{1}{c} \frac{e}{\tau}$, e perciò, detta S l'area dell'ellisse,

$$(342) \quad \mu = \frac{Se}{c\tau}.$$

Ora, per la legge delle aree,

$$dS = \frac{1}{2} r^2 d\omega = \frac{p}{2m} dt,$$

e perciò, integrando per un intero periodo

$$(343) \quad S = \frac{p}{2m} \tau.$$

Sostituendo nella (342), si vede che il momento magnetico μ risulta, in grandezza, legato al momento angolare p da

$$(344) \quad \mu = \frac{e}{2mc} p.$$

Se si considera poi che i vettori μ e p sono normali al piano dell'orbita, e diretti (come si vede facilmente) in senso opposto, si può anche scrivere la relazione vettoriale

$$(344') \quad \mu = -\frac{e}{2mc} p$$

(dove e rappresenta la carica dell'elettrone in valore assoluto). Questo risultato si potrebbe estendere ai sistemi con quanti si vogliono elettroni.

Si osservi ora che p è sempre multiplo intero di $h/2\pi$, secondo la (329) o meglio la (329'), perciò si potrà scrivere

$$\mu = l \frac{he}{4\pi mc},$$

e si riconosce che: il momento magnetico dovuto al moto orbitale degli elettroni è sempre un multiplo intero di

$$(345) \quad \mu_0 = \frac{he}{4\pi mc}.$$

Questo momento magnetico elementare chiamasi *magnitone di Bohr*: il suo valore risulta, in unità CGS,

$$\mu_0 = 0,9158 \cdot 10^{-20}.$$

Si osservi inoltre che, per un risultato trovato al § 56 (quantizzazione spaziale) che si estende immediatamente anche ad atomi non idrogenoidi, un atomo posto in un campo magnetico si orienta in modo che la componente di p sulla direzione del campo sia $m \frac{h}{2\pi}$, dove m è un intero che abbiamo chiamato *quanto magnetico*: dalla (344') si vede allora che la componente, sulla direzione del campo, del momento magnetico μ risulta anch'essa quantizzata, e cioè

$$(346) \quad \mu_H = m\mu_0.$$

Di questo fatto, STERN e GERLACH hanno dato una notevole dimostrazione sperimentale, che ha permesso anche di misurare direttamente i momenti magnetici di diversi atomi (1)

Osserviamo incidentalmente che al risultato (346) si giunge anche con la meccanica ondulatoria, come è stato indicato dal FERMI. Difatti ricordiamo dal § 31 l'espressione (137) della densità media di corrente elettrica j , e sostituiamovi per ψ l'espressione trovata al § 46 per l'elettrone soggetto ad un campo centrale e cioè

$$\psi = \frac{1}{\sqrt{2\pi}} R(r) \Theta(\theta) e^{im^*\varphi}$$

(dove R e Θ sono due funzioni di cui qui non interessa l'espressione: basta ritenere che sono reali e sono normalizzate secondo le formule (244) e (252)): si trova subito che le componenti di j secondo il raggio vettore e secondo il meridiano sono nulle, men-

(1) V. p. es., bibl. n. 27, p. 168.

tre quella secondo il parallelo è

$$j_{\varphi} = \frac{he}{4\pi im} \left(\psi^* \frac{\partial \psi}{\partial \varphi} - \psi \frac{\partial \psi^*}{\partial \varphi} \right) \frac{1}{r \sin \theta} = \frac{he}{4\pi^2 m} \frac{R^2 \Theta^2}{r \sin \theta} m^* .$$

L'elemento d'area $r dr d\theta$ del piano meridiano è dunque attraversato da una corrente di intensità $j_{\varphi} r dr d\theta$, che descrive un cerchio di raggio $r \sin \theta$ e quindi equivale a una lamina magnetica di momento

$$d\mu_H = \frac{\pi}{c} r^2 \sin^2 \theta j_{\varphi} r dr d\theta .$$

Integrando su tutto il semipiano meridiano, si ottiene il momento magnetico totale nella direzione dell'asse polare, che è

$$\mu_H = \frac{he}{4\pi mc} m^* \int_0^{\infty} R^2 r^2 dr \int_0^{\pi} \Theta^2 \sin \theta d\theta ,$$

e poichè, per le citate condizioni di normalizzazione, i due integrali valgono 1, si ritrova, ricordando la (345), il risultato (346).

§ 62. - L'ELETTRONE ROTANTE E IL QUANTO INTERNO.

Le considerazioni del § precedente indussero per molto tempo a ritenere che tutti i fenomeni, i quali dipendono dal momento magnetico degli atomi (paramagnetismo, ferromagnetismo, ecc.) si potessero spiegare col momento derivante dai moti orbitali degli elettroni. Ma vi sono vari fenomeni (effetto Zeeman, effetti giromagnetici) i quali permettono di determinare il rapporto μ/p tra il momento magnetico e quello meccanico (cioè momento angolare) dell'atomo, rapporto che dovrebbe risultare, a norma del § precedente, costante ed uguale a $\frac{e}{2mc}$: invece si è trovato che esso ha tal-

volta valori diversi, ma sempre in rapporto razionale semplice con questo. La spiegazione più ovvia è che qualche parte dell'atomo abbia un momento meccanico ed uno magnetico di origine diversa da quella dei moti orbitali, e che il rapporto di questi momenti sia diverso da quello ora detto. E si è trovato che le difficoltà anzidette si appianano mirabilmente se si ammette che ciascun elettrone possieda un momento angolare intrinseco (*spin*) uguale a mezza unità quan-

tistica, cioè a

$$p_0 = \frac{1}{2} \frac{h}{2\pi}$$

(come se ruotasse su sè stesso a guisa di trottola) ed un momento magnetico intrinseco avente direzione opposta ed il valore di un magnetone di Bohr, cioè

$$\mu_0 = \frac{eh}{4\pi mc} .$$

È questa l'ipotesi dell'*elettrone rotante*, cui si è accennato al § 25, p. I, dove si è spiegata la ragione di questa impropria denominazione. Si osservi che il rapporto μ_0/p_0 è la metà di quello analogo relativo ai moti orbitali.

Come si orienti l'asse dell'elettrone rotante in un campo magnetico non può naturalmente venire dedotto dalle ipotesi precedenti, ma richiede un'ipotesi nuova, la quale viene suggerita dal risultato del § 56 sulla quantizzazione spaziale delle orbite, e ne rappresenta la naturale generalizzazione. Per enunciare la regola di quantizzazione spaziale nella forma più generale, consideriamo un sistema il cui momento angolare totale sia rappresentato da un vettore j : usando, come faremo sempre in questo §, unità quantistiche, j sia un numero intero o semi-intero (cioè un numero intero più $1/2$). *Ammetteremo che questo sistema, posto in un campo magnetico, si orienti in modo che la proiezione di j sulla direzione del campo abbia uno dei $2j + 1$ valori discreti compresi tra $-j$ e $+j$ (gli estremi inclusi) e spazati di una unità, cioè:*

$$(347) \quad -j, -(j-1), \dots, (j-1), j .$$

I valori estremi corrispondono evidentemente, nel modello intuitivo, a j antiparallelo, o parallelo, al campo. Naturalmente, se j è intero, anche i valori della serie (347) risultano interi: se j è semi-intero, risultano semi-interi.

Questa regola, applicata al momento orbitale di un elettrone, riconduce al risultato del § 56: applicata invece al momento derivante dal solo spin, nel qual caso $j = 1/2$, dà evidentemente il risultato seguente: la proiezione dello spin nella direzione del campo può assumere solo i due valori $\pm 1/2$.

Cioè: lo spin dell'elettrone è sempre o parallelo o antiparallelo al campo.

Si vedrà nel cap. V, p. III come dalla meccanica quantistica si possano dedurre i valori di p_0 e μ_0 e delle loro proiezioni sulla direzione del campo senza introdurre ipotesi nuove: per ora però li ammetteremo come ipotesi giustificate a posteriori dall'esperienza (1).

Si osservi ora che l'elettrone dei sistemi idrogenoidi si trova, a causa del suo moto orbitale, immerso in un campo magnetico perpendicolare al piano dell'orbita: per comprenderlo si pensi che un osservatore il quale accompagnasse l'elettrone nel suo movimento vedrebbe il nucleo girare intorno all'elettrone stesso e quindi (poichè una carica elettrica in moto equivale ad una corrente) si vedrebbe circondato da una corrente elettrica e perciò si troverebbe immerso in un campo magnetico normale al piano di questa corrente (2). Per effetto di questo campo, lo spin si orienterà sempre normalmente al piano dell'orbita, in un verso o nell'altro, e perciò il momento angolare intrinseco dell'elettrone verrà ad aggiungersi (o a sottrarsi) a quello del moto orbitale che, nella nostra approssimazione, è l : il momento angolare risultante sarà dunque,

$$j = l \pm \frac{1}{2} \quad \text{se } l \neq 0 \text{ (orbite } p, d, f, \text{ ecc.)},$$

$$j = \frac{1}{2} \quad \text{se } l = 0 \text{ (orbite } s).$$

Per distinguere tra loro i due stati dell'atomo, corrispondenti alle due orientazioni dello spin, basta aggiungere alla indicazione dei tre numeri quantici n, l, m che caratteriz-

(1) Si vedrà inoltre che il « modello vettoriale » di cui qui ci serviamo (vettore suscettibile di orientazioni discrete) non rappresenta del tutto fedelmente le proprietà dello spin: tuttavia esso è, come tutti i modelli, molto utile per aiutare l'intuizione e semplificare il linguaggio.

(2) Si giunge alla stessa conclusione ricordando dalla teoria della relatività che, se rispetto ad un certo sistema di riferimento, che diremo fisso, esiste un campo elettrico, rispetto ad un altro sistema, in moto rispetto a quello, esiste un campo elettrico e magnetico: nel caso nostro, il campo elettrico del sistema fisso è quello prodotto dal nucleo.

ziano l'orbita, quella del momento angolare risultante, che, espresso in unità quantistiche, si indica con j e si chiama *quanto interno*: esso può assumere, nel caso che consideriamo, soltanto i due valori (semi interi) $j = l \pm \frac{1}{2}$ se $l \neq 0$, e $j = \frac{1}{2}$ se $l = 0$: in quest'ultimo caso, del resto, non vi è luogo a considerare le due orientazioni dello spin come corrispondenti a due stati distinti dell'atomo, perchè, essendo nullo il vettore \mathbf{p} (ed anche, in conseguenza, il campo magnetico), le due orientazioni che può assumere lo spin non sono fisicamente distinte.

Ciò che abbiamo detto si applica non solo ai sistemi idrogenoidi, ma anche all'elettrone ottico dei metalli alcalini, poichè in questi atomi i campi magnetici generati dai $Z - 1$ elettroni del nocciolo si elidono: invece negli altri atomi vi è da tener conto degli spin e dei campi magnetici generati anche dagli elettroni del nocciolo (o da alcuni di essi) e quindi le possibilità sono più numerose. Per tali argomenti rimandiamo al volume sulla spettroscopia: ci limitiamo a dire che il momento angolare complessivo di tutti gli elettroni dell'atomo (dovuto sia ai loro moti orbitali, sia ai loro spin) è un vettore costante la cui grandezza risulta generalmente espressa (in unità quantistiche) da un numero intero o semi-intero J , che chiamasi *quanto interno* dell'atomo: esso si identifica, nei metalli alcalini e nei sistemi idrogenoidi, col quanto interno j dell'elettrone ottico.

Tornando ora ai sistemi idrogenoidi ed ai metalli alcalini, si osservi che ai due valori del quanto interno j corrispondono due livelli energetici leggermente diversi. Difatti, assumendo l'elettrone a un magnete di momento μ_0 si riconosce che esso, trovandosi nel campo \mathbf{H} per effetto del suo moto orbitale, possiede per questo fatto l'energia potenziale $\mu_0 \times \mathbf{H}$: quindi alle due orientazioni che può assumere lo spin corrispondono i due valori

$$\varepsilon_1 = \mu_0 \bar{H} \quad , \quad \varepsilon_2 = -\mu_0 \bar{H} \quad ,$$

dell'energia magnetica, da aggiungere all'energia cinetica e potenziale dell'atomo: con \bar{H} abbiamo indicato il valore medio del campo H lungo l'orbita. Però un esame più appro-

fondito, fatto tenendo conto della teoria della relatività, mostra che i valori precedenti devono essere ridotti a metà ⁽¹⁾. Con un calcolo, che non riportiamo, basato sulle ordinarie leggi dell'elettromagnetismo, e sulla legge del moto kepleriano, si trova, utilizzando le espressioni note del numero di Rydberg R e della costante di struttura fina α (form. 340):

$$(348) \quad \varepsilon_1 = \frac{1}{2} \alpha^2 Z^4 \frac{Rhc}{n^3 l^2}, \quad \varepsilon_2 = -\frac{1}{2} \alpha^2 Z^4 \frac{Rhc}{n^3 l^2}.$$

Ogni livello energetico quindi si sdoppia, per effetto dell'esistenza dello spin, in due livelli vicini: la differenza di essi è dell'ordine di $\alpha^2 Z^4 Rhc$, come quella prodotta dalla correzione relativistica (v. form. 341). Non insistiamo tuttavia sull'aspetto quantitativo di questa teoria, perchè la meccanica quantistica, come si vedrà, modifica sensibilmente le formule (348).

§ 63. - PRINCIPIO DI CORRISPONDENZA.

Vogliamo ora stabilire un confronto tra lo spettro che un sistema emette in base alla teoria di Bohr e Sommerfeld (spettro *quantistico*) e lo spettro (che chiameremo *classico*) che esso emetterebbe se, pur potendosi trovare solo negli stati quantici, esso irradiasse, in ciascuno di questi, secondo le leggi ordinarie dell'elettromagnetismo. Ricerchiamo perciò anzitutto quali sarebbero le frequenze emesse in quest'ultimo caso (che è, naturalmente, puramente fittizio).

Ricordiamo che secondo le leggi elettromagnetiche l'emissione di radiazione da parte di un sistema di cariche elettriche è determinata dalla variazione del suo « momento elettrico » il quale è funzione delle coordinate lagrangiane q , che sono funzioni periodiche del tempo, ciascuna con la frequenza $\nu_i = \frac{1}{T_i}$: segue che ciascuna delle componenti X , Y , Z del momento si può sviluppare in *serie multipla* di Fourier, sotto

⁽¹⁾ V. L. H. THOMAS, *Phil. Mag.* 3, (1927) p. 1; J. FRENKEL, *ZS. f. Phys.* 37, (1926) p. 243.

la forma

$$(349) \quad X(t) = \sum_{-\infty}^{\infty} A_{\tau_1, \tau_2, \dots, \tau_f} e^{2\pi i(\tau_1 \nu_1 + \dots + \tau_f \nu_f)t}$$

(e analogamente per le componenti Y e Z).

Difatti consideriamo per un momento X come funzione della sola coordinata q_1 , e teniamo costanti le altre coordinate: la X sarà allora una funzione periodica del tempo, con la frequenza ν_1 , e potrà svilupparsi nella serie semplice di Fourier,

$$X = \sum_{-\infty}^{\infty} a_{\tau_1} e^{2\pi i \tau_1 \nu_1 t},$$

dove i coefficienti a_{τ_1} sono funzioni di q_2, q_3, \dots, q_f . A ciascuno di questi coefficienti possiamo ora applicare lo stesso procedimento, considerandolo funzione della sola q_2 e riguardando come costanti q_3, \dots, q_f : avremo allora

$$a_{\tau_1} = \sum_{-\infty}^{\infty} b_{\tau_1 \tau_2} e^{2\pi i \tau_2 \nu_2 t},$$

e quindi, sostituendo nell'espressione di X , questa diventa una serie doppia

$$X = \sum_{-\infty}^{\infty} b_{\tau_1 \tau_2} e^{2\pi i(\tau_1 \nu_1 + \tau_2 \nu_2)t},$$

dove i coefficienti $b_{\tau_1 \tau_2}$ dipendono da q_3, \dots, q_f .

Applicando successivamente lo stesso procedimento si giunge evidentemente allo sviluppo (349).

Ora, se le componenti del momento si decompongono in una somma di termini sinusoidali, anche ciascuna componente del campo elettrico e magnetico generato si decompone nello stesso modo, e quindi la radiazione emessa dal sistema nello stato considerato consta della sovrapposizione di radiazioni monocromatiche, le cui frequenze (che indicheremo con ν_{cl}) sono date da tutti i valori che assume l'espressione

$$(350) \quad \nu_{cl} = \tau_1 \nu_1 + \tau_2 \nu_2 + \dots + \tau_f \nu_f,$$

dando a $\tau_1, \tau_2, \dots, \tau_f$ tutti i valori interi (positivi o negativi) che non rendono negativo il secondo membro. L'intensità

di ciascuna di queste componenti monocromatiche risulta proporzionale a

$$(351) \quad (A\tau_1\tau_2\dots\tau_f)^2 + (B\tau_1\tau_2\dots\tau_f)^2 + (C\tau_1\tau_2\dots\tau_f)^2,$$

dove con $B_{\tau_1\tau_2\dots\tau_f}$ e $C_{\tau_1\tau_2\dots\tau_f}$ si sono indicati i coefficienti dello sviluppo di Y e Z analogo a (349). In particolare, se per qualcuno dei gruppi di valori $\tau_1\tau_2\dots\tau_f$ si annullano tutti e tre i coefficienti A, B, C , manca nella radiazione emessa la corrispondente componente monocromatica.

Lo spettro fittizio che abbiamo convenuto di chiamare classico si compone, perciò, di righe che sono individuate da due gruppi di indici: n_1, n_2, \dots, n_f (che definiscono lo stato del sistema, e quindi le frequenze $\nu_1, \nu_2, \dots, \nu_f$) e $\tau_1, \tau_2, \dots, \tau_f$. Perciò, a ciascuna riga dello spettro classico si può far corrispondere una riga dello spettro quantistico, e precisamente quella emessa nel salto dallo stato di numeri quantici n_1, n_2, \dots, n_f allo stato di numeri $n_1 - \tau_1, n_2 - \tau_2, \dots, n_f - \tau_f$. Vediamo ora quale è la frequenza di questa riga, calcolata secondo la teoria dei quanti.

Ricordiamo perciò anzitutto dalla meccanica razionale (v. § 52) che l'energia E di un sistema multiplamente periodico si può esprimere nella forma

$$(352) \quad E(J_1, J_2, \dots, J_f),$$

cioè in funzione degli integrali di fase J_i (che sostituiscono le f costanti di integrazione a_i). Si può poi dimostrare che le derivate parziali di questa funzione sono eguali alle frequenze ν_i corrispondenti ai singoli gradi di libertà:

$$(353) \quad \frac{\partial E}{\partial J} = \nu_i; \quad (i = 1, 2, \dots, f).$$

Ciò premesso, nel salto quantico di cui sopra gli integrali di fase passano dai valori $J_i = n_i h$ ai valori $(n_i - \tau_i) h$: perciò l'energia emessa è

$$\Delta E = E(J_1, J_2, \dots, J_f) - E(J_1 - \tau_1 h, J_2 - \tau_2 h, \dots, J_f - \tau_f h),$$

ovvero, applicando il teorema del valor medio

$$\Delta E = \frac{\partial E}{\partial J_1} \tau_1 h + \frac{\partial E}{\partial J_2} \tau_2 h + \dots + \frac{\partial E}{\partial J_f} \tau_f h,$$

dove la sopralineatura sta ad indicare che le derivate si riferiscono a uno stato (anche non quantico) intermedio tra quello iniziale e quello finale. Tenendo presente la (353) si può scrivere (con analogo significato della sopralineatura)

$$\Delta E = \bar{\nu}_1 \tau_1 h + \bar{\nu}_2 \tau_2 h + \dots + \bar{\nu}_f \tau_f h.$$

Quindi la frequenza emessa nel salto quantico considerato è

$$(354) \quad \nu_{qu} = \frac{\Delta E}{h} = \bar{\nu}_1 \tau_1 + \bar{\nu}_2 \tau_2 + \dots + \bar{\nu}_f \tau_f.$$

Confrontando questa formula con la (350), si osservi che, se i due stati, iniziale e finale, differiscono di poco (cioè se i numeri τ_i sono piccoli rispetto agli n_i , per il che questi necessariamente debbono essere grandi) le frequenze $\bar{\nu}_i$ si avvicinano molto alle frequenze ν_i corrispondenti allo stato iniziale, e quindi la frequenza ν_{qu} della riga dello spettro quantistico diviene approssimativamente uguale alla frequenza ν_{cl} della riga corrispondente nello spettro classico.

Si suol dire perciò che, *al limite per numeri quantici elevati, le righe « corrispondenti »* (nel senso spiegato sopra) *hanno la stessa frequenza* ⁽¹⁾. Questo teorema, che abbiamo dedotto dai postulati fondamentali della teoria di Bohr e Sommerfeld, prova che questa teoria soddisfa (per ciò che riguarda la previsione delle frequenze spettrali) al requisito essenziale che si deve richiedere da ogni meccanica atomica, e cioè di avere per limite la meccanica e l'elettromagnetismo ordinari.

Quanto abbiamo detto ora si riferisce solo alle frequenze delle righe spettrali, non alla loro intensità ed al loro stato di polarizzazione, che, mentre per lo spettro classico si sanno ricavare dai coefficienti dello sviluppo di Fourier, sono indeterminati nella teoria quantistica di Sommerfeld. Questa

(1) Si osservi che il teorema vale con la stessa approssimazione anche se si fa corrispondere la riga dello spettro quantistico, di frequenza (354), alla riga dello spettro classico avente gli stessi indici τ_i , ma le ν_i relative allo stato finale anziché a quello iniziale (o anche, relative ad un qualunque stato intermedio). Vi è cioè una certa arbitrarietà nella scelta del criterio di corrispondenza.

lacuna è stata parzialmente colmata da BOHR, aggiungendo alla teoria un nuovo postulato, suggerito dalla considerazione che il requisito di cui sopra deve essere soddisfatto non solo nei riguardi della frequenza ma anche delle altre proprietà della radiazione emessa. Tale postulato, noto col nome di *principio di corrispondenza*, è il seguente: *le righe corrispondenti (nel senso spiegato sopra) hanno intensità e stato di polarizzazione simili, e la somiglianza tende a diventare identità col crescere dei numeri quantici*. Calcolando quindi coi metodi dell'elettromagnetismo le intensità e le condizioni di polarizzazione delle righe dello spettro classico, si hanno delle indicazioni approssimative sugli analoghi elementi delle righe effettivamente emesse: perciò l'intensità della riga di frequenza (354) deve risultare, secondo il principio di corrispondenza, all'incirca proporzionale all'espressione (351).

Questo principio, nonostante il suo carattere soltanto qualitativo, ha fornito risultati brillantissimi ed ha avuto importanza fondamentale nello sviluppo della teoria dei quanti, tanto più che in certi casi (come si vedrà nel § seguente) esso permette di arrivare a risultati quantitativi e non solo qualitativi. Aggiungeremo che l'attuale teoria dell'irraggiamento, fondata sulla meccanica quantistica, permette non solo di giustificare il principio ora enunciato, ma anche di sostituirlo con una valutazione esatta dell'intensità e dello stato di polarizzazione delle righe, la quale si trova d'accordo coi risultati sperimentali come si è accennato al § 32.

§ 64. — PRINCIPIO DI SELEZIONE.

Una delle più importanti applicazioni del principio di corrispondenza si ha nel caso in cui, nello spettro classico, risultano nulle le intensità delle righe di indici τ_i che sarebbero emesse sia nello stato di numeri quantici n_i , sia in quello di numeri quantici $(n_i - \tau_i)$, sia in tutti quelli intermedi. Allora la riga quantistica emessa nel salto $n_i \rightarrow n_i - \tau_i$ corrisponde sempre ad una riga classica di intensità nulla, qualunque sia il criterio adottato per definire la corrispondenza:

si può in tal caso concludere che la sua intensità è nulla, cioè, che il salto quantico stesso non può avvenire, ossia è *proibito*. È questo, come si vede, uno dei casi in cui il principio di corrispondenza conduce ad una affermazione precisa e non solo qualitativa, e questa si trova confermata dall'esperienza. In tale caso particolare il principio di corrispondenza prende il nome di *principio di selezione*.

Aggiungiamo che quando nello spettro classico risulta nulla, sia nello stato iniziale che in quello finale ed in quelli intermedi, una sola (o due) delle componenti X , Y , Z del momento corrispondente ad una data riga, si possono trarre conclusioni sullo stato di polarizzazione della radiazione quantistica, come apparirà dagli esempi successivi.

Vediamo ora alcune applicazioni del principio di selezione.

a) *Oscillatore armonico*. — Il momento elettrico ha in tal caso la sola componente X , data (se la particella mobile porta la carica e) da

$$X = e\dot{x} = eA \sin(2\pi\nu_0 t - \varphi) = \frac{eA}{2i} [e^{i\varphi} e^{2\pi i\nu_0 t} - e^{-i\varphi} e^{-2\pi i\nu_0 t}] .$$

Lo sviluppo di Fourier si riduce dunque ai soli due termini di frequenza ν_0 , e cioè mancano tutti i termini in cui l'indice τ non è uguale a ± 1 . Saranno dunque possibili solo i salti quantici in cui n varia di ± 1 , e tutti gli altri saranno vietati. Si osservi che, essendo $E_n = \left(n + \frac{1}{2}\right) h\nu_0$, la frequenza che viene emessa (o assorbita) in questi salti è ν_0 , cioè lo spettro quantistico si riduce ad una sola riga, la cui frequenza è esattamente uguale a quella dello spettro classico, ossia alla frequenza propria dell'oscillatore.

Siccome poi sono evidentemente sempre nulle le componenti Y e Z del momento elettrico, la luce emessa nel salto quantico risulta sempre polarizzata lungo l'asse x , cioè parallelamente alla direzione delle oscillazioni.

b) *Rotatore*. — Consideriamo un rotatore contenente delle cariche elettriche e_i , poste a distanza r_i dall'asse: prendendo come asse z l'asse di rotazione, le componenti del momento elettrico sono

$$\begin{aligned} X &= \sum_i e_i r_i \cos(2\pi\nu_0 t - \varphi_i) , \\ Y &= \sum_i e_i r_i \sin(2\pi\nu_0 t - \varphi_i) , \\ Z &= 0 , \end{aligned}$$

dove ν_0 è la frequenza del moto di rotazione e le φ_i sono delle costanti: le somme contengono tanti termini quante sono le cariche. Alle prime due componenti si può applicare la considerazione già fatta per l'oscillatore, e si riconosce che, essendo presente nello sviluppo di Fourier la sola frequenza ν_0 , sono possibili solo i salti quantici in cui n varia di ± 1 . Tenendo presente l'espressione (315) dei livelli energetici, si vede che la frequenza emessa nel salto $n \rightarrow n - 1$ è

$$(355) \quad \nu_n = (2n + 1) \frac{h}{8\pi^2 J} .$$

Lo spettro emesso consta quindi di infinite righe, equidistanti (nella scala delle frequenze). Tale risultato è assai importante per la teoria degli spettri di bande.

Riguardo allo stato di polarizzazione, poichè il momento elettrico ruota nel piano xy , nello spettro classico ogni riga apparirebbe polarizzata circolarmente se la luce si osservasse nel verso dell'asse z , e rettilineamente (e cioè parallelamente al piano xy) se si osservasse in direzione perpendicolare a questa, mentre nelle direzioni intermedie si avrebbe polarizzazione ellittica. Poichè ciò vale per tutte le righe dello spettro classico, si può dire altrettanto delle righe effettivamente emesse.

c) Regola di selezione per il quanto azimutale. — Appliciamo ora il principio di selezione al moto centrale di un elettrone ottico, supponendo che il campo generato dal nocciolo sia abbastanza vicino al tipo newtoniano per poter decomporre il movimento in un moto kepleriano più una precessione uniforme, come si è fatto al § 59: quanto diremo si applica, in particolare, ai sistemi idrogenoidi in cui la precessione è dovuta solo alla lieve correzione relativistica.

Assumiamo un sistema di coordinate cartesiane con gli assi x ed y nel piano (fisso) dell'orbita: il loro legame con le coordinate polari r , ω si può riassumere nella formula

$$(356) \quad x + iy = r e^{i\omega} .$$

Ora si osservi che, a causa del movimento di precessione, r ed ω non hanno lo stesso periodo: indicando con ν_r e ν_ω le rispettive frequenze, e con ε la frequenza del lento moto di precessione, si ha

$$(357) \quad \nu_\omega = \nu_r + \varepsilon .$$

Ora introduciamo un sistema di assi x' , y' , ruotanti in modo da accompagnare il moto di precessione: rispetto a questi il moto sarà periodico e quindi l'angolo ω' del raggio vettore con

l'asse x' sarà, al pari di r , una funzione periodica con frequenza ν_r , cosicchè si potrà scrivere lo sviluppo di Fourier

$$(358) \quad r e^{i\omega'} = \sum_{-\infty}^{\infty} A_\tau e^{2\pi i \tau \nu_r t}$$

Siccome poi

$$\omega = \omega' + 2\pi \varepsilon t ,$$

dalla (356) e dalla (358), si ricava

$$(359) \quad x + iy = r e^{i\omega'} e^{2\pi i \varepsilon t} = \sum_{-\infty}^{\infty} A_\tau e^{2\pi i (\tau \nu_r + \varepsilon) t} ,$$

ovvero, ricavando ε dalla (357) e ponendo $\tau_r = \tau - 1$,

$$x + iy = \sum_{-\infty}^{\infty} A_{\tau_r} e^{2\pi i (\tau_r \nu_r + \nu_\omega) t} ,$$

Per ottenere gli sviluppi di x ed y separatamente, basterebbe scrivere l'espressione coniugata della precedente, ed operare per addizione e sottrazione. Si troveranno evidentemente due serie contenenti tutti termini del tipo

$$e^{\pm 2\pi i (\tau_r \nu_r + \nu_\omega) t} .$$

Si è così ottenuto lo sviluppo delle coordinate x , y (e quindi delle componenti del momento elettrico) nella forma (349), e si riconosce che, mentre τ_r può assumere tutti i valori interi, il coefficiente di ν_ω assume solo i valori ± 1 . Applicando dunque il principio di selezione, e ricordando che il numero quantico legato alla coordinata ω è il quanto azimutale l , si vede che sono possibili solo i salti quantici in cui il quanto azimutale l (e quindi k) varia di ± 1 : il che esprimeremo scrivendo

$$(360) \quad \Delta l = \pm 1 ,$$

È questa la regola di selezione per il quanto azimutale, che abbiamo già trovato, con la meccanica ondulatoria, al § 50, e che ha importanza fondamentale. Difatti, riferendosi allo schema dei termini rappresentato in fig. 45, essa esprime che sono possibili solo i salti quantici tra due colonne contigue dello schema: ciò che riduce enormemente la complicazione degli spettri. In altre parole, i termini s si possono combinare solo coi termini p , i termini p solo con gli s e coi d , ecc.

Non vi è invece nessuna limitazione nelle variazioni del quanto radiale n_r , e quindi nemmeno del quanto totale n .

Si tenga presente però che il ragionamento precedente vale

nell'ipotesi che il moto dell'elettrone sia poco diverso da quello kepleriano, e quindi sono da aspettarsi delle eccezioni alla regola nel caso di termini fortemente diversi dai balmeriani, come pure se l'atomo è soggetto a forti perturbazioni (campo elettrico o magnetico): in questi casi spesso si osservano difatti nello spettro delle righe proibite, ma esse sono generalmente, come è da aspettarsi, di debole intensità.

d) *Regole di selezione per il quanto magnetico e per il quanto interno.* — Un ragionamento analogo al precedente può farsi per il quanto magnetico: è necessario però, affinché questo intervenga effettivamente, supporre che l'atomo si trovi in un campo magnetico, sia pur debolissimo, così che resti fissata una direzione privilegiata nello spazio (v. § 53), che si assume come direzione dell'asse z : il campo magnetico produce un lento movimento di precessione intorno a questo asse (teor. di Larmor). Si dimostra allora, in modo simile al precedente, che per il quanto magnetico m vale la regola di selezione

$$\Delta m = 0, \pm 1,$$

che abbiamo già dimostrato al § 50 mediante la meccanica ondulatoria: inoltre, si ritrovano le regole di polarizzazione enunciate al § 50.

Infine, per il quanto interno j si trova, con considerazioni analoghe, la stessa regola di selezione

$$\Delta j = 0, \pm 1,$$

con l'aggiunta però che è vietato il passaggio da $j = 0$ a $j = 0$.

PARTE III.

METODI GENERALI DELLA MECCANICA QUANTISTICA

CAPITOLO I.

INTRODUZIONE MATEMATICA

§ 1. - LO SPAZIO FUNZIONALE.

È noto che conviene spesso designare un insieme di N numeri f_1, f_2, \dots, f_N come un *punto* P in uno spazio a N dimensioni (riferito ad assi cartesiani numerati da 1 ad N), ovvero come un *vettore* $\mathbf{f} = OP$ nello stesso spazio, chiamando O l'origine degli assi ⁽¹⁾. Si può allora, con ovvia generalizzazione di nozioni elementari, definire la *lunghezza* o *modulo* $|\mathbf{f}|$ del vettore \mathbf{f} mediante la formula:

$$|\mathbf{f}| = \sqrt{f_1^2 + f_2^2 + \dots + f_N^2},$$

si può definire il *prodotto scalare* di due vettori \mathbf{f}, \mathbf{g} :

$$\mathbf{f} \times \mathbf{g} = f_1 g_1 + f_2 g_2 + \dots + f_N g_N.$$

Si possono inoltre definire in modo ovvio le operazioni di somma, differenza, ecc.

Se poi si vogliono estendere queste considerazioni anche al caso che $f_1, f_2 \dots f_N$ siano numeri complessi, è opportuno sostituire le formule precedenti con le seguenti che, nel caso reale, si riducono a quelle (come al solito, l'asterisco denota il complesso coniugato):

$$(1) \quad |\mathbf{f}| = \sqrt{f_1 f_1^* + f_2 f_2^* + \dots + f_N f_N^*}$$

$$(2) \quad \mathbf{f} \times \mathbf{g} = f_1 g_1^* + f_2 g_2^* + \dots + f_N g_N^*.$$

⁽¹⁾ In tutto questo capitolo si tratterà solo di vettori uscenti dall'origine: perciò ad ogni punto corrisponde un vettore, e viceversa.

Possiamo dunque dire che: assegnare un vettore nello spazio a N dimensioni, significa far corrispondere ad ogni intero r (da 1 ad N) un numero (reale o complesso) f_r , che è la componente r -esima del vettore.

Vogliamo ora estendere queste considerazioni introducendo uno spazio con infinite dimensioni. Consideriamo perciò, invece degli N valori dell'indice r da 1 ad N , gli infiniti valori che può assumere una variabile reale x in un intervallo (a, b) : potremo dire che consideriamo, invece di N assi, una *infinità (continua)* di assi coordinati, corrispondenti ciascuno a un valore di x . Assegnare un vettore \mathbf{f} in questo spazio, significa far corrispondere ad ogni valore di x (da a a b) un numero (reale o complesso), cioè significa assegnare una funzione $f(x)$, monodroma da a a b . Ciò suggerisce di considerare ogni funzione $f(x)$, definita e monodroma in un intervallo a, b (eventualmente infinito) come un *vettore f* in uno spazio a infinite dimensioni, in cui ognuno dei valori di x da a a b caratterizza un asse coordinato e il valore corrispondente assunto dalla funzione rappresenta la proiezione del vettore su quell'asse (componente x -esima del vettore).

Questo spazio si chiama perciò *spazio funzionale*. Si può anche dire che la funzione $f(x)$ è rappresentata da un *punto* nello spazio funzionale e, viceversa, ogni punto di questo spazio rappresenta una funzione $f(x)$; ma spesso è più utile la interpretazione vettoriale.

Quello che abbiamo detto ora per una funzione di *una* variabile x , si può estendere senza difficoltà ad una funzione di p variabili $f(x_1, x_2, \dots, x_p)$, definita e monodroma entro un certo campo S (eventualmente infinito): lo spazio funzionale avrà in questo caso ∞^p dimensioni, ed ogni suo asse sarà denominato mediante un gruppo di p numeri: la relativa componente del vettore \mathbf{f} sarà il valore che assume la funzione f in corrispondenza di quei valori delle variabili indipendenti. In tutto quel che segue ci riferiremo, per maggiore generalità, a una funzione di p variabili, ma spesso indicheremo il loro complesso con la sola lettera x , e scriveremo $f(x)$ invece di $f(x_1, x_2, \dots, x_p)$.

Ai vettori dello spazio funzionale si trasportano imme-

diatamente le consuete definizioni di somma e differenza, nonchè di prodotto di un vettore per uno scalare ⁽¹⁾: p. es. *somma* dei due vettori \mathbf{f} e \mathbf{g} è il vettore la cui componente x -esima (o, nel caso di più variabili, la cui componente sull'asse x_1, x_2, \dots, x_p) è $f(x) + g(x)$, cioè il vettore rappresentante la funzione $f(x) + g(x)$. *Prodotto* del vettore \mathbf{f} per la costante c è il vettore $c\mathbf{f}$ rappresentante la funzione $cf(x)$.

Se poi è dato un certo numero n di vettori nello spazio funzionale, $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$, tutti i vettori ottenibili da essi mediante una combinazione lineare a coefficienti (costanti) arbitrari

$$c_1\mathbf{f}_1 + c_2\mathbf{f}_2 + \dots + c_n\mathbf{f}_n$$

(i quali rappresentano tutte le funzioni esprimibili come combinazioni lineari di $f_1(x), f_2(x) \dots$) si dice che formano una *varietà (o sottospazio) lineare ad n dimensioni*. È questa la naturale generalizzazione del concetto di piano o di retta (passanti per l'origine) dell'ordinario spazio tridimensionale, i quali si possono ritenere individuati mediante due vettori, o rispettivamente un vettore, uscenti dall'origine. Naturalmente, la stessa varietà lineare può essere individuata in infiniti modi, scegliendo per definirla un'altra ennupla di vettori ad essa appartenenti.

Restano da estendere le formule relative alla lunghezza e al prodotto scalare, per il che occorre restringere le nostre considerazioni a una classe più limitata di funzioni, come si dirà nel § seguente.

§ 2. — LO SPAZIO HILBERTIANO.

Per estendere al caso dello spazio a infinite dimensioni la (1), la (2) e le formule analoghe, si dovranno evidentemente sostituire le sommatorie da 1 ad N con integrali rispetto a x da a a b nel caso di una sola variabile, e, nel caso generale di p variabili, con integrali multipli estesi a tutto il campo S . Si sarà così condotti a definire come *modulo* del vettore \mathbf{f} rappresentante la funzione f (o, come anche si dice,

⁽¹⁾ In questo ordine di considerazioni, uno scalare significa una quantità costante (rispetto a x_1, x_2, \dots, x_p).

come *norma* della funzione f) il numero positivo $|f|$ il cui quadrato è dato dalla formula, analoga alla (1):

$$(3) \quad |f|^2 = \int f(x)f^*(x)dS,$$

dove si è indicato, come faremo sempre, con un semplice segno di integrazione l'integrale, generalmente multiplo, esteso a tutto il campo S , e con dS l'elemento di volume del campo, cioè $dS = dx_1 dx_2 \dots dx_p$. (Nel caso di una sola variabile, $dS = dx$ e l'integrale è semplice). Si noti che l'integrale (3) non è sempre convergente: siamo perciò condotti a considerare d'ora innanzi *solo quelle funzioni f , per le quali l'integrale (3) è convergente (funzioni a quadrato sommabile)*, cioè solo quei punti dello spazio funzionale per i quali la distanza dall'origine ha un valore determinato e finito. L'insieme di questi punti sarà chiamato *spazio hilbertiano*, e costituisce una parte dello spazio funzionale. Chiameremo poi *funzioni normalizzate* ⁽¹⁾ quelle la cui norma è 1, ossia quelle rappresentate da un vettore unitario (o *versore*); ogni funzione a quadrato sommabile si può normalizzare dividendola per la sua norma.

Estendiamo ora allo spazio hilbertiano la formula (2): *prodotto scalare* di due vettori f, g , rappresentanti le funzioni $f(x), g(x)$, o *prodotto scalare delle due funzioni* è il numero

$$(4) \quad f \times g = \int f(x)g^*(x)dS$$

(si può dimostrare che l'integrale è sempre convergente, in conseguenza della convergenza degli integrali che definiscono $|f|$ e $|g|$). Si osservi che il prodotto scalare (a differenza di ciò che avviene per vettori reali) *non* è in genere commutativo: mutando l'ordine dei fattori esso si cambia nel suo coniugato.

Si osservi anche che, se c è una costante

$$(5) \quad (cf) \times g = cf \times g,$$

mentre invece

$$(5') \quad f \times (cg) = c^*f \times g.$$

⁽¹⁾ Si noti che questa definizione coincide con quella già data al § 4 p. II per le autofunzioni normalizzate.

Alla definizione (3) del modulo di un vettore f o norma di una funzione f si può ora anche dare la forma seguente: essa è la radice quadrata di $f \times f$.

Proiezione del vettore f sul vettore g è il numero

$$(6) \quad f_g = f \times \text{vers } g = f \times \frac{g}{|g|}.$$

Condizione di ortogonalità dei due vettori f, g (o delle funzioni $f(x), g(x)$) è che sia $f \times g = 0$, cioè, in conseguenza della definizione (4), che sia

$$\int f(x)g^*(x)dS = 0.$$

Questa definizione di ortogonalità tra funzioni è già stata introdotta al § 5, p. II: ora si vede la ragione della denominazione. (Si noti che la condizione è indipendente dall'ordine dei due vettori).

Lo sviluppo di una funzione in serie di funzioni ortogonali (v. § 9, p. II) ha una notevole interpretazione nello spazio hilbertiano. Consideriamo dapprima, per semplicità, il caso di una sola variabile, e osserviamo che ognuna delle autofunzioni ortogonali normalizzate $y_n(x)$ (derivanti da una equazione differenziale del tipo già considerato al § 3, p. II:

$$(7) \quad \frac{d}{dx} \left(P \frac{dy}{dx} \right) + Qy = 0,$$

i cui autovalori supponiamo discreti) è rappresentata da un versore y_n , e che questi infiniti versori sono ortogonali tra loro: possiamo quindi dire che essi definiscono nello spazio hilbertiano un sistema di ∞ assi coordinati ortogonali (uno per ogni valore di n), allo stesso modo come una terna di versori i, j, k nello spazio ordinario definisce ⁽¹⁾ un sistema di assi cartesiani. Ora, con la notazione (4), la (32) della p. II si scrive

$$(8) \quad f_n = f \times y_n$$

e si interpreta così: il coefficiente f_n dello sviluppo della funzione f mediante le funzioni ortogonali y_n , è la proiezione

⁽¹⁾ Si sottintende che l'origine è fissata una volta per tutte.

del vettore \mathbf{f} sul versore \mathbf{y}_n , ossia è la componente di \mathbf{f} secondo l'asse coordinato n -esimo. E allora lo sviluppo (31, p. II), che si può scrivere vettorialmente

$$(8') \quad \mathbf{f} = \sum_1^{\infty} f_n \mathbf{y}_n,$$

acquista lo stesso significato che ha nello spazio ordinario la nota relazione

$$\mathbf{V} = V_x \mathbf{i} + V_y \mathbf{j} + V_z \mathbf{k}$$

tra un vettore \mathbf{V} e le sue componenti V_x, V_y, V_z .

Siamo stati così condotti a considerare, nello spazio hilbertiano, oltre al primitivo sistema di assi corrispondenti agli infiniti valori di x (che chiameremo « assi continui », essendo i valori di x una infinità continua), anche un altro sistema di assi (in numero infinito, ma discreti) definito dai versori \mathbf{y}_n ($n = 1, 2, \dots$) (e quindi, dipendente dall'equazione differenziale di cui le y_n sono le autofunzioni). Ogni funzione $f(x)$ a quadrato sommabile resta individuata ⁽¹⁾ assegnando i coefficienti c_n del suo sviluppo in serie di autofunzioni y_n : ciò significa che ogni vettore (o ogni punto) dello spazio hilbertiano si può individuare mediante le sue componenti (o le sue coordinate) relative agli assi definiti dai versori \mathbf{y}_n , anzichè quelle relative agli assi continui usati in principio. E ciò mostra che lo spazio hilbertiano ha una infinità *numerabile* di dimensioni, e non una infinità *continua*, come lo spazio funzionale di cui è una parte.

Riferendosi agli assi \mathbf{y}_n , la lunghezza $|\mathbf{f}|$ del vettore \mathbf{f} può essere calcolata mediante la formula

$$(9) \quad |\mathbf{f}|^2 = \sum_1^{\infty} f_n f_n^*$$

⁽¹⁾ Veramente, due funzioni i cui valori siano uguali dappertutto, tranne in alcuni punti x costituenti un aggregato di misura nulla, hanno evidentemente gli stessi coefficienti c_n : perciò è opportuno considerare due funzioni siffatte come rappresentate dallo stesso vettore (o punto) dello spazio hilbertiano.

Aggiungiamo inoltre che la $f(x)$ resterebbe individuata (nel senso chiarito sopra) dalle c_n anche se la serie $\sum c_n y_n$ non fosse convergente assolutamente, ma solo *in media*, nel senso spiegato nella nota a pag. 105.

che, per la formula di Parseval (v. p. 106), è equivalente alla (3). Similmente, si dimostra facilmente che il prodotto scalare $\mathbf{f} \times \mathbf{g}$ si può calcolare mediante le componenti f_n di \mathbf{f} e g_n di \mathbf{g} con la formula

$$(10) \quad \mathbf{f} \times \mathbf{g} = \sum_1^{\infty} f_n g_n^*,$$

e quindi la condizione di ortogonalità si può scrivere

$$(11) \quad \sum_1^{\infty} f_n g_n^* = 0.$$

Passando ora a considerare il caso di p variabili, dovremo modificare le considerazioni precedenti nel senso che ogni autofunzione è individuata non più da un solo indice n ma da p indici, cosicchè si dovrà scrivere $y_{n_1 n_2 \dots n_p}$: resta così definito nello spazio hilbertiano un sistema di ∞^p assi coordinati ortogonali (ciascuno dei quali è individuato mediante un gruppo di p numeri interi). Tutte le formule precedenti vanno allora modificate ponendo al posto di n il gruppo $n_1 n_2 \dots n_p$, e sostituendo le sommatorie semplici con sommatorie multiple: p. es. la (9) diverrà

$$(9') \quad |\mathbf{f}|^2 = \sum_1^{\infty} y_{n_1 n_2 \dots n_p} f_{n_1 n_2 \dots n_p} f_{n_1 n_2 \dots n_p}^*.$$

Le formule divengono quindi formalmente un po' più complicate, ma non si introducono difficoltà nuove. Perciò nel seguito continueremo a scrivere le formule relative al caso di $p = 1$, avvertendo che per passare al caso generale basta sostituire ogni indice con un gruppo di p indici, e ogni sommatoria semplice con una multipla.

§ 3. — OPERATORI LINEARI.

Chiameremo *operatore* qualsiasi simbolo che premesso ad una funzione (di una o più variabili ⁽¹⁾), la muti in un'altra funzione delle stesse variabili, secondo una legge determinata. La matematica elementare e il calcolo infinitesimale ne forniscono parecchi esempi, tra cui i seguenti:

⁽¹⁾ Talvolta un operatore è definito solo per certe determinate classi di funzioni, mentre per altre non ha senso. P. es., l'operatore $\frac{d}{dx}$ ha senso per le sole funzioni derivabili.

a) Ogni numero k si può riguardare come un operatore, perchè premesso ad una $f(x, y, \dots)$ la muta nel prodotto $kf(x, y, \dots)$. Ciò vale, naturalmente, anche se k è a sua volta una funzione. In particolare, 1 è un operatore che muta ogni funzione in sè stessa, e dicesi *identità*.

b) I simboli \log , \sin , \cos , ecc. sono altrettanti operatori, che mutano la funzione $f(x_1, x_2, \dots)$ nella funzione $\log f(x_1, x_2, \dots)$, $\sin f(x_1, x_2, \dots)$ ecc.

c) Il simbolo $\frac{d}{dx}$ è un operatore che muta ogni funzione derivabile $f(x)$ nella sua derivata. Similmente (per le funzioni di più variabili) sono operatori i simboli $\frac{\partial}{\partial x_1}$, $\frac{\partial}{\partial x_2}$ ecc.: così pure i simboli di derivata seconda, terza, ecc.

d) Il simbolo $\int_{x_0}^{x_1} dx_1$ (con x_0 costante) è un operatore che muta ogni funzione integrabile $f(x_1, x_2, \dots, x_p)$ nella funzione $\int_{x_0}^{x_1} f(x_1, x_2, \dots, x_p) dx_1$.

Un operatore generico viene indicato con una lettera: noi useremo di regola per questo scopo le lettere gotiche. Per esempio scriveremo $F = \mathfrak{A}f$ per indicare che l'operatore \mathfrak{A} applicato alla funzione f la muta nella funzione F .

Usando il linguaggio geometrico, possiamo dire che un operatore definisce una corrispondenza tra punti (o tra vettori) dello spazio funzionale: perciò talvolta scriveremo anche, p. es., $F = \mathfrak{A}f$. Noi supporremo sempre tale corrispondenza biunivoca.

Nel seguito ci occuperemo solo degli *operatori lineari* (o. l.) cioè di quelli che godono le due proprietà seguenti:

$$\alpha) \quad \mathfrak{A}(f + g) = \mathfrak{A}f + \mathfrak{A}g$$

dove f, g sono due funzioni qualunque (1);

$$\beta) \quad \mathfrak{A}(cf) = c\mathfrak{A}f$$

(1) Purchè, beninteso, siano tali che abbia senso l'applicazione dell'operatore \mathfrak{A} ad esse. Questa condizione si sottintenderà sempre nel seguito.

dove c è una costante ed f una funzione qualunque. Per esempio, tra gli operatori citati sopra, sono lineari gli operatori k , $\frac{d}{dx}$, $\frac{\partial}{\partial x_1}$, $\frac{\partial}{\partial x_2}$, $\int_{x_0}^{x_1} dx$, mentre non sono lineari gli operatori \log , \sin , \cos , ecc.

Nello spazio funzionale un o. l. stabilisce una corrispondenza tra vettori, che è la naturale generalizzazione di una *omografia vettoriale* dell'ordinario spazio tridimensionale.

La classe degli operatori lineari gode delle proprietà di alto interesse matematico (v. bibl. n. 33): noi ci limiteremo qui alle nozioni essenziali della loro teoria.

§ 4. — ALGEBRA DEGLI OPERATORI LINEARI.

Si possono definire delle operazioni di combinazione tra operatori lineari analoghe alle operazioni di somma, differenza, ecc. con cui si possono combinare tra loro le quantità algebriche, ciò che permette di costruire un'algebra degli o. l. analoga (benchè non identica) all'algebra ordinaria.

Dati due o. l. \mathfrak{A} e \mathfrak{B} , chiamasi loro *somma*, e si indica con $\mathfrak{A} + \mathfrak{B}$, l'operatore (lineare) definito da (1)

$$(\mathfrak{A} + \mathfrak{B})f = \mathfrak{A}f + \mathfrak{B}f.$$

In modo analogo si definisce la *differenza* di due o. l., e la somma di un numero qualunque di essi.

P. es., l'o. l. Δ , detto laplaciano, di cui si fa grande uso in fisica matematica, definito da

$$\Delta f(x, y, z) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2},$$

è la somma dei tre o. l. $\frac{\partial^2}{\partial x^2}$, $\frac{\partial^2}{\partial y^2}$, $\frac{\partial^2}{\partial z^2}$: si può scrivere cioè

$$(12) \quad \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

(1) Qui, e nel seguito, f è una funzione qualunque cui si possano applicare gli operatori in questione.

Un altro esempio notevole è l'operatore che figura nel primo membro dell'equazione di Schrödinger (131) p. II, la quale si può scrivere

$$(13) \quad \left[\Delta - \frac{8\pi^2 m}{h^2} (E - U) \right] \psi = 0.$$

Chiamasi *prodotto* dell'o. l. \mathfrak{A} per l'o. l. \mathfrak{B} , e si indica con $\mathfrak{A}\mathfrak{B}$, l'operatore esprimente l'operazione di applicare prima l'operazione \mathfrak{B} e poi, sulla funzione ottenuta, l'operazione \mathfrak{A} : vale a dire

$$\mathfrak{A}\mathfrak{B}f = \mathfrak{A}(\mathfrak{B}f).$$

Evidentemente, il prodotto di due o. l. è anch'esso lineare.

In generale, per tale prodotto *non* vale la proprietà commutativa, cioè l'operatore $\mathfrak{A}\mathfrak{B}$ non coincide con l'operatore $\mathfrak{B}\mathfrak{A}$: è questo che rende l'algebra degli operatori non identica all'algebra ordinaria. Quando avviene che i due o. l. $\mathfrak{A}\mathfrak{B}$ e $\mathfrak{B}\mathfrak{A}$ coincidano si dice che \mathfrak{A} e \mathfrak{B} sono *permutabili*.

Esempi. — Due fattori numerici k_1, k_2 (costanti o no) sono sempre operatori permutabili, perchè $k_1 k_2 f = k_2 k_1 f$. Così pure sono permutabili — di regola — gli o. l. $\frac{\partial}{\partial x}$ e $\frac{\partial}{\partial y}$, il cui prodotto nell'ordine indicato è $\frac{\partial^2}{\partial x \partial y}$ e nell'ordine inverso è $\frac{\partial^2}{\partial y \partial x}$.

Invece i due o. l.

$$(14) \quad \mathfrak{A} = x \quad \mathfrak{B} = \frac{\partial}{\partial x}$$

non sono permutabili, perchè

$$\mathfrak{A}\mathfrak{B}f = x \frac{\partial f}{\partial x}$$

$$\mathfrak{B}\mathfrak{A}f = \frac{\partial}{\partial x} (xf) = f + x \frac{\partial f}{\partial x}.$$

L'ultima eguaglianza si può anche scrivere

$$\mathfrak{B}\mathfrak{A}f = f + \mathfrak{A}\mathfrak{B}f = (1 + \mathfrak{A}\mathfrak{B})f$$

il che significa che per gli operatori \mathfrak{A} e \mathfrak{B} definiti dalle (14) vale, invece della proprietà commutativa, la seguente *formula di permutazione*:

$$(15) \quad \mathfrak{B}\mathfrak{A} - \mathfrak{A}\mathfrak{B} = 1.$$

La *potenza n-esima* (con n intero e positivo) di un o. l. \mathfrak{A} è definita ovviamente come il prodotto di n fattori uguali ad \mathfrak{A} . Si conviene poi che $\mathfrak{A}^0 = 1$.

Esempio. — La potenza n -esima dell'operatore $\frac{\partial}{\partial x}$ è $\frac{\partial^n}{\partial x^n}$: cioè

$$\left(\frac{\partial}{\partial x} \right)^n = \frac{\partial^n}{\partial x^n}.$$

Perciò la (12) si può anche scrivere

$$\Delta = \left(\frac{\partial}{\partial x} \right)^2 + \left(\frac{\partial}{\partial y} \right)^2 + \left(\frac{\partial}{\partial z} \right)^2.$$

Dato un o. l. \mathfrak{A} , se esiste un o. l. \mathfrak{B} tale che

$$\mathfrak{A}\mathfrak{B} = \mathfrak{B}\mathfrak{A} = 1,$$

si dice che \mathfrak{B} è il *reciproco* o *l'inverso* di \mathfrak{A} , e viceversa, e si scrive

$$\mathfrak{B} = \mathfrak{A}^{-1}, \quad \mathfrak{A} = \mathfrak{B}^{-1}.$$

Applicando successivamente a una funzione un operatore e il suo inverso, le due operazioni si elidono e si ritrova la funzione primitiva.

Se esiste l'inverso di \mathfrak{A} , si possono definire le potenze di \mathfrak{A} ad esponente negativo, ponendo

$$(16) \quad \mathfrak{A}^{-n} = (\mathfrak{A}^{-1})^n.$$

È evidente che valgono per gli operatori gli ordinari teoremi sulle potenze, p. es. $\mathfrak{A}^n \mathfrak{A}^m = \mathfrak{A}^{n+m}$ (n, m interi, positivi, nulli o negativi), ecc.

Si osservi che se \mathfrak{A} e \mathfrak{B} ammettono entrambi un reciproco, anche il loro prodotto $\mathfrak{A}\mathfrak{B}$ lo ammette, ed è $\mathfrak{B}^{-1}\mathfrak{A}^{-1}$ (si noti l'inversione dei fattori): difatti

$$(\mathfrak{A}\mathfrak{B})(\mathfrak{B}^{-1}\mathfrak{A}^{-1}) = \mathfrak{A}(\mathfrak{B}\mathfrak{B}^{-1})\mathfrak{A}^{-1} = \mathfrak{A}\mathfrak{A}^{-1} = 1.$$

Definite le potenze di un o. l. \mathfrak{A} , si possono definire altri o. l. detti *funzioni* di esso nel modo seguente. Sia $F(a)$ il simbolo di una funzione, sviluppabile in serie di potenze, di una variabile a , cioè sia, entro un certo cerchio di convergenza,

$$F(a) = a_0 + a_1 a + a_2 a^2 + \dots$$

Definiremo allora come $F(\mathfrak{A})$ l'o. l. ottenuto sostituendo materialmente, nella serie precedente, il simbolo α col simbolo \mathfrak{A} (con che ogni termine della serie acquista il significato di un o. l. ben determinato).

Esempio. — Prendiamo come \mathfrak{A} l'o. l. $\lambda \frac{\partial}{\partial x}$ (dove λ è una costante), e definiamo l'o. l. $e^{\mathfrak{A}}$ ossia $e^{\lambda \frac{\partial}{\partial x}}$. Poichè la funzione e^{α} è definita dalla serie

$$e^{\alpha} = 1 + \alpha + \frac{1}{2!} \alpha^2 + \frac{1}{3!} \alpha^3 + \dots$$

avremo, sostituendo α con $\lambda \frac{\partial}{\partial x}$,

$$e^{\lambda \frac{\partial}{\partial x}} = 1 + \lambda \frac{\partial}{\partial x} + \frac{\lambda^2}{2!} \frac{\partial^2}{\partial x^2} + \frac{\lambda^3}{3!} \frac{\partial^3}{\partial x^3} + \dots$$

L'operatore a secondo membro ha un'interpretazione assai notevole: esso, muta $f(x)$ in $f(x + \lambda)$, poichè, per la formula di Taylor,

$$f + \lambda \frac{\partial f}{\partial x} + \frac{\lambda^2}{2!} \frac{\partial^2 f}{\partial x^2} + \frac{\lambda^3}{3!} \frac{\partial^3 f}{\partial x^3} + \dots = f(x + \lambda).$$

Quindi scriveremo

$$e^{\lambda \frac{\partial}{\partial x}} f(x) = f(x + \lambda).$$

È evidente che un o. l. \mathfrak{A} è permutabile con qualunque propria potenza \mathfrak{A}^n , e quindi anche con una qualunque $F(\mathfrak{A})$.

Inoltre si vede immediatamente che: se un operatore \mathfrak{B} è permutabile con \mathfrak{A} , lo è anche con qualunque $F(\mathfrak{A})$.

Passiamo ora a definire una funzione di più o. l. $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}$, limitandoci (per semplicità di scrittura) al caso di due. Data una funzione sviluppabile di due variabili

$$(17) \quad F(\alpha, \beta) = \sum_{ik} a_{ik} \alpha^i \beta^k,$$

l'estensione immediata del caso precedente porterebbe a definire come $F(\mathfrak{A}, \mathfrak{B})$ l'o. l.

$$(18) \quad \mathfrak{F} = \sum_{ik} a_{ik} \mathfrak{A}^i \mathfrak{B}^k.$$

Si deve però osservare che, mentre questa definizione si può senz'altro adottare se \mathfrak{A} e \mathfrak{B} sono permutabili, richiede invece qualche avvertenza se non lo sono. Infatti, nel termine generale della (17) l'ordine dei fattori α e β è indifferente, talchè si potrebbe scrivere invece di $\alpha^i \beta^k$, p. es. $\alpha^{i-1} \beta^k \alpha$, o $\alpha^{i-2} \beta^k \alpha^2$, o anche in molti altri modi: a tutte queste scritte corrisponde sempre la stessa funzione $F(\alpha, \beta)$; invece, sostituendo α e β con due o. l. non permutabili \mathfrak{A} e \mathfrak{B} , si ottengono da queste serie tanti o. l. diversi. Bisogna dunque far corrispondere l'o. l. $\mathfrak{F} = F(\mathfrak{A}, \mathfrak{B})$ non alla funzione analitica $F(\alpha, \beta)$, ma a un particolare modo di scrivere questa funzione.

Si riconosce poi immediatamente che un o. l., funzione di uno o più o. l. $\mathfrak{A}, \mathfrak{B}, \dots$ permutabili tra loro, è permutabile con ciascuno di essi.

§ 5. — RAPPRESENTAZIONE DI UN OPERATORE LINEARE MEDIANTE UNA MATRICE.

È facile dimostrare che, se \mathfrak{A} è un o. l. che opera tra vettori dello spazio hilbertiano, le componenti del vettore $\mathfrak{A}\mathbf{f}$ (che indicheremo con \mathbf{F}) sono combinazioni lineari delle componenti di \mathbf{f} . Vale a dire: qualunque o. l. equivale a una sostituzione lineare sulle componenti del vettore cui è applicato (analogamente a ciò che avviene per le omografie vettoriali dello spazio ordinario).

Difatti, detti al solito \mathbf{y}_n i versori degli assi (autofunzioni di un'equazione differenziale) ogni vettore \mathbf{f} si può scrivere nella forma

$$(19) \quad \mathbf{f} = \sum_n f_n \mathbf{y}_n,$$

dove le f_n sono le componenti di \mathbf{f} , e la somma si intende estesa da 1 all' ∞ (come si sottintenderà anche nelle formule successive). Applicando l'operatore \mathfrak{A} , poichè questo è lineare, si ha

$$(20) \quad \mathbf{F} = \mathfrak{A}\mathbf{f} = \sum_n f_n \mathfrak{A}\mathbf{y}_n,$$

cosicchè basta saper applicare l'o. l. \mathfrak{A} ai versori fondamentali \mathbf{y}_n per saperlo applicare ad una \mathbf{f} qualunque. Ora, il vettore $\mathfrak{A}\mathbf{y}_n$ sarà individuato dalle sue componenti che in-

dicheremo con A_{mn} (dove il primo indice caratterizza la componente, mentre il secondo specifica il vettore di cui si tratta): potremo scrivere perciò

$$(21) \quad \mathfrak{A}y_n = \sum_m A_{mn} y_m.$$

Sostituendo nella (20) abbiamo

$$\mathbf{F} = \sum_{mn} f_n A_{mn} y_m,$$

ovvero, posto

$$(22) \quad F_m = \sum_n A_{mn} f_n,$$

avremo

$$(20') \quad \mathbf{F} = \sum_m F_m y_m.$$

Si vede di qui che le componenti del vettore $\mathfrak{A}f$ sono le F_m , le quali si ottengono dalle componenti f_n di \mathbf{f} mediante il sistema di (infinite) relazioni lineari (22), i cui coefficienti sono le A_{mn} . Basta quindi la conoscenza di questi coefficienti per permettere di ricavare da ogni \mathbf{f} il corrispondente $\mathfrak{A}f$: perciò l'insieme di questi coefficienti caratterizza completamente l'operatore \mathfrak{A} .

I coefficienti A_{mn} costituiscono una doppia infinità (discreta) di numeri (generalmente complessi), che si possono disporre in uno specchio (*matrice*) di infinite righe, caratterizzate dal primo indice, e infinite colonne, caratterizzate dal secondo indice:

$$\left\| \begin{array}{cccc} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & A_{23} & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \dots & \dots & \dots & \dots \end{array} \right\|.$$

Ad ogni o. l. \mathfrak{A} corrisponde così una matrice, che lo individua perfettamente, e che si indica generalmente con lo stesso simbolo \mathfrak{A} dell'operatore: spesso, quando sia necessario mettere in evidenza che si tratta della matrice, la indicheremo con $\{\mathfrak{A}\}$.

Gli *elementi* A_{mn} di questa matrice si possono calcolare, osservando che A_{mn} rappresenta la componente m -esima del vettore $\mathfrak{A}y_n$, e quindi (v. form. (8)):

$$(23) \quad A_{mn} = (\mathfrak{A}y_n) \times y_m.$$

Ovvero, ricordando la (4) e scrivendo, per evitare equivoci, prima il fattore con l'asterisco

$$(23') \quad A_{mn} = \int y_m^* \mathfrak{A} y_n dx$$

Esempio. — Sia \mathfrak{A} l'identità ($\mathfrak{A} = 1$). Allora la (23) ci dà

$$(24) \quad A_{mn} = \int y_n^* y_n dx = \delta_{mn}.$$

e quindi la matrice che rappresenta l'identità è

$$(25) \quad \left\| \begin{array}{cccc} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots \end{array} \right\|.$$

Questa si chiama *matrice unità* e si indica con $\{1\}$.

Osservazione. — Nel caso generale di p variabili, ogni elemento della matrice dovrà essere scritto, per quanto si è detto al § 2, nella forma $A_{m_1 m_2 \dots m_p, n_1 n_2 \dots n_p}$: si avrà cioè una matrice in cui le righe e le colonne sono caratterizzate non da due indici ma da due gruppi di p indici. Questo non porta nessuna difficoltà concettuale, ma solo una maggiore complicazione nella scrittura.

§ 6. — ALGEBRA DELLE MATRICI.

Poichè a ogni o. l. corrisponde (fissato il sistema di riferimento) una matrice, e viceversa, è evidente che dalle operazioni di somma, differenza, ecc. definite tra o. l., restano definite altrettante operazioni tra matrici. P. es., si chiamerà *somma* delle matrici $\{\mathfrak{A}\}$ e $\{\mathfrak{B}\}$, e si indicherà con $\{\mathfrak{A}\} + \{\mathfrak{B}\}$, la matrice $\{\mathfrak{A} + \mathfrak{B}\}$, cioè quella corrispondente all'o. l. somma di \mathfrak{A} e \mathfrak{B} . Similmente sarà definito il *prodotto* di due matrici, da cui deriverà anche la *potenza* n -esima (n intero e positivo) di una matrice. Del pari, si definirà una qualunque *funzione analitica* $F(\{\mathfrak{A}\})$ di una matrice, come la matrice corrispondente all'o. l. $F(\mathfrak{A})$, definito al § 4. E analogamente per una funzione di più matrici. Se poi \mathfrak{A} ammette un inverso \mathfrak{A}^{-1} , le matrici $\{\mathfrak{A}\}$ e $\{\mathfrak{A}^{-1}\}$ si chiameranno pure *inverse*, o *reciproche*: il loro prodotto (in qualunque ordine) è la matrice unità (25). La matrice $\{\mathfrak{A}^{-1}\}$ verrà

indicata con $\{\mathfrak{A}\}^{-1}$ e le sue potenze saranno considerate come potenze di $\{\mathfrak{A}\}$ ad esponente negativo.

È preferibile però dare, di queste operazioni tra matrici, una definizione equivalente a questa, ma indipendente dai rispettivi operatori: definiremo cioè le operazioni tra matrici mediante operazioni da eseguirsi sui loro elementi. Ecco come.

Gli elementi della matrice *somma* $\{\mathfrak{A}\} + \{\mathfrak{B}\}$ sono la somma degli elementi corrispondenti delle due matrici $\{\mathfrak{A}\}$ e $\{\mathfrak{B}\}$: cioè

$$\{\mathfrak{A}\} + \{\mathfrak{B}\} = \begin{vmatrix} A_{11} + B_{11} & A_{12} + B_{12} & A_{13} + B_{13} & \dots \\ A_{21} + B_{21} & A_{22} + B_{22} & A_{23} + B_{23} & \dots \\ A_{31} + B_{31} & A_{32} + B_{32} & A_{33} + B_{33} & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix}.$$

Dalla (23) risulta subito che la matrice somma, così definita, è effettivamente la matrice corrispondente all'operatore $\mathfrak{A} + \mathfrak{B}$.

Analogamente per la *differenza* di due matrici, e per la somma di quante si vogliono di esse.

La *moltiplicazione* di una matrice $\{\mathfrak{A}\}$ per una costante k si esegue moltiplicando ogni elemento della matrice per k : anche questo risulta immediatamente dalla (23).

Passiamo al *prodotto* di due matrici \mathfrak{A} , \mathfrak{B} . Chiamiamo \mathfrak{P} l'o. l.

$$(26) \quad \mathfrak{P} = \mathfrak{A}\mathfrak{B}$$

e calcoliamo mediante la (23) l'elemento generico della matrice prodotto $\{\mathfrak{P}\} = \{\mathfrak{A}\}\{\mathfrak{B}\}$:

$$(27) \quad P_{mn} = (\mathfrak{P}\mathbf{y}_n) \times \mathbf{y}_m = (\mathfrak{A}\mathfrak{B}\mathbf{y}_n) \times \mathbf{y}_m.$$

Calcoliamo il fattore entro parentesi: si ha, per la (21)

$$\mathfrak{B}\mathbf{y}_n = \sum_i B_{in} \mathbf{y}_i$$

e quindi

$$\mathfrak{A}\mathfrak{B}\mathbf{y}_n = \sum_i B_{in} \mathfrak{A}\mathbf{y}_i = \sum_i B_{in} \sum_k A_{ki} \mathbf{y}_k = \sum_k (\sum_i A_{ki} B_{in}) \mathbf{y}_k.$$

Sostituendo nella (27), e ricordando che le \mathbf{y} sono ortogonali e normalizzate, si ha

$$(28) \quad P_{mn} = \sum_i A_{mi} B_{in}.$$

(si osservi, sotto il punto di vista mnemonico, che scrivendo, come si è fatto, i due fattori A_{mi} e B_{in} nell'ordine corrispondente a quello della (26), i due indici m e n figurano nello stesso ordine nei due membri, e l'indice di sommatoria resta in mezzo ad essi).

La formula (28) equivale, come si vede facilmente, alla seguente regola: « il prodotto di due matrici si effettua con la nota regola del prodotto di due determinanti, ma moltiplicando sempre le righe della prima matrice per le colonne della seconda, non viceversa ».

Naturalmente il prodotto di due matrici *non* è commutativo, eccettuato il caso che i due operatori corrispondenti siano permutabili, nel qual caso pure le due matrici sono *permutabili*.

Si verifica subito, applicando la regola precedente alla matrice unità (25) e a un'altra matrice $\{\mathfrak{A}\}$ qualunque, che

$$\{1\}\{\mathfrak{A}\} = \{\mathfrak{A}\}\{1\} = \{\mathfrak{A}\}.$$

§ 6 bis. — RAPPRESENTAZIONE DI UNA FUNZIONE MEDIANTE UNA MATRICE DI UNA SOLA COLONNA.

Alle formule del § 5 si può dare un'interpretazione espressiva se si conviene di considerare le componenti f_n (che caratterizzano una funzione f rispetto a un certo sistema di funzioni ortogonali) come elementi di una matrice di una sola colonna (e di infinite righe), cioè se si scrive

$$f = \begin{vmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \end{vmatrix}$$

Allora non solo gli operatori, ma anche le funzioni sono rappresentate da matrici, e si verifica subito che per applicare a una funzione f un operatore \mathfrak{A} basta fare il prodotto delle corrispondenti matrici con la regola ordinaria, cioè « righe per colonne »: difatti applicando questa regola e te-

nendo presente la (22) si trova facilmente

$$(29) \quad \begin{vmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & A_{23} & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix} \cdot \begin{vmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \end{vmatrix} = \begin{vmatrix} F_1 \\ F_2 \\ F_3 \\ \vdots \end{vmatrix}.$$

§ 7. — CAMBIAMENTI DI ASSI NELLO SPAZIO HILBERTIANO.
MATRICI DI TRASFORMAZIONE.

Consideriamo ora, oltre alle autofunzioni y_n definite dall'equazione (7), un altro sistema completo di autofunzioni y'_n ⁽¹⁾ definite da un'altra equazione della stessa forma (relativa allo stesso intervallo e con le stesse condizioni agli estremi): esse definiranno un altro sistema di assi ortogonali, individuati dai versori y'_n . Rispetto a questi assi il vettore \mathbf{f} avrà certe componenti f'_n espresse, similmente alle f_n , da

$$(30) \quad f'_n = \mathbf{f} \times \mathbf{y}'_n,$$

e si potrà scrivere, analogamente alla (8'),

$$(31) \quad \mathbf{f} = \sum_n f'_n \mathbf{y}'_n.$$

Vogliamo ora cercare che relazione intercede tra le componenti del vettore \mathbf{f} rispetto ai nuovi e agli antichi assi, cioè tra le f'_n e le f_n . Cominciamo perciò con l'osservare che ciascuno dei versori \mathbf{y}'_n potrà essere individuato dalle sue componenti (rispetto agli assi antichi) che designeremo con S_{mn} (il primo indice specifica la componente, il secondo il versore di cui si tratta), talchè

$$(32) \quad \mathbf{y}'_n = \sum_m S_{mn} \mathbf{y}_m.$$

Sarà, per la (8)

$$(33) \quad S_{mn} = \mathbf{y}'_n \times \mathbf{y}_m$$

e le S_{mn} saranno legate dalle relazioni (v. form. (10)):

$$(34) \quad \sum_i S_{im} S_{in}^* = \delta_{mn},$$

esprimenti che le \mathbf{y}'_n sono ortogonali e normalizzate.

⁽¹⁾ È superfluo avvertire che l'apice qui (e in tutto questo §) non ha il significato di derivazione.

Ciò posto, dalla (30) e dalla (32) si ricava per le nuove componenti l'espressione seguente (si badi alla (5')):

$$f'_n = \sum_m S_{mn}^* \mathbf{f} \times \mathbf{y}_m,$$

ossia, per la (8)

$$(35) \quad f'_n = \sum_m S_{mn}^* f_m.$$

Per ottenere le formule inverse si potrebbero risolvere queste rispetto alle f_m , ma è più conveniente operare in modo simmetrico, cioè considerare le componenti dei versori \mathbf{y} rispetto agli assi \mathbf{y}' : le indicheremo con \tilde{S}_{mn} ponendo

$$(36) \quad \tilde{S}_{mn} = \mathbf{y}_n \times \mathbf{y}'_m,$$

cosicchè sia

$$(37) \quad \mathbf{y}_n = \sum_m \tilde{S}_{mn} \mathbf{y}'_m.$$

Si troverà allora, analogamente alla (35):

$$(38) \quad f_n = \sum_m \tilde{S}_{mn}^* f'_m.$$

Si osservi ora che

$$(39) \quad \tilde{S}_{mn} = (\mathbf{y}'_m \times \mathbf{y}_n)^* = S_{mn}^*,$$

cosicchè si può anche scrivere

$$(39'') \quad f_n = \sum_m S_{mn} f'_m.$$

Conviene considerare le S_{mn} come gli elementi di una matrice $\{\mathcal{S}\}$: diremo allora che dalle componenti di un vettore rispetto agli assi \mathbf{y} si passa alle sue componenti rispetto agli assi \mathbf{y}' mediante la matrice $\{\mathcal{S}\}$ (cioè mediante la sostituzione lineare (32)). Similmente, la matrice $\{\tilde{\mathcal{S}}\}$ fa passare dalle componenti \mathbf{y}' alle \mathbf{y} (mediante la (37)). E la (37) esprime che la matrice $\{\tilde{\mathcal{S}}\}$ è ottenuta dalla $\{\mathcal{S}\}$ cambiandone le righe in colonne (cioè sostituendo ogni elemento col suo simmetrico rispetto alla diagonale principale) e prendendo di ogni elemento il coniugato. Se facciamo il prodotto di queste due matrici ($\{\tilde{\mathcal{S}}\} \{\mathcal{S}\}$) con la regola del § 6, troviamo, per il suo elemento generico di indici m ,

n , ricordando la (37) e la (34),

$$\sum_i \widetilde{S}_{mi} S_{in} = \sum_i S_{im}^* S_{in} = \delta_{mn} :$$

potremo dunque scrivere

$$\{\widetilde{\mathfrak{S}}\}\{\mathfrak{S}\} = \{1\}.$$

Similmente si proverebbe che $\{\mathfrak{S}\}\{\widetilde{\mathfrak{S}}\} = 1$: quindi la matrice $\{\widetilde{\mathfrak{S}}\}$ è l'inversa della $\{\mathfrak{S}\}$:

$$(38) \quad \{\widetilde{\mathfrak{S}}\} = \{\mathfrak{S}\}^{-1}.$$

Questa proprietà, equivalente alla (34), caratterizza le matrici che diconsi *unitarie*.

Le matrici così introdotte non si considerano come rappresentanti di operatori, poichè non servono a passare da un vettore a un altro, ma invece dalle componenti di un vettore rispetto a un sistema di riferimento, alle componenti dello stesso vettore rispetto a un altro sistema di riferimento. Si chiamano perciò *matrici di trasformazione*.

Supponiamo che, dopo essere passati dal riferimento y al riferimento y' mediante la matrice $\{\mathfrak{S}\}$, si passi ad un terzo riferimento (completo e ortogonale) di versori y'' , mediante un'altra matrice di trasformazione $\{\mathfrak{T}\}$: possiamo allora dimostrare che si può passare direttamente dal riferimento y al riferimento y'' , mediante la matrice di trasformazione

$$\{\mathfrak{P}\} = \{\mathfrak{S}\}\{\mathfrak{T}\}.$$

Difatti, il passaggio dalle f' alle f'' sarà espresso dalla formula, analoga alla (35),

$$(39) \quad f''_l = \sum_n T_{nl}^* f'_n.$$

Sostituendo per f'_n la (35) si ha

$$f''_l = \sum_{mn} T_{nl}^* S_{mn}^* f_m = \sum_m (\sum_n S_{mn}^* T_{nl}^*) f_m$$

e ricordando la regola di moltiplicazione delle matrici (28):

$$(40) \quad f''_l = \sum_m P_{ml}^* f_m,$$

la quale, confrontata con la (35), mostra che si passa dalle f alle f'' mediante la matrice $\{\mathfrak{P}\}$ nel modo stesso con cui la matrice $\{\mathfrak{S}\}$ fa passare dalle f alle f' .

§ 8. — RAPPRESENTAZIONE DI UNO STESSO OPERATORE IN DIVERSI SISTEMI DI RIFERIMENTO.

Si è visto che un operatore \mathfrak{A} , fissato un sistema di assi (individuati dai versori y_n) nello spazio hilbertiano, è rappresentato da una matrice $\{\mathfrak{A}\}$ i cui elementi sono dati dalla (20). Se ora si considera un altro sistema di assi y'_n , come al § precedente, si pone il problema di trovare la matrice $\{\mathfrak{A}'\}$ che rappresenta lo stesso operatore nel nuovo sistema di riferimento, cioè la matrice i cui elementi A'_{mn} consentono di esprimere le componenti F'_m di $\mathfrak{A}f$ (rispetto agli assi y'_n) mediante le componenti f'_n di f con la formula, corrispondente alla (22):

$$(41) \quad F'_m = \sum_n A'_{mn} f'_n.$$

L'elemento A'_{mn} sarà dato dalla formula, corrispondente alla (23),

$$(42) \quad A'_{mn} = (\mathfrak{A}y'_n) \times y'_m.$$

Ora si sostituiscano per y'_n e y'_m le loro espressioni mediante le y , cioè (v. (32))

$$y'_n = \sum_i S_{in} y_i, \quad y'_m = \sum_k S_{km} y_k$$

e si avrà, ricordando la (23),

$$(43) \quad A'_{mn} = \sum_{ik} S_{km}^* A_{ki} S_{in},$$

ovvero, per la (37),

$$A'_{mn} = \sum_{ik} \widetilde{S}_{mk} A_{ki} S_{in}.$$

Ma, per la regola di moltiplicazione, questo non è che l'elemento (m, n) della matrice $\{\widetilde{\mathfrak{S}}\}\{\mathfrak{A}\}\{\mathfrak{S}\}$, ossia, per la (38), $\{\mathfrak{S}\}^{-1}\{\mathfrak{A}\}\{\mathfrak{S}\}$: quindi scriveremo

$$(44) \quad \{\mathfrak{A}'\} = \{\mathfrak{S}\}^{-1}\{\mathfrak{A}\}\{\mathfrak{S}\}.$$

Questa è la legge con cui si trasforma la matrice $\{\mathfrak{A}\}$ nel passaggio dagli assi y agli assi y' .

È superfluo rilevare che le relazioni algebriche tra matrici conservano la stessa forma in qualunque sistema di riferimento: se p. es. nel primo sistema si ha

$$(45) \quad \{\mathfrak{A}\} = F(\{\mathfrak{B}\}),$$

(F simbolo di funzione analitica), si ha anche nel secondo sistema

$$(45') \quad \{\mathfrak{A}\}' = F(\{\mathfrak{B}\}').$$

Ciò discende immediatamente dal fatto che sia la (45) che la (45') traducono la stessa relazione $\mathfrak{A} = F(\mathfrak{B})$ tra gli operatori. Del resto, sarebbe facile provarlo anche direttamente, verificando che la trasformazione (44) non altera le relazioni di somma di prodotto e di inversa con cui si costruisce qualsiasi relazione analitica.

§ 9. - OPERATORI E MATRICI HERMITIANI.

Hanno particolare interesse nella meccanica quantistica quegli o. l. \mathfrak{A} che godono la proprietà seguente: per qualunque funzione f , il prodotto $\mathbf{f} \times \mathfrak{A}f$ è reale, cioè

$$(46) \quad \mathbf{f} \times \mathfrak{A}f = (\mathbf{f} \times \mathfrak{A}f)^*.$$

Tali operatori diconsi *hermitiani*.

Come esempio notevole, si consideri l'operatore che intervenne al § 1, p. II, cioè (A, B, C reali):

$$(47) \quad \mathfrak{Q} = A \frac{d^2}{dx^2} + B \frac{d}{dx} + C$$

e si ricerchi la condizione perchè \mathfrak{Q} sia hermitiano. Applicando la (46), si vede che deve essere, per qualunque f ,

$$\int A (f''f^* - f^{*''}f) dx + \int B (f'f^* - f^{*'}f) dx = 0.$$

Osservando che la prima delle differenze in parentesi è la derivata della seconda, e calcolando per parti il primo integrale si ha (se la f si annulla agli estremi)

$$\int (A' - B) (f'f^* - f^{*'}f) dx = 0,$$

che richiede, essendo f arbitraria, $A' = B$. Si trova dunque la condizione che abbiamo già espresso dicendo che l'equazione $\mathfrak{Q}f = 0$ era *autoaggiunta* (v. § 3, p. II).

La proprietà di un operatore \mathfrak{A} di essere hermitiano si traduce in una proprietà notevole della matrice che lo rappresenta (in un qualunque sistema di assi). Infatti, si noti

che (ponendo, al solito, $F = \mathfrak{A}f$) si ha, per la (10) e la (22)

$$(48) \quad \mathbf{f} \times \mathfrak{A}f = \sum_m f_m F_m^* = \sum_{mn} f_m A_{mn}^* f_n^*$$

e, prendendo il coniugato,

$$(\mathbf{f} \times \mathfrak{A}f)^* = \sum_{mn} f_m^* A_{mn} f_n$$

che si può anche scrivere, scambiando gli indici di sommatoria,

$$(48') \quad (\mathbf{f} \times \mathfrak{A}f)^* = \sum_{mn} f_n^* A_{nm} f_m.$$

Sostituendo questa, insieme alla (48), nella condizione di hermiticità (46), si ricava

$$(46') \quad A_{mn}^* = A_{nm},$$

cioè: gli elementi simmetrici rispetto alla diagonale principale sono coniugati. (In particolare, gli elementi della diagonale principale risulteranno reali). Con la notazione spiegata al § 7, la (46') si può scrivere: $\{\overline{\mathfrak{A}}\} = \{\mathfrak{A}\}$. Una matrice siffatta dicesi *hermitiana*. È ovvio che, viceversa, una matrice hermitiana rappresenti sempre un operatore hermitiano.

Di qui ricaviamo facilmente un'altra proprietà degli operatori hermitiani: per due funzioni qualunque f e g , si ha, se \mathfrak{A} è hermitiano (e solo se è tale):

$$(49) \quad \mathbf{f} \times \mathfrak{A}g = \mathfrak{A}f \times g.$$

Difatti, dette g_n le componenti di g , si ha, analogamente alla (48),

$$\begin{aligned} \mathbf{f} \times \mathfrak{A}g &= \sum_{mn} f_m A_{mn}^* g_n^*, \\ g \times \mathfrak{A}f &= \sum_{mn} g_m A_{mn}^* f_n^*. \end{aligned}$$

Prendendo il coniugato della seconda sommatoria e scambiando tra loro gli indici m ed n , si riconosce che, in virtù della (46'), questa sommatoria si identifica con la prima, cioè è verificata la (49).

Si osservi che se $\{\mathfrak{A}\}$ è una matrice hermitiana, è tale anche la matrice $\{\mathfrak{A}'\}$ che corrisponde ad essa in un qualsiasi altro sistema di riferimento: ciò si può riconoscere sia mediante la (44), sia osservando che se $\{\mathfrak{A}\}$ è hermitiana,

è tale l'operatore che essa rappresenta, e quindi esso è rappresentato, anche in un altro sistema di riferimento, da una matrice hermitiana.

Nel seguito, ci occuperemo soltanto di operatori hermitiani e di matrici hermitiane.

Dimostriamo ora che, se \mathfrak{A} e \mathfrak{B} sono o. l. hermitiani, sono tali anche i due o. l.

$$(50) \quad \mathfrak{C} = \mathfrak{A}\mathfrak{B} + \mathfrak{B}\mathfrak{A}$$

$$(50') \quad \mathfrak{D} = i(\mathfrak{A}\mathfrak{B} - \mathfrak{B}\mathfrak{A}).$$

e in particolare, se \mathfrak{A} e \mathfrak{B} sono permutabili, il loro prodotto è hermitiano.

Infatti, dalla (49) si ha, ponendo al posto di \mathfrak{g} (vettore arbitrario) $\mathfrak{B}\mathfrak{g}$,

$$(51) \quad \mathfrak{f} \times \mathfrak{A}\mathfrak{B}\mathfrak{g} = \mathfrak{A}\mathfrak{f} \times \mathfrak{B}\mathfrak{g}.$$

Similmente si troverà la formula (ottenibile dalla precedente con lo scambio di \mathfrak{A} con \mathfrak{B} e di \mathfrak{f} con \mathfrak{g})

$$\mathfrak{g} \times \mathfrak{B}\mathfrak{A}\mathfrak{f} = \mathfrak{B}\mathfrak{g} \times \mathfrak{A}\mathfrak{f},$$

ovvero

$$\mathfrak{B}\mathfrak{A}\mathfrak{f} \times \mathfrak{g} = \mathfrak{A}\mathfrak{f} \times \mathfrak{B}\mathfrak{g}.$$

Confrontando con la (51) si trova

$$\mathfrak{f} \times \mathfrak{A}\mathfrak{B}\mathfrak{g} = \mathfrak{B}\mathfrak{A}\mathfrak{f} \times \mathfrak{g}.$$

Similmente (scambiando \mathfrak{A} con \mathfrak{B}),

$$\mathfrak{f} \times \mathfrak{B}\mathfrak{A}\mathfrak{g} = \mathfrak{A}\mathfrak{B}\mathfrak{f} \times \mathfrak{g}.$$

Sommando le ultime due, e badando alla definizione (50) si ha

$$\mathfrak{f} \times \mathfrak{C}\mathfrak{g} = \mathfrak{C}\mathfrak{f} \times \mathfrak{g}$$

che esprime che \mathfrak{C} è hermitiano. Sottraendole invece, e badando alla (50'), si trova

$$(52) \quad \mathfrak{f} \times i\mathfrak{D}\mathfrak{g} = -i\mathfrak{D}\mathfrak{f} \times \mathfrak{g}$$

e poichè, per la (5') $\mathfrak{f} \times i\mathfrak{D}\mathfrak{g} = -i\mathfrak{f} \times \mathfrak{D}\mathfrak{g}$, la (52) si può scrivere

$$\mathfrak{f} \times \mathfrak{D}\mathfrak{g} = \mathfrak{D}\mathfrak{f} \times \mathfrak{g}$$

ed esprime che \mathfrak{D} è hermitiano.

Corollario del teorema precedente è che se \mathfrak{A} è hermitiano, sono tali tutte le sue potenze, e quindi qualunque sua funzione analitica (a coefficienti reali). Inoltre, un o. l., funzione (a coefficienti reali) di più o. l. hermitiani e permutabili, è evidentemente hermitiano anch'esso (se però gli o. l. di cui è funzione non sono permutabili, può non essere hermitiano; p. es., l'o. l. $\mathfrak{A}\mathfrak{B} - \mathfrak{B}\mathfrak{A}$, se non è nullo, non è hermitiano, perchè altrimenti non sarebbe tale \mathfrak{D} , dato dalla (50')).

§ 10. - ASSI PRINCIPALI DI UN OPERATORE.

Dato un o. l. (hermitiano) \mathfrak{A} , proponiamoci la seguente questione: esistono vettori (dello spazio hilbertiano) che vengano dall'operatore \mathfrak{A} mutati di grandezza ma non di direzione? Ciò significa ricercare se esistono funzioni f tali che

$$(51) \quad \mathfrak{A}f = Af,$$

dove A è una costante moltiplicativa, da determinarsi.

Questo problema, nel caso in cui \mathfrak{A} è l'operatore \mathfrak{Q} della (47) (con $A' = B$) consiste nella ricerca delle soluzioni a quadrato sommabile dell'equazione autoaggiunta

$$(52) \quad Af'' + A'f' + (C - L)f = 0,$$

dove L funge da « parametro »: come si è visto al § 2, p. II esistono infinite soluzioni indipendenti (autofunzioni) $f = y_n$, a ciascuna delle quali corrisponde un valore L_n di L (autovalore) (alcuni degli L_n potendo anche essere uguali tra loro).

In generale, chiameremo *autovalori* dell'o. l. \mathfrak{A} i numeri A_n e *autofunzioni* le funzioni u_n tali che

$$(53) \quad \mathfrak{A}u_n = A_n u_n.$$

e supporremo sempre le u_n normalizzate. Esse rappresentano dunque altrettanti vettori unitari u_n che possiamo anche supporre tutti ortogonali tra loro ⁽¹⁾ cosicchè sarà

$$(54) \quad u_n \times u_m = \delta_{nm}.$$

⁽¹⁾ La dimostrazione di questo si fa per un o. l. \mathfrak{A} generico (purchè hermitiano) come fu fatta al § 5 p. II per l'o. l. (47). Se u_m e u_n

Le direzioni di questi vettori si chiamano *assi principali* dell'o. l. \mathfrak{A} , e qualunque vettore che giaccia lungo uno di questi assi viene dall'operatore \mathfrak{A} mutato solo di grandezza e non di direzione.

Gli autovalori di un o. l. hermitiano sono (come si dimostrerà più avanti) sempre reali: essi costituiscono un sistema completo di funzioni ortogonali, e possono essere continui o discreti: generalmente ci riferiremo a quest'ultimo caso, intendendo che nel caso degli autovalori continui si debbano sostituire, in tutte le formule, le sommatorie con opportuni integrali, e che come criterio di normalizzazione e di ortogonalità si debba adottare quello spiegato al § 10 p. II.

Osservazione sugli operatori incompleti. — In un problema in cui intervengano più variabili indipendenti x, y, \dots , un o. l. si dirà *completo*, se nella sua espressione intervengono tutte le variabili, mentre si dirà *incompleto* se ne manca qualcuna (se, cioè, p. es., non vi figura nè la y nè operazioni di derivazione rispetto a y). Chiamiamo genericamente x le variabili che intervengono nell'o. l. \mathfrak{A} , y quelle che non vi intervengono: è evidente che risolvendo la (53) si otterranno per le u_n delle funzioni $\chi_n(x)$ delle sole x e non delle y . Tuttavia una qualunque di esse può

appartengono a due autovalori distinti A_m e A_n , essi sono ortogonali: difatti si ha

$$\mathfrak{A} \mathbf{u}_n = A_n \mathbf{u}_n, \quad \mathfrak{A} \mathbf{u}_m = A_m \mathbf{u}_m$$

e moltiplicando scalarmente la prima per \mathbf{u}_m , a destra, la seconda per \mathbf{u}_n a sinistra e sottraendo membro a membro, si ha

$$(\mathfrak{A} \mathbf{u}_n) \times \mathbf{u}_m - \mathbf{u}_n \times (\mathfrak{A} \mathbf{u}_m) = (A_n - A_m) \mathbf{u}_n \times \mathbf{u}_m.$$

Ma se A è hermitiano, il primo membro è nullo e quindi segue (essendo $A_n \neq A_m$): $\mathbf{u}_n \times \mathbf{u}_m = 0$, cioè l'ortogonalità.

Se poi A_n è un autovalore multiplo, a cui corrispondono le p autofunzioni indipendenti $u_n^{(1)}, u_n^{(2)}, \dots, u_n^{(p)}$, un'autofunzione generica appartenente a questo autovalore ha la forma

$$(55) \quad u_n = \sum_1^p c_n^j u_n^j,$$

con i coefficienti c arbitrari. Naturalmente, alle u_n^j si possono sostituire p loro combinazioni lineari, ortogonali tra loro (ciò si dimostra come al § 6); vi è anzi una larga arbitrarietà nella scelta di queste.

essere moltiplicata per una funzione arbitraria $f_n(y)$ delle y , senza cessare di soddisfare la (53), per cui anche $f_n(y)\chi_n(x)$ si potrà considerare come un'autofunzione $\varphi_n(x, y)$ di \mathfrak{A} , la cui espressione generale sarà dunque

$$(56) \quad \varphi_n(x, y) = f_n(y) \chi_n(x).$$

Perciò, considerando lo spazio hilbertiano delle funzioni di x e y , diremo che in questo spazio gli assi principali dell'o. l. incompleto \mathfrak{A} non sono determinati (anche a prescindere da eventuali degenerazioni), ma che si possono scegliere con una larga arbitrarietà scegliendo arbitrariamente le funzioni $f_n(y)$. Generalmente converrà supporre queste ortogonali e normalizzate rispetto alle y , e allora (supposte le χ_n normalizzate rispetto alle x) le $\varphi_n(x, y) = f_n \chi_n$ costituiranno un sistema di funzioni ortogonali e normalizzate (ma non un sistema completo).

Questa indeterminazione nelle autofunzioni di un o. l. incompleto si può in certo modo assimilare ad una degenerazione, considerando l'autovalore A_n di \mathfrak{A} come *multiplo d'ordine infinito*. Difatti, sia $w^j(y)$ un qualunque sistema completo di assi ortogonali nello spazio delle funzioni di y : la $f_n(y)$ di cui sopra si potrà sviluppare nella forma $f_n(y) = \sum_j f_n^j w^j$ e quindi la (56) può scriversi

$$(57) \quad \varphi_n(x, y) = \sum_1^\infty f_n^j w^j \chi_n = \sum_1^\infty f_n^j \varphi_n^j(x, y)$$

ponendo $\varphi_n^j(x, y) = w^j(y) \chi_n(x)$. Questa formula è analoga alla (55); essa esprime la più generale autofunzione di \mathfrak{A} appartenente all'autovalore A_n , come combinazione lineare (a coefficienti f_n^j arbitrari) delle autofunzioni fondamentali φ_n^j .

È importante per le applicazioni il seguente teorema: se φ_n è un'autofunzione di \mathfrak{A} , appartenente all'autovalore A_n , essa è anche un'autofunzione di $F(\mathfrak{A})$ (F simbolo di funzione analitica qualunque, v. § 4) appartenente all'autovalore $F(A_n)$.

Dimostrazione. — Per ipotesi si ha

$$\mathfrak{A} \varphi_n = A_n \varphi_n.$$

Applicando ai due membri l'o. l. \mathfrak{A} si ottiene

$$\mathfrak{A}^2 \varphi_n = A_n \mathfrak{A} \varphi_n = A_n^2 \varphi_n.$$

Ripetendo il procedimento, si riconosce che per qualunque potenza di \mathfrak{A} vale

$$(58) \quad \mathfrak{A}^r \varphi_n = A_n^r \varphi_n.$$

Sia ora la funzione F definita dalla serie

$$F(\alpha) = \sum_r a_r \alpha^r$$

e denotiamo con \mathfrak{F} l' o. l. $F(\mathfrak{A})$. Sarà

$$\mathfrak{F} \varphi_n = \sum_r a_r \mathfrak{A}^r \varphi_n$$

e per la (58)

$$(59) \quad \mathfrak{F} \varphi_n = \sum_r a_r A_n^r \cdot \varphi_n = F(A_n) \varphi_n$$

che dimostra l'asserto.

È evidente poi che, se la funzione F è invertibile (cioè se si può scrivere $\mathfrak{A} = G(\mathfrak{F})$ con G simbolo di funzione analitica), vale anche il reciproco di questo teorema, cosicchè si può dire in tal caso: gli o. l. \mathfrak{A} ed $\mathfrak{F} = F(\mathfrak{A})$ hanno gli stessi autovalori e le stesse autofunzioni.

Nel caso degli autovalori continui è importante osservare che se $\varphi_{A'}$ è un'autofunzione di \mathfrak{A} appartenente all'autovalore A' , e *normalizzata* (col criterio del § 10), essa è anche un'autofunzione di $F(\mathfrak{A})$, appartenente all'autovalore $F(A')$, ma non è normalizzata: perchè lo sia, occorre dividerla per $\sqrt{\frac{dF}{dA'}}$, come si può verificare facilmente.

Questo teorema suggerisce una importante generalizzazione del concetto di funzione di un o. l. (che fin qui era limitato alle funzioni analitiche). Sia difatti $F(\alpha)$ il simbolo di una funzione della variabile α (anche non sviluppabile in serie), e sia \mathfrak{A} un o. l. con gli autovalori A_n e le autofunzioni φ_n : definiremo l' o. l. $F(\mathfrak{A})$ come quello che, applicato a una φ_n , la trasforma in $F(A_n)\varphi_n$, conformemente alla (59), e quindi applicato a una f qualunque avente lo sviluppo

$$f = \sum a_n \varphi_n,$$

la trasforma in

$$(60) \quad \mathfrak{A} f = \sum a_n F(A_n) \varphi_n.$$

Unica condizione richiesta alla funzione $F(\alpha)$ è di essere univocamente definita per tutti i valori A_n di α .

§ 11. — TEOREMA FONDAMENTALE SUGLI OPERATORI PERMUTABILI.

Condizione necessaria e sufficiente perchè due o. l. \mathfrak{A} e \mathfrak{B} ammettano un sistema completo di autofunzioni (e quindi di assi principali) in comune, è che essi siano permutabili.

Dimostriamo dapprima che la condizione è necessaria. Supponiamo perciò che esista un sistema completo di funzioni ortogonali $\varphi_1, \varphi_2, \dots$ che siano ad un tempo autofunzioni di \mathfrak{A} e di \mathfrak{B} , cosicchè si possa scrivere (ordinando convenientemente gli indici degli autovalori)

$$\mathfrak{A} \varphi_i = A_i \varphi_i, \quad \mathfrak{B} \varphi_i = B_i \varphi_i.$$

Applicando alla prima l' o. l. \mathfrak{B} , alla seconda \mathfrak{A} , si ottiene rispettivamente

$$\mathfrak{B} \mathfrak{A} \varphi_i = A_i \mathfrak{B} \varphi_i = A_i B_i \varphi_i$$

$$\mathfrak{A} \mathfrak{B} \varphi_i = B_i \mathfrak{A} \varphi_i = B_i A_i \varphi_i$$

e quindi

$$\mathfrak{A} \mathfrak{B} \varphi_i = \mathfrak{B} \mathfrak{A} \varphi_i.$$

Poichè le φ_i formano un sistema completo, qualunque funzione f si può sviluppare in serie delle φ_i , e quindi per qualunque f varrà

$$\mathfrak{A} \mathfrak{B} f = \mathfrak{B} \mathfrak{A} f,$$

il che significa che

$$(61) \quad \mathfrak{A} \mathfrak{B} = \mathfrak{B} \mathfrak{A}.$$

Dimostriamo ora che la condizione è sufficiente. Supponiamo perciò che valga la (61) e chiamiamo φ_i una generica autofunzione di \mathfrak{A} appartenente all'autovalore A_i , talchè sia

$$\mathfrak{A} \varphi_i = A_i \varphi_i,$$

e sia ψ_j un sistema completo di autofunzioni di \mathfrak{B} , per cui

$$\mathfrak{B} \psi_j = B_j \psi_j.$$

Sviluppiamo la φ_i in serie delle ψ_j :

$$\varphi_i = \sum_j c_j \psi_j$$

e applichiamo ai due membri l'operatore $\mathfrak{A} F(\mathfrak{B})$, dove $F(\mathfrak{B})$ è un o. l. funzione qualunque di \mathfrak{B} : sarà, ricordando il teorema del § 10,

$$(62) \quad \mathfrak{A} F(\mathfrak{B}) \varphi_i = \sum_j c_j F(B_j) \mathfrak{A} \psi_j.$$

D'altra parte, \mathfrak{A} , essendo permutabile con \mathfrak{B} , lo è anche con $F(\mathfrak{B})$, quindi

$$\begin{aligned} \mathfrak{A} F(\mathfrak{B}) \varphi_i &= F(\mathfrak{B}) \mathfrak{A} \varphi_i = F(\mathfrak{B}) A_i \varphi_i = \\ &= A_i \sum_j c_j F(\mathfrak{B}) \psi_j = A_i \sum_j c_j F(B_j) \psi_j \end{aligned}$$

Confrontando questa con la (62), si ricava

$$(63) \quad \sum_j c_j F(B_j) \mathfrak{A} \psi_j = A_i \sum_j c_j F(B_j) \psi_j$$

ed essendo F una funzione arbitraria, i coefficienti $c_j F(B_j)$ si debbono riguardare come numeri del tutto arbitrari: ne segue che la (63) non può sussistere se non è, per ciascun termine,

$$\mathfrak{A} \psi_j = A_i \psi_j \quad (j = 1, 2, \dots),$$

il che prova che ψ_j è anche un'autofunzione di \mathfrak{A} (e precisamente appartenente all'autovalore A_i). Resta così provato che le ψ_j costituiscono un sistema completo di autofunzioni comuni ad \mathfrak{A} e \mathfrak{B} .

Osservazione. — Vogliamo richiamare l'attenzione sul significato del teorema precedente nei casi in cui esistono autovalori multipli o in cui uno degli operatori (o eventualmente anche l'altro) è incompleto. Se p. es. A_i è un autovalore di \mathfrak{A} multiplo di ordine p , si possono scegliere in infiniti modi p autofunzioni ortogonali e indipendenti che ad esso appartengono (i cui vettori formano una varietà piana V a p dimensioni): ma per una sola di queste scelte, in generale, si ottengono autofunzioni comuni anche all'altro operatore (e, si noti, all'unico autovalore A_i corrispondono p autovalori di \mathfrak{B} , cioè vi sono p assi principali di \mathfrak{B} entro la varietà V). Ciò si applica anche al caso della molteplicità infinita, cioè se uno degli operatori, p. es. \mathfrak{A} , è incompleto: il teorema ha in tal caso il significato seguente. Vi siano p. es. due variabili x, y , di cui solo la x interviene in \mathfrak{A} : sappiamo già che, dette $\chi_i(x)$ le autofunzioni di \mathfrak{A} nello spazio delle funzioni di x , si possono considerare come autofunzioni di \mathfrak{A} nello spazio delle funzioni di x e y tutte le funzioni della forma $\varphi_i(x, y) = f(y) \chi_i(x)$ con f funzione arbitraria delle sole y . Ma soltanto alcune di queste autofunzioni sono comuni ad \mathfrak{A} e a \mathfrak{B} : e precisamente, la $f(y)$ dev'essere scelta entro un insieme discreto (o eventualmente continuo) di funzioni $f_k(y)$, talchè, dette ψ_{ik} le autofunzioni di \mathfrak{B} (le caratterizziamo esplicitamente con due indici, perchè dipendono dalle due variabili x ed y) si ha $\psi_{ik} = f_k(x) \chi_i(y)$: ad ogni χ_i corrispondono infinite autofunzioni di \mathfrak{B} .

§ 12. — MATRICI DIAGONALI.

Assumiamo gli assi principali di \mathfrak{A} (di versori \mathbf{u}_n) come assi coordinati nello spazio hilbertiano, e ricerchiamo la forma che assume la matrice che rappresenta l'o. l. \mathfrak{A} rispetto a questi assi: designeremo questa matrice con $\{\mathfrak{A}\}'$, mentre seguiranno a indicare con $\{\mathfrak{A}\}$ la matrice che rappresenta \mathfrak{A} rispetto agli assi, generici, y_n .

L'elemento generico A'_{mn} della matrice $\{\mathfrak{A}\}'$ sarà, conformemente alla (23),

$$A'_{mn} = \mathfrak{A} \mathbf{u}_n \times \mathbf{u}_m$$

e per la (53) e la (54),

$$(64) \quad A'_{mn} = A_n \mathbf{u}_n \times \mathbf{u}_m = A_n \delta_{mn}.$$

Vale a dire, la matrice $\{\mathfrak{A}\}'$ sarà

$$\{\mathfrak{A}\}' = \begin{vmatrix} A_1 & 0 & 0 & \dots \\ 0 & A_2 & 0 & \dots \\ 0 & 0 & A_3 & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix}.$$

Una matrice, come questa, in cui tutti gli elementi sono nulli tranne quelli sulla diagonale principale (*elementi diagonali*) dicesi *matrice diagonale*. Diremo dunque che: « un o. l. è rappresentato, rispetto ai suoi assi principali, da una matrice diagonale: gli elementi diagonali di questa sono gli autovalori dell'operatore ».

Si osservi che se l'operatore è hermitiano tale è anche la matrice, il che per una matrice diagonale significa che i suoi elementi sono reali. Dunque: gli autovalori di un o. l. hermitiano sono sempre reali.

Sia ora dato un o. l. \mathfrak{A} mediante la matrice $\{\mathfrak{A}\}$ che lo rappresenta rispetto ad assi generici y_n : vogliamo trovare la matrice (diagonale) $\{\mathfrak{A}\}'$ che rappresenta \mathfrak{A} rispetto ai suoi assi principali: questa operazione si dice talvolta « riduzione della matrice $\{\mathfrak{A}\}$ a forma diagonale ».

Ricordiamo dal § 8 che la $\{\mathfrak{A}\}'$ si ricava dalla $\{\mathfrak{A}\}$ con la formula (44): si tratta dunque di trovare la matrice di trasformazione $\{\mathfrak{S}\}$. A tal uopo, osserviamo che la (44)

può anche scriversi

$$\{\in\} \{A\} = \{A\} \{\in\}$$

ossia, uguagliando l'elemento generico (m, k) nei due membri,

$$\sum_r S_{mr} A'_{rk} = \sum_r A_{mr} S_{rk};$$

ma, essendo $A'_{rk} = A_k \delta_{rk}$, la prima sommatoria si riduce al solo termine in cui $r = k$, cosicchè l'equazione diviene

$$\sum_r A_{mr} S_{rk} - A_k S_{mk} = 0.$$

Fissiamo k , e diamo ad m i successivi valori $1, 2, \dots$: avremo le equazioni

$$(65) \quad \begin{cases} (A_{11} - A_k) S_{1k} + A_{12} S_{2k} + A_{13} S_{3k} + \dots = 0 \\ A_{21} S_{1k} + (A_{22} - A_k) S_{2k} + A_{23} S_{3k} + \dots = 0 \\ A_{31} S_{1k} + A_{32} S_{2k} + (A_{33} - A_k) S_{3k} + \dots = 0 \\ \dots \dots \dots \end{cases}$$

Queste costituiscono un sistema di infinite equazioni lineari ed omogenee, nelle infinite incognite $S_{1k}, S_{2k}, S_{3k}, \dots$

Per comprendere la natura di questo problema, si consideri dapprima il caso che si tratti di uno spazio ordinario a tre dimensioni: allora il sistema (65) si riduce a un sistema di tre equazioni lineari ed omogenee nelle tre incognite S_{1k}, S_{2k}, S_{3k} (si potrebbero scrivere tre di tali sistemi, corrispondenti a $k = 1, k = 2, k = 3$), ed ammette soluzioni non nulle solo se

$$(66) \quad \begin{vmatrix} A_{11} - A_k & A_{12} & A_{13} \\ A_{21} & A_{22} - A_k & A_{23} \\ A_{31} & A_{32} & A_{33} - A_k \end{vmatrix} = 0.$$

Questa equazione in A_k , detta *equazione secolare* (che risulta la stessa per tutti e tre i sistemi), ammette, come si sa dall'algebra, tre radici sempre reali A_1, A_2, A_3 : trovate queste, si sostituiscono nei tre sistemi e si trovano facilmente le S_{mk} , a meno di un fattore che si determina con le condizioni (34). Tutto ciò coincide, come si riconosce, col procedimento classico per determinare gli assi principali di una quadrica a centro (i cui coefficienti dei termini di secondo grado siano le A_{ij}): le S_{mk} danno i coseni di questi assi, mentre le lunghezze dei tre semiassi sono date da $\pm 1/\sqrt{A_k}$.

Si vede da ciò, che il problema di ridurre una matrice a forma diagonale (ossia, di trovare gli autovalori e le auto-

funzioni di un o. l. \mathfrak{Q}) è la generalizzazione, al caso di infinite dimensioni, del problema di ridurre una quadrica ai suoi assi principali. Generalizzando al caso di n infinito ciò che è noto per un sistema di n equazioni lineari omogenee ad n incognite, si è condotti a concludere che il sistema (65) ammette soluzioni non nulle solo se A_k ha certi valori determinati (formanti una successione infinita e generalmente discreta) che sono appunto gli autovalori richiesti.

§ 13. - MATRICI A INDICI CONTINUI.

Nei §§ precedenti abbiamo sempre supposto che le autofunzioni che definiscono gli assi coordinati nello spazio hilbertiano formassero una successione discreta: ma, come si è detto nel § 10 p. II, quando il campo entro cui si deve integrare l'equazione differenziale è infinito possono presentarsi degli autovalori formanti spettro continuo (oltre, eventualmente, ad autovalori discreti). Si è così condotti a considerare casi in cui gli indici, che entrano nelle formule precedenti, anzichè assumere solo valori interi variano con continuità. L'estensione delle considerazioni precedenti a casi siffatti presenta dei punti delicati, e per una trattazione rigorosa di essi rimandiamo ad opere speciali ⁽¹⁾; qui ci limitiamo a delle indicazioni generali di carattere intuitivo.

Cominciamo con l'osservare che se a_n è una qualsiasi successione di grandezze corrispondenti ai valori interi dell'indice n , quando questo indice diventa una variabile continua λ dovremo considerare a_λ come il simbolo di una ordinaria funzione di λ , che quindi si potrà scrivere anche $a(\lambda)$. Così, come le autofunzioni y_n sono sostituite dalla y_λ (λ indice continuo), ogni vettore \mathbf{f} nello spazio hilbertiano è rappresentato, rispetto agli assi y_λ , invece che dalle componenti f_n , dalla funzione di λ :

$$f_\lambda = \mathbf{f} \times y_\lambda$$

(componente di \mathbf{f} sull'asse y_λ) e analogamente tutte le for-

⁽¹⁾ Si veda per es., oltre ai lavori del Dirac, la memoria di E. H. Kennard, *ZS. f. Phys.*, 44 (1927) p. 326.

mule della teoria dei vettori dello spazio hilbertiano verranno modificate nel senso di sostituire alle sommatorie rispetto a un indice degli integrali. P. es., la (10) diviene

$$\mathbf{f} \times \mathbf{g} = \int_a^b f_\lambda g_\lambda^* d\lambda.$$

Vediamo ora che cosa divengono in questo caso le matrici. Un'espressione a due indici A_{mn} , quando m ed n diventano due variabili continue

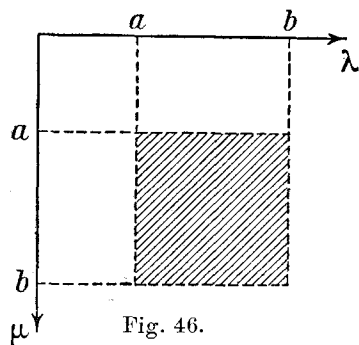


Fig. 46.

μ, λ nell'intervallo (a, b) , diventerà un'ordinaria funzione di due variabili che si potrà anche scrivere, invece che $A_{\mu\lambda}$, nel modo usuale $A(\mu, \lambda)$. Quindi il concetto di « matrice ad indici continui » (« o matrice continua ») è equivalente al concetto di funzione di due variabili. Se poi si vuole conservare l'imma-

gine materiale della matrice, si devono interpretare μ, λ come coordinate cartesiane nel piano, prendendo gli assi come nella fig. 46: allora ad ogni punto del quadrato tratteggiato si può far corrispondere un valore di $A(\mu, \lambda)$: si può dire che gli elementi di tale matrice « riempiono con continuità » il quadrato tratteggiato (il quale, eventualmente, può estendersi sino all'infinito, diventando un « quadrante », se $b = \infty$).

A queste matrici continue si estendono tutte le definizioni già date: p. es. il prodotto di due matrici $A(\mu, \lambda)$, $B(\varrho, \sigma)$ è la matrice

$$(67) \quad P(\mu, \sigma) = \int_a^b A(\mu, \lambda) B(\lambda, \sigma) d\lambda.$$

Una matrice continua si dirà *hermitiana* se

$$A^*(\mu, \lambda) = A(\lambda, \mu).$$

Un operatore hermitiano \mathfrak{A} è rappresentato, rispetto agli assi individuati dalle autofunzioni y_λ , appunto da una

matrice siffatta, il cui elemento generico è, analogamente a (23) e (23'),

$$A(\mu, \lambda) = (\mathfrak{A}y_\lambda) \times y_\mu = \int y_\mu^* \mathfrak{A}y_\lambda dx$$

e, mediante questa matrice continua, si ottengono le componenti del vettore $\mathbf{F} = \mathfrak{A}\mathbf{f}$ da quelle di \mathbf{f} con la formula, analoga alla (22),

$$F_\mu = \int_a^b A(\mu, \lambda) f_\lambda d\lambda.$$

Per passare da un sistema di assi y_λ a un altro sistema y'_λ si dovrà introdurre una « matrice di trasformazione » (continua) definita da (v. (33))

$$S(\mu, \lambda) = y'_\lambda \times y_\mu.$$

Mediante questa matrice, si passa dalle componenti f_λ del vettore \mathbf{f} alle componenti f'_λ rispetto ai nuovi assi dello stesso vettore, mediante la formula, analoga a (35),

$$(68) \quad f'_\lambda = \int_a^b S^*(\mu, \lambda) f_\mu d\mu.$$

Il passaggio inverso è dato (v. (35'')) da

$$(68') \quad f_\lambda = \int_a^b S(\mu, \lambda) f'_\mu d\mu.$$

In questo cambiamento di assi, la matrice $A(k, j)$ che rappresenta un operatore \mathfrak{A} rispetto agli assi y_λ , si cambia nella matrice $A'(\mu, \lambda)$ che rappresenta lo stesso operatore rispetto ai nuovi assi, mediante la formula (44), che, esplicitata, diviene la seguente (che è analoga alla (43))

$$(69) \quad A'(\mu, \lambda) = \int_a^b \int_a^b S^*(k, \mu) A(k, j) S(j, \lambda) dk dj.$$

Accenneremo infine al caso in cui l'equazione che definisce le y ha, oltre allo spettro continuo tra a e b , anche degli autovalori discreti. Allora le formule precedenti si devono modificare nel senso che ad ogni integrale fatto tra

a e b si deve aggiungere una sommatoria rispetto agli autovalori discreti. Ciò porta maggiore complicazione nella scrittura (che del resto si potrebbe evitare con l'uso degli integrali di STIELTJES, v. bibl. n. 14, pag. 122) ma non introduce difficoltà concettuali.

§ 14. - FUNZIONE DI DIRAC.

L'estensione al caso di indici continui della matrice unità introdotta al § 5 (cioè la generalizzazione del simbolo δ_{mn}) conduce a introdurre un simbolo $\delta(x)$, assai comodo nei calcoli, chiamato spesso *funzione di Dirac*. Esso rappresenta una funzione che goda le proprietà seguenti:

$$(70) \quad \delta(x) = 0 \text{ per } x \neq 0,$$

mentre nel punto 0 la δ è infinita, e precisamente tale che sia

$$(71) \quad \int_a^b \delta(x) dx = 1,$$

dove a, b sono due limiti qualunque ($a < b$), comprendenti tra loro lo 0. Non esiste una funzione propriamente detta che goda queste proprietà, e perciò la $\delta(x)$ si chiama una *funzione impropria*: ad essa però ci si può approssimare quanto si vuole mediante funzioni analitiche ⁽¹⁾. P. es., la funzione gaussiana $\frac{h}{\sqrt{\pi}} e^{-\frac{x^2}{h^2}}$ per $h \rightarrow \infty$ tende al comportamento della $\delta(x)$: quindi ci si può immaginare la curva che rappresenta la $\delta(x)$ come una curva a campana infinitamente stretta e infinitamente alta, così da racchiudere l'area 1.

Faremo uso generalmente della funzione $\delta(x - x_0)$, che presenta nel punto $x = x_0$ la stessa singolarità che la $\delta(x)$ presenta in $x = 0$: essa ha la proprietà fondamentale che,

⁽¹⁾ L'uso della funzione impropria può essere evitato sostituendola con il concetto di integrale di Stieltjes, come ha fatto sistematicamente il Neumann (v. bibl. n. 13): tale uso però contribuisce a rendere più semplici ed espresse le formule, e può essere considerato come l'indicazione abbreviata di un passaggio al limite.

se $f(x)$ è una funzione qualunque (purchè limitata entro l'intervallo che si considera e continua in x_0), è

$$(72) \quad \int_a^b f(x) \delta(x - x_0) dx = \begin{cases} 0 & \text{se l'intervallo } (a, b) \text{ non contiene } x_0 \\ f(x_0) & \text{» » » contiene } x_0. \end{cases}$$

Infatti, se l'intervallo non contiene x_0 , la funzione δ è nulla in tutto l'intervallo, e quindi l'integrale è nullo: se poi l'intervallo (a, b) contiene x_0 , l'integrale si può estendere solo a un tratto $x_0 - \varepsilon, x_0 + \varepsilon$ (con ε piccolo quanto si vuole): si avrà quindi, chiamando f_1 un valore compreso tra il limite inferiore e il limite superiore di f nell'intervallo $(x_0 - \varepsilon, x_0 + \varepsilon)$:

$$\int_a^b f(x) \delta(x - x_0) dx = \int_{x_0 - \varepsilon}^{x_0 + \varepsilon} f(x) \delta(x - x_0) dx = f_1 \int_{x_0 - \varepsilon}^{x_0 + \varepsilon} \delta(x - x_0) dx = f_1$$

e poichè questo deve valere per qualunque ε , deve essere $f_1 = f(x_0)$.

Una proprietà della δ , di cui si fa spesso uso, è quella espressa dalla formula

$$(73) \quad (x - x_0) \delta(x - x_0) = 0$$

che si suol giustificare dicendo che per tutti i valori di x per cui non è nullo il primo fattore, è nullo il secondo. Più esattamente diremo che la (73) va intesa nel senso che l'integrale del primo membro rispetto ad x , esteso a un intervallo qualunque (anche piccolo quanto si vuole) è nullo.

L'introduzione della funzione impropria δ ci permette di considerare formalmente gli assi dello spazio hilbertiano che abbiamo chiamati « continui » al § 12, come assi principali di un o. l., e cioè come un caso particolare degli assi y_n considerati fin qui. Difatti, si consideri l'o. l. $\xi = x$ e si ricerchino i suoi assi principali, ponendo l'equazione

$$(74) \quad x \varphi_{x'}(x) = x' \varphi_{x'}(x) \quad , \text{ ossia } (74') \quad (x - x') \varphi_{x'} = 0,$$

dove si è indicato con x' l'autovalore (trattandosi, come si vedrà, di autovalori continui). Ora, la (74') è soddisfatta prendendo x' qualunque e

$$(75) \quad \varphi_{x'}(x) = \delta(x - x'),$$

come mostra la (73). Inoltre le autofunzioni sono ortogonali e normalizzate perchè (v. § 10, p. II) detti $\Delta_1 x'$, $\Delta_2 x'$ due intervalli infinitesimi, si ha, come si riconosce facilmente tenendo presente la (72):

$$(76) \quad \int_{-\infty}^{\infty} dx \int_{\Delta_1 x'} \delta(x-x') dx' \int_{\Delta_2 x'} \delta(x-x') dx' =$$

$$= \begin{cases} 0, & \text{se } \Delta_1 x', \Delta_2 x' \text{ non hanno punti comuni;} \\ \Delta x, & \text{se } \Delta_1 x', \Delta_2 x' \text{ hanno in comune un tratto } \Delta x. \end{cases}$$

Dunque: gli autovalori dell'operatore x sono tutti i numeri reali x' , e ad ognuno di essi corrisponde un asse individuato dalla funzione (75). Tali assi sono proprio gli « assi continui » introdotti al § 1: difatti la proiezione di una qualunque funzione $f(x)$ sull'asse corrispondente all'autovalore x' è

$$(77) \quad f_{x'} = f \times \delta(x-x') = \int f(x) \delta(x-x') dx = f(x').$$

Si osservi che, dal punto di vista formale, il simbolo δ_{nm} più volte usato ($= 0$ se $n \neq m$, $= 1$ se $n = m$) trova, nel caso degli indici continui, il suo analogo nella funzione $\delta(\lambda - \mu)$. Per esempio, la condizione di ortogonalità e normalizzazione delle autofunzioni, espressa dalla (46) e dalla (46') del capitolo I, p. II, si può scrivere simbolicamente

$$(78) \quad \int y_{\lambda'} y_{\lambda''}^* dx = \delta(\lambda'' - \lambda').$$

Infatti, se questa è soddisfatta si ha (detti $\Delta\lambda'$, $\Delta\lambda''$ due tratti infinitesimi dello spettro continuo di autovalori)

$$\int dx \int_{\Delta\lambda'} y_{\lambda'} d\lambda' \int_{\Delta\lambda''} y_{\lambda''}^* d\lambda'' = \int d\lambda' \int_{\Delta\lambda''} d\lambda'' \int_{\Delta\lambda'} y_{\lambda'} y_{\lambda''}^* dx = \int d\lambda' \int_{\Delta\lambda''} d\lambda'' \delta(\lambda'' - \lambda').$$

L'integrale rispetto a λ'' è uguale a 1 se λ' è interno all'intervallo $\Delta\lambda''$, altrimenti è nullo: perciò, detto $\Delta\lambda$ il tratto comune (eventualmente nullo) ai due intervalli $\Delta\lambda'$, $\Delta\lambda''$, l'espressione precedente si riduce a

$$\int_{\Delta\lambda} d\lambda',$$

cioè a $\Delta\lambda$: si ritrova così la condizione di ortogonalità e normalizzazione introdotta al § 10 p. II.

CAPITOLO II

TEORIA GENERALE DELLA MECCANICA QUANTISTICA

§ 15. — CONCETTO DI « OSSERVABILE ».

Chiameremo *osservazione* (o anche *misura*) una serie di operazioni fisiche il cui risultato sia esprimibile mediante un numero (includendo tra le operazioni della misura anche le eventuali operazioni matematiche da compiersi sui risultati diretti dell'esperienza). Un'osservazione sarà definita quando siano descritte queste operazioni e sia indicato l'istante nel quale esse devono compiersi. P. es. uno dei dispositivi sperimentali del § 23 p. II, con l'indicazione dell'istante t nel quale si deve prendere la fotografia, definisce una « osservazione ».

Nella meccanica ordinaria, un'osservazione serve a conoscere il valore che ha in un dato istante una certa « grandezza meccanica », cioè una coordinata, o un impulso, o una qualunque funzione $g(q, p)$ delle coordinate e degli impulsi, come p. es. il momento angolare o l'energia. Tale valore però viene pensato come esistente anche se non si compie l'osservazione destinata a conoscerlo: la g è insomma pensata come una *variabile* che assume successivamente diversi valori numerici durante il moto del sistema, i quali possono essere calcolati a priori, anche senza osservarli, mediante le leggi della meccanica, se si conoscono p. es. le coordinate e le velocità iniziali del sistema.

Nella meccanica quantistica (come abbiamo già osservato a proposito delle coordinate di una particella) si adotta invece un punto di vista profondamente diverso: il valore

numerico di una grandezza fisica non ha nessun senso finchè non se ne compie l'osservazione, e il risultato di questa osservazione non può in generale venire previsto se non probabilisticamente. Vale a dire, si afferma che l'osservazione potrà dare come risultato uno qualunque di certi numeri $(^1) G_1, G_2, \dots$ con le rispettive probabilità $(^2) P_1, P_2, \dots$ e la meccanica quantistica ha appunto per oggetto di determinare questi possibili risultati e le rispettive probabilità. Esempi di ciò si sono visti a proposito dell'osservazione di una coordinata o di un impulso (§ 25, P. II) o dell'energia (§ 29, P. II).

Enti come le coordinate, le componenti dell'impulso, l'energia, ecc. (in un dato istante) si considerano dunque in meccanica quantistica sotto un aspetto diverso da quello della meccanica ordinaria, e cioè come enti definiti soltanto dal tipo di « osservazione » che ad essi corrisponde. Per mettere in evidenza ciò, il DIRAC ha proposto per essi il nome di *osservabili*. (Ricordiamo che nel concetto di osservazione, e quindi anche in quello di osservabile, è implicito anche l'istante nel quale si esegue la misura: p. es., l'ascissa di una particella al tempo 7^{sec} è un'osservabile $(^3)$, mentre l'ascissa della stessa particella al tempo 8^{sec} è un'altra osservabile).

Nella meccanica ordinaria, il concetto di osservabile si confonde con quello di « valore numerico di una variabile »: invece nella meccanica quantistica in generale un'osservabile non ha un valore numerico, ma ad essa corrispondono infiniti valori numerici (continui o discreti) G_1, G_2, \dots , con le rispettive probabilità P_1, P_2, \dots . Solo in casi particolari può avvenire che queste probabilità siano tutte nulle tranne una,

⁽¹⁾ Si considera qui, per semplicità di scrittura, il caso di valori discreti, ma le G' possono anche costituire un sistema continuo, come p. es. nella misura di una coordinata.

⁽²⁾ Il concetto di probabilità si deve intendere qui precisato nel modo spiegato nella nota al § 25 p. II.

⁽³⁾ A rigore, si dovrebbe precisare il dispositivo di misura: p. es. la camera oscura descritta al § 23 p. II (1° met.) definisce una osservabile che possiamo chiamare ascissa: che il 2° ed il 3° metodo ivi descritti definiscano la stessa osservabile, è un postulato che si ammette generalmente.

p. es. P_1 (che è quindi = 1) e solo allora si può dire che l'osservabile G ha il valore G_1 . Questo caso si verifica, p. es., per l'energia, quando il sistema è in uno degli stati che abbiamo chiamato « stazionari » o « semplici » o « a energia definita » (§ 27 p. II), cioè quando la ψ è una autofunzione dell'equazione di SCHRÖDINGER. Quando invece la ψ è una combinazione lineare di autofunzioni (v. § 29 p. II) l'energia non ha nessun valore numerico: un'osservazione diretta a misurarla può dare per risultato uno qualunque degli autovalori E_1, E_2, \dots con le rispettive probabilità $c_1 c_1^*, c_2 c_2^* \dots$

Altro esempio: la x di una particella (per un determinato t) è un'osservabile che in generale non ha nessun valore numerico, poichè osservandola si possono trovare tutti i valori da $-\infty$ a $+\infty$, con densità di probabilità $P(x) = \psi\psi^*$: la curva che rappresenta la $P(x)$ può avere qualsiasi forma ed alcuni esempi si sono visti al § 36, P. II (in particolare può la $P(x)$ essere nulla dovunque, tranne che per un determinato x , e solo in questo caso l'osservabile x ha un valore definito).

§ 16. - COMPATIBILITÀ DELLE OSSERVABILI.

Supponiamo che la misura di una osservabile A abbia dato un certo risultato A' , e misuriamo, immediatamente dopo A , un'altra osservabile B ottenendo p. es. B' : con ciò mettiamo il sistema in un nuovo stato in cui la B ha il valore definito B' , ma la A , in generale, non avrà più un valore definito (cioè, se subito dopo B si tornasse ad osservare A , non si sarebbe più certi di trovare A'). Vi sono però certi casi in cui questo non si verifica, e cioè si può osservare B subito dopo A senza che A cessi di avere il valore risultato dalla misura precedente: si dice allora che le due osservazioni sono *compatibili*. In tal caso, subito dopo che l'osservazione di A ha dato il risultato A' e quella di B il risultato B' , il sistema si trova in uno stato tale, che tanto A quanto B hanno un valore definito, e cioè, rispettivamente, A' e B' .

La definizione precedente, come si vede, si applica solo a osservazioni *contemporanee*. Si può porre allora la questione, se sia indifferente eseguire prima l'osservazione di A e im-

mediatamente dopo quella di B o viceversa: vedremo in seguito che le leggi della meccanica quantistica richiedono questa simmetria, che per ora ammetteremo senz'altro.

Due osservabili si dicono *compatibili* se le loro osservazioni sono sempre compatibili, qualunque siano i loro risultati: questo concetto è uno dei più importanti della meccanica quantistica. Un esempio di osservabili compatibili è dato da due coordinate (p. es. x, y) di una particella nel medesimo istante ⁽¹⁾: invece una coordinata x ed il momento coniugato p_x non sono compatibili, come si è visto nel § 22, P. II.

Si dirà poi che più osservabili contemporanee A, B, C, \dots sono compatibili, se ognuna di esse è compatibile con tutte le altre.

§ 17. - OSSERVABILI FUNZIONI DI ALTRE.

Data un'osservabile X , e data una funzione $f(\lambda)$, l'osservabile $G = f(X)$ è definita dal seguente insieme di operazioni: eseguire l'osservazione X e sul risultato X_r , eseguire le operazioni matematiche indicate dalla funzione f : il numero $G_r = f(X_r)$ è per definizione il risultato della misura dell'osservabile G . Si può anche dire che la G si misura con lo stesso apparecchio che misura la X salvo la sostituzione della scala graduata su cui si leggevano i numeri X_r , con un'altra scala in cui, nelle stesse posizioni, si leggono i numeri $f(X_r)$.

Evidentemente, se la f è funzione univalente come supporremo, e se le G_r sono discrete, le loro probabilità sono uguali a quelle delle corrispondenti X_r : se invece la X ha uno spettro continuo di valori X' , con densità di probabilità $P(X')$, i valori G' della osservabile G formano uno spettro continuo con densità di probabilità $Q(G') = P(X') \left| \frac{df}{dX'} \right|$.

Risulta senz'altro dalla definizione che un'osservabile X è compatibile con qualunque $f(X)$.

Passiamo ora alla definizione di una funzione di più

⁽¹⁾ Almeno se i processi di misura con cui si definiscono dette osservabili soddisfano certe condizioni: v. § 24.

osservabili X, Y, Z, \dots (relative allo stesso istante). Se queste sono compatibili tra loro, il procedimento di definizione di $F(X, Y, Z, \dots)$ è l'immediata estensione di quello relativo al caso di una sola osservabile: misura simultanea di X, Y, Z, \dots e, sui risultati numerici X_r, Y_r, Z_r, \dots , operazioni matematiche espresse dalla funzione $F(X_r, Y_r, Z_r, \dots)$.

Risulta poi evidente che la osservabile F così definita è compatibile con ciascuna delle osservabili date X, Y, Z, \dots

Nel caso che invece X, Y, Z, \dots non siano compatibili tra loro, questo procedimento evidentemente non è più applicabile. Tuttavia, data una funzione di più variabili $F(x, y, z, \dots)$, si può definire (almeno sotto condizioni assai larghe) una osservabile $F(X, Y, Z, \dots)$, nel modo seguente.

Consideriamo anzitutto il caso della somma di più osservabili X, Y, Z, \dots non compatibili. Data un'osservabile G (definita con non importa quale procedimento) daremo il seguente criterio per decidere se essa è uguale alla somma $X + Y + Z + \dots$. Si consideri un gran numero di esemplari del sistema considerato (tutti egualmente costituiti, ma eventualmente in stati differenti), se ne estragga a caso un gruppo molto numeroso, su ognuno dei sistemi di questo gruppo si misuri X , e sia \bar{X} la media dei risultati ottenuti; se ne estragga poi a caso un altro gruppo molto numeroso, su cui si misuri Y e si calcoli \bar{Y} , e così via, infine un ultimo gruppo molto numeroso su cui si misuri G e si calcoli \bar{G} . Se avviene (per qualsiasi stato iniziale dell'insieme dei sistemi) che

$$\bar{G} = \bar{X} + \bar{Y} + \bar{Z} + \dots$$

diremo che G è la somma di X, Y, Z, \dots , e scriveremo ⁽¹⁾

$$G = X + Y + Z + \dots$$

⁽¹⁾ Questa definizione non permette di costruire il procedimento di misura di G conoscendo quelli di X, Y, Z, \dots ; essa lascia quindi aperta la questione, se esista in ogni caso un'osservabile G che goda la proprietà indicata. Tuttavia, nei casi che praticamente si presentano ciò si verifica effettivamente, e noi ammetteremo per postulato che ciò avvenga sempre. P. es., l'energia cinetica di un elettrone è definita da $X = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2)$ e quindi ricavabile da misure di impulso,

Si osservi che questa definizione, se X, Y, Z, \dots sono compatibili tra loro, non è in contrasto con quella data sopra per una funzione di più osservabili compatibili. Questa definizione di somma conserva poi evidentemente tutte le proprietà ordinarie di questa operazione.

Dalla definizione di somma si può passare a quella di « prodotto simmetrizzato », cioè di $\frac{1}{2}(XY + YX)$. Difatti, supposto che esista un'osservabile G tale che $G = X + Y$, si conserveranno le proprietà formali dell'algebra se potremo scrivere $G^2 = X^2 + Y^2 + XY + YX$: poichè X^2, Y^2, G^2 rappresentano osservabili già definite, ciò conduce a introdurre la definizione seguente del prodotto simmetrizzato

$$\frac{1}{2}(XY + YX) = \frac{1}{2}(G^2 - X^2 - Y^2):$$

diremo cioè che un'osservabile F è uguale a $\frac{1}{2}(XY + YX)$ se

essa è uguale (nel senso specificato sopra) a $\frac{1}{2}(G^2 - X^2 - Y^2)$:

Se poi le osservabili X, Y sono compatibili, il loro prodotto simmetrizzato si identifica col prodotto XY o YX . Se invece sono incompatibili, non si può dare un significato alle scritture XY, YX , ma solo a $\frac{1}{2}(XY + YX)$. Con procedimento analogo si possono definire i prodotti simmetrizzati di quanti si vogliono fattori.

Data ora una funzione di più variabili $F(x, y, z, \dots)$ sviluppabile in serie di potenze, si può sempre scrivere ciascun termine della serie in forma simmetrizzata, e quindi attribuire un significato alla scrittura che si ottiene sostituendo x, y, z ecc. con altrettante osservabili (anche non compatibili), X, Y, Z, \dots : resta così definito il significato di $F(X, Y, Z, \dots)$.

l'energia potenziale da $Y = eV(x, y, z)$ ed è ricavabile da misure di posizione: esse sono osservabili incompatibili, ma l'energia totale E (definibile con riferimento ai termini spettroscopici), gode la proprietà $\bar{E} = \bar{X} + \bar{Y}$ e quindi si può scrivere $E = X + Y$.

§ 18. — OSSERVAZIONI MASSIME. STATO DI UN SISTEMA.

Nella meccanica classica, quando siano assegnate le posizioni e le velocità di tutti i punti di un sistema in un dato istante, si è definito *completamente* il suo « stato » nell'istante considerato: con ciò si vuol dire che qualunque ulteriore condizione (p. es., l'assegnazione di un dato valore all'energia) sarebbe o incompatibile con le precedenti, o automaticamente soddisfatta.

Vediamo ora come si può trasportare questa nozione nella meccanica quantistica. Se noi assegnamo, p. es., le coordinate di un sistema di particelle in un dato istante, abbiamo già raggiunto una descrizione completa dello stato del sistema, nel senso precisato sopra: infatti una ulteriore condizione sarebbe o automaticamente soddisfatta, o incompatibile con le precedenti (p. es., se si assegnasse anche il valore di una componente di velocità, questa condizione sarebbe incompatibile con le precedenti in virtù del principio di indeterminazione). Si dirà allora che la misura di tutte le coordinate delle particelle costituisce un « gruppo completo di osservazioni » o anche una « osservazione massima » perchè fornisce, per così dire, un massimo di notizie ricavabili da misure eseguite sul sistema. Invece delle coordinate, si potrebbero misurare tutte le componenti dell'impulso, e si avrebbe un altro gruppo completo di osservazioni. In generale, diremo che un gruppo di osservazioni è *completo*, o che costituisce un'osservazione massima, se non esiste nessun'altra osservazione che sia indipendente da esse e con esse compatibile.

I risultati di un'osservazione massima definiscono completamente lo stato del sistema (supposto isolato). Resta così precisata la nozione quantistica di « stato ». Si noti che lo stato così definito non si riferisce a un particolare istante, ma a tutto l'intervallo di tempo in cui il sistema resta isolato: durante questo tempo naturalmente il sistema si evolve secondo un processo ben determinato e caratteristico dello stato in questione. Così, p. es., se si sono misurate le tre coordinate di una particella al tempo $t = 0$, lo stato della particella per tutto il tempo successivo (finchè essa entri in rela-

zione con altri sistemi, eventualmente col campo di radiazione) resta lo stesso, ed è definito dal fatto che per $t = 0$ le coordinate avevano quei dati valori (naturalmente, ciò non significa che le coordinate li conservino: anzi, esse cessano subito di avere valori determinati). Ad ogni stato (come vedremo più avanti generalizzando quanto abbiamo già detto per una sola particella) corrisponde una certa funzione ψ che lo caratterizza.

Se una delle osservazioni che servono a definire lo stato è una misura di energia, si ha uno di quegli stati che nel § 27, P. II abbiamo chiamato *stati semplici o stazionari* ⁽¹⁾ o *quantici* e a cui corrispondono le autofunzioni della equazione di SCHRÖDINGER.

Quanto precede si applica a sistemi isolati: si può però, almeno in molti casi, estendere la nozione di stato anche a sistemi soggetti ad azioni esterne, definendo lo stato del sistema in un dato istante t come lo stato in cui esso resterebbe se al tempo t cessasse l'azione esterna. Si capisce che in generale lo stato di un sistema non isolato è variabile col tempo.

Tornando al caso di un sistema isolato, si noti che, per eseguire su di esso un'osservazione qualunque, si deve necessariamente farlo reagire con un altro sistema (cioè con l'apparecchio di misura) e quindi far cessare, sia pure momentaneamente, l'isolamento: perciò *una osservazione eseguita su un sistema ne altera generalmente lo stato*. Vi sono però dei casi in cui ciò non avviene: p. es., se il sistema è in uno stato stazionario si può misurare l'energia (ritrovando, naturalmente, il valore iniziale) senza alterarne lo stato.

Confrontiamo la nozione quantistica di « osservazione massima » e di « stato » con le corrispondenti classiche. Osserviamo anzitutto che, nella meccanica classica, quando si conosce lo

⁽¹⁾ Questa denominazione, di cui si vedrà la ragione al § 24, non deve far credere che questi siano i soli stati che non variano col tempo. P. es., sovrapponendo due stati stazionari col prendere come ψ una combinazione lineare di due autofunzioni di Schrödinger, (v. § 29, p. II) si ha uno stato *non* stazionario, ma tuttavia invariabile nel tempo.

stato del sistema al tempo 0 (nel senso specificato al principio di questo §) si può calcolare il valore di qualsiasi grandezza meccanica ad esso relativa: quindi la conoscenza p. es. delle coordinate e velocità al tempo 0 equivale alla conoscenza di tutte le variabili meccaniche al tempo 0 (e anche in qualsiasi altro istante). Invece, nella meccanica quantistica, quando si è eseguita una osservazione massima non si può in generale calcolare il risultato di una ulteriore osservazione (sia pure eseguita nello stesso istante) ma si possono soltanto assegnare i possibili risultati e le relative probabilità (con un metodo la cui forma generale sarà spiegata in questo capitolo): la conoscenza completa dello « stato » del sistema non implica dunque affatto la conoscenza dell'insieme di tutte le grandezze ad esso relative (anzi, non si può attribuire a questo insieme nessun significato fisico): vi sono in generale soltanto alcune osservabili che hanno un valore definito, mentre sulle altre non si possono dare altro che indicazioni probabilistiche.

Un'altra differenza essenziale tra la meccanica classica e quella quantistica sta nel fatto che, nella prima, se si sostituisce al gruppo dei valori delle coordinate e delle velocità quello di altrettante loro funzioni indipendenti, si ha una rappresentazione dello stato in tutto equivalente alla precedente; nella meccanica quantistica invece vi sono infiniti modi, non equivalenti tra loro, di dare una descrizione completa dello stato del sistema, a seconda del gruppo completo di osservazioni che si sceglie: p. es. assegnare i valori delle coordinate, oppure assegnare i valori degli impulsi, sono due modi, altrettanto completi, di definire lo stato del sistema, ma non sono equivalenti, poichè le osservabili che hanno un valore definito nel primo caso non lo hanno nel secondo, e viceversa. Si veda anche, a questo proposito, il § 32.

§ 19. — INTERPRETAZIONE DEL METODO DI SCHRÖDINGER NELLO SPAZIO HILBERTIANO.

Vogliamo ora ricapitolare brevemente il procedimento della meccanica ondulatoria di SCHRÖDINGER, enunciandolo col linguaggio geometrico dello spazio hilbertiano: si scorderà in questo modo la via per importanti generalizzazioni.

Cominciamo con l'osservare una analogia formale tra l'equazione di SCHRÖDINGER per gli stati stazionari, che scriveremo nella forma

$$(79) \quad \left[-\frac{\hbar^2}{8\pi^2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + U \right] \psi_n = E_n \psi_n,$$

ed il teorema della forza viva della meccanica classica, che scriveremo

$$(80) \quad \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + U = E.$$

L'analogia consiste in questo: se nella (80) si sostituiscono materialmente le variabili

$$(S) \quad \left\{ \begin{array}{l} p_x, \quad p_y, \quad p_z, \\ \text{rispettivamente con gli operatori (hermitiani)} \\ p_x = \frac{h}{2\pi i} \frac{\partial}{\partial x}, \quad p_y = \frac{h}{2\pi i} \frac{\partial}{\partial y}, \quad p_z = \frac{h}{2\pi i} \frac{\partial}{\partial z} \end{array} \right.$$

il primo membro si muta nell'operatore che è applicato a ψ_n nel primo membro della (79).

Ricordando (v. nota al § 52, P. II) che l'espressione dell'energia in funzione delle q e delle p si è indicata genericamente con $\mathcal{H}(q, p)$ e si è chiamata *hamiltoniana*, diremo che l'equazione di SCHRÖDINGER si ottiene trasformando l'hamiltoniana \mathcal{H} in un operatore \mathfrak{H} (che chiameremo *bramiltoniano*) mediante la sostituzione (S) e scrivendo

$$(81) \quad \mathfrak{H}\psi_n = E_n\psi_n,$$

ossia (v. § 10) ricercando gli autovalori e le autofunzioni di questo operatore (che risulta hermitiano): i suoi autovalori E_n rappresentano i possibili valori dell'energia, e le sue autofunzioni ψ_n , che si possono interpretare come vettori unitari e ortogonali, diretti secondo gli assi principali dell'operatore \mathfrak{H} , hanno la seguente interpretazione fisica: quando il sistema è nello stato stazionario n -esimo, se si esegue una osservazione delle coordinate, la densità di probabilità di trovare i valori x, y, z è data da $|\psi_n(x, y, z)|^2$ cioè, in linguaggio geometrico, dal quadrato del modulo della proiezione del vettore ⁽¹⁾ ψ_n sull'asse (del sistema di assi continui, v. § 2) individuato dai valori x, y, z . Diremo brevemente che « l'ampiezza di probabilità » dei valori x, y, z è uguale a questa proiezione.

Quando poi il sistema non è in uno stato stazionario,

⁽¹⁾ Per semplicità useremo la stessa lettera per indicare una funzione ψ e il vettore corrispondente nello spazio hilbertiano (anziché usare per quest'ultimo il grassetto, come nel cap. prec.).

cioè la ψ ha la forma generale $\psi = \sum_n c_n \psi_n$, ossia il vettore ψ non è diretto secondo uno degli assi principali dell'operatore \mathfrak{H} , vale ancora lo stesso significato per le componenti di questo vettore sugli assi continui. Inoltre, se si esegue una misura di energia, la probabilità di trovare il valore E_n è $|c_n|^2$ (v. § 29 p. II): cioè l'ampiezza di questa probabilità è c_n , proiezione del vettore ψ sull' n -esimo asse principale dell'operatore \mathfrak{H} .

Il vettore ψ determina dunque la probabilità dei risultati di qualunque misura di coordinate o di energia (e anche, come si vedrà poi, di qualunque altra osservazione che si esegua sul sistema): esso è generalmente funzione del tempo, e la sua evoluzione nel tempo è regolata dall'equazione temporale di SCHRÖDINGER (v. § 30 p. II) che ora possiamo scrivere nella forma

$$(82) \quad \mathfrak{H}\psi = -\frac{h}{2\pi i} \frac{\partial \psi}{\partial t}.$$

Essa si può considerare ottenuta dalla equazione classica della conservazione dell'energia $\mathfrak{H} = E$, operando, oltre alla sostituzione (S), anche la sostituzione di

$$(S') \quad E \quad \text{con} \quad -\frac{h}{2\pi i} \frac{\partial}{\partial t},$$

e applicando gli operatori ottenuti alla ψ .

Il vettore ψ , considerato come funzione del tempo, caratterizza lo « stato » del sistema e verrà chiamato nel seguito « *vettore di stato* ».

§ 20. — ESTENSIONE A UN SISTEMA DI N PARTICELLE DISTINTE.

La forma data nel § prec. al problema di SCHRÖDINGER per una sola particella suggerisce, per ovvia generalizzazione, il modo di trattare il problema più generale di quante si vogliano particelle *distinte* ⁽¹⁾.

⁽¹⁾ " Distinte " vuol dire che supponiamo che ogni particella abbia una propria individualità, che cioè si possa distinguere dalle altre: se si trattasse di particelle identiche (p. es., elettroni) si dovrebbero fare altre considerazioni, che rimandiamo al cap. VI.

Sia N il numero delle particelle: le loro $3N = f$ coordinate saranno indicate talvolta con $x_1, y_1, z_1, x_2, y_2, z_2, \dots, z_N$, e talvolta, se farà comodo, con q_1, q_2, \dots, q_f . Si dovrà (generalizzando il criterio di impostazione seguito al § 25, P. II), introdurre la probabilità che una osservazione simultanea di tutte le particelle al tempo t dia per le coordinate valori compresi tra q_1, q_2, \dots, q_f e $q_1 + dq_1, q_2 + dq_2, \dots, q_f + dq_f$ rispettivamente: questa probabilità sarà indicata con

$$P(q_1, \dots, q_f, t) dq_1 dq_2 \dots dq_f.$$

Se invece si esegue una osservazione della sola particella k -esima, senza occuparsi delle altre, la probabilità di trovarla nel volume elementare definito da $x_k, y_k, z_k, x_k + dx_k, y_k + dy_k, z_k + dz_k$ sarà

$$P_k(x_k, y_k, z_k, t) dx_k dy_k dz_k,$$

dove P_k sarà ottenuta dalla P integrandola rispetto a tutte le coordinate, tranne x_k, y_k, z_k , e per tutti i valori che quelle coordinate possono assumere.

È evidente che le N funzioni P_k restano determinate quando sia nota la P , ma in generale la P non è determinata dalla conoscenza delle P_k . Ciò si comprende, oltrechè per via analitica, con la seguente considerazione intuitiva. Vi sia una scatola con due scompartimenti uguali A e B entro cui vengono gettate a caso due palline a, b assai piccole e *indipendenti tra loro*: se si osserva una sola di esse, la probabilità di trovarla in A o in B è $\frac{1}{2}$: se si osservano entrambe, si possono presentare i quattro casi,

a in A , b in B :	probabilità	$\frac{1}{4}$
a » A , b » A	»	$\frac{1}{4}$
a » B , b » B	»	$\frac{1}{4}$
a » B , b » A	»	$\frac{1}{4}$

Supponiamo ora invece che le due palline siano così grandi che in ogni scompartimento ve ne entri una sola: allora per ciascuna di esse vi è ancora la probabilità $\frac{1}{2}$ di trovarla in A o in B , ma per l'insieme delle due, le probabilità dello specchietto precedente diventano rispettivamente $\frac{1}{2}, 0, 0, \frac{1}{2}$. Ciò mostra che la conoscenza delle probabilità relative all'osservazione di ciascuna pallina separatamente non basta a determinare le probabilità relative all'osservazione delle due: è infatti necessario sapere anche come la posizione di una pallina influenza la probabilità

della posizione dell'altra. Se quest'influenza è nulla (come nel primo caso) si dice che le due palline sono « statisticamente indipendenti ».

Tornando al caso delle N particelle elementari, diremo che esse sono « statisticamente indipendenti » se la $P(q_1, q_2, \dots, q_f, t)$ ha la forma seguente ⁽¹⁾

$$(83) \quad P = P_1(x_1, y_1, z_1, t) P_2(x_2, y_2, z_2, t) \dots P_N(x_N, y_N, z_N, t).$$

È questo il solo caso in cui la P è determinata dalla conoscenza delle P_k . Si noti che le particelle possono non essere statisticamente indipendenti, anche se tra esse non agiscono forze: è questa una considerazione assai importante sulla quale torneremo in seguito.

Analogamente a quanto fu fatto per una sola particella, introdurremo una funzione (complessa)

$$\psi(q_1, q_2, \dots, q_f, t)$$

tale che

$$P(q_1, q_2, \dots, q_f, t) = |\psi|^2$$

e imporremo alla ψ di soddisfare un'equazione, che sia la generalizzazione di quella di SCHRÖDINGER per una sola particella: questa generalizzazione sarà fatta fondandosi sull'analogia rilevata al § 19. Partiamo perciò dall'espressione classica dell'hamiltoniana di un sistema di N particelle in coordinate cartesiane: supponiamo che su di esse agiscano delle forze ⁽²⁾ derivate da un potenziale $U(q_1, q_2, \dots, q_f)$ e indichiamo con $p_x^{(k)}, p_y^{(k)}, p_z^{(k)}$ i momenti coniugati a x_k, y_k, z_k (oppure con p_j il momento coniugato a q_j): la massa della k -esima particella sarà indicata con $m^{(k)}$ e la sua carica con $e^{(k)}$. Allora l'hamiltoniana del sistema si scrive

$$\mathcal{H} = \sum_k^N \frac{1}{2m^{(k)}} (p_x^{(k)^2} + p_y^{(k)^2} + p_z^{(k)^2}) + U(x_1, y_1, \dots, z_N).$$

⁽¹⁾ Si verifica immediatamente che integrando questa P rispetto a tutte le variabili meno x_k, y_k, z_k , per tutto il loro campo di variabilità, si ottiene $P_k(x_k, y_k, z_k)$.

⁽²⁾ Trascuriamo le azioni magnetiche tra le particelle del sistema le quali sono intimamente legate alle correzioni relativistiche che saranno introdotte al cap. V.

ovvero, con notazioni più comode,

$$(84) \quad \mathcal{H} = \sum_1^{3N} \frac{1}{2m_j} p_j^2 + U(q),$$

dove $m_1 = m_2 = m_3$ rappresentano $m^{(1)}$, $m_4 = m_5 = m_6$ rappresentano $m^{(2)}$ ecc.

Trasformiamo questa espressione in operatore, come si è fatto nel caso di una sola particella, sostituendo

$$(85) \quad p_j \text{ con } \mathfrak{p}_j = \frac{\hbar}{2\pi i} \cdot \frac{\partial}{\partial q_j}$$

e avremo

$$\mathfrak{H} = -\frac{\hbar^2}{8\pi^2} \sum_1^f \frac{1}{m_j} \frac{\partial^2}{\partial q_j^2} + U(q),$$

ovvero, indicando con $\Delta_k = \frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2}$ l'operatore di LAPLACE relativo alla particella k -esima,

$$(86) \quad \mathfrak{H} = -\sum_1^N \frac{\hbar^2}{8\pi^2 m_k} \Delta_k + U(q).$$

Ammetteremo ora che la ψ del sistema soddisfi l'equazione seguente, generalizzazione dell'equazione temporale di SCHRÖDINGER (v. (136) P. II):

$$(87) \quad \mathfrak{H}\psi = -\frac{\hbar}{2\pi i} \frac{\partial \psi}{\partial t},$$

ossia

$$(87') \quad \sum_1^N \frac{1}{m_k} \Delta_k \psi - \frac{8\pi^2}{\hbar^2} U\psi = -\frac{4\pi i}{\hbar} \frac{\partial \psi}{\partial t}.$$

Gli stati stazionari o a energia definita saranno quelli per cui la ψ è un'autofunzione ψ_n dell'operatore \mathfrak{H} , tale cioè che sia ⁽¹⁾

$$(88) \quad \mathfrak{H}\psi_n = E_n \psi_n.$$

La dipendenza dal tempo di queste ψ_n si ottiene confrontando la (88) con la (87), il che dà

$$\frac{\partial \psi_n}{\partial t} = -\frac{2\pi i}{\hbar} E_n \psi_n$$

⁽¹⁾ Naturalmente l'indice n tiene qui il luogo di un gruppo di f indici.

e quindi

$$(89) \quad \psi_n = u_n(q) e^{-\frac{2\pi i}{\hbar} E_n t},$$

con la $u_n(q)$ soddisfacente l'equazione

$$(88') \quad \mathfrak{H}u_n = E_n u_n,$$

come nel caso di una particella sola.

La ψ più generale si può naturalmente sviluppare in serie delle (89), cioè qualunque stato del sistema si può considerare come una sovrapposizione di stati stazionari.

È interessante verificare che l'equazione (87), da noi ammessa per induzione, è soddisfatta, in particolare, nel caso di N particelle *dinamicamente indipendenti*, cioè che non esercitano forze tra loro, ma sono soggette solo a forze esterne. In tal caso evidentemente la U si spezza nella somma di N funzioni U_k , ciascuna contenente le coordinate di una sola particella e quindi anche l'operatore (86) si spezza nella somma di N operatori \mathfrak{H}_k , ciascuno operante sulle coordinate di una sola particella, e cioè

$$\mathfrak{H}_k = -\frac{\hbar^2}{8\pi^2 m_k} \Delta_k + U_k;$$

così la (87) diviene

$$(87'') \quad \sum_1^N \mathfrak{H}_k \psi = -\frac{\hbar}{2\pi i} \frac{\partial \psi}{\partial t}.$$

Ora, si vede subito che questa equazione può essere soddisfatta prendendo

$$(90) \quad \psi = \psi^{(1)} \psi^{(2)} \dots \psi^{(N)}$$

dove $\psi^{(k)}$ è una funzione delle sole x_k, y_k, z_k, t , che soddisfa l'equazione

$$\mathfrak{H}_k \psi^{(k)} = -\frac{\hbar}{2\pi i} \frac{\partial \psi^{(k)}}{\partial t},$$

cioè l'ordinaria equazione di Schrödinger relativa alla particella k -esima. Si può dunque prendere come ψ del sistema il prodotto delle ψ delle singole particelle, e in tal caso evidentemente è soddisfatta la (83), cioè le particelle sono statisticamente indipendenti. E se ognuna delle particelle è in uno stato stazionario, (di indice n_k), cioè se

$$\mathfrak{H}_k \psi_{n_k}^{(k)} = E_{n_k}^{(k)} \psi_{n_k}^{(k)},$$

si verifica subito che la ψ , prodotto di tutte le $\psi_{n_k}^{(k)}$, soddisfa l'equazione

$$\sum_k \mathfrak{H}_k \psi = E_{n_1 n_2 \dots n_N} \psi,$$

con

$$E_{n_1 n_2 \dots n_N} = \sum_k E_{n_k}^{(k)},$$

vale a dire: il sistema è in uno stato stazionario, e la sua energia è la somma delle energie delle singole particelle.

Dunque lo spezzarsi dell'hamiltoniana nella somma di N termini ciascuno dei quali dipende dalle coordinate di una sola particella porta con sé la possibilità di scindere la ψ nel prodotto di N fattori, ciascuno corrispondente a una particella, e di spezzare ogni autovalore nella somma di N termini, ciascuno rappresentante l'energia di una particella.

Ma si badi che la (87'') (stabilita nella sola ipotesi della indipendenza dinamica delle particelle) ammette, oltre alle soluzioni del tipo (90), anche qualsiasi combinazione lineare di soluzioni siffatte: queste soluzioni rappresentano casi in cui le particelle non sono statisticamente indipendenti (benchè indipendenti dinamicamente).

Si riconosce facilmente che, se in un dato istante t_0 le particelle sono statisticamente indipendenti, esse lo sono anche in qualunque altro istante (purchè, beninteso, non agiscano forze tra loro): infatti, per la (87''), la ψ è univocamente determinata dai suoi valori per $t = t_0$ in tutto lo spazio, quindi se per $t = t_0$ vale la soluzione (90), essa vale anche per qualunque altro t .

Si noti che, mentre nel caso di una sola particella la ψ rappresenta delle onde, fittizie, ma nello spazio ordinario, nel caso di N particelle non si può in generale conservare una tale immagine, poichè la ψ contiene $3N$ coordinate (oltre t): essa può quindi essere interpretata solo mediante onde in uno spazio a $3N$ dimensioni.

§ 21. - IL PROBLEMA DEI DUE CORPI.

Come applicazione, consideriamo il problema che si presenta nella trattazione dei sistemi idrogenoidi quando si vuol tener conto del fatto che il nucleo non è fisso come lo si è supposto al § 48 p. II: tale problema fu già trattato, dal punto di vista di BOHR e SOMMERFELD, nel § 58, P. II. Esso corri-

sponde al noto problema dei due corpi nella meccanica classica.

Siano x_1, y_1, z_1 le coordinate del nucleo, x_2, y_2, z_2 quelle dell'elettrone (rispetto ad assi fissi qualunque) e siano $p_{1x}, p_{1y}, p_{1z}, p_{2x}, p_{2y}, p_{2z}$ i momenti rispettivamente coniugati a queste coordinate: sia poi M la massa del nucleo, m quella dell'elettrone. L'hamiltoniana classica è

$$(91) \quad \mathcal{H} = \frac{1}{2M} (p_{1x}^2 + p_{1y}^2 + p_{1z}^2) + \frac{1}{2m} (p_{2x}^2 + p_{2y}^2 + p_{2z}^2) + U,$$

dove U è l'energia potenziale, che, dipendendo solo dalla posizione relativa, sarà funzione di $x_2 - x_1, y_2 - y_1, z_2 - z_1$.

La $\psi(x_1, y_1, z_1, x_2, y_2, z_2, t)$ dovrà dunque soddisfare l'equazione

$$(92) \quad \mathfrak{H}\psi = -\frac{h}{2\pi i} \cdot \frac{\partial \psi}{\partial t},$$

con

$$(93) \quad \mathfrak{H} = -\frac{h^2}{8\pi^2 M} \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2} \right) - \frac{h^2}{8\pi^2 m} \left(\frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial z_2^2} \right) + U(x_2 - x_1, y_2 - y_1, z_2 - z_1).$$

Esprimiamo questo operatore, invece che con le sei variabili $x_1, y_1, z_1, x_2, y_2, z_2$, con le tre coordinate del baricentro

$$\xi = \frac{Mx_1 + mx_2}{M + m}, \quad \eta = \frac{My_1 + my_2}{M + m}, \quad \zeta = \frac{Mz_1 + mz_2}{M + m}$$

e le tre coordinate dell'elettrone rispetto al nucleo

$$x = x_2 - x_1, \quad y = y_2 - y_1, \quad z = z_2 - z_1.$$

Si ha evidentemente

$$\frac{\partial}{\partial x_1} = \frac{\partial \xi}{\partial x_1} \frac{\partial}{\partial \xi} + \frac{\partial x}{\partial x_1} \frac{\partial}{\partial x} = \frac{M}{M + m} \frac{\partial}{\partial \xi} - \frac{\partial}{\partial x}$$

e quindi

$$\frac{\partial^2}{\partial x_1^2} = \left(\frac{\partial}{\partial x_1} \right)^2 = \left(\frac{M}{M + m} \right)^2 \frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial x^2} - \frac{2M}{M + m} \frac{\partial^2}{\partial \xi \partial x};$$

similmente si trova

$$\frac{\partial^2}{\partial x_2^2} = \left(\frac{\partial}{\partial x_2} \right)^2 = \left(\frac{m}{M + m} \right)^2 \frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial x^2} + \frac{2m}{M + m} \frac{\partial^2}{\partial \xi \partial x}.$$

Sostituendo in (93) si vede che i termini con le derivate miste si elidono, e analogamente per le altre coordinate, e resta

$$\mathfrak{H} = -\frac{h^2}{8\pi^2(M+m)}\left(\frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial\eta^2} + \frac{\partial^2}{\partial\zeta^2}\right) - \frac{h^2}{8\pi^2}\left(\frac{1}{m} + \frac{1}{M}\right)\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + U(x, y, z).$$

Poichè la U non contiene ξ, η, ζ , l'operatore si spezza nella somma di uno, \mathfrak{H}_0 , contenente solo ξ, η, ζ , l'altro, \mathfrak{H}' , che contiene solo x, y, z :

$$\mathfrak{H}_0 = -\frac{h^2}{8\pi^2(M+m)}\left(\frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial\eta^2} + \frac{\partial^2}{\partial\zeta^2}\right)$$

$$\mathfrak{H}' = -\frac{h^2}{8\pi^2 m'}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + U(x, y, z),$$

dove si è posto

$$\frac{1}{m'} = \frac{1}{m} + \frac{1}{M} \quad \text{cioè} \quad m' = \frac{mM}{M+m} = \frac{m}{1+m/M};$$

(m' prende il nome di *massa ridotta*). Corrispondentemente la ψ potrà spezzarsi (v. § 20) nel prodotto

$$\psi = \chi(\xi, \eta, \zeta) \Psi(x, y, z),$$

dove le due funzioni χ e Ψ soddisfano alle

$$(94) \quad \mathfrak{H}_0 \chi = -\frac{h}{2\pi i} \frac{\partial \chi}{\partial t},$$

$$(95) \quad \mathfrak{H}' \Psi = -\frac{h}{2\pi i} \frac{\partial \Psi}{\partial t}.$$

Come si vede, l'operatore \mathfrak{H}' coincide con quello che si presenta nel problema del moto dell'elettrone attorno al nucleo supposto *fisso*, salvo la sostituzione della massa m con la m' (leggermente inferiore): la Ψ dunque coinciderà con la ψ della teoria svolta al § 46, P. II, purchè si sostituisca la massa m con la massa ridotta m' (questa modificazione è la stessa che si deve fare nella meccanica classica ovvero nella teoria di BOHR e SOMMERFELD per tener conto del movimento del nucleo: v. § 58, P. II).

Quanto all'operatore \mathfrak{H}_0 , esso è quello che si presenta nel moto di una particella di massa $M+m$ non soggetta

a forze: la χ è dunque un'autofunzione del punto libero (v. § 44, P. II): ciò significa che il moto del baricentro si può trattare, anche in meccanica ondulatoria, come il moto di un punto di massa uguale alla massa totale del sistema (¹).

Quanto ai livelli energetici, ciascuno di essi risulta (v. § prec.) dalla somma di un autovalore della (94) e uno della (95): ciò significa che l'energia dell'atomo in uno stato stazionario si scinde in un'energia di traslazione e un'energia interna, come nella meccanica ordinaria. Per quest'ultima energia (che sola interessa la spettroscopia) valgono dunque gli autovalori trovati al § 48, p. II, con la sola lieve correzione della sostituzione della massa m con la massa ridotta m' .

§ 22. - PRINCIPIO GENERALE DELLA MECCANICA QUANTISTICA. CORRISPONDENZA TRA OSSERVABILI E OPERATORI.

Nel procedimento enunciato nei §§ 19 e 20 l'osservabile « energia » ha una parte privilegiata: si può cercare di generalizzare queste considerazioni, dando invece ad un'altra osservabile generica G il ruolo tenuto fin qui dall'energia: la generalizzazione si farà per analogia col caso precedente, salvo a verificarne la legittimità confrontando con l'esperienza le conseguenze dedotte.

Ammettiamo perciò il seguente *postulato*: ad ogni osservabile G si può far corrispondere un operatore lineare hermitiano \mathfrak{G} che ha le seguenti proprietà: *i suoi autovalori G_r rappresentano (²) i possibili risultati di una misura di G , e i suoi assi principali, individuati dalle autofunzioni φ_r dell'equazione*

$$(96) \quad \mathfrak{G}\varphi_r = G_r\varphi_r,$$

godono la proprietà che: la proiezione del vettore di stato ψ sull'asse principale r -esimo fornisce (supposto che G_r non sia

(¹) Questo risultato autorizza ad applicare l'equazione di SCHRÖDINGER al moto d'insieme di un sistema complesso come un atomo o una molecola. Le esperienze sulla diffrazione di H, H_2, He confermano tale risultato (v. § 29 p. I).

(²) Questi autovalori risultano sempre reali, perchè l'operatore è hermitiano.

un autovalore multiplo) l'ampiezza di probabilità che una misura di G dia il risultato G_r . Vale a dire, chiamando ψ_{G_r} questa componente, cioè ponendo (supposte le φ_r normalizzate)

$$\psi_{G_r} = \psi \times \varphi_r = \int \psi \varphi_r^* dS,$$

la probabilità del valore G_r è $|\psi_{G_r}|^2$. (Si noti che, essendo il vettore ψ unitario si ha $\sum_r |\psi_{G_r}|^2 = 1$, come deve essere per poter attribuire alle $|\psi_{G_r}|^2$ il detto significato di probabilità). Se l'operatore ha invece uno spettro continuo di autovalori, indicheremo uno generico degli autovalori continui con G' , anzichè con G_r , e con $\varphi_{G'}$, la corrispondente autofunzione (normalizzata col criterio del § 10, p. II): allora $|\varphi_{G'}|^2 dG'$ sarà la probabilità che una osservazione di G dia un risultato compreso tra G' e $G' + dG'$. Generalmente, per comodità di scrittura, ci riferiremo al caso di autovalori discreti, intendendo che nel caso continuo si facciano alle formule gli ovvii adattamenti.

Se, in particolare, il vettore di stato giace su uno degli assi principali di \mathcal{G} (cioè se $\psi = \varphi_r$), il sistema è in uno stato tale che una misura di G dà con certezza il valore G_r : ciò avviene in particolare per l'energia quando lo stato è uno « stato stazionario ».

Osservazione sui casi di degenerazione. — Se G_r è un autovalore multiplo di ordine p , cui corrispondono p autofunzioni ortogonali $\varphi_r^1, \varphi_r^2, \dots, \varphi_r^p$, dette ψ_r^i ($i = 1, 2, \dots, p$) le proiezioni di ψ su questi assi, come probabilità del valore G_r si deve prendere

$$(97) \quad P_r = \sum_i^p |\psi_r^i|^2$$

(che è, si noti, invariante rispetto a qualsiasi sostituzione ortogonale delle φ_r^i). Ciò si può giustificare considerando l'operatore degenero \mathcal{G} come limite di un operatore non degenero \mathcal{H} , ponendo p. es. $\mathcal{H} = \mathcal{G} + \varepsilon \mathcal{F}$ (dove \mathcal{F} è scelto in modo che \mathcal{H} non sia degenero, ed ε è una quantità che si fa tendere a 0). Allora all'autovalore G_r di \mathcal{G} corrispondono p autovalori H_r^i ($i = 1, 2, \dots, p$) di \mathcal{H} , i quali vengono a coincidere con G_r quando ε tende a 0, e ad essi corrispondono altrettante autofunzioni $\bar{\varphi}_r^i$, e altrettante proiezioni $\bar{\psi}_r^i$ di ψ : la probabilità del valore H_r^i dell'osservabile

H è $|\bar{\psi}_r^i|^2$, e quindi la probabilità totale dei p valori $H_r^1, H_r^2, \dots, H_r^p$, che per $\varepsilon \rightarrow 0$ confluiscono in G_r , è $\sum_i |\bar{\psi}_r^i|^2$: facendo tendere ε a 0 questa espressione ha per limite la (97).

Un caso analogo si presenta se l'operatore \mathcal{G} è incompleto. Siano p. es. x e y le variabili del problema, e l'operatore \mathcal{G} contenga solo la x : sappiamo allora (v. § 10) che all'autovalore G_r corrispondono le infinite autofunzioni $\varphi_r^j = w^j(y) \chi_r(x)$ (dove le $w^j(y)$ rappresentano un qualsiasi sistema completo di autofunzioni ortogonali e normalizzate nello spazio delle funzioni di y): perciò, conformemente alla (97), assumeremo ⁽¹⁾ come probabilità del valore G_r

$$(97') \quad P_r = \sum_j^\infty |\psi_r^j|^2.$$

dove $\psi_r^j = \psi \times \varphi_r^j$. Questa espressione è indipendente dalla scelta delle w^j : prendendo come tali le funzioni $\delta(y - y')$, dove y' fa le veci dell'indice j (v. § 14), si ha

$$\psi_r^{y'} = \psi \times \delta(y - y') \chi_r(x) = \iint \psi(x, y) \delta(y - y') \chi_r^*(x) dx dy = \int \psi(x, y') \chi_r^*(x) dx$$

e quindi la (97') si trasforma nell'integrale seguente (dove scriviamo y in luogo di y'):

$$(98) \quad P_r = \int |\psi_r^y|^2 dy.$$

Donde la regola: « per avere la probabilità P_r , si calcolano le autofunzioni $\chi_r(x)$ dell'operatore \mathcal{G} nello spazio delle funzioni della sola x , e si sviluppa la $\psi(x, y)$, considerandola solo come funzione di x , in serie di queste autofunzioni: i coefficienti sono naturalmente funzioni di y (e si possono perciò indicare con ψ_r^y): il quadrato del modulo dell' r -esimo coefficiente, integrato rispetto a y , dà P_r ».

Naturalmente, se \mathcal{G} avesse degli autovalori continui G' , nella (98) si dovrebbe sostituire l'indice discreto r con la variabile continua G' , e si dovrebbe interpretare $P_{G'}$ come densità di probabilità; cosicchè la probabilità di un valore compreso fra G' e dG' sarebbe

$$(99) \quad dG' \int |\psi_{G'}^y|^2 dy.$$

Resta da vedere come si determina l'operatore \mathcal{G} corrispondente a una data osservabile G . Procediamo a tal uopo per via di generalizzazioni successive.

⁽¹⁾ Ciò potrebbe giustificarsi direttamente con un passaggio al limite del tutto analogo a quello fatto per il caso degli autovalori multipli d'ordine p .

a) Caso di $\mathcal{G} = x$. Se l'osservabile G è una delle coordinate della particella, p. es. x , l'operatore \mathcal{G} che ad essa corrisponde è, come ora mostreremo, la moltiplicazione per x , cioè $\mathfrak{x} = x$ (ed è incompleto, salvo il caso unidimensionale). Difatti, come si è visto al § 14, tale operatore ha per autovalori tutti i numeri reali x' , e per autofunzioni $\delta(x - x')$: la proiezione del vettore di stato ψ su questo asse principale (calcolata nello spazio delle funzioni di x) è

$$\psi_{x'} = \int \psi(x, y, z, t) \delta(x - x') dx = \psi(x', y, z, t)$$

e quindi la probabilità di trovare per la x un valore compreso fra x' e $x' + dx'$ è, a norma della (99),

$$dx' \iint |\psi(x', y, z, t)|^2 dy dz,$$

in pieno accordo con il significato di $|\psi|^2$ come densità di probabilità.

b) Caso di $G = p_x$. Se l'osservabile G è una componente dell'impulso, p. es. p_x , possiamo verificare che l'operatore corrispondente è

$$\mathfrak{p}_x = \frac{\hbar}{2\pi i} \frac{\partial}{\partial x}.$$

Difatti la (96) si scrive allora, chiamando p'_x un autovalore generico,

$$\frac{\hbar}{2\pi i} \frac{\partial \varphi_{p'_x}}{\partial x} = p'_x \varphi_{p'_x}.$$

Questa equazione ha per autovalore qualunque valore di p'_x , e dà, con immediata integrazione,

$$(100) \quad \varphi_{p'_x} = \frac{1}{\sqrt{\hbar}} e^{\frac{2\pi i}{\hbar} p'_x x}$$

dove il fattore $1/\sqrt{\hbar}$ è stato determinato mediante la condizione di normalizzazione.

La proiezione di ψ sull'asse principale p'_x è

$$(101) \quad \psi_{p'_x}(y, z, t) = \frac{1}{\sqrt{\hbar}} \int_{-\infty}^{\infty} \psi(x, y, z, t) e^{-\frac{2\pi i}{\hbar} p'_x x} dx$$

e la probabilità che la componente x dell'impulso sia compresa tra p'_x e $p'_x + dp'_x$ è, a norma della (99),

$$(102) \quad dp'_x \iint |\psi_{p'_x}|^2 dy dz.$$

Calcoliamo ora la stessa probabilità mediante il principio di sovrapposizione: se si decompone la ψ in integrale di FOURIER (considerandola solo come funzione di x , ossia supponendo fissati y e z) si ottiene

$$\psi = \frac{1}{\sqrt{\hbar}} \int \psi_{p'_x} e^{\frac{2\pi i}{\hbar} p'_x x} dp'_x,$$

dove $\psi_{p'_x}$ è proprio l'espressione (101): dunque $\psi_{p'_x}$ misura l'ampiezza della componente monocromatica, di lunghezza d'onda \hbar/p'_x , delle onde di DE BROGLIE corrispondenti a una particella di dati y e z (caso unidimensionale, v. § 36, p. II) e quindi, per il principio di sovrapposizione, $|\psi_{p'_x}|^2 dp'_x$ è la probabilità che la particella di dati y e z abbia una componente x dell'impulso compresa fra p'_x e $p'_x + dp'_x$: lasciando ora del tutto indeterminati y e z si ottiene evidentemente per la probabilità di una p_x compresa tra p'_x e $p'_x + dp'_x$ proprio il valore (102).

c) Caso di un'osservabile definita come funzione delle coordinate o dei momenti. Sia A un'osservabile a cui si sappia che corrisponde un certo operatore \mathfrak{A} , con autovalori A_r e autofunzioni φ_r . Sia poi G un'altra osservabile definita come funzione di A , cioè $G = F(A)$: i possibili risultati di una misura di G saranno $F(A_r)$, e ciascuno di essi avrà la stessa probabilità del valore A_r di A . Ora, è facile vedere che questo risultato si ritrova applicando il procedimento di pag. 347 e facendo corrispondere a G l'operatore $\mathcal{G} = F(\mathfrak{A})$. Difatti, come si è visto al § 10, questo operatore ha gli stessi assi principali di \mathfrak{A} (individuati dai vettori φ_r), con gli autovalori $G_r = F(A_r)$: perciò la proiezione del vettore di stato ψ sull'asse r -esimo di \mathcal{G} è la stessa ψ_r che rappresenta la proiezione di ψ sull'asse r -esimo di A , e quindi la probabilità del valore G_r risulta $|\psi_r|^2$, come quella del valore A_r di A . Il teorema si estende immediatamente a una funzione

$F(A, B, \dots)$ di più osservabili compatibili tra loro e si dimostra allo stesso modo.

Ora, avendo già riconosciuto che gli operatori corrispondenti alle coordinate q_r sono le q_r stesse, e quelli corrispondenti ai momenti p_r sono $\frac{\hbar}{2\pi i} \frac{\partial}{\partial q_r}$, possiamo concludere che a un'osservabile definita come $F(q)$ corrisponde l'operatore $F(q)$, e a una $F(p)$ corrisponde $F\left(\frac{\hbar}{2\pi i} \frac{\partial}{\partial q}\right)$, cioè l'operatore ottenuto dalla data funzione sostituendovi ogni p_r con $\frac{\hbar}{2\pi i} \frac{\partial}{\partial q_r}$.

d) *Caso generale.* Sia ora G un'osservabile qualunque, definita direttamente (cioè mediante operazioni che non implicano la misura delle q e delle p): è questo il caso, p. es., dell'energia. Per trovare l'operatore \mathcal{G} che ad essa corrisponde non esiste nessun procedimento generale: però, in quasi tutti i casi che in pratica si presentano, riesce il seguente metodo euristico, fondato sul postulato di una profonda analogia formale tra la meccanica ordinaria e quella quantistica. Esso, naturalmente, va giustificato a posteriori dal successo delle sue conseguenze, il che si è verificato in tutti i casi in cui è stato applicato.

La meccanica ordinaria permette di esprimere il valore della grandezza G al tempo t come funzione $G = F(q, p)$ delle coordinate cartesiane q_i e dei momenti p_i (ed eventualmente del tempo t , che consideriamo fissato). Supponiamo questa funzione sviluppabile in serie di potenze, e scriviamola (se vi sono termini del tipo $q_r^m p_r^n$) in forma simmetrizzata, come si è spiegato al § 17. Ciò fatto, la naturale generalizzazione di quanto si è visto poc'anzi nel caso di una $F(q)$ o di una $F(p)$, conduce ad ammettere il seguente postulato ⁽¹⁾: *all'osservabile G , avente in meccanica classica l'espressione (eventualmente simmetrizzata) $F(q, p)$, corrispon-*

⁽¹⁾ Nel seguito avremo bisogno di applicare questo postulato solo a funzioni della forma $F(q, p) = Q(q) + P(p) + \sum_r A_r(q)p_r$, dove solo l'ultima parte richiede simmetrizzazione.

de l'operatore $\mathcal{G} = F\left(q, \frac{\hbar}{2\pi i} \frac{\partial}{\partial q}\right)$. In questa regola sono evidentemente compresi anche i casi già esaminati sotto (a), (b) e (c).

Vi sono infine alcuni casi in cui non è possibile far corrispondere alla osservabile in questione una espressione della meccanica classica, e in tali casi l'operatore \mathcal{G} che la rappresenta viene introdotto mediante una ipotesi « ad hoc », da giustificarsi mediante le sue conseguenze: tale è il caso, p. es., delle componenti dello spin nella teoria di PAULI (v. § 45).

Riassumendo, il principio generale della meccanica quantistica si può enunciare così. Una volta determinato, o mediante la regola data sopra o mediante una ipotesi speciale, l'operatore \mathcal{G} corrispondente alla osservabile G , si pone l'equazione

$$\mathcal{G}\varphi = G'\varphi$$

(dove G' è un parametro): gli autovalori G_r di questa equazione danno i possibili risultati di una misura dell'osservabile G , e, dette φ_r le corrispondenti autofunzioni, la probabilità del valore G_r nello stato ψ è $|\psi \times \varphi_r|^2$. Per i casi di degenerazione o di autovalori continui, v. pag. 348.

Osserviamo subito che se si applica questa regola generale al caso in cui l'osservabile G è l'energia di una particella, o di un sistema di particelle, soggette a forze derivanti da un potenziale (nel qual caso la funzione $F(q, p)$ è la hamiltoniana $\mathcal{H}(q, p)$) si ritrova il procedimento del § 19 per dedurre l'equazione di SCHRÖDINGER degli stati stazionari, o la sua generalizzazione (88). Tale equazione dunque ci appare ora come un caso particolare del problema generale della ricerca dei valori possibili di un'osservabile G , e delle rispettive probabilità: il caso particolare in questione occupa una posizione di privilegio nella meccanica quantistica a causa della importanza fondamentale dell'osservabile « energia », e della sua proprietà di mantenersi costante.

§ 23. — CENNO SULL'USO DI COORDINATE CURVILINEE.

La regola data al § precedente per trovare l'operatore corrispondente a una data osservabile G suppone che questa venga espressa mediante le coordinate cartesiane q_r e i momenti coniu-

gati p_r . In molti casi però è più comodo procurarsi l'espressione classica di G mediante coordinate lagrangiane qualunque, che indicheremo con Q_r , chiamando P_r i momenti ad esse coniugati: sia $F(Q, P)$ questa espressione. Per ottenere l'operatore \mathcal{G} , bisogna passare alle coordinate cartesiane, applicare la sostituzione (8) e poi trasformare di nuovo l'operatore ottenuto esprimendolo mediante le Q : questo processo non è in generale equivalente alla semplice sostituzione di P_r con $\frac{h}{2\pi i} \frac{\partial}{\partial Q_r}$ nella $F(Q, P)$.

(Un esempio di questo procedimento si è già avuto al § 21).

P. es., nel caso di un punto nel piano, non soggetto a forze, usando le coordinate polari $Q_1 = r$, $Q_2 = \theta$ e i rispettivi momenti $P_1 = mr\dot{\theta}$, $P_2 = mr^2\dot{\theta}$, l'espressione dell'hamiltoniana risulta

$$(105) \quad \mathcal{H} = \frac{1}{2m} \left(P_1^2 + \frac{P_2^2}{r^2} \right).$$

In coordinate cartesiane invece è $\mathcal{H} = \frac{1}{2m} (p_x^2 + p_y^2)$ e l'operatore corrispondente è, come è ben noto,

$$\mathfrak{H} = -\frac{h^2}{8\pi^2 m} \Delta$$

che, espresso in coordinate polari, si scrive:

$$\mathfrak{H} = -\frac{h^2}{8\pi^2 m} \left[\frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} \right].$$

Questa espressione si ottiene non dalla (105), ma dalla seguente (che algebricamente equivale a quella):

$$(105') \quad \mathcal{H} = \frac{1}{2m} \left[\frac{1}{r} P_1 r P_1 + \frac{1}{r^2} P_2^2 \right],$$

con la sostituzione di P_r con $\frac{h}{2\pi i} \frac{\partial}{\partial Q_r}$.

§ 24. — ALCUNE CONSEGUENZE DEL PRINCIPIO GENERALE DELLA MECCANICA QUANTISTICA.

Supponiamo che la misura di una osservabile G abbia fornito il risultato G_r : se indichiamo con ψ^- il vettore di stato immediatamente prima della osservazione, è ovvio che, immediatamente dopo questa, lo stato del sistema sarà in-

dividuato da un vettore ψ^+ in generale diverso da ψ^- , cioè la perturbazione prodotta sul sistema dall'osservazione avrà prodotto un brusco cambiamento del vettore di stato. Si riconosce subito che ψ^+ deve giacere lungo un asse principale dell'operatore \mathfrak{G} , asse corrispondente all'autovalore G_r : difatti l'immediata ripetizione dell'osservazione G darebbe con certezza il risultato G_r , e quindi le proiezioni di ψ^+ sugli assi principali di \mathfrak{G} corrispondenti agli autovalori diversi da G_r devono essere nulle, ossia ψ^+ deve essere ortogonale a tutti questi altri assi. Ora, supponiamo dapprima che all'autovalore G_r corrisponda un solo asse (cioè che l'operatore \mathfrak{G} sia completo e l'autovalore G_r semplice): allora resta individuato senz'altro il vettore ψ^+ (di cui basta conoscere la direzione, sapendosi già che è unitario), vale a dire, lo stato del sistema dopo l'osservazione è pienamente individuato dal risultato di questa, senza che occorra conoscere nulla circa lo stato anteriore all'osservazione (1). Osservazioni siffatte rinnovano *completamente*, per così dire, lo stato del sistema, e permettono di definirlo con un numero: esempio tipico, la misura dell'energia in un sistema non degenere, nel qual caso gli stati così individuati sono quelli che abbiamo chiamato « stati stazionari » (2) o « semplici » o « a energia definita »: possiamo ora caratterizzarli con la proprietà che il vettore di stato giace su uno degli assi principali dell'operatore hamiltoniano \mathfrak{H} . Gli stati che invece abbiamo chiamato « a energia non definita » e che abbiamo caratterizzato al § 29, p. II prendendo per ψ una combinazione lineare di autofunzioni ψ_n , sono rappresentati da un vettore ψ che non giace su nessuno degli assi principali di \mathfrak{H} : può esistere

(1) Si noti che, determinato il vettore ψ^+ nell'istante immediatamente successivo all'osservazione, l'ulteriore evoluzione di ψ col tempo resta definita dall'equazione temporale di Schrödinger, fino a che non interviene una nuova osservazione a perturbare il sistema.

(2) La ragione di questo nome si comprende ora immediatamente osservando che per uno di tali stati il vettore ψ ha la forma $\psi_n = u_n e^{-\frac{2\pi i E_n}{h} t}$, e quindi conserva direzione e modulo invariati nel tempo (benchè vari la sua « fase »).

però (anzi ne esistono infiniti ⁽¹⁾) un altro operatore hermitiano completo \mathcal{G} , di cui uno degli assi principali coincida con la direzione di ψ , ossia (ammesso che ad ogni operatore hermitiano corrisponda un'osservabile), esiste un'osservabile reale G (diversa dall'energia) tale che lo stato considerato (a energia non definita) si possa considerare uno stato « a G definita » e possa essere caratterizzato dal valore G_r di G .

Consideriamo ora brevemente i casi in cui all'autovalore G_r corrisponde più di un asse principale, o perchè G_r è un autovalore multiplo, o perchè \mathcal{G} è un operatore incompleto. In questi casi la ψ^+ non resta determinata dal risultato dell'osservazione, cioè dipende anche da alcune modalità dell'osservazione che non sono rispecchiate nel suo risultato numerico ⁽²⁾. Per precisare queste modalità, faremo l'ipotesi che l'osservazione venga eseguita in modo « da perturbare il sistema il meno possibile », con che vogliamo intendere quanto segue. Si sviluppi ψ^- in serie delle autofunzioni φ_r di \mathcal{G} : in questo sviluppo entrerà, per ogni autovalore G_r , una determinata delle infinite autofunzioni che ad esso appartengono: diremo che l'osservazione è fatta con la minima perturbazione possibile se ψ^+ si identifica con questa determinata autofunzione. Poichè supporremo sempre le osser-

(1) Nel caso unidimensionale, p. es., posto $\psi = \rho e^{i\theta}$ (con $\rho(x)$ e $\theta(x)$ reali) si verifica subito che questa condizione è soddisfatta dall'operatore

$$G(x, p) = \left(p - \frac{h}{2\pi} \theta' \right)^2 + \frac{h^2}{4\pi^2} \frac{\rho''}{\rho}$$

e che la ψ corrisponde all'autovalore 0. Naturalmente anche qualunque funzione di questa G soddisfa la condizione voluta. (v. E. FERMI, *N. Cim.*, VII, 10, p. 361, (1930).

(2) Per esempio, se si misura la x di una particella senza determinare contemporaneamente y e z (o p_y e p_z), il risultato della misura non basta a determinare lo stato successivo del sistema, perchè in generale nel misurare la x (p. es. col metodo della camera oscura, § 23, p. II) si comunica anche un impulso lungo le direzioni y e z . Ciò è rispecchiato nel fatto che l'operatore x è incompleto. Si può però disporre l'esperienza in modo da essere certi di non alterare p_y e p_z , ma solo p_x , e l'osservazione sarà allora fatta « con la minima perturbazione possibile ».

vazioni eseguite con questa condizione, potremo affermare che in ogni caso la ψ^- si può sviluppare in serie di tutte le possibili ψ^+ .

L'interpretazione geometrica di quanto sopra è la seguente. Se l'autovalore G_r è multiplo d'ordine p , ad esso corrispondono infiniti assi principali che formano una varietà lineare a p dimensioni (ortogonale a tutti gli altri assi principali) che chiameremo V : in essa deve in ogni caso giacere il vettore ψ^+ . Il caso che abbiamo chiamato della perturbazione minima è quello in cui ψ^+ si identifica con la proiezione di ψ^- sulla varietà V .

Il caso dell'operatore incompleto si può far rientrare nel precedente, considerando p infinito.

§ 25. - CRITERIO DI COMPATIBILITÀ DELLE OSSERVABILI.

Dimostriamo ora un teorema della massima importanza e cioè: *condizione necessaria e sufficiente perchè due osservabili siano compatibili è che i loro operatori siano permutabili.*

Difatti, siano A e B due osservabili (relative allo stesso istante) compatibili: se si misura A ottenendo p. es. A_r , e B col risultato B_s , il sistema resta in uno stato in cui un'ulteriore (immediata) osservazione di A darebbe con certezza il risultato A_r , e una di B il risultato B_s : quindi il vettore di stato ψ^+ dopo la duplice osservazione giace su un asse principale comune ai due operatori \mathfrak{A} e \mathfrak{B} . E poichè questo deve verificarsi qualunque sia lo stato antecedente, ciò significa che il vettore ψ^- , qualunque esso sia, deve potersi sviluppare in autofunzioni comuni ad \mathfrak{A} e \mathfrak{B} , ossia che queste formano un sistema completo. Perciò, per il teor. del § 11, \mathfrak{A} e \mathfrak{B} sono permutabili.

Viceversa, se \mathfrak{A} e \mathfrak{B} sono permutabili, essi ammettono un sistema completo di assi principali comuni (v. § 11). Supponiamo dapprima che nè l'uno nè l'altro siano degeneri: allora ogni asse principale di \mathfrak{A} è anche un asse principale di \mathfrak{B} , cui potremo attribuire lo stesso indice, e perciò, dopo eseguita l'osservazione di A col risultato A_r , il vettore di stato viene a trovarsi lungo l'asse principale corrispondente al valore A_r di \mathfrak{A} e B_r di \mathfrak{B} : perciò osservando la B si ottiene con certezza B_r , e il vettore di stato non è alterato da questa nuova osservazione. Lo stesso dicasi se si osserva prima

B e poi A . Perciò le due osservabili sono compatibili: anzi, poichè a ogni valore di una corrisponde un determinato valore dell'altra, esse sono funzioni l'una dell'altra.

Supponiamo ora invece che \mathfrak{B} sia un operatore degenere (o incompleto). Allora l'osservazione di A porta il vettore di stato nella direzione dell'asse A_r (quindi ψ diviene φ_r) il quale è anche uno degli infiniti assi di \mathfrak{B} appartenenti a un certo autovalore B_s , multiplo d'ordine p (eventualmente ∞), che costituiscono una varietà piana V ortogonale a tutti gli altri assi di \mathfrak{B} appartenenti ad altri autovalori: perciò la successiva osservazione di B deve necessariamente dare il risultato B_s , e farà al più ruotare il vettore di stato entro la V . Poichè però supponiamo che l'osservazione sia fatta con la minima perturbazione possibile (nel senso spiegato sopra), il nuovo vettore di stato ψ^+ deve essere la proiezione ortogonale di φ_r su V , la quale evidentemente si identifica con φ_r stesso: perciò ψ^+ sarà ad un tempo su un asse principale di \mathfrak{A} e di \mathfrak{B} . Se poi si osservasse prima B (col risultato B_s) e poi A (col risultato A_r), la prima osservazione porterebbe il vettore di stato entro la V , e poichè (v. § 11) entro a questa V vi sono p assi principali di \mathfrak{A} , mentre tutti gli altri assi principali di \mathfrak{A} sono perpendicolari a V , necessariamente l'ulteriore osservazione di A porterebbe il vettore di stato su uno dei p assi contenuti in V , e quindi comuni ad A ed a B . Si noti che a un valore di B possono corrispondere p valori di A .

In modo analogo si ragiona per il caso che entrambi gli operatori siano degeneri o incompleti, nel qual caso il legame tra i risultati delle due misure è ancora meno stretto, o manca del tutto.

Da questo teorema discende, in particolare, che la condizione di compatibilità di due osservazioni è simmetrica, come si è detto al § 16.

§ 26. - RELAZIONI DI PERMUTAZIONE.

Possiamo ritrovare facilmente, mediante il teorema ora dimostrato, il fatto ben noto che una coordinata cartesiana

$q_i (i = 1, 2, 3)$ e la corrispondente componente dell'impulso p_i , sono osservabili incompatibili: difatti i loro operatori sono rispettivamente, come si è visto,

$$q_i = q_i, \quad p_i = \frac{h}{2\pi i} \frac{\partial}{\partial q_i},$$

e non sono evidentemente permutabili, poichè per qualunque funzione f si ha

$$p_i q_i f - q_i p_i f = \frac{h}{2\pi i} \frac{\partial}{\partial q_i} (q_i f) - \frac{h}{2\pi i} q_i \frac{\partial f}{\partial q_i} = \frac{h}{2\pi i} f,$$

ossia

$$(106) \quad p_i q_i - q_i p_i = \frac{h}{2\pi i}.$$

Invece una q_i e una $p_j (j \neq i)$ sono evidentemente permutabili, cioè

$$(106') \quad p_j q_i - q_i p_j = 0,$$

come pure sono evidentemente permutabili due q o due p

$$(107) \quad q_i q_j - q_j q_i = 0$$

$$(108) \quad p_i p_j - p_j p_i = 0.$$

La (106) e la (106') hanno importanza fondamentale in meccanica quantistica, e prendono il nome di *relazioni di permutazione*.

Molto spesso nella meccanica quantistica si indicano con lo stesso simbolo una osservabile e il suo operatore (o la matrice corrispondente), anzichè distinguerle usando un diverso carattere come abbiamo fatto noi: ciò porta a scrivere le relazioni di permutazione (106) sotto la forma

$$p_i q_i - q_i p_i = \frac{h}{2\pi i}.$$

L'aspetto paradossale di queste equazioni scompare quando si tenga presente che esse si riferiscono non alle grandezze fisiche p_i e q_i ma ai loro operatori o alle corrispondenti matrici.

Vogliamo ora stabilire delle altre importanti relazioni di permutazione. Sia $Q(q_1, q_2, \dots, q_f)$ una funzione delle sole q ,

e consideriamola come un operatore $\mathfrak{Q} = Q(q_1, q_2, \dots, q_f)$: si ha per qualunque funzione (scalare) f

$$\begin{aligned} p_i \mathfrak{Q} f &= \frac{\hbar}{2\pi i} \frac{\partial}{\partial q_i} (Qf) = \frac{\hbar}{2\pi i} \frac{\partial Q}{\partial q_i} f + \frac{\hbar}{2\pi i} Q \frac{\partial f}{\partial q_i} \\ \mathfrak{Q} p_i f &= \frac{\hbar}{2\pi i} Q \frac{\partial f}{\partial q_i} \end{aligned}$$

e quindi si può scrivere la relazione di permutazione

$$(109) \quad p_i \mathfrak{Q} - \mathfrak{Q} p_i = \frac{\hbar}{2\pi i} \frac{\partial Q}{\partial q_i}.$$

Per una funzione (razionale e intera) $P(p_1, p_2, \dots, p_f)$ delle sole p vale una relazione analoga, e cioè, chiamando \mathfrak{P} l'operatore $P(p_1, p_2, \dots, p_f)$,

$$(110) \quad \mathfrak{P} q_i - q_i \mathfrak{P} = \frac{\hbar}{2\pi i} \frac{\partial \mathfrak{P}}{\partial p_i},$$

dove $\frac{\partial \mathfrak{P}}{\partial p_i}$ sta a designare l'operatore che si ottiene dalla funzione $\frac{\partial P}{\partial p_i}$ sostituendovi le p_i con gli operatori p_i . Questa formula si dimostra osservando che \mathfrak{P} sarà composto di termini della forma $a p_1^{n_1} p_2^{n_2} \dots p_f^{n_f}$, e che basta quindi dimostrare la (110) per un'espressione di questa forma: ora questi fattori sono tutti permutabili con q_i tranne il fattore $p_i^{n_i}$, per il quale vale la seguente regola di commutazione (che si trova applicando successivamente n_i volte la (106)):

$$p_i^{n_i} q_i - q_i p_i^{n_i} = n_i \frac{\hbar}{2\pi i} p_i^{n_i-1} = \frac{\hbar}{2\pi i} \frac{\partial p_i^{n_i}}{\partial p_i};$$

resta così dimostrata la (110).

Dalla (109) e dalla (110) si ricavano, per una G della forma $G = P(p) + Q(q)$, le relazioni di permutazione seguenti (che comprendono come casi particolari tutte quelle date precedentemente):

$$(111) \quad p_i \mathfrak{G} - \mathfrak{G} p_i = \frac{\hbar}{2\pi i} \frac{\partial \mathfrak{G}}{\partial q_i}$$

$$(112) \quad \mathfrak{G} q_i - q_i \mathfrak{G} = \frac{\hbar}{2\pi i} \frac{\partial \mathfrak{G}}{\partial p_i}.$$

§ 27. - VALORI MEDI.

La meccanica quantistica è sorta per inquadrare i cosiddetti *fenomeni elementari*, cioè quelli che riguardano una sola particella elementare o un piccolo numero di esse. È importante però far vedere che quando essa si applica a sistemi *macroscopici*, cioè composti di un grandissimo numero di particelle elementari, come sono i corpi ordinari, i suoi risultati vengono a coincidere con quelli della meccanica classica, la quale perciò non è in contrasto con quella quantistica, ma ne rappresenta solo un caso limite.

Chiameremo « osservazione macroscopica » un'osservazione che equivalga a misurare una stessa osservabile G su un grandissimo numero di sistemi uguali (p. es. di atomi) e a prendere il valor medio dei risultati ottenuti. La maggior parte delle misure ordinarie della fisica sono appunto di questo genere: p. es., determinare l'intensità di una corrente elettrica equivale a fare la media delle velocità dei singoli elettroni che la costituiscono (e moltiplicarla per un opportuno fattore); determinare le coordinate del baricentro di un corpo equivale a misurare le coordinate delle singole molecole e prenderne la media, e così via. Da ciò si vede l'interesse che presenta lo studio del « valore medio » di una osservabile in un insieme di sistemi (di eguale costituzione).

Se tutti i sistemi dell'insieme che consideriamo sono nello stesso « stato », si dice che l'insieme rappresenta un *caso puro*, altrimenti si dirà che è un *miscuglio*, e si potrà decomporlo in tanti insiemi parziali, ciascuno dei quali rappresenta un caso puro. Nel seguito ci riferiremo, salvo indicazione contraria, a un caso puro.

Se in ciascun sistema dell'insieme la misura dell'osservabile G può dare i risultati G_r con le rispettive probabilità P_r , il valore medio \bar{G} di tutti i risultati ottenuti misurando G in tutti i sistemi dell'insieme sarà

$$(113) \quad \bar{G} = \sum_r G_r P_r.$$

Ora, se lo stato di ciascun sistema è rappresentato dal vettore ψ (lo stesso per tutti i sistemi) e se si chiama φ_r

l'autofunzione dell'operatore \mathcal{G} corrispondente all'autovalore ⁽¹⁾ G_r , e

$$(114) \quad \psi_r = \psi \times \varphi_r$$

è la proiezione di ψ sul vettore φ_r , si ha, per il principio generale della meccanica quantistica, $P_r = \psi_r \psi_r^*$, e quindi

$$\bar{G} = \sum_r G_r \psi_r \psi_r^*.$$

Questa formula si può trasformare, poichè dalla (114) si ha che $\psi_r^* = \varphi_r \times \psi$: inoltre essendo $\mathcal{G}\varphi_r = G_r \varphi_r$, e ψ_r una costante, si ha

$$\bar{G} = \sum_r G_r \psi_r (\varphi_r \times \psi) = \sum_r \psi_r (\mathcal{G}\varphi_r) \times \psi = (\mathcal{G} \sum_r \psi_r \varphi_r) \times \psi,$$

ossia

$$(115) \quad \bar{G} = (\mathcal{G}\psi) \times \psi,$$

o, in forma esplicita,

$$(115') \quad \bar{G} = \int \psi^* \mathcal{G} \psi dS,$$

dove, al solito, $dS = dq_1 dq_2 \dots dq_f$, e l'integrale si intende esteso a tutto lo spazio delle q . Come si vede, a un determinato stato ψ dei sistemi corrisponde un determinato *valor medio* per qualunque osservabile, quindi *per le osservazioni macroscopiche non vi è principio di indeterminazione*.

Se l'insieme è un miscuglio, lo si decomporrà in insiemi parziali, in ciascuno dei quali lo stato dei sistemi è rappresentato da un vettore ψ^j , si calcolerà per ciascuno di essi il valore medio \bar{G}^j mediante la (115), e il valore medio richiesto sarà

$$(116) \quad \bar{G} = \sum c^j \bar{G}^j,$$

dove $c^j = N^j/N$ è il rapporto tra il numero N^j dei sistemi nello stato ψ^j e il numero totale N .

La (115') dà, in particolare, se G è una coordinata cartesiana q_i ,

$$(116) \quad \bar{q}_i = \int q_i \psi^* \psi dS$$

⁽¹⁾ Per includere anche i casi di degenerazione, bisogna ad ogni autovalore multiplo di ordine p far corrispondere nella (113) p termini separati.

e, se G è una componente cartesiana dell'impulso p_i ,

$$(117) \quad \bar{p}_i = \frac{h}{2\pi i} \int \psi^* \frac{\partial \psi}{\partial q_i} dS,$$

formule che coincidono con quelle del cap. I, p. II, che definiscono il centro d'un pacchetto d'onde e il suo vettore di propagazione medio.

§ 28. — DERIVATE RISPETTO AL TEMPO. LA MECCANICA CLASSICA COME CASO LIMITE.

Finora abbiamo considerato un'osservabile come definita solo per un determinato istante $t = t_0$. Ma è ovvia l'opportunità di introdurre qualcosa di analogo al concetto classico di una grandezza fisica funzione di t . Pensiamo perciò che lo stesso processo fisico che, messo in opera al tempo t_0 , definisce l'osservabile G_{t_0} , messo in opera ad un altro istante t definirà un'altra osservabile G_t : siamo così condotti a considerare una successione continua di osservabili del tutto analoga alla successione continua dei valori di una grandezza fisica funzione di t nella meccanica classica: diremo talvolta, per analogia, che si tratta di una « osservabile funzione di t ». Una generica delle osservabili G_t sarà rappresentata dallo stesso operatore $\mathcal{G}(q, p)$ che rappresenta la G_{t_0} , poichè il processo fisico che la definisce è lo stesso. Si potrà però anche generalizzare il concetto di variabilità nel tempo, supponendo che alcuni dei parametri che entrano a definire il processo di osservazione siano funzioni di t : allora l'operatore \mathcal{G} conterrà, oltre le q e p , anche t , ossia dipenderà esplicitamente dal tempo (si dirà allora che l'osservabile G dipende *esplicitamente* da t). In tal caso gli autovalori e le autofunzioni di G saranno in genere funzioni di t , mentre nel caso che G non contenga esplicitamente t gli autovalori sono costanti, e le autofunzioni indipendenti da t , cioè gli assi principali sono fissi nello spazio hilbertiano.

Vediamo ora come si può introdurre per le osservabili l'analogo della derivata rispetto al tempo di una grandezza fisica classica. È chiaro che non si può definire come deri-

vata \dot{G} dell'osservabile G l'espressione $\lim_{dt \rightarrow 0} \frac{G_{t+dt} - G_t}{dt}$, poichè tale scrittura non ha senso, essendochè G_{t+dt} e G_t non rappresentano due numeri ma due procedimenti di osservazione, relativi a istanti diversi, ciascuno dei quali può dare infiniti risultati numerici. Bisogna invece ricorrere alla considerazione dei valori medi, e definire la \dot{G} come un'osservabile tale, che il suo valor medio (definito come nel § prec.) sia uguale alla derivata del valor medio dell'osservabile G (1).

Deriviamo dunque la (115), supponendo, per maggior generalità, che l'operatore \mathcal{G} dipenda esplicitamente da t : avremo

$$\dot{G} = \frac{\partial}{\partial t}(\mathcal{G}\psi) \times \psi + (\mathcal{G}\psi) \times \frac{\partial \psi}{\partial t}$$

e, indicando con $\frac{\partial \mathcal{G}}{\partial t}$ l'operatore ottenuto dall'espressione di \mathcal{G} con l'operazione formale di derivazione rispetto a t (quindi $\frac{\partial \mathcal{G}}{\partial t} = 0$ se \mathcal{G} non contiene esplicitamente t):

$$\dot{G} = \left(\frac{\partial \mathcal{G}}{\partial t} \psi + \mathcal{G} \frac{\partial \psi}{\partial t} \right) \times \psi + \mathcal{G}\psi \times \frac{\partial \psi}{\partial t}.$$

Sostituendo la derivata di ψ con la sua espressione (87) si ha (ricordando la (5')):

$$\dot{G} = \left(\frac{\partial \mathcal{G}}{\partial t} \psi \right) \times \psi - \frac{2\pi i}{h} (\mathcal{G}\mathcal{H}\psi) \times \psi + \frac{2\pi i}{h} (\mathcal{G}\psi) \times (\mathcal{H}\psi).$$

(1) Ciò equivale a dire che da un insieme numerosissimo di esemplari del sistema come definito nel § precedente (anche un miscuglio), si estrae a caso un certo numero (assai grande) di sistemi, si osserva su essi la G_t , e se ne prende il valor medio; poi si estrae dall'insieme un altro gruppo assai grande di sistemi, si osserva su di essi la G_{t+dt} , e se ne prende il valor medio: si calcola poi la quantità numerica $\lim \frac{G_{t+dt} - G_t}{dt}$. Si postula poi che esista un'osservabile \dot{G} tale che il suo valor medio (preso su un terzo gruppo di esemplari tratto dallo stesso insieme) sia uguale a questa quantità, e ciò qualunque sia la composizione dell'insieme. Così, invece di misurare G_t e G_{t+dt} sullo stesso sistema (il che dà luogo alla difficoltà che la seconda misura è fatta sul sistema in uno stato diverso da quello della prima) si eseguono le due misure su sistemi diversi.

Ma, essendo \mathcal{H} hermitiano, si ha (v. § 9):

$$(\mathcal{G}\psi) \times (\mathcal{H}\psi) = (\mathcal{H}\mathcal{G}\psi) \times \psi$$

e quindi

$$\dot{G} = \left[\frac{\partial \mathcal{G}}{\partial t} + \frac{2\pi i}{h} (\mathcal{H}\mathcal{G} - \mathcal{G}\mathcal{H}) \right] \psi \times \psi.$$

Confrontando questa con la (115) si vede che \dot{G} si calcola formalmente come il valor medio di un'osservabile il cui operatore sia l'espressione contenuta nella parentesi quadra: rappresenteremo perciò simbolicamente la derivata dell'osservabile G con la formula

$$(118) \quad \dot{G} = \frac{\partial G}{\partial t} + \frac{2\pi i}{h} (G\mathcal{H} - \mathcal{H}G),$$

a cui corrisponde la relazione analoga tra gli operatori (indicando con $\dot{\mathcal{G}}$ l'operatore che corrisponde all'osservabile \dot{G}):

$$(118') \quad \dot{\mathcal{G}} = \frac{\partial \mathcal{G}}{\partial t} + \frac{2\pi i}{h} (\mathcal{H}\mathcal{G} - \mathcal{G}\mathcal{H}).$$

Se G non dipende esplicitamente da t , nel secondo membro mancherà il primo termine.

Alla derivata \dot{G} definita dalla (118), o meglio al differenziale $\dot{G}dt$, si può dare un'interpretazione espressiva con la considerazione seguente. Si consideri l'osservabile $G + Gdt$, che chiameremo g , e si supponga di misurarla (al tempo t), trovando il valore g' : dimostreremo che dopo tale osservazione il sistema rimane in uno stato tale che, se al tempo $t + dt$ si misura l'osservabile \dot{G} , si trova pure il valore g' : si può dunque dire che la misura (al tempo t) dell'osservabile $G + Gdt$ serve a determinare il valore che avrà la G al tempo $t + dt$; si noti però che tale misura è incompatibile con quella della G al tempo t . Per dimostrare l'asserto chiamiamo ψ la funzione rappresentante lo stato in cui rimane il sistema dopo che la misura di g ha dato il risultato g' : si avrà

$$(119) \quad \mathcal{G}\psi(t) = g'\psi(t),$$

ossia, scrivendo semplicemente ψ per $\psi(t)$:

$$(119') \quad \mathcal{G}\psi + \left[\frac{\partial \mathcal{G}}{\partial t} + \frac{2\pi i}{h} (\mathcal{H}\mathcal{G} - \mathcal{G}\mathcal{H}) \right] \psi dt = g'\psi.$$

Si tratta di verificare che, lasciando evolvere questa ψ per un tempo dt , si ottiene una $\psi(t + dt)$ che è un'autofunzione dell'operatore \mathfrak{G}_{t+dt} corrispondente all'autovalore g' , cioè che

$$\mathfrak{G}_{t+dt}\psi(t + dt) = g'\psi(t + dt),$$

ossia

$$\left(\mathfrak{G} + \frac{\partial \mathfrak{G}}{\partial t} dt\right) \left(\psi + \frac{\partial \psi}{\partial t} dt\right) = g' \left(\psi + \frac{\partial \psi}{\partial t} dt\right).$$

Sviluppando, e sostituendo per la derivata di ψ il valore dato dall'equazione di Schrödinger (87), si ha

$$\mathfrak{G}\psi + \left(\frac{\partial \mathfrak{G}}{\partial t} - \frac{2\pi i}{h} \mathfrak{G}\mathfrak{H}\right)\psi dt = g'\psi - \frac{2\pi i}{h} \mathfrak{H}g'\psi dt$$

e, per la (119),

$$\mathfrak{G}\psi + \left(\frac{\partial \mathfrak{G}}{\partial t} - \frac{2\pi i}{h} \mathfrak{G}\mathfrak{H}\right)\psi dt = g'\psi - \frac{2\pi i}{h} \mathfrak{H}g'\psi dt.$$

Nell'ultimo termine si può sostituire g con \mathfrak{G} , (a meno di termini in dt^2): con ciò l'equazione viene a coincidere con la (119'), ed è quindi verificata.

Applichiamo questo risultato per ritrovare, generalizzando e precisandolo, il principio che un pacchetto d'onde si muove come un punto nella meccanica ordinaria, principio che (limitato a pacchetti d'onde sufficientemente piccoli da potersi considerare puntiformi) ha costituito nella p. II il nostro punto di partenza per stabilire l'equazione di SCHRÖDINGER.

Identificando G con una coordinata q_i si ha dalla (118')

$$\dot{q}_i = \frac{2\pi i}{h} (\mathfrak{H}q_i - q_i\mathfrak{H})$$

e identificandola invece con il momento p_i :

$$\dot{p}_i = \frac{2\pi i}{h} (\mathfrak{H}p_i - p_i\mathfrak{H}).$$

Essendo l'hamiltoniana \mathfrak{H} della forma $P(p) + Q(q)$, si possono applicare le (111), (112) e si trova così

$$\dot{q}_i = \frac{\partial \mathfrak{H}}{\partial p_i} \quad \dot{p}_i = -\frac{\partial \mathfrak{H}}{\partial q_i}.$$

Secondo quanto si è detto a proposito della (118) queste relazioni tra operatori traducono le seguenti relazioni tra i valori medi delle corrispondenti osservabili

$$(120) \quad \dot{\bar{q}}_i = \frac{\partial \bar{\mathfrak{H}}}{\partial p_i} \quad \dot{\bar{p}}_i = -\frac{\partial \bar{\mathfrak{H}}}{\partial q_i}.$$

Valgono dunque, *in media*, le equazioni di HAMILTON. Per esempio, per un punto in coordinate cartesiane, si ha

$$\mathfrak{K} = \frac{1}{2m} \sum_1^3 p_i^2 + U(q)$$

e quindi le equazioni precedenti danno

$$\dot{\bar{q}}_i = \frac{1}{m} \bar{p}_i, \quad \dot{\bar{p}}_i = -\frac{\partial \bar{U}}{\partial q_i} = \bar{F}_i$$

(indicando con F_i le componenti della forza). Se tra queste si elimina p_i si ha

$$(121) \quad m\ddot{\bar{q}}_i = \bar{F}_i,$$

cioè: « il centro del pacchetto d'onde si muove come un punto materiale obbediente alla meccanica classica e soggetto ad una forza che si calcola prendendo il valor medio della forza su tutto il pacchetto d'onde » (Teorema di EHRENFEST). Si noti che questo teorema si applica a un pacchetto comunque esteso, mentre nelle considerazioni che ci hanno condotto all'equazione di SCHRÖDINGER nella p. II ci riferivamo al caso limite di un pacchetto praticamente puntiforme. Inoltre, si può vedere facilmente, applicando la (116), che il teorema vale anche per i valori medi di un miscuglio.

§ 29. - INTEGRALI PRIMI.

Nella meccanica classica si chiama *integrale primo* di un problema una espressione $G(q, p)$ tale che si riduca a una costante se le q e le p variano col tempo in modo da soddisfare le equazioni della dinamica.

Analogamente, in meccanica quantistica definiremo come *integrale primo* un'osservabile G tale che la sua derivata \dot{G} definita da (118) sia identicamente zero, cioè tale che

$$(122) \quad \frac{\partial \mathfrak{G}}{\partial t} + \frac{2\pi i}{h} (\mathfrak{H}\mathfrak{G} - \mathfrak{G}\mathfrak{H}) = 0.$$

Si vede subito che l'energia \mathcal{H} è un integrale primo se (e solo se) $\frac{\partial \mathcal{H}}{\partial t} = 0$, cioè se l'hamiltoniana non contiene esplicitamente il tempo: si dirà in tal caso che il sistema è *conservativo*.

Si può dimostrare che se G è un integrale primo, i suoi autovalori sono costanti (anche se G contiene esplicitamente t) e inoltre, sebbene gli assi principali di \mathcal{G} non siano fissi, la loro rotazione nello spazio hilbertiano è legata alla rotazione del vettore di stato ψ (qualunque sia questo) in modo tale, che le proiezioni di ψ su questi assi conservano modulo costante: ciò significa che le probabilità dei singoli autovalori non variano col tempo (1).

In generale ci si limita a considerare integrali primi che non contengono esplicitamente t . Allora la (122) diviene

$$\mathfrak{G} = \mathfrak{G}\mathfrak{G}$$

ed esprime che: *condizione necessaria e sufficiente perchè una osservabile G (non contenente t) sia un integrale primo è che il suo operatore sia permutabile con l'hamiltoniano, ossia che l'osservazione di G sia compatibile con quella simultanea dell'energia.*

§ 30. — MOMENTI ANGOLARI E LORO OPERATORI.

Come applicazione dei §§ precedenti, consideriamo le tre osservabili M_x, M_y, M_z , momenti dell'impulso (o *momenti angolari*) di una particella rispetto agli assi x, y, z rispettivamente, e ricerchiamo anzitutto gli operatori $\mathfrak{M}_x, \mathfrak{M}_y, \mathfrak{M}_z$ ad esse corrispondenti.

Considerando p. es. M_z , osserviamo che la sua espressione in meccanica classica è

$$(123) \quad M_z = xp_y - yp_x$$

e quindi, secondo la regola del § 22, l'operatore ad essa corrispondente è

$$(124) \quad \mathfrak{M}_z = xp_y - yp_x = \frac{h}{2\pi i} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right)$$

(1) Per la dimostrazione v. p. es. bibl. n. 4, p. 229.

e similmente per M_x e M_y . Notiamo innanzi tutto che queste tre osservabili sono incompatibili due a due: difatti si ha p. es.

$$\begin{aligned} \mathfrak{M}_x \mathfrak{M}_y &= (yp_z - zp_y)(zp_x - xp_z) = ypzp_xz - yxp_z^2 - z^2p_y p_x + zxp_y p_z \\ \mathfrak{M}_y \mathfrak{M}_x &= xp_z p_y z - xyp_z^2 - z^2p_x p_y + zyp_x p_z \end{aligned}$$

e quindi, per la (106) e la (124),

$$\mathfrak{M}_x \mathfrak{M}_y - \mathfrak{M}_y \mathfrak{M}_x = (yp_x - xp_y)(p_z z - zp_z) = -\frac{h}{2\pi i} \mathfrak{M}_z.$$

Valgono dunque le seguenti formule di permutazione per i momenti

$$(125) \quad \left\{ \begin{aligned} \mathfrak{M}_y \mathfrak{M}_z - \mathfrak{M}_z \mathfrak{M}_y &= -\frac{h}{2\pi i} \mathfrak{M}_x \\ \mathfrak{M}_z \mathfrak{M}_x - \mathfrak{M}_x \mathfrak{M}_z &= -\frac{h}{2\pi i} \mathfrak{M}_y \\ \mathfrak{M}_x \mathfrak{M}_y - \mathfrak{M}_y \mathfrak{M}_x &= -\frac{h}{2\pi i} \mathfrak{M}_z \end{aligned} \right.$$

che, introducendo le notazioni vettoriali anche per gli operatori, si riassumono nella formula

$$(125') \quad \mathfrak{M} \wedge \mathfrak{M} = -\frac{h}{2\pi i} \mathfrak{M}.$$

Ricerchiamo ora le autofunzioni e gli autovalori di questi operatori. Prendiamo p. es. \mathfrak{M}_z : osserviamo che, se si introducono coordinate polari r, θ, φ , con l'asse z per polo, si ha

$$\frac{\partial}{\partial \varphi} = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}$$

e quindi la (124) diviene

$$(126) \quad \mathfrak{M}_z = \frac{h}{2\pi i} \frac{\partial}{\partial \varphi}.$$

Perciò l'equazione delle autofunzioni $\psi(r, \theta, \varphi)$ diviene, detto M'_z un autovalore generico

$$\frac{h}{2\pi i} \frac{\partial \psi}{\partial \varphi} = M'_z \psi,$$

da cui

$$\psi = f(r, \theta) e^{\frac{2\pi i}{h} M'_z \varphi},$$

con f funzione arbitraria di r e θ (la sua presenza è dovuta al fatto che \mathfrak{M}_z è un operatore incompleto). Questa ψ è evidentemente periodica in φ a periodo h/M'_z : ma affinché ψ

abbia un sol valore in ogni punto dello spazio, essa deve essere periodica in φ a periodo 2π : quindi dovrà aversi $m\hbar/M'_z = 2\pi$ con m intero, ossia

$$M'_z = m \frac{\hbar}{2\pi}.$$

Sono questi gli autovalori cercati, e gli stessi si troverebbero per M_x ed M_y .

Una misura del momento angolare rispetto ad un asse dà dunque come risultato sempre un multiplo (positivo, nullo o negativo) di $\hbar/2\pi$. È questo un risultato che estende e precisa la «quantizzazione spaziale» della teoria di SOMMERFELD (v. § 56 p. II).

Si vede poi subito che, se le forze hanno momento nullo rispetto all'asse z , la M_z è (come in meccanica classica) un integrale primo. Difatti in tal caso il potenziale U , espresso in coordinate polari, deve risultare indipendente da φ , e quindi l'hamiltoniano § non contiene φ ed è perciò permutabile con l'operatore (126).

Consideriamo ora l'osservabile M , modulo del momento angolare della particella rispetto all'origine. Classicamente si ha $M^2 = M_x^2 + M_y^2 + M_z^2$: perciò assumeremo come operatore corrispondente a M^2

$$(127) \quad \mathfrak{M}^2 = \mathfrak{M}_x^2 + \mathfrak{M}_y^2 + \mathfrak{M}_z^2,$$

dove \mathfrak{M}_x ecc. sono dati dalla (124) e dalle analoghe. Sostituendovi queste espressioni, e tenendo conto delle (106), si trova con facili calcoli

$$\mathfrak{M}^2 = r^2 \Sigma p_x^2 - (\Sigma x p_x)^2 - \frac{\hbar}{2\pi i} \Sigma x p_x$$

(dove Σ indica una somma fatta cambiando x successivamente in y e z). Ora, tenendo presente il significato degli operatori p_x, p_y, p_z , si vede che

$$\Sigma p_x^2 = -\frac{\hbar^2}{4\pi^2} \Delta$$

$$\Sigma x p_x = \frac{\hbar}{2\pi i} r \frac{\partial}{\partial r},$$

e quindi

$$(128) \quad \mathfrak{M}^2 = -\frac{\hbar^2}{4\pi^2} \left[r^2 \Delta - r \frac{\partial}{\partial r} \left(r \frac{\partial}{\partial r} \right) - r \frac{\partial}{\partial r} \right].$$

Si rammenti ora che l'operatore di LAPLACE Δ in coordinate polari è espresso da

$$(129) \quad \Delta = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \mathcal{A},$$

dove con \mathcal{A} si è indicato l'operatore, indipendente da r ,

$$(130) \quad \mathcal{A} = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \psi^2}.$$

Sostituendo nella (128), si trova

$$(131) \quad \mathfrak{M}^2 = -\frac{\hbar^2}{4\pi^2} \mathcal{A}.$$

I valori dell'osservabile M^2 sono dunque dati da $-\frac{\hbar^2}{4\pi^2} \mathcal{A}'$, dove \mathcal{A}' è un autovalore dell'equazione

$$(132) \quad \mathcal{A}\psi = \mathcal{A}'\psi.$$

Questa equazione non è altro che la (223') del § 46, p. II, cioè l'equazione differenziale delle funzioni sferiche (\mathcal{A}' corrisponde a $-C$), e i suoi autovalori sono, come si è visto, $\mathcal{A}' = -l(l+1)$ con $l = 0, 1, 2, \dots$. Perciò i valori che può assumere l'osservabile M , momento dell'impulso, sono dati da

$$(133) \quad M = \sqrt{l(l+1)} \frac{\hbar}{2\pi}.$$

Questo risultato fu già enunciato nel § 46, p. II.

Si osservi che l'operatore \mathfrak{M}^2 , e quindi anche \mathfrak{M} , è permutabile con ciascuno degli operatori $\mathfrak{M}_x, \mathfrak{M}_y, \mathfrak{M}_z$. Difatti si ha, per le (125),

$$\begin{aligned} \mathfrak{M}^2 \mathfrak{M}_x - \mathfrak{M}_x \mathfrak{M}^2 &= \mathfrak{M}_y^2 \mathfrak{M}_x - \mathfrak{M}_x \mathfrak{M}_y^2 + \mathfrak{M}_z^2 \mathfrak{M}_x - \mathfrak{M}_x \mathfrak{M}_z^2 = \\ &= \mathfrak{M}_y \left(\mathfrak{M}_x \mathfrak{M}_y + \frac{\hbar}{2\pi i} \mathfrak{M}_z \right) - \left(\mathfrak{M}_y \mathfrak{M}_x - \frac{\hbar}{2\pi i} \mathfrak{M}_z \right) \mathfrak{M}_y + \\ &+ \mathfrak{M}_z \left(\mathfrak{M}_x \mathfrak{M}_z - \frac{\hbar}{2\pi i} \mathfrak{M}_y \right) - \left(\mathfrak{M}_z \mathfrak{M}_x + \frac{\hbar}{2\pi i} \mathfrak{M}_y \right) \mathfrak{M}_z = 0 \end{aligned}$$

e similmente si dimostra la permutabilità di \mathfrak{M}^2 con \mathfrak{M}_y e con \mathfrak{M}_z . Dunque la misura del momento angolare totale è compatibile con la misura della sua proiezione su una direzione qualunque, mentre non sono compatibili le misure di due componenti del detto momento.

Si riconosce poi che, se la forza è centrale, M^2 è un inte-

grale primo (come in meccanica classica). Difatti l'operatore \mathfrak{H} per una particella è (v. § 19)

$$(134) \quad \mathfrak{H} = -\frac{h^2}{8\pi^2m} \Delta + U = -\frac{h^2}{8\pi^2m} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \Delta \right] + U$$

e poichè Δ è un operatore che non coinvolge r , esso è sempre permutabile coi primi due termini di questa espressione: se poi la forza è centrale, U è funzione solo di r e quindi Δ è permutabile anche con l'ultimo termine: in tal caso dunque Δ è permutabile con \mathfrak{H} , e quindi, per la (131), lo è anche \mathfrak{M}^2 , il che significa che M^2 è un integrale primo.

§ 31. - FORZE MAGNETICHE.

In tutte le considerazioni fatte fin qui si sono trattate solo particelle soggette a forze derivanti da un potenziale U : il caso, fisicamente assai importante, di particelle elettrizzate muovendosi in un campo magnetico non rientra nelle considerazioni precedenti e richiede perciò una apposita estensione di esse. Tale estensione verrà fatta, come le precedenti, prendendo a guida le considerazioni del § 19.

Se una particella di carica e si muove con velocità \mathbf{v} in un campo elettrico \mathbf{E} e in un campo magnetico \mathbf{H} , su di essa agisce la forza

$$(135) \quad \mathbf{F} = e \left(\mathbf{E} + \frac{1}{c} \mathbf{v} \wedge \mathbf{H} \right),$$

di cui la parte di origine elettrica deriva da un potenziale U (posto $e\mathbf{E} = -\text{grad } U$) mentre la parte magnetica non ammette potenziale. Si osservi che le ordinarie equazioni della dinamica della particella si possono anche in questo caso mettere sotto forma hamiltoniana introducendo ⁽¹⁾ il *potenziale vettore* \mathbf{A} e prendendo come hamiltoniana la funzione

$$(138) \quad \mathfrak{H} = \sum_r \frac{1}{2m} \left(p_r - \frac{e}{c} A_r \right)^2 + U(q_1, q_2, q_3).$$

⁽¹⁾ Ricordiamo che il campo elettrico \mathbf{E} e quello magnetico \mathbf{H} si deducono dal potenziale scalare V e da quello vettoriale \mathbf{A} con le note formule

$$(136) \quad \mathbf{E} = -\text{grad } V - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{H} = \text{rot } \mathbf{A}.$$

Si ricordi inoltre che il potenziale U da noi usato è uguale a eV .

Difatti le equazioni di Hamilton che se ne ricavano sono

$$\begin{aligned} \dot{q}_s &= \frac{\partial \mathfrak{H}}{\partial p_s} = \frac{1}{m} \left(p_s - \frac{e}{c} A_s \right) \\ \dot{p}_s &= -\frac{\partial \mathfrak{H}}{\partial q_s} = \frac{e}{c} \sum_r \frac{1}{m} \left(p_r - \frac{e}{c} A_r \right) \frac{\partial A_r}{\partial q_s} - \frac{\partial U}{\partial q_s}. \end{aligned}$$

La prima dà ⁽¹⁾

$$(139) \quad p_s = m\dot{q}_s + \frac{e}{c} A_s$$

e la seconda, tenendo conto di queste, delle (136) e della (135), dà l'equazione della dinamica

$$m\ddot{q}_s = eF_s.$$

Volendo dunque conservare l'analogia rilevata nel § 19, l'equazione della meccanica ondulatoria, quando esiste il campo magnetico, si otterrà trasformando la hamiltoniana (138) in operatore mediante la solita sostituzione (S) di pagina 338, cioè ponendo

$$(140) \quad \mathfrak{H} = \sum_r \frac{1}{2m} \left(\frac{h}{2\pi i} \frac{\partial}{\partial q_r} - \frac{e}{c} A_r \right)^2 + U(q),$$

e scrivendo l'equazione per ψ_n o per ψ nella solita forma (81) o (82). Si può anche dire che l'operatore \mathfrak{H} corrispondente alla presenza di un campo magnetico si ottiene dall'hamiltoniana senza campo sostituendo p_r con l'operatore

$$(141) \quad \mathfrak{P}_r = p_r - \frac{e}{c} A_r = \frac{h}{2\pi i} \frac{\partial}{\partial q_r} - \frac{e}{c} A_r.$$

L'equazione di SCHRÖDINGER per gli stati stazionari è dunque per una particella nel campo magnetico:

$$(142) \quad \frac{1}{2m} \sum_r \left(\frac{h}{2\pi i} \frac{\partial}{\partial q_r} - \frac{e}{c} A_r \right)^2 \psi_n = (E_n - U) \psi_n$$

e quella temporale, per uno stato qualunque:

$$(142') \quad \frac{1}{2m} \sum_r \left(\frac{h}{2\pi i} \frac{\partial}{\partial q_r} - \frac{e}{c} A_r \right)^2 \psi + U\psi = -\frac{h}{2\pi i} \frac{\partial \psi}{\partial t}.$$

⁽¹⁾ Si ponga mente al fatto espresso da questa formula, che, in presenza del campo magnetico, i *momenti* p_1, p_2, p_3 non sono più le componenti della velocità moltiplicate per m . Una particella *ferma* in un campo magnetico ha momenti diversi da zero.

La giustificazione di questa estensione dell'equazione di SCHRÖDINGER è data dal teorema seguente che ci limitiamo ad enunciare ⁽¹⁾: se con le ψ_n , soddisfacenti la (142) si costruisce un pacchetto d'onde abbastanza piccolo, esso si muove come si muoverebbe un corpuscolo di carica e , sotto l'azione delle forze derivanti dal potenziale U e del campo magnetico derivante dal potenziale vettore A .

Quanto sopra si estende immediatamente a un sistema di N particelle distinte: l'operatore hamiltoniano è in tal caso (usando le stesse notazioni della (84), e indicando con A_j la componente del potenziale vettore corrispondente alla coordinata q_j):

$$\mathcal{H} = \sum_j^{3N} \frac{1}{2m_j} \left(\frac{\hbar}{2\pi i} \frac{\partial}{\partial q_j} - \frac{e_j}{c} A_j \right)^2 + U(q).$$

§ 32. — IL DETERMINISMO E LA MECCANICA QUANTISTICA ⁽²⁾.

In meccanica razionale, come si sa, è univocamente determinata la soluzione del seguente problema: date, in un certo istante $t = 0$, le posizioni e le velocità di tutti i punti di un sistema (soggetti a forze dipendenti in modo noto dalle posizioni e dalle velocità), calcolare il valore di qualunque coordinata, o componente di velocità, del sistema, o di qualsiasi loro funzione, in un altro istante qualunque t_1 . È questa una proprietà analitica delle equazioni fondamentali della meccanica, che dà espressione precisa, nel campo meccanico, a quella convinzione filosofica generale che si designa col nome di *determinismo*.

Mettiamoci ora dal punto di vista della meccanica quantistica. Come abbiamo detto più volte, non si può attribuire nessun significato fisico alla locuzione «insieme delle posizioni e delle velocità dei punti di un sistema in un dato istante», e quindi l'enunciato citato sopra, valido in mec-

⁽¹⁾ Per la dimostrazione. v. p. es. bibl. n. 14. p. 109.

⁽²⁾ Per una discussione più approfondita di questo argomento vedasi p. es. il n. 21 della bibl. e inoltre: A. EDDINGTON, *Sur le problème du déterminisme (Actualités Scient. et Ind.)*, n. 112; ed. Hermann, Parigi 1934.

canica classica, perde qualsiasi significato in meccanica quantistica. Vale in suo luogo la proprietà seguente.

Supponiamo che al tempo 0 si sia eseguita un'osservazione massima: i suoi risultati rappresentano una descrizione completa del sistema all'istante considerato (ossia, ne definiscono completamente lo « stato »). Le leggi della meccanica quantistica permettono allora come si è visto ⁽¹⁾ di calcolare, per qualunque istante t_1 (eventualmente lo stesso), i possibili risultati della misura di una qualsiasi osservabile G , e le rispettive probabilità. In generale, questi possibili risultati costituiscono una infinità (discreta o continua) il che si suol esprimere dicendo che vi è una certa « indeterminazione » nel valore della grandezza osservata (indeterminazione che deriva, è bene ripeterlo, non da imperfezione delle misure o da incompletezza dei dati iniziali, ma dalla natura stessa dei fenomeni elementari, come diffusione dei fotoni, urto di particelle, ecc., a cui si applicano le leggi della meccanica quantistica).

⁽¹⁾ Non sarà inutile riassumere lo schema di questo calcolo, che risulta dai §§ precedenti. Siano $g_1(q, p)$, $g_2(q, p)$... le osservabili misurate al tempo 0 (costituenti un'osservazione massima) e g'_1 , g'_2 ... i valori trovati: la ψ che caratterizza lo stato del sistema è determinata, per $t = 0$, dalle equazioni

$$(143) \quad g_1 \psi_0 = g'_1 \psi_0, \quad g_2 \psi_0 = g'_2 \psi_0, \dots$$

La ψ si evolve poi col tempo obbedendo l'equazione differenziale di Schrödinger

$$(144) \quad \mathfrak{H}\psi = -\frac{\hbar}{2\pi i} \frac{\partial \psi}{\partial t}$$

che, insieme con il valore iniziale ψ_0 dato dalle (143), definisce la ψ a un tempo t qualunque, e in particolare la $\psi(t_1)$. Si ponga poi l'equazione

$$(145) \quad \mathfrak{G}\varphi = G'\varphi:$$

i suoi autovalori G'_n danno i possibili risultati della misura di G , e le rispettive probabilità sono date da $|\psi(t_1) \times \varphi_n|^2$. Geometricamente, diremo che le osservazioni al tempo 0 definiscono la posizione iniziale del vettore di stato ψ , la (144) definisce il modo con cui esso si evolve nel tempo e quindi permette di calcolare il vettore di stato al tempo t_1 : proiettandolo sugli assi principali dell'operatore \mathfrak{G} , dati dalla (145), si hanno le ampiezze di probabilità richieste.

Può avvenire però, per qualche particolare osservabile ⁽¹⁾, che vi sia un solo valore con probabilità 1, e tutti gli altri abbiano probabilità 0; cioè che l'osservabile *abbia* al tempo t_1 un valore determinato: in tal caso evidentemente la meccanica quantistica permette di calcolare tale valore, partendo dai dati iniziali (risultati dell'osservazione massima), senza alcuna indeterminazione.

Per meglio chiarire la cosa, si consideri l'esempio dell'oscillatore lineare (v. § 39, p. II), e si supponga di averne misurato, al tempo $t = 0$, l'energia, trovando il risultato E_n . È questa una osservazione massima, che determina completamente lo stato del sistema (infatti, la sua ψ sarà l'autofunzione ψ_n dell'equazione di SCHRÖDINGER). Se poi, in un istante t_1 qualsiasi (anche, eventualmente, lo stesso) si esegue una osservazione di posizione, cioè della osservabile x , si può trovare un valore qualsiasi da $-\infty$ a $+\infty$; tutto ciò che si può dire a priori sul risultato di questa misura è che la sua probabilità è distribuita secondo le densità $|\psi_n|^2$ (cioè secondo le curve della fig. 29). Se però, invece di misurare l'osservabile x , misuriamo al tempo t_1 l'osservabile E (energia), siamo sicuri di ritrovare il valore E_n : per questa osservabile dunque non vi è alcuna indeterminazione. Lo stesso varrebbe naturalmente per qualsiasi altra osservabile compatibile con E .

Si noti che l'osservazione massima da farsi al tempo 0 può scegliersi con ampia arbitrarietà, e questi diversi modi di definire lo stato del sistema, come si è detto al § 18, non sono equivalenti: sono infatti diverse nei vari casi le osservabili che hanno valore definito, e ciò vale non solo per le osservabili relative al tempo 0, ma anche per quelle relative a un istante qualunque t_1 . P. es., nel caso dell'oscillatore si potrebbe misurare, invece di E , la x , o l'impulso p , o una loro qualsiasi funzione $g(x, p)$. In tal caso però non sarebbe più l'energia ad avere al tempo t_1 un valore determinabile

⁽¹⁾ Precisamente, ciò avviene per tutte quelle osservabili il cui operatore ha un asse principale nella direzione del vettore di stato ψ all'istante t_1 .

a priori, ma (eventualmente) un'altra osservabile G . Si può anzi dimostrare ⁽¹⁾ che, data un'osservabile $G(x, p)$ relativa a un dato istante t_1 , è sempre possibile trovare un'osservazione massima da farsi al tempo 0 tale che il suo risultato permetta di calcolare — senza indeterminazione — il valore di G al tempo t_1 . Però, se si volesse poter calcolare senza indeterminazione il valore di un'altra osservabile G' (che può anche essere la stessa G a un istante diverso) si dovrebbe eseguire al tempo 0 un'osservazione massima diversa, e incompatibile con l'altra se G e G' sono incompatibili tra loro. Si può dire, in conclusione, che lo stato del sistema al tempo t_1 si può determinare non soltanto con un gruppo (completo) di osservazioni al tempo t_1 , ma anche con un gruppo (completo) di osservazioni al tempo 0: vale a dire, lo stato del sistema al tempo t_1 è determinato univocamente dallo stato al tempo 0. Questa proposizione ha però, nella meccanica quantistica, un significato diverso da quello della meccanica classica, perchè, come si è visto, il concetto di « stato » in meccanica quantistica non implica l'insieme delle posizioni e delle velocità dei punti, come in meccanica classica.

Si noti che quanto sopra si applica solo ai fenomeni « elementari », in cui cioè interviene un limitato numero di atomi, elettroni, fotoni, ecc., mentre per i fenomeni « macroscopici », in cui interviene un grandissimo numero di particelle elementari (cioè, per la maggior parte dei fenomeni comunemente osservati), valgono, come è naturale, gli ordinari rapporti di determinismo. Ciò non è affatto in contraddizione con la meccanica quantistica, anzi ne è una conseguenza necessaria. Difatti, una osservazione « macroscopica » (v. § 27) equivale alla misura di un valore medio su un grandissimo numero di sistemi elementari: ma in meccanica quantistica il valore medio di qualsiasi osservabile è univocamente determinato, come si è detto al § 27, quando è noto lo stato del sistema, e quindi per i valori medi vale il determinismo nel senso classico.

⁽¹⁾ V. E. FERMI, *Rend. Acc. Linc.*, XI, serie 6^a, 1^o sem. 1930, p. 980, o anche *N. Cim.*, VII (1930), p. 361.

CAPITOLO III.

METODO DELLE MATRICI

§ 33. — CENNI GENERALI.

Nel cap. I si è visto che, fissato un sistema completo qualsiasi di funzioni ortogonali y_n (cioè, in linguaggio geometrico, un sistema ortogonale di assi nello spazio hilbertiano), ogni vettore di tale spazio può essere rappresentato dalle sue componenti f_n rispetto ai detti assi, e similmente ogni operatore lineare \mathfrak{A} può essere rappresentato da una certa matrice $\{\mathfrak{A}\}$. Vogliamo ora applicare questa rappresentazione al « vettore di stato » ψ (di cui ci siamo ampiamente occupati nel cap. precedente) e agli operatori lineari che operano su di esso e che, come abbiamo visto, corrispondono alle diverse osservabili: dovremo dunque ora considerare lo stato del sistema definito dall'insieme delle (infinite) componenti del vettore ψ , secondo il prefissato sistema di assi y_n , e ad ogni osservabile far corrispondere, anziché un operatore, una matrice. Pertanto le relazioni algebriche tra osservabili si tradurranno ora in altrettante relazioni tra matrici. Perciò questo metodo di trattare i problemi della meccanica quantistica prende il nome di *metodo delle matrici*, ed è, naturalmente, del tutto equivalente, dal punto di vista teorico, al metodo degli operatori e a quello della meccanica ondulatoria che ci è servito fin qui: solo ragioni pratiche possono consigliare nei vari casi la preferenza di uno o dell'altro metodo.

Per meglio comprendere la situazione rispettiva dei vari metodi, si può paragonare il metodo degli operatori a quello che nella meccanica razionale è la trattazione puramente vettoriale senza uso di sistemi di riferimento, mentre il

metodo della meccanica ondulatoria e quello delle matrici si possono paragonare alla trattazione della meccanica con l'uso di un particolare sistema di coordinate. E precisamente, nel caso della meccanica ondulatoria lo spazio hilbertiano è riferito a quel particolare sistema di assi che abbiamo chiamato « continui » (v. § 2) (individuato ciascuno da un gruppo di valori delle « coordinate » del sistema), mentre nel metodo delle matrici lo si riferisce ad un sistema generico di assi, per lo più discreti.

Il metodo delle matrici, come si è detto nella parte I, è stato ideato da HEISENBERG ed è stata la prima forma data alla meccanica quantistica: tuttavia il punto di vista dal quale era presentato allora era notevolmente diverso da quello che abbiamo ora accennato, al quale ci atterremo nel seguito di questo capitolo.

Esponiamo ora brevemente l'idea fondamentale di questo metodo.

Fissiamo anzitutto un sistema di assi di riferimento nello spazio hilbertiano, scegliendo ⁽¹⁾ una determinata osservabile K e assumendo come assi di riferimento gli assi principali del suo operatore \mathfrak{K} ; cioè le direzioni individuate dalle autofunzioni y_n della equazione

$$\mathfrak{K}y_n = K_n y_n:$$

diremo allora che le matrici che usiamo sono riferite allo « schema K ». Particolare interesse ha, come vedremo, lo schema \mathcal{H} (dove \mathcal{H} è l'energia): in questo schema evidentemente le direzioni degli assi di riferimento sono date dalle autofunzioni ψ_n della equazione di SCHRÖDINGER.

Fissato lo « schema », ad ogni osservabile A corrisponde una matrice hermitiana

$$(146) \quad \{\mathfrak{A}\} = \begin{vmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & A_{23} & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix}.$$

⁽¹⁾ Se il sistema è a più gradi di libertà, si dovrà intendere che K rappresenta una osservabile massima, ossia un gruppo completo di osservabili (v. § 18) e le y_n sono le autofunzioni comuni a tutti i loro operatori.

Se vogliamo legare questa rappresentazione al metodo degli operatori, ricorderemo (v. § 5) che gli elementi di questa matrice si ottengono dall'operatore \mathfrak{A} (corrispondente secondo la regola del § 22 all'osservabile A) mediante la formula

$$(147) \quad A_{mn} = (\mathfrak{A}y_n) \times y_m = \int y_m^* \mathfrak{A}y_n dS$$

e che tali elementi sono i coefficienti della sostituzione lineare che fa passare dalle componenti di un qualsiasi vettore f alle componenti di $\mathfrak{A}f$ ossia, brevemente, della sostituzione che rappresenta l'effetto dell'operatore. Nel metodo delle matrici però si considerano le A_{mn} come elementi caratteristici dell'osservabile, senza ricollegarli alla espressione (147).

Ricordiamo dal § 12 che, in particolare, la matrice che nello schema K rappresenta l'osservabile K , cioè la stessa che serve a definire lo schema, è una matrice diagonale

$$\{\mathfrak{K}\} = \begin{vmatrix} K_1 & 0 & 0 & \dots \\ 0 & K_2 & 0 & \dots \\ 0 & 0 & K_3 & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix}$$

i cui elementi non sono altro che gli autovalori dell'operatore \mathfrak{K} (v. § 10), e quindi rappresentano i possibili risultati di una misura di K . Gli elementi A_{mn} della matrice $\{\mathfrak{A}\}$ non hanno invece, in generale, alcun significato fisico semplice, tranne quelli della diagonale principale: difatti A_{nn} rappresenta (v. form. (115), § 27) il valor medio dell'osservabile A nello stato individuato dal vettore y_n , cioè nello stato definito dal valore K_n di K . Gli elementi non diagonali $A_{mn}(m \neq n)$, sebbene non interpretabili in modo semplice, sono tuttavia legati in modo essenziale alla statistica dei risultati sperimentali: difatti, il valor medio di una qualsiasi potenza A^k di A (nello stato definito da K_n) è evidentemente dato dall' n -esimo termine diagonale della matrice $\{\mathfrak{A}^k\}$ e nel calcolo di esso intervengono anche elementi non diagonali di $\{\mathfrak{A}\}$.

Particolare interesse hanno poi gli elementi delle tre matrici $\{\mathfrak{X}\}$, $\{\mathfrak{Y}\}$, $\{\mathfrak{S}\}$, rappresentanti le componenti del mo-

mento elettrico del sistema nello schema \mathcal{H} , ossia le espressioni

$$X_{mn} = \int \psi_m^* X \psi_n dS \text{ etc.}$$

difatti, come si è accennato al § 32, p. II, la radiazione emessa (o assorbita) nel passaggio dallo stato m allo stato n corrisponde qualitativamente a quella che sarebbe emessa da un oscillatore il cui momento elettrico avesse, sui tre assi, le componenti X_{mn} , Y_{mn} , Z_{mn} .

§ 34. — RELAZIONI DI PERMUTAZIONE.

Le relazioni algebriche tra osservabili si tradurranno in relazioni della stessa forma tra le matrici che le rappresentano, intendendosi naturalmente le operazioni di somma e prodotto tra matrici definite con le regole del § 6. In particolare, tra le matrici $\{q_k\}$ e $\{p_k\}$ rappresentanti le coordinate e i momenti varranno, in qualunque schema, le relazioni di permutazione

$$(148) \quad \{p_k\}\{q_k\} - \{q_k\}\{p_k\} = \frac{\hbar}{2\pi i} \{1\}$$

$$(148') \quad \{p_k\}\{q_l\} - \{q_k\}\{p_l\} = 0 \quad \text{per } (k \neq l)$$

e, per k ed l qualunque,

$$(149) \quad \{q_k\}\{q_l\} - \{q_l\}\{q_k\} = 0$$

$$(150) \quad \{p_k\}\{p_l\} - \{p_l\}\{p_k\} = 0.$$

Se poi G è una funzione delle q e delle p della forma

$$G = P(p) + Q(q)$$

(dove P è simbolo di funzione razionale intera e Q di funzione qualunque), ad essa corrisponderà una matrice $\{G\}$ per la quale varranno (in qualunque schema) le relazioni di permutazione seguenti, che discendono immediatamente dalle (111) e (112) del § 26:

$$(151) \quad \{p_i\}\{G\} - \{G\}\{p_i\} = \frac{\hbar}{2\pi i} \left\{ \frac{\partial G}{\partial q_i} \right\}$$

$$(152) \quad \{G\}\{q_i\} - \{q_i\}\{G\} = \frac{\hbar}{2\pi i} \left\{ \frac{\partial G}{\partial p_i} \right\},$$

in cui le precedenti rientrano come casi particolari.

§ 35. — RICERCA DEGLI AUTOVALORI COL METODO DELLE MATRICI.

Ciò premesso, vediamo come si imposta col metodo delle matrici la ricerca degli autovalori di un'osservabile G (che, in particolare, può essere l'energia).

Si deve partire, come nel § 22, dall'espressione analitica dell'osservabile G in funzione delle q e delle p , espressione che tiene luogo di definizione della G e che, come si è detto, si costruisce di solito per analogia con la meccanica classica tenendo presente, ove sia necessario, la necessità di « simmetrizzare » i prodotti $p_k q_k$ affinché la matrice corrispondente a G risulti hermitiana. Costruita questa espressione $G(q, p)$, essa si potrà interpretare come una relazione tra le matrici corrispondenti alle osservabili G, q, p :

$$(153) \quad \{G\} = G(\{q\}, \{p\})$$

e ciò in qualunque schema di riferimento. In particolare, nello schema G , varrà questa stessa relazione, con in più la condizione che la matrice $\{G\}$ sia diagonale: se dunque ci riferiamo a questo schema, si tratta di determinare gli elementi delle matrici $\{q\}$ e $\{p\}$ in modo che queste soddisfino le relazioni di permutazione (148) a (150), e che inoltre la matrice $\{G\}$ calcolata mediante la (153) risulti diagonale. Risolto questo problema (per la cui soluzione non si possono indicare metodi generali) gli elementi della matrice diagonale $\{G\}$ danno gli autovalori richiesti. Mostriamo al § seguente un esempio di questo procedimento.

§ 36. — APPLICAZIONE ALL'OSCILLATORE.

Prendiamo il caso di un oscillatore armonico, di massa m e forza di richiamo $-Kx$, trattato con la meccanica ondulatoria al § 39, p. II, e proponiamoci di cercare, col metodo delle matrici, i valori che può assumere la sua energia.

L'espressione dell'energia in funzione di x e di $p = m\dot{x}$ (hamiltoniana) è, analogamente alla meccanica classica,

$$\mathcal{H} = \frac{1}{2m} p^2 + \frac{1}{2} Kx^2.$$

Si tratta dunque di determinare gli elementi delle matrici $\{x\}$ e $\{p\}$ (riferite allo schema \mathcal{H}) in modo che valga la relazione di permutazione

$$(154) \quad \{p\}\{x\} - \{x\}\{p\} = \frac{\hbar}{2\pi i} \{1\}$$

ed inoltre che la matrice

$$(155) \quad \{\mathcal{H}\} = \frac{1}{2m} \{p\}^2 + \frac{K}{2} \{x\}^2$$

risulti diagonale. Traducendo queste uguaglianze tra matrici in uguaglianze tra gli elementi corrispondenti, e indicando con E_n gli elementi diagonali (H_{nn}) della matrice $\{\mathcal{H}\}$, cioè gli autovalori cercati, si ha ⁽¹⁾

$$(156) \quad \sum_r (p_{jr}x_{rk} - x_{jr}p_{rk}) = \frac{\hbar}{2\pi i} \delta_{jk}$$

$$(157) \quad E_k \delta_{jk} = \frac{1}{2m} \sum_r p_{jr}p_{rk} + \frac{K}{2} \sum_r x_{jr}x_{rk}.$$

Applicando poi alla matrice $\{\mathcal{H}\}$ le regole di permutazione (151) e (152) si trova

$$\{p\}\{\mathcal{H}\} - \{\mathcal{H}\}\{p\} = \frac{\hbar}{2\pi i} K\{x\}$$

$$\{x\}\{\mathcal{H}\} - \{\mathcal{H}\}\{x\} = -\frac{\hbar}{2\pi i} \frac{1}{m} \{p\}$$

e queste si traducono nelle seguenti relazioni tra gli elementi delle matrici

$$\sum_r (p_{jr}H_{rk} - H_{jr}p_{rk}) = \frac{\hbar}{2\pi i} Kx_{jk}$$

$$\sum_r (x_{jr}H_{rk} - H_{jr}x_{rk}) = -\frac{\hbar}{2\pi i} \frac{1}{m} p_{jk},$$

ossia

$$(158) \quad p_{jk}(E_k - E_j) = \frac{\hbar}{2\pi i} Kx_{jk}$$

$$(158') \quad x_{jk}(E_k - E_j) = -\frac{\hbar}{2\pi i} \frac{1}{m} p_{jk},$$

(1) In questo problema, numeriamo le righe e le colonne delle matrici a partire da 0 anzichè da 1, per conformarci alla convenzione adottata nella trattazione ondulatoria dello stesso problema in cui abbiamo numerato gli autovalori E_0, E_1 , etc.

dalle quali eliminando p_{jk} si ricava

$$x_{jk} \left[1 - \frac{4\pi^2 m}{\hbar^2 K} (E_k - E_j)^2 \right] = 0,$$

ovvero, introducendo la frequenza classica dell'oscillatore,

$$v_0 = \frac{1}{2\pi} \sqrt{\frac{K}{m}},$$

$$(159) \quad x_{jk} \left[1 - \left(\frac{E_k - E_j}{\hbar v_0} \right)^2 \right] = 0,$$

da cui si vede che gli elementi x_{jk} sono tutti nulli, tranne al più quelli i cui indici j, k sono tali che

$$(160) \quad E_k - E_j = \pm \hbar v_0.$$

Analoga osservazione si può fare per gli elementi p_{jk} . Adunque nelle matrici $\{x\}$ e $\{p\}$ vi sono in ogni linea e in ogni colonna al più due elementi diversi da zero.

Ora, la relazione di permutazione (156) dà in particolare, per un elemento diagonale ($j = k$),

$$\sum_r (p_{kr}x_{rk} - x_{kr}p_{rk}) = \frac{\hbar}{2\pi i}$$

e, ricavando le p dalla (158), e notando che $x_{kr} = x_{rk}^*$,

$$(161) \quad 2K \sum_r \frac{|x_{kr}|^2}{E_r - E_k} = 1 \quad (k = 0, 1, 2, \dots).$$

Di qui si vede intanto che in ogni linea della matrice $\{x\}$ vi è almeno un elemento non nullo (altrimenti, se la k -esima linea fosse tutta di zeri, la sommatoria della (161) sarebbe nulla e la relazione non potrebbe sussistere per quel valore di k): quindi, se un certo numero E_k è un autovalore, uno almeno dei numeri $E_k \pm \hbar v_0$ è pure un autovalore. Questa condizione è soddisfatta se gli autovalori formano una progressione aritmetica di ragione $\hbar v_0$, cioè se sono dati dalla formula

$$(162) \quad E_k = \varepsilon + k\hbar v_0,$$

dove ε è una costante che determineremo in seguito. Numerando in tal modo gli autovalori, la condizione (160) è soddisfatta solo per $j = k \pm 1$: quindi le matrici $\{x\}$ e $\{p\}$ hanno diversi da zero solo gli elementi del tipo $x_{k, k \pm 1}, p_{k, k \pm 1}$

(che formano due file oblique, parallele e contigue, da una parte e dall'altra, alla diagonale principale). La (161) dà allora, per $k = 0$

$$(163) \quad 2K \frac{|x_{01}|^2}{\hbar v_0} = 1$$

e per $k = 1, 2, \dots$

$$(164) \quad 2K \left[\frac{|x_{k,k+1}|^2}{\hbar v_0} - \frac{|x_{k,k-1}|^2}{\hbar v_0} \right] = 1,$$

la quale si può anche scrivere

$$|x_{k,k+1}|^2 - |x_{k-1,k}|^2 = \frac{\hbar v_0}{2K}$$

e ci dice che le quantità $|x_{k,k+1}|^2$ formano una successione aritmetica di ragione $\frac{\hbar v_0}{2K}$: il primo elemento di questa successione è dato dalla (163) ed è

$$(163') \quad |x_{01}|^2 = \frac{\hbar v_0}{2K};$$

l'elemento generico è dunque

$$|x_{k,k+1}|^2 = (k+1) \frac{\hbar v_0}{2K} = (k+1) \frac{\hbar}{8\pi^2 m v_0}.$$

Per gli elementi $x_{k,k+1}$ della matrice $\{\mathfrak{x}\}$ si può dunque prendere (1)

$$(165) \quad x_{k,k+1} = \sqrt{k+1} \sqrt{\frac{\hbar}{8\pi^2 m v_0}}.$$

Gli elementi della forma $x_{k,k-1}$ si deducono da questi osservando che, dovendo la matrice essere hermitiana, $x_{k,k-1} = x_{k-1,k}^*$, e che quest'ultima quantità si ricava dalla (165) semplicemente cambiando k in $k-1$: quindi

$$(165') \quad x_{k,k-1} = \sqrt{k} \sqrt{\frac{\hbar}{8\pi^2 m v_0}}.$$

(1) Si potrebbe naturalmente aggiungere a queste espressioni un fattore della forma $e^{i\delta_k}$, con δ_k arbitraria, ma gli autovalori risulterebbero, come si riconosce subito, gli stessi. Ciò corrisponderebbe a moltiplicare i versori, che individuano gli assi nello spazio hilbertiano, per fattori di modulo 1, il che non modifica nulla di essenziale.

Da queste formule, mediante la (158), o la (158'), si ricavano le espressioni degli elementi non nulli della matrice $\{\mathfrak{p}\}$, che risultano

$$(166) \quad p_{k,k+1} = -i \sqrt{k+1} \sqrt{\frac{m v_0 \hbar}{2}}$$

$$(166') \quad p_{k,k-1} = i \sqrt{k} \sqrt{\frac{m v_0 \hbar}{2}}.$$

Così sono completamente determinate le matrici $\{\mathfrak{x}\}$ e $\{\mathfrak{p}\}$. Per trovare i livelli energetici, non resta che da determinare la costante ε della (162): ciò si fa immediatamente particolarizzando la (157) per $j = k = 0$: si ha

$$E_0 = \frac{1}{2m} |p_{01}|^2 + \frac{K}{2} |x_{01}|^2$$

e prendendo $|x_{01}|^2$ e $|p_{01}|^2$ dalla (163') e dalla (166), si trova

$$E_0 = \frac{\hbar v_0}{2}.$$

Quindi nella (162) deve prendersi $\varepsilon = \frac{\hbar v_0}{2}$, e l'espressione degli autovalori dell'energia diviene

$$(167) \quad E_k = \left(k + \frac{1}{2}\right) \hbar v_0$$

che coincide con quella trovata con la meccanica ondulatoria al § 39, p. II.

Si può poi facilmente verificare che le espressioni trovate per gli elementi delle matrici $\{\mathfrak{x}\}$ e $\{\mathfrak{p}\}$ soddisfano le relazioni (156) e (157) (che abbiamo utilizzato solo particolarizzando per $j = k$) anche per $j \neq k$.

Gli elementi delle matrici $\{\mathfrak{x}\}$ e $\{\mathfrak{p}\}$ che abbiamo calcolato (e che intervengono anche in problemi di teoria della radiazione), si potrebbero anche calcolare mediante la loro espressione ondulatoria (v. (147)):

$$(168) \quad x_{jk} = \int \psi_j^* x \psi_k dx$$

$$(168') \quad p_{jk} = \frac{\hbar}{2\pi i} \int x_j^* \frac{\partial \psi_k}{\partial x} dx,$$

inserendovi per le autofunzioni ψ le espressioni trovate nel § 39, p. II: tuttavia questo procedimento porterebbe a calcoli assai più lunghi di quelli svolti in questo paragrafo. Nel caso dell'oscillatore, dunque, il metodo delle matrici presenta alcuni vantaggi sul metodo di SCHRÖDINGER.

CAPITOLO IV.

TEORIA DELLE PERTURBAZIONI

§ 37. — GENERALITÀ.

È ben noto che in meccanica celeste il problema del moto dei pianeti sotto l'azione delle attrazioni reciproche e di quella del sole sarebbe praticamente insolubile se tutte queste forze fossero dello stesso ordine di grandezza: fortunatamente invece l'attrazione del Sole è molto preponderante sulle attrazioni tra pianeti, e questo permette di calcolare il moto di un pianeta con un metodo di approssimazioni successive, che dicesi « metodo delle perturbazioni » e che permette di raggiungere un'esattezza più che sufficiente per la pratica. Esso consiste essenzialmente nel considerare dapprima ogni pianeta come soggetto alla sola attrazione solare, nella quale ipotesi, come si sa, il problema è rigorosamente solubile, e poi nel calcolare, con approssimazioni successive, come viene modificato il movimento per effetto delle forze di attrazione reciproca prima trascurate (forze perturbatrici). Un procedimento analogo si segue in fisica atomica per risolvere quei problemi nei quali le particelle si possono considerare soggette a certe forze preponderanti, che, se fossero sole, consentirebbero la completa risoluzione del problema (problema « imperturbato ») ed inoltre ad altre forze « perturbatrici » più lievi. Così, p. es., nello studio di un atomo sottoposto a un campo magnetico (effetto Zeeman) si può considerare l'azione del campo sugli elettroni dell'atomo come una forza perturbatrice, e, supposto di aver saputo risolvere il problema (cioè determinare auto-

funzioni e autovalori) per l'atomo in assenza di campo, si può, col metodo che esporremo, determinare la perturbazione provocata dal campo magnetico sulle autofunzioni e sugli autovalori.

L'importanza della teoria delle perturbazioni in fisica atomica è grandissima, poichè molti problemi, di grande interesse dal lato sperimentale, condurrebbero, sia col metodo della meccanica ondulatoria che con quello delle matrici, a difficoltà matematiche grandissime od anche praticamente insuperabili se non potessero trattarsi col metodo delle perturbazioni. Tale metodo si può adattare tanto alla trattazione ondulatoria che a quella con le matrici, come mostreremo nei §§ seguenti.

§ 38. — PERTURBAZIONE DEGLI STATI STAZIONARI (CASO NON DEGENERE).

Consideriamo dapprima il sistema imperturbato, e diciamo $\mathcal{H}^0(q, p)$ l'hamiltoniana che lo caratterizza (distingueremo in genere con uno zero in alto le quantità che si riferiscono al problema imperturbato): allora l'equazione di SCHRÖDINGER per lo stato stazionario i -esimo del sistema imperturbato sarà

$$(169) \quad \mathfrak{S}^0 \psi_i^0 = E_i^0 \psi_i^0,$$

dove \mathfrak{S}^0 designa come al solito, l'operatore ottenuto da $\mathcal{H}^0(q, p)$ sostituendovi ogni p_r con $\frac{h}{2\pi i} \frac{\partial}{\partial q_r}$, secondo le norme del § 22. Supponiamo che il problema imperturbato si sappia completamente risolvere cioè che si conoscano tutti gli autovalori E_i^0 (che supponiamo discreti ⁽¹⁾) e le rispettive autofunzioni ψ_i^0 . Fissiamo l'attenzione su uno determinato degli stati e sia l' n -esimo, e proponiamoci di determinare

(1) Se gli autovalori sono in parte discreti ed in parte continui, e se l'autovalore su cui si fissa l'attenzione appartiene ai primi, le formule di questo § e del successivo continuano a valere, purchè si sostituiscano certe sommatorie con integrali, in modo abbastanza ovvio. Se invece l'autovalore considerato appartiene allo spettro continuo, si richiede un procedimento alquanto diverso (v. p. es. bibl. n. 14, p. 157).

l'effetto su di esso delle forze perturbatrici, cioè la modificazione prodotta da queste su E_n^0 e su ψ_n^0 . Supporremo in questo § che l'autovalore E_n^0 non sia multiplo: il caso contrario (*degenerazione*) richiede una trattazione a parte, che sarà fatta nel § seguente; gli altri autovalori invece possono anche essere multipli, ma in tal caso ciascuno di essi, se è multiplo d'ordine p , va contato come p autovalori coincidenti (denotati con indici distinti).

Le forze perturbatrici saranno rappresentate da un termine $\mathfrak{L}(q, p)$ aggiunto all'hamiltoniana, che diverrà $\mathcal{H} = \mathcal{H}^0 + \mathfrak{L}$: quindi l'operatore corrispondente diventerà $\mathfrak{S}^0 + \mathfrak{L}$, dove \mathfrak{L} è ottenuto da $\mathfrak{L}(q, p)$ con la solita sostituzione (previa eventuale simmetrizzazione dei prodotti $p_r q_r$), e risulta hermitiano. Se p. es. le forze perturbatrici derivano da un potenziale $u(q)$, sarà (ricordando che $\mathcal{H}^0 = T + U$)

$$\mathfrak{L} = \mathfrak{L} = u(q):$$

in casi più generali, p. es. se si tratta di forze magnetiche, \mathfrak{L} dipenderà anche dalle p oltrechè dalle q , e quindi \mathfrak{L} conterrà anche dei simboli di derivazione.

Supporremo però in ogni caso che \mathfrak{L} non contenga esplicitamente t , il che si esprime dicendo che la perturbazione è « indipendente dal tempo »: il caso contrario verrà trattato al § 41.

L'equazione di SCHRÖDINGER relativa al sistema perturbato sarà, per lo stato stazionario n -esimo (chiamando ⁽¹⁾ $E_n = E_n^0 + \varepsilon$ e $\psi_n = \psi_n^0 + \chi$ l'autovalore e l'autofunzione perturbati)

$$(170) \quad (\mathfrak{S}^0 + \mathfrak{L})\psi_n = E_n \psi_n.$$

La funzione χ , che rappresenta l'effetto della perturbazione su ψ_n , si potrà poi sviluppare in serie mediante le auto-

(1) Si dovrebbe scrivere ε_n , χ_n , e, più oltre, a_{rn} perchè per ogni valore di n vi è una ε , una χ ed un sistema di coefficienti a : per semplicità di scrittura tralasciamo l'indice n , che si può ritenere fisso in tutto il ragionamento. Così faremo spesso anche in seguito.

funzioni imperturbate (che formano, come si sa, un sistema ortogonale completo) e quindi si scriverà

$$(171) \quad \psi_n = \psi_n^0 + \sum_1^{\infty} a_s \psi_s^0,$$

dove le a_s sono coefficienti costanti che nel loro insieme definiscono l'effetto della perturbazione su ψ_n^0 . Sostituendo questa espressione nella (170) e tenendo conto della (169) (per $i = n$) si ha

$$(-\varepsilon + \mathcal{Q}) \psi_n^0 + \sum_s a_s (E_s^0 - E_n^0 - \varepsilon + \mathcal{Q}) \psi_s^0 = 0.$$

Moltiplichiamo (a sinistra) i due membri per ψ_n^{0*} e integriamo rispetto a tutto lo spazio delle q , tenendo presenti le condizioni di ortogonalità e normalizzazione delle ψ e indicando, al solito, con L_{rs} gli elementi della matrice che rappresenta l'operatore \mathcal{Q} nello schema \mathcal{H}^0 , (matrice di perturbazione), cioè ponendo

$$(172) \quad L_{rs} = \int \psi_r^{0*} \mathcal{Q} \psi_s^0 dS.$$

Avremo allora

$$(173) \quad \varepsilon(1 + a_n) = L_{nn} + \sum_s a_s L_{ns}.$$

Similmente, moltiplicando per ψ_σ^{0*} (indichiamo con σ un indice che assume tutti i valori interi e positivi tranne n) e integrando, avremo

$$(174) \quad a_\sigma (E_n^0 - E_\sigma^0 + \varepsilon) = L_{cn} + \sum_s a_s L_{cs} \quad (\sigma \neq n).$$

Tutte le formule dedotte fin qui valgono rigorosamente, cioè qualunque sia l'ordine di grandezza di \mathcal{Q} . Ora sfruttiamo la circostanza che l'effetto della perturbazione è piccolo, cioè consideriamo le L_{rs} , ε , e le a_s come quantità piccole del I ordine ⁽¹⁾. Se allora nella (173) trascuriamo le quantità del secondo ordine, essa ci dà per ε un valore di prima approssimazione che indicheremo con ε' e che è

$$(175) \quad \varepsilon' = L_{nn}.$$

⁽¹⁾ Più precisamente supponiamo tutte le L_{ns} piccole del primo ordine rispetto alle differenze $E_n^0 - E_\sigma^0$: da ciò consegue che anche ε e le a sono piccole del primo ordine (rispetto a E_n^0 e ad I rispettivamente).

Dunque: la perturbazione dell'autovalore n -esimo, in prima approssimazione, è data dall' n -esimo termine diagonale della matrice di perturbazione, o anche, se si vuole (v. § 27) dal valor medio di \mathcal{Q} calcolato per l' n -esimo stato stazionario del sistema imperturbato. Questo risultato è del tutto analogo a un noto teorema di meccanica classica secondo cui la correzione da apportarsi all'energia di un sistema per effetto di una forza perturbatrice è uguale alla media del potenziale di questa forza, calcolata sul movimento imperturbato.

Si noti che per calcolare (in prima approssimazione) gli autovalori perturbati non è stato necessario conoscere le autofunzioni perturbate ψ_n : molte volte queste non interessano e si può fare a meno di occuparsene. Volendole conoscere in prima approssimazione bisogna ricavare a_σ dalla (174) che, trascurando i termini del secondo ordine, e indicando con a'_σ la prima approssimazione di a_σ , dà

$$(176) \quad a'_\sigma = \frac{L_{cn}}{E_n^0 - E_\sigma^0} \quad (\sigma \neq n).$$

Per avere tutti i termini dello sviluppo (171), manca ancora la conoscenza di a_n : questa si determina imponendo a ψ_n la condizione di normalizzazione, che (trascurando quantità del secondo ordine e indicando, al solito, con l'apice la prima approssimazione) dà

$$\int \left[\psi_n^{0*} \sum_1^{\infty} a'_s \psi_s^0 + \psi_n^0 \sum_1^{\infty} a_s'^* \psi_s^{0*} \right] dS = 0,$$

cioè

$$a_n' + a_n'^* = 0,$$

il che significa che a_n' deve essere un numero puramente immaginario, del resto arbitrario (purchè piccolo del primo ordine): prendendolo ⁽¹⁾ uguale a zero, la (171) si scrive

$$(177) \quad \psi_n' = \psi_n^0 + \sum_{\sigma \neq n} \frac{L_{cn}}{E_n^0 - E_\sigma^0} \psi_\sigma^0.$$

⁽¹⁾ L'arbitrarietà di a_n rispecchia l'arbitrarietà della « fase » di ψ_n^0 , e non ha conseguenze pratiche. Difatti, posto $a_n = i\delta_n$ (con δ_n reale e piccolo del I ordine), si ha in prima approssimazione $1 + a_n = e^{i\delta_n}$, e, poichè, nella (171) la ψ_n^0 figura appunto moltiplicata per $1 + a_n$, basta sostituire ψ_n^0 con $\psi_n^0 e^{-i\delta_n}$ per ridursi al caso di $a_n = 0$.

Sotto l'aspetto intuitivo si può rilevare che l'effetto della perturbazione sullo stato stazionario n -esimo è quello di « mescolare » alla autofunzione ψ_n^0 ciascuna delle altre ψ_σ^0 , in misura tanto maggiore quanto più i rispettivi livelli energetici sono vicini a quello considerato, e quanto più è rilevante l'elemento $L_{\sigma n}$ della matrice di perturbazione.

Passiamo ora al calcolo degli autovalori in seconda approssimazione. Per ricavare dalla (173) ε in seconda approssimazione è necessario inserirvi i valori delle a , ma poichè queste figurano moltiplicate per quantità del primo ordine sarà sufficiente introdurre i valori di prima approssimazione, cioè (176) e $a'_n = 0$: si ottiene allora (indicando con ε'' i termini del secondo ordine di ε)

$$(178) \quad \varepsilon' + \varepsilon'' = L_{nn} + \sum_{\sigma} \frac{|L_{\sigma n}|^2}{E_n^0 - E_\sigma^0} \quad (\sigma \neq n).$$

Proseguendo in modo analogo si calcolerebbero le autofunzioni di seconda approssimazione, e mediante queste la terza approssimazione ($\varepsilon' + \varepsilon'' + \varepsilon'''$) di ε , e così via. Tralasciamo di scrivere queste formule che raramente trovano applicazione, essendo per lo più sufficienti la prima o la seconda approssimazione per gli autovalori, e la prima per le autofunzioni.

§ 39. - PERTURBAZIONE DEGLI STATI STAZIONARI (CASO DEGENERARE E QUASI-DEGENERARE).

Per la validità delle approssimazioni svolte al § prec. è necessario, come si è detto, che sia

$$(179) \quad |L_{\sigma n}| \ll |E_n^0 - E_\sigma^0|;$$

ciò è reso manifesto dalla formula (176), che mostra che, se per qualche σ questa condizione non è soddisfatta, la corrispondente a'_σ non è più una quantità piccola del primo ordine. La condizione (179) cessa di essere soddisfatta se lo spostamento del livello n -esimo provocato dalla perturbazione è dello stesso ordine di grandezza della distanza tra il livello stesso e quelli vicini: perciò, se il livello in questione è assai vicino a qualche altro livello, anche a una perturba-

zione leggera non è più applicabile il ragionamento precedente. Come limite di questo caso, si può considerare quello in cui il livello E_n è multiplo (degenerazione) poichè a tale condizione si può pensare di giungere gradatamente facendo avvicinare a E_n , fino a coincidere, uno o più degli altri autovalori E_σ . Perciò ricercheremo in questo § come si modifica il metodo delle perturbazioni nel caso in cui vi è un gruppo di autovalori (che chiameremo E_1, E_2, \dots, E_p) molto vicini tra loro, in modo da formare un gruppo assai serrato detto « multipletto » (*quasi-degenerazione*), o addirittura coincidenti in modo da formare un autovalore multiplo (*degenerazione*): quest'ultimo caso sarà considerato come un caso limite del primo. Si osservi che queste circostanze, che sotto l'aspetto matematico sembrano eccezionali, si verificano invece nella maggior parte dei problemi di interesse fisico: in particolare, i livelli energetici sono di regola multipli in tutti i problemi a simmetria sferica o cilindrica (come si è visto, p. es., nel caso dell'idrogeno). Quando poi si tenga conto delle correzioni di spin e relatività che verranno introdotte in seguito, i livelli energetici coincidenti si separano (in parte) lievemente, dando origine alla « struttura fina », ossia a multipletti: si passa cioè dalla degenerazione completa al caso quasi-degenerare (ciascuna componente del multipletto può però a sua volta presentare degenerazione completa).

Conveniamo di indicare in tutto questo § con una lettera latina i, j, k, l, \dots ogni indice (assumendo i valori 1, 2, ... p) che serve a distinguere tra le varie componenti del multipletto, mentre le lettere greche ρ, σ, \dots serviranno per gli indici (assumendo tutti i valori interi e positivi tranne 1, 2, ... p) che si riferiscono agli stati non facenti parte del multipletto. Indicheremo poi (conformemente alla notazione adottata sopra) con E_i^0 e E_σ^0 gli autovalori imperturbati ⁽¹⁾, con E_i e E_σ quelli perturbati, e le autofunzioni ad essi corrispon-

(1) I livelli E_σ^0 possono anche essere raggruppati, tutti o in parte, in multipletti o in livelli multipli: ciò non muta nulla alle formule seguenti, purchè ad ogni livello di molteplicità n si facciano corrispondere n valori dell'indice σ , come se si trattasse di livelli distinti.

denti con ψ_i^0 , ψ_σ^0 e rispettivamente ψ_i , ψ_σ . È opportuno poi introdurre il livello energetico *medio* del multipletto, cioè la quantità

$$(180) \quad E_c = \frac{\sum_i E_i^0}{p}$$

e porre

$$(181) \quad E_i^0 = E_0 + \varepsilon_i^0, \quad E_i = E_0 + \varepsilon_i.$$

con che le ε risultano quantità piccole rispetto alle E (del primo ordine). Si noti che, nel caso della degenerazione completa, le ε_i^0 sono tutte nulle.

Ciò premesso, l'equazione di SCHRÖDINGER per gli stati imperturbati si scriverà (indicando come prima con \mathfrak{H}^0 l'operatore hamiltoniano imperturbato):

$$(182) \quad \mathfrak{H}^0 \psi_i^0 = E_i^0 \psi_i^0, \quad \mathfrak{H}^0 \psi_\sigma^0 = E_\sigma^0 \psi_\sigma^0.$$

Per lo stato perturbato i -esimo si avrà invece

$$(183) \quad (\mathfrak{H}^0 + \mathfrak{D}) \psi_i = E_i \psi_i.$$

Ora si noti che, a differenza di quanto accadeva nel caso precedente, l'autofunzione perturbata ψ_i in generale non è prossima alla autofunzione imperturbata ψ_i^0 , ma differisce da essa (e dalle altre ψ_k^0) per termini che non si possono riguardare come piccoli: ciò si può prevedere intuitivamente dal § precedente, osservando che la vicinanza al livello E_i di altri livelli energetici dà luogo nella sommatoria della (177) a dei termini che non sono più piccoli rispetto a ψ_n^0 : in altre parole, le varie autofunzioni del multipletto si « mescolano » tra loro, senza che ve ne sia una che prevale sulle altre. Se dunque sviluppiamo la ψ_i in serie mediante le funzioni ortogonali ψ_j^0 , ψ_σ^0 (che, tutte insieme, formano un sistema completo) avremo

$$(184) \quad \psi_i = \sum_j c_{ij} \psi_j^0 + \sum_\sigma a_{i\sigma} \psi_\sigma^0$$

e potremo considerare le a come piccole del primo ordine rispetto all'unità, mentre le c sono da considerarsi in gene-

rale dell'ordine di grandezza di 1. Sostituendo nella (183), e badando alle (182), otteniamo

$$\sum_j c_{ij} (E_j^0 - E_i + \mathfrak{D}) \psi_j^0 + \sum_\sigma a_{i\sigma} (E_\sigma^0 - E_i + \mathfrak{D}) \psi_\sigma^0 = 0$$

o anche

$$\sum_j c_{ij} (\varepsilon_j^0 - \varepsilon_i + \mathfrak{D}) \psi_j^0 + \sum_\sigma a_{i\sigma} (\varepsilon_\sigma^0 - \varepsilon_i + \mathfrak{D}) \psi_\sigma^0 = 0.$$

Moltiplicando per ψ_k^{0*} , e integrando su tutto il campo di variabilità delle coordinate si ha (ricordando l'ortogonalità e la normalizzazione delle ψ^0 , e introducendo le (172))

$$\sum_j c_{ij} [(\varepsilon_j^0 - \varepsilon_i) \delta_{jk} + L_{kj}] + \sum_\sigma a_{i\sigma} L_{k\sigma} = 0.$$

Poichè le L sono, al pari delle a , piccole del primo ordine, la seconda sommatoria sarà, in prima approssimazione, trascurabile, e resterà (introducendo per le c_{ij} e ε_i i valori di prima approssimazione che chiameremo c_{ij}^0 , ε_i')

$$(185) \quad \sum_j c_{ij}^0 [(\varepsilon_j^0 - \varepsilon_i') \delta_{jk} + L_{kj}] = 0.$$

Fissato un valore di i , dando a k i p valori 1, 2, ..., p , si ha da questa formula un sistema di p equazioni lineari ed omogenee nelle p incognite $c_{i1}^0, c_{i2}^0, \dots, c_{ip}^0$: nei coefficienti di questo sistema figura, oltre alle quantità note L_{kj} , ε_j^0 , la quantità ancora incognita ε_i' .

Ora, affinché il sistema ammetta soluzioni non tutte nulle, bisogna che si annulli il determinante dei coefficienti, ossia dovrà essere

$$(186) \quad \begin{vmatrix} L_{11} + \varepsilon_1^0 - \varepsilon_i' & L_{12} & L_{13} & \dots \\ L_{21} & L_{22} + \varepsilon_2^0 - \varepsilon_i' & L_{23} & \dots \\ L_{31} & L_{32} & L_{33} + \varepsilon_3^0 - \varepsilon_i' & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix} = 0$$

Da questa equazione di grado p possiamo ricavare ε_i' . Essa è della forma detta « equazione secolare » (v. § 12), ed essendo $L_{jk}^* = L_{kj}$, le sue p radici sono tutte reali (e per ora le supporremo distinte): si noti poi che i p sistemi di equazioni che si ottengono da (185) dando ad i i p valori di cui è suscettibile, conducono tutti alla stessa equazione secolare, cosicchè le p radici di questa si possono assumere

come i valori di $\varepsilon'_1, \varepsilon'_2, \dots, \varepsilon'_p$: per stabilire a quale radice spettino ciascuno degli indici 1, 2, ... p , basta osservare che se tendono a 0 tutte le L_{ik} , l'equazione precedente tende a

$$(\varepsilon_1^0 - \varepsilon'_i)(\varepsilon_2^0 - \varepsilon'_i) \dots = 0$$

e quindi una delle radici tende a ε_1^0 , una a ε_2^0 ecc.: con questo criterio si fa il coordinamento.

Così, dalla risoluzione dell'equazione secolare (186), si hanno, in prima approssimazione, le perturbazioni degli autovalori del multipletto.

Da ciascuno dei sistemi (185) si hanno poi, a meno di un fattore costante, le $c_{i1}^0, c_{i2}^0, \dots, c_{ip}^0$: il detto fattore si determina imponendo la condizione

$$(187) \quad \sum_j |c_{ij}^0|^2 = 1.$$

Con ciò le c_{ij}^0 vengono ad avere la proprietà (1)

$$(188) \quad \sum_j c_{ij}^{0*} c_{kj}^0 = \delta_{ik}$$

che si esprime dicendo che esse formano i coefficienti di una « sostituzione ortogonale ».

Se ora poniamo

$$c_{ij} = c_{ij}^0 + c'_{ij}$$

(dove il primo termine è dell'ordine dell'unità, e il secondo è una correzione, piccola del primo ordine), potremo scrivere la (184):

$$\psi_i = \sum_j c_{ij}^0 \psi_j^0 + \sum_j c'_{ij} \psi_j^0 + \sum_{\sigma} a_{i\sigma} \psi_{\sigma}^0,$$

o anche

$$(189) \quad \psi_i = \bar{\psi}_i + \sum_j c'_{ij} \psi_j^0 + \sum_{\sigma} a_{i\sigma} \psi_{\sigma}^0,$$

dove si è posto

$$(190) \quad \bar{\psi}_i = \sum_j c_{ij}^0 \psi_j^0.$$

Nella (189), la $\bar{\psi}_i$ rappresenta il termine principale: come si vede, l'autofunzione imperturbata si approssima (a meno di termini del primo ordine) non a ψ_i^0 ma a $\bar{\psi}_i$. Le $\bar{\psi}_i$ si pos-

(1) Si può infatti dimostrare, servendosi delle (185) e della relazione $L_{ik} = L_{ki}^*$, che si ha $\sum_j c_{ij}^{0*} c_{kj}^0 = 0$ per $i \neq k$.

sono chiamare le autofunzioni di approssimazione *d'ordine zero*.

Merita considerazione separata il caso limite della degenerazione completa, cioè che si tratti non di un multipletto ma addirittura di un livello multiplo d'ordine p . In tal caso le (185) divengono, essendo nulle le ε_j^0 ,

$$(185') \quad \sum_j c_{ij}^0 [L_{kj} - \varepsilon'_i \delta_{kj}] = 0$$

e l'equazione secolare da risolvere diviene

$$(200) \quad \begin{vmatrix} L_{11} - \varepsilon'_i & L_{12} & L_{13} & \dots \\ L_{21} & L_{22} - \varepsilon'_i & L_{23} & \dots \\ L_{31} & L_{32} & L_{33} - \varepsilon'_i & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix} = 0$$

e alle sue p radici reali possono venire attribuiti gli indici 1, 2, ... p in un ordine qualunque: l'effetto della perturbazione è quello di « scindere » il livello energetico E_0 in un gruppo di p livelli vicini, dati, in prima approssimazione, da $E_0 + \varepsilon'_1, E_0 + \varepsilon'_2, \dots$. Al livello E_0 corrispondevano p stati stazionari, differenti quanto alle ψ , ma identici quanto all'energia: per effetto della perturbazione ciascuno di questi stati acquista un'energia livemente diversa. Si può dire perciò che la perturbazione *toglie* la degenerazione (1).

Si può rendersi ragione intuitivamente del fatto che le ψ_i perturbate non si approssimano, in generale, alle ψ_i^0 , ma a certe loro combinazioni lineari $\bar{\psi}_i$, con la considerazione seguente. È noto (v. § 6 p. II) che ad un autovalore multiplo (d'ordine p) si possono attribuire infiniti sistemi di p autofunzioni (ortogonali e normalizzate), i quali sistemi si possono tutti ottenere da uno qualunque di essi mediante una sostituzione lineare ortogonale:

(1) Se l'equazione secolare avesse due o più radici coincidenti, la degenerazione verrebbe solo parzialmente tolta, vale a dire, il livello E_0 , multiplo d'ordine p , si scinderebbe in un gruppo di livelli, alcuni dei quali sarebbero ancora multipli (d'ordine $< p$) malgrado la perturbazione (almeno in prima approssimazione). È questo un caso che si verifica abbastanza spesso.

le ψ_i^0 rappresentano uno qualunque (scelto a caso) di questi sistemi. Ora si pensi allo stato perturbato, che non è degenero (le ψ_i quindi rappresentano un sistema ben definito di autofunzioni): se si fa tendere a 0 la perturbazione, queste ψ_i tendono a un sistema ben definito di autofunzioni imperturbate $\bar{\psi}_i$, e non vi è alcuna ragione che questo sia proprio il sistema delle ψ_i^0 : esso sarà, in generale, legato a quello da una certa sostituzione lineare ortogonale. I sistemi di equazioni (185) servono appunto a trovare i coefficienti c_{ij}^0 di questa sostituzione, e a determinare quindi, mediante la (190), le $\bar{\psi}_i$, le quali risultano in virtù della (188), normalizzate e ortogonali tra loro (nonchè alle ψ_σ^0).

Riprendiamo ora il caso generale, e occupiamoci della ricerca (in prima approssimazione) delle autofunzioni perturbate date dalla (189). È opportuno trasformare ancora questa formula esprimendo anche il secondo termine mediante le $\bar{\psi}$ invece delle ψ^0 : con ciò la formula diverrà

$$(201) \quad \psi_i = \bar{\psi}_i + \sum_l \eta_{il} \bar{\psi}_l + \sum_\sigma a_{i\sigma} \psi_\sigma^0,$$

dove i coefficienti η_{il} (piccoli del primo ordine) sono legati ai c_{ij}^0 dalle relazioni lineari seguenti, che si trovano subito usando la (190),

$$(202) \quad c_{ij}^0 = \sum_l \eta_{il}^0 \eta_{lj}.$$

Scriviamo ora che la ψ_i soddisfa l'equazione di SCHRÖDINGER (183), cioè che

$$(\mathfrak{S}^0 - E_i + \mathfrak{S})(\bar{\psi}_i + \sum_l \eta_{il} \bar{\psi}_l + \sum_\sigma a_{i\sigma} \psi_\sigma^0) = 0.$$

Nello sviluppare questa espressione si osservi che, per la (190) e la prima delle (182),

$$\mathfrak{S}^0 \bar{\psi}_i = \sum_j c_{ij}^0 E_j^0 \psi_j^0 = E_i^0 \bar{\psi}_i + \sum_j \varepsilon_j^0 c_{ij}^0 \psi_j^0$$

e similmente per $\mathfrak{S}^0 \bar{\psi}_l$: inoltre si ricordi la seconda delle (182) e si tenga conto della (202); si ottiene allora

$$(203) \quad \left\{ \begin{array}{l} -\varepsilon_i \bar{\psi}_i + \sum_j \varepsilon_j^0 c_{ij}^0 \psi_j^0 + L \bar{\psi}_i - \varepsilon_i \sum_l \eta_{il} \bar{\psi}_l + \sum_j \varepsilon_j^0 c_{ij}^0 \psi_j^0 + \\ + \sum_l \eta_{il} L \bar{\psi}_l + \sum_\sigma a_{i\sigma} (E_\sigma^0 - E_i) \psi_\sigma^0 + \sum_\sigma a_{i\sigma} L \psi_\sigma^0 = 0. \end{array} \right.$$

Di qui possiamo anzitutto ricavare le a , moltiplicando l'equazione per ψ_σ^{0*} e integrando: si ottiene allora (poichè ψ_σ^0 è ortogonale a $\bar{\psi}_i$, alle ψ_j^0 , alle $\bar{\psi}_l$ e a tutte le ψ_σ^0 per cui $\sigma \neq \varrho$):

$$(204) \quad \bar{L}_{\varrho i} + \sum_l \eta_{il} \bar{L}_{\varrho l} + a_{i\varrho} (E_\varrho^0 - E_i) + a_{i\varrho} L_{\varrho\sigma} = 0,$$

dove si è introdotta la notazione, analoga alla (172),

$$(205) \quad \bar{L}_{\varrho l} = \int \psi_\varrho^{0*} \mathfrak{S} \bar{\psi}_l dS.$$

La (204) è stata ottenuta senza approssimazioni: il suo primo e il suo terzo termine sono piccoli del primo ordine, gli altri del secondo. Trascurando questi ultimi, e sostituendo E_i con E_i^0 , si ricava per le $a_{i\varrho}$ il valore di prima approssimazione

$$(206) \quad a'_{i\varrho} = \frac{\bar{L}_{\varrho i}}{E_i^0 - E_\varrho^0}.$$

Per ricavare le η , e la seconda approssimazione delle ε operiamo ora analogamente, moltiplicando la (203) per $\bar{\psi}_k^*$ e integrando: osserviamo però prima che

$$\int \bar{\psi}_k^* \psi_j^0 dS = \sum_l c_{kl}^{0*} \int \psi_l^{0*} \psi_j^0 dS = c_{kl}^{0*}.$$

e poniamo

$$(205') \quad \bar{L}_{ki} = \int \bar{\psi}_k^* \mathfrak{S} \bar{\psi}_i dS.$$

Otteniamo dunque dalla (203)

$$(207) \quad -\varepsilon_i \delta_{ik} + \sum_j \varepsilon_j^0 c_{ij}^0 c_{kj}^{0*} + \bar{L}_{ki} - \varepsilon_i \eta_{ik} + \sum_j \varepsilon_j^0 c_{ij}^0 c_{kj}^{0*} + \\ + \sum_l \eta_{il} \bar{L}_{kl} + \sum_\sigma a_{i\sigma} \bar{L}_{k\sigma} = 0.$$

Ora si badi che, per la (190), la (205') si scrive:

$$\bar{L}_{ki} = \int \sum_l c_{kl}^{0*} \psi_l^{0*} c_{ij}^0 \mathfrak{S} \psi_j^0 dS = \sum_j c_{kl}^{0*} c_{ij}^0 L_{lj}$$

e poichè la (185) si può scrivere (cambiando l'indice k in l)

$$\sum_j c_{ij}^0 L_{lj} = c_{il}^0 (\varepsilon'_i - \varepsilon_l^0),$$

si ha

$$\bar{L}_{ki} = \sum_l c_{kl}^{0*} c_{il}^0 (\varepsilon'_i - \varepsilon_l^0),$$

e cambiando l'indice di sommatoria l in j e badando alla (188):

$$(208) \quad \bar{L}_{ki} = \varepsilon'_i \delta_{ik} - \sum_j \varepsilon_j^0 c_{kj}^{0*} c_{ij}^0.$$

Sostituendo questa espressione in luogo del terzo termine della (207), si vede che il secondo termine di questa si elide con la sommatoria della (208). Tenendo poi conto della (208) e della (202), il penultimo termine della (207) si trasforma così

$$\begin{aligned} \sum \eta_{il} \bar{L}_{kl} &= \sum \eta_{il} (\varepsilon'_i \delta_{ik} - \sum_j \varepsilon_j^0 c_{kj}^0 c_{lj}^0) = \varepsilon'_k \eta_{ik} - \sum_j \varepsilon_j^0 c_{kj}^0 \sum \eta_{il} c_{lj}^0 = \\ &= \varepsilon'_k \eta_{ik} - \sum_j \varepsilon_j^0 c_{kj}^0 c_{ij}^0. \end{aligned}$$

Così, in definitiva, la (207) si riduce a

$$(209) \quad -(\varepsilon_i - \varepsilon'_i) \delta_{ik} - (\varepsilon_i - \varepsilon'_i) \eta_{ik} + \sum_{\sigma} a_{i\sigma} \bar{L}_{k\sigma} = 0.$$

Da questa ricaviamo, per $k \neq i$,

$$\eta_{ik} = \frac{1}{\varepsilon_i - \varepsilon'_k} \sum_{\sigma} a_{i\sigma} \bar{L}_{k\sigma},$$

ovvero, in prima approssimazione, utilizzando la (206)

$$(210) \quad \eta_{ik} = \frac{1}{\varepsilon'_i - \varepsilon'_k} \sum_{\sigma} \frac{\bar{L}_{\sigma i} \bar{L}_{k\sigma}}{E_i^0 - E_{\sigma}^0} \quad (i \neq k).$$

Per completare la conoscenza in prima approssimazione delle autofunzioni perturbate, mancano solo i coefficienti η_{ii} , che si determinano (come a_{nn} nel § precedente) con la condizione di normalizzazione di ψ_i , e si trova anche qui che si possono prendere nulli.

Passiamo alla determinazione (in pratica assai più importante) degli autovalori perturbati in seconda approssimazione: questi si ricavano dalla (209) per $k = i$. Difatti, indicando con ε_i'' , i termini del secondo ordine, e trascurando quelli d'ordine superiore, cioè ponendo $\varepsilon_i = \varepsilon'_i + \varepsilon_i''$, la (209) dà, per $k = i$,

$$-\varepsilon_i'' - \varepsilon_i'' \eta_{ii} + \sum_{\sigma} a_{i\sigma} \bar{L}_{i\sigma} = 0,$$

ossia, trascurando il secondo termine (che si è preso nullo ed è, in ogni caso, del terzo ordine) e utilizzando la (206):

$$(211) \quad \varepsilon_i'' = \sum_{\sigma} \frac{|\bar{L}_{i\sigma}|^2}{E_i^0 - E_{\sigma}^0}.$$

Quindi la seconda approssimazione degli autovalori del multipletto o del livello multiplo è data da

$$(212) \quad E_i = E_0 + \varepsilon'_i + \sum_{\sigma} \frac{|\bar{L}_{i\sigma}|^2}{E_i^0 - E_{\sigma}^0}.$$

§ 40. — TEORIA DELLE PERTURBAZIONI COL METODO DELLE MATRICI.

La teoria delle perturbazioni degli stati stazionari svolta nei §§ precedenti si può naturalmente presentare anche dal punto di vista del metodo delle matrici, che, come sappiamo, conduce a risultati equivalenti a quelli della meccanica ondulatoria. Lo mostreremo ora a titolo d'esempio, limitandoci alla prima approssimazione ed al caso non degenere.

Assumendo come sistema di riferimento nello spazio hilbertiano quello definito dalle ψ_n^0 , cioè riferendoci allo « schema \mathcal{H} », v. § 33, l'operatore hamiltoniano imperturbato \mathfrak{H}^0 è rappresentato dalla matrice diagonale

$$\mathfrak{H}^0 = \begin{vmatrix} E_1^0 & 0 & 0 & \dots \\ 0 & E_2^0 & 0 & \dots \\ 0 & 0 & E_3^0 & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix}.$$

L'operatore hamiltoniano perturbato $\mathfrak{H} = \mathfrak{H}^0 + \mathfrak{Q}$ sarà invece rappresentato rispetto agli stessi assi da una matrice $\{\mathfrak{H}\}$ non diagonale (in cui però gli elementi non diagonali sono piccoli rispetto a quelli diagonali): si tratta di operare un cambiamento di assi di riferimento tale, che rispetto ai nuovi assi questo operatore sia rappresentato da una matrice diagonale $\{\mathfrak{H}'\}$ (i cui elementi daranno i livelli energetici perturbati E_n). Ricordiamo perciò dal § 7 che a un cambiamento di assi corrisponde una matrice di trasformazione $\{\mathfrak{C}\}$ e che la trasformata $\{\mathfrak{H}'\}$ della matrice $\{\mathfrak{H}\}$ è data da $\{\mathfrak{H}'\} = \{\mathfrak{C}\}^{-1} \{\mathfrak{H}\} \{\mathfrak{C}\}$: si ha cioè

$$(213) \quad \{\mathfrak{C}\} \{\mathfrak{H}'\} = \{\mathfrak{H}\} \{\mathfrak{C}\}.$$

Nel caso attuale, essendo la matrice $\{\mathfrak{H}\}$ quasi diagonale, la matrice di trasformazione che la rende diagonale sarà poco diversa dalla matrice unità $\{1\}$ (ossia, definirà una rotazione piccolissima degli assi di riferimento), e perciò la scriveremo nella forma

$$(214) \quad \{\mathfrak{C}\} = \{1\} + \{\alpha\}.$$

dove gli elementi a_{rs} della matrice $\{\alpha\}$ sono quantità piccole del primo ordine, che si tratta di determinare. La (213) diviene allora, ponendovi $\{\mathfrak{S}\} = \{\mathfrak{S}^0\} + \{\mathfrak{Q}\}$, trascurando il prodotto del secondo ordine $\{\mathfrak{Q}\}\{\alpha\}$, e sostituendo perciò $\{\alpha\}$ con la sua prima approssimazione $\{\alpha'\}$,

$$\{\mathfrak{S}'\} = \{\mathfrak{S}^0\} + \{\mathfrak{Q}\} + \{\mathfrak{S}^0\}\{\alpha'\} - \{\alpha'\}\{\mathfrak{S}'\},$$

ovvero, poichè nell'ultimo termine si può sostituire $\{\mathfrak{S}\}$, con $\{\mathfrak{S}^0\}$ commettendo un errore del secondo ordine,

$$(215) \quad \{\mathfrak{S}'\} = \{\mathfrak{S}^0\} + \{\mathfrak{Q}\} + \{\mathfrak{S}^0\}\{\alpha'\} - \{\alpha'\}\{\mathfrak{S}^0\}.$$

Questa relazione si traduce nella seguente relazione tra gli elementi (ricordando che gli elementi di $\{\mathfrak{S}^0\}$ sono della forma $H_{rs}^0 = E_r^0 \delta_{rs}$, e quelli di $\{\mathfrak{S}'\}$ devono risultare pure della forma $H'_{rs} = E_r \delta_{rs}$):

$$(216) \quad E_n \delta_{nm} = E_n^0 \delta_{nm} + L_{nm} + (E_n^0 - E_m^0) a'_{nm}.$$

Da questa relazione si traggono intanto gli autovalori perturbati, anche senza determinare le a'_{nm} : difatti, per $m = n$ essa diviene

$$(217) \quad E_n = E_n^0 + L_{nn}$$

e si trova così il risultato (175). Per $m \neq n$ si ricava invece

$$(218) \quad a'_{nm} = \frac{-L_{nm}}{E_n^0 - E_m^0}.$$

Gli elementi diagonali a_{nn} restano arbitrari (purchè reali) e si possono prendere nulli.

Determinata così la matrice di trasformazione $\{\mathfrak{S}\}$, ricordiamo che i versori ψ_n degli assi ruotati si ottengono da quelli ψ_n degli assi primitivi mediante la formula (v. § 7)

$$\psi_n = \sum_m S_{mn} \psi_m^0,$$

ossia, essendo $S_{nn} = \delta_{nn} + a_{nn}$,

$$\psi_n = \psi_n^0 + \sum_m a_{mn} \psi_m^0.$$

Poichè $a_{nn} = 0$, converrà usare invece di m l'indice di sommatoria σ (che esclude il valore n) e con ciò, introducendo

per le a i valori di prima approssimazione (218), la formula si identifica con la (177). Si osservi che i coefficienti a del § 38 (in cui si è trascurato di scrivere l'indice n) si identificano con gli elementi della matrice $\{\alpha\}$ della (214), e definiscono quindi la rotazione infinitesima che porta dagli assi, in cui è diagonale $\{\mathfrak{S}^0\}$, a quelli in cui è diagonale $\{\mathfrak{S}\}$. È così messa in luce l'interpretazione da darsi nello spazio hilbertiano al procedimento di approssimazione svolto al § 38.

§ 41. - PERTURBAZIONE DI STATI NON STAZIONARI. PERTURBAZIONI DIPENDENTI DAL TEMPO.

Vogliamo ora trattare il problema delle perturbazioni in modo più generale, così da includere anche il caso di stati risultanti dalla sovrapposizione di più stati stazionari, e di perturbazioni dipendenti dal tempo: il metodo che useremo è noto in matematica col nome di « metodo della variazione delle costanti ».

Indichiamo, come prima, con ψ_r^0 le autofunzioni del sistema imperturbato, le quali hanno la forma

$$(219) \quad \psi_r^0 = u_r^0(q) e^{-\frac{2\pi i}{h} E_r^0 t}$$

e soddisfano l'equazione

$$(220) \quad \mathfrak{S}^0 \psi_r^0 = -\frac{h}{2\pi i} \frac{\partial \psi_r^0}{\partial t}.$$

Introduciamo ora una perturbazione, dipendente eventualmente anche dal tempo, per la quale l'hamiltoniana divenga

$$\mathfrak{H}(q, p) + \mathfrak{Q}(q, p, t).$$

Ci troviamo qui per la prima volta di fronte ad un'hamiltoniana dipendente da t : postuliamo che anche per essa valga l'equazione temporale di SCHRÖDINGER nella stessa forma usata fin qui. L'autofunzione perturbata ψ soddisferà dunque l'equazione

$$(220') \quad (\mathfrak{S}^0 + \mathfrak{Q})\psi = -\frac{h}{2\pi i} \frac{\partial \psi}{\partial t}.$$

Questa ψ si potrà sviluppare mediante le funzioni ortogonali ψ_r^0 : avremo

$$(221) \quad \psi = \sum_r c_r \psi_r^0,$$

dove i coefficienti c_r in generale saranno funzioni di t . Sostituendo questo sviluppo nella (220') (e indicando, come faremo sempre, col punto la derivazione rispetto a t) si ottiene, tenendo conto della (220),

$$\sum_r c_r \Omega \psi_r^0 = -\frac{\hbar}{2\pi i} \sum_r \dot{c}_r \psi_r^0.$$

Di qui si ricavano le \dot{c} , moltiplicando i due membri per ψ_s^{0*} e integrando su tutto lo spazio delle q : si ottiene così

$$(222) \quad \dot{c}_s = -\frac{2\pi i}{\hbar} \sum_r c_r L_{sr},$$

dove, come nei §§ precedenti, le

$$(223) \quad L_{sr} = \int \psi_s^{0*} \Omega \psi_r^0 dS$$

rappresentano gli elementi della « matrice di perturbazione », e si possono anche scrivere, in virtù della (219),

$$(224) \quad L_{sr} = \lambda_{sr} e^{\frac{2\pi i}{\hbar} (E_s^0 - E_r^0)t},$$

con

$$(225) \quad \lambda_{sr} = \int u_s^0 \Omega u_r^0 dS.$$

Supponiamo che per $t = 0$ lo stato del sistema sia rappresentato da una certa $\psi(0)$ da considerarsi nota, che, sviluppata in serie mediante le u_r^0 , sia

$$(226) \quad \psi(0) = \sum_r c_r^0 u_r^0.$$

Confrontando con la (221) e notando che le u_r^0 sono i valori delle ψ_r^0 per $t = 0$, si vede che i coefficienti c_r^0 dello sviluppo (226) (che sono da riguardarsi noti) rappresentano i valori iniziali delle c_r : essi vanno associati alle equazioni differenziali (222) per ottenere le c_r in funzione di t , ossia l'evoluzione dello stato del sistema. In assenza di perturbazione, le c_r conserverebbero evidentemente i valori iniziali, e quindi lo stato del sistema (imperturbato) al tempo t sarebbe dato da

$$(227) \quad \psi^0(t) = \sum_r c_r^0 \psi_r^0.$$

Tutte queste formule sono rigorose qualunque sia l'entità della perturbazione. Ora supponiamo che lo stato perturbato al tempo t differisca poco dallo stato imperturbato, ossia che $\psi(t)$ differisca poco da $\psi^0(t)$ e precisamente per termini del primo ordine (questa approssimazione sarà dunque valida per un tempo tanto più lungo, a partire da $t = 0$, quanto più lieve è la perturbazione). Allora le $c_r(t)$ differiscono dalle c_r^0 per termini del primo ordine, cosicchè nel secondo membro della (222) si può sostituire $c_r(t)$ con c_r^0 commettendo un errore del secondo ordine: si ottengono allora per le c_s le espressioni

$$\dot{c}_s' = -\frac{2\pi i}{\hbar} \sum_r c_r^0 L_{sr}$$

(dove l'apice indica che si tratta di prima approssimazione). Integrando tra 0 e t si hanno i valori di prima approssimazione delle c_s

$$(228) \quad c_s'(t) = c_s^0 - \frac{2\pi i}{\hbar} \sum_r c_r^0 \int_0^t L_{sr} dt,$$

che, sostituiti nella (221), ci danno lo stato perturbato in prima approssimazione.

Sostituendo poi le (228) nei secondi membri delle (222), e integrando fra 0 e t si otterrebbe facilmente la seconda approssimazione, e così per le successive; noi però ci limiteremo alla prima.

§ 42. - PROBABILITÀ DI TRANSIZIONE.

Supponiamo ora che la perturbazione duri soltanto per un certo intervallo di tempo, da 0 a t_1 , mentre per $t < 0$ e $t > t_1$ sia $\Omega = 0$: supponiamo inoltre che prima dell'istante $t = 0$ il sistema si trovi nello stato stazionario ψ_n^0 , di energia E_n^0 , cioè che nella (226) sia

$$(229) \quad c_n^0 = 1, \quad c_q^0 = 0, \quad \text{per } q \neq n.$$

Da quando comincia ad agire la perturbazione, il detto stato non è più stazionario, e la ψ del sistema nel tempo da 0 a t_1 si può scrivere nella forma (221), dove le c_r , tenuto

conto delle (228) e dei valori iniziali (229), sono date in prima approssimazione da

$$c'_n(t) = 1 - \frac{2\pi i}{h} \int_0^t L_{nn} dt$$

$$c'_\varrho(t) = - \frac{2\pi i}{h} \int_0^t L_{\varrho n} dt \quad (\varrho \neq n).$$

Dall'istante t_1 in poi i coefficienti c_n , c_ϱ tornano a diventare costanti, ma anzichè i valori (229) hanno i valori ottenuti dalle formule precedenti sostituendovi t con t_1 : li chiameremo c_n^1 , c_ϱ^1 , cioè porremo

$$(230) \quad c_n^1 = 1 - \frac{2\pi i}{h} \int_0^{t_1} L_{nn} dt$$

$$(230') \quad c_\varrho^1 = - \frac{2\pi i}{h} \int_0^{t_1} L_{\varrho n} dt \quad (\varrho \neq n).$$

Ciò significa che se, dopo l'istante t_1 , si fa una nuova determinazione dello stato, c'è una certa probabilità, data in prima approssimazione da $|c_\varrho^1|^2$, di trovare il sistema nello stato ψ_ϱ^0 anzichè nello stato iniziale ψ_n^0 . Se l'osservazione dello stato dà questo risultato, si dirà che si è prodotta una *transizione* dallo stato n -esimo al ϱ -esimo: l'effetto della perturbazione è dunque quello di indurre una certa « *probabilità di transizione* » tra uno stato (stazionario in assenza di perturbazione) e gli altri, e l'ampiezza di questa probabilità per il passaggio $n \rightarrow \varrho$ è determinata, come mostra la (230'), dall'elemento $L_{\varrho n}$ della matrice di perturbazione. La probabilità del passaggio inverso $\varrho \rightarrow n$ è evidentemente la stessa, essendo $L_{n\varrho} = L_{\varrho n}^*$.

Si osservi la differenza tra questo risultato della meccanica quantistica ed il punto di vista della antica teoria di Bohr e Sommerfeld, la quale postulava che il sistema restasse nello stato stazionario n -esimo fino a che, con un processo brusco, « saltava » nello stato ϱ -esimo: la meccanica quantistica invece, ammettendo, oltre agli stati stazionari puri, l'esistenza di stati risultanti dalla sovrapposizione di diversi stati stazionari, concepisce l'effetto della perturbazione come un processo continuo per cui allo stato n -esimo puro si sovrappongono gradualmente « dosi » va-

riabili degli altri stati: la transizione brusca in uno di questi avviene poi solo al momento dell'osservazione, per effetto dell'inevitabile azione degli strumenti di osservazione sopra il sistema osservato.

§ 43. - PERTURBAZIONE DI TIPO SINUSOIDALE. RISONANZA (1).

Applichiamo i risultati del § precedente al caso in cui la forza perturbatrice è funzione sinusoidale del tempo, di frequenza ν : tale caso si presenta, p. es., quando un atomo è investito da una radiazione monocromatica.

Supponiamo dunque che il termine perturbatore dell'hamiltoniana sia della forma

$$\mathcal{L} = A(q, p) \cos 2\pi\nu t:$$

gli elementi della matrice di perturbazione risulteranno (v. form. 224 e 225):

$$L_{sr} = A_{sr} e^{\frac{2\pi i}{h} (E_s^0 - E_r^0)t} \cos 2\pi\nu t,$$

con A_{sr} indipendente dal tempo, ovvero anche, ponendo

$$\frac{E_s^0 - E_r^0}{h} = \nu_{sr},$$

$$(231) \quad L_{sr} = \frac{1}{2} A_{sr} \left[e^{2\pi i(\nu_{sr} + \nu)t} + e^{2\pi i(\nu_{sr} - \nu)t} \right].$$

L'ampiezza della probabilità di transizione dallo stato n allo stato ϱ , dopo un tempo t_1 di azione della perturbazione, ci è data dalla (230'), che fornisce in questo caso:

$$(232) \quad c_\varrho^1 = - \frac{A_{\varrho n}}{2h} \left[e^{\frac{2\pi i(\nu_{\varrho n} + \nu)t_1 - 1}{\nu_{\varrho n} + \nu}} + \frac{e^{\frac{2\pi i(\nu_{\varrho n} - \nu)t_1 - 1}{\nu_{\varrho n} - \nu}} \right] =$$

$$= - \frac{iA_{\varrho n}}{h} \left[e^{i\pi(\nu_{\varrho n} + \nu)t_1} \frac{\sin \pi(\nu_{\varrho n} + \nu)t_1}{\nu_{\varrho n} + \nu} + e^{i\pi(\nu_{\varrho n} - \nu)t_1} \frac{\sin \pi(\nu_{\varrho n} - \nu)t_1}{\nu_{\varrho n} - \nu} \right].$$

(1) La parola « risonanza » è qui usata nel senso classico. In meccanica quantistica essa ha anche un altro significato, che verrà illustrato nel cap. VI.

Ora è noto che la espressione $\left| \frac{\sin \pi a t_1}{a} \right|$, considerata come funzione di a , ha un massimo assoluto uguale a πt_1 , per $a = 0$, mentre gli altri massimi sono notevolmente inferiori (l'andamento qualitativo è lo stesso della curva della fig. 20): perciò i due termini della parentesi quadra potranno acquistare valori notevoli solo per $\nu = -\nu_{en}$ (il primo) e per $\nu = \nu_{en}$ (il secondo) (si potrà verificare il primo di questi casi se $E_q^0 < E_n^0$, il secondo se $E_q^0 > E_n^0$): in entrambi i casi si ha, trascurando l'altro termine,

$$c_q^1 = - \frac{i\pi A_{en}}{h} t_1.$$

Ciò significa che la probabilità di transizione dallo stato n allo stato q è rilevante solo se la differenza di energia tra i due stati, $|E_q^0 - E_n^0|$, è molto vicina al valore $h\nu$, ossia se $|\nu_{en}|$ è assai vicina a ν : praticamente, il massimo è così acuto che si deve ritenere possibile la transizione solo se $|\nu_{en}| = \nu$ (condizione di *risonanza*). Questa transizione, se $E_q^0 > E_n^0$, richiede assorbimento di energia a spese della radiazione: ne segue che *l'atomo nello stato n può assorbire solo radiazioni di frequenza $\nu = \frac{1}{h}(E_q^0 - E_n^0)$, e le assorbe per quanti $h\nu$* : così il postulato di BOHR sulla relazione tra energia e frequenza resta dimostrato (per il caso dell'assorbimento) come conseguenza della meccanica ondulatoria.

Se $E_q^0 < E_n^0$, la transizione $n \rightarrow q$ si compie con emissione di radiazione: la teoria qui esposta non dà gli elementi per calcolare la frequenza di questa radiazione, ma la teoria completa dell'interazione tra campo di radiazione e atomo, dovuta al DIRAC (alla quale si è accennato al § 32, p. II) consente di dimostrare anche in questo caso la validità del postulato di BOHR. Inoltre essa dimostra la possibilità di transizioni « spontanee » (con emissione) dallo stato n ad uno inferiore, mentre quelle di cui qui ci occupiamo sono transizioni indotte dalla radiazione esterna.

CAPITOLO V.

RELATIVITÀ E SPIN

§ 44. - CONSIDERAZIONI GENERALI.

La meccanica quantistica svolta nei capitoli precedenti è stata costruita partendo dalla meccanica classica (non relativistica) del punto materiale (e ad essa si riduce, come sappiamo, ogni volta che un pacchetto d'onde si può considerare puntiforme). Ma è noto che la meccanica classica rappresenta solo una prima approssimazione, valida per moti non troppo veloci rispetto alla velocità della luce c , della meccanica relativistica, che è valida per moti comunque veloci: è quindi da ritenersi che anche la meccanica quantistica che abbiamo svolto fin qui sia valida con le stesse limitazioni, e che per ottenere invece la meccanica quantistica più generale e rigorosa si debba partire dalla meccanica relativistica del punto anziché da quella classica. La necessità di questo perfezionamento risulta evidente se si considera che i risultati della meccanica ondulatoria di Schrödinger non sono invarianti rispetto alle trasformazioni di Lorentz.

Una seconda circostanza lasciata da parte nei capitoli precedenti è l'esistenza di un momento angolare intrinseco (*spin*) e di un momento magnetico, sia nell'elettrone che nel protone e presumibilmente in altre particelle, esistenza che, come si è detto al § 25, p. I e al § 62, p. II, è provata da molti fatti d'ordine spettroscopico e magnetico, e fu postulata per la prima volta da UHLENBECK e GOUDSMIT col nome di « ipotesi dell'elettrone rotante ».

A queste due cause di inesattezza della meccanica quantistica si tentò dapprima di ovviare separatamente. Così, estendendo l'analogia con la meccanica ordinaria, fu cercata da vari autori una modificazione dell'equazione di Schrödinger per tener conto della correzione relativistica: questo tentativo, che esporremo brevemente al § 46, ebbe un successo soltanto parziale. D'altra parte, il PAULI riuscì a introdurre nella meccanica quantistica (non relativistica) l'ipotesi dello spin, costruendo una notevole teoria che sarà esposta al § 45. Ma la soluzione più soddisfacente di entrambe le questioni fu trovata dal DIRAC, il quale mostrò che le due modifiche — quella per la relatività e quella per l'elettrone rotante — si riducono concettualmente ad una sola, perchè, data alla meccanica ondulatoria una conveniente forma relativistica, ne discende come conseguenza l'esistenza dello spin e del momento magnetico, coi loro giusti valori e con le loro regole di quantizzazione spaziale, senza necessità di introdurlì con ipotesi *ad hoc*. Da questa teoria di Dirac si può poi ricavare quella di Pauli come una prima approssimazione non relativistica: noi però premetteremo quest'ultima, perchè ci fornirà un esempio semplice di taluni algoritmi nuovi, che nella teoria di Dirac sono usati su più vasta scala ⁽¹⁾.

§ 45. — FONDAMENTI DELLA TEORIA NON RELATIVISTICA DELLO SPIN (PAULI).

Ricordiamo che il punto essenziale dell'ipotesi di UHLENBECK e GOUDSMIT (dal quale derivano le conseguenze spettroscopiche che hanno dato valore all'ipotesi) è che *la proiezione dello spin su una direzione qualunque* ha sempre uno

(1) La teoria del DIRAC si riferisce essenzialmente al caso di un solo elettrone: manca finora una teoria relativistica rigorosa e completa di un sistema di più elettroni. Sulle difficoltà che sorgono in tale problema, come pure su varie altre questioni di principio ancora insolute o controverse nella attuale teoria dell'elettrone, vedasi il n. 14 della bibl., a pag. 214 e 269, e inoltre vari lavori recenti, tra cui: W. PAULI e V. WEISSKOPF, *Helv. Phys. Acta*, 7 (1934), p. 709; M. BORN, *Proc. Roy. Soc.*, 143 (1934), p. 410; M. BORN e L. INFELD, *Proc. Roy. Soc.*, 144 (1934), p. 425; V. WEISSKOPF, *Naturw.* 23 (1935), p. 669 e segg.

dei due valori $\pm h/4\pi$, e corrispondentemente la proiezione del momento magnetico sulla stessa direzione ha il valore $\mp \mu_0$ (chiamando μ_0 il valore assoluto del magnetone di Bohr). Introdurremo perciò tre nuove osservabili $\sigma_x, \sigma_y, \sigma_z$, rappresentanti le proiezioni dello spin sui tre assi, *misurate in unità $h/4\pi$* e, per rispecchiare l'ipotesi di Uhlenbeck e Goudsmit, ammetteremo che ciascuna di esse abbia come soli autovalori ± 1 . Le chiameremo senz'altro « componenti dello spin » (sottintendendo: in unità $h/4\pi$) e, secondo l'uso, denoteremo con gli stessi simboli gli operatori che ad esse corrispondono. Inoltre, introdurremo le componenti μ_x, μ_y, μ_z del momento magnetico, come tre osservabili (o tre operatori) legate alle precedenti dalle formule:

$$(233) \quad \mu_x = -\mu_0\sigma_x, \quad \mu_y = -\mu_0\sigma_y, \quad \mu_z = -\mu_0\sigma_z.$$

Studiamo ora le proprietà degli operatori $\sigma_x, \sigma_y, \sigma_z$ così definiti. Osserviamo anzitutto che, poichè essi hanno i soli due autovalori ± 1 , i loro quadrati avranno l'unico autovalore $+1$, cosicchè sarà lecito scrivere ⁽¹⁾:

$$(234) \quad \sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1.$$

Bisogna poi stabilire le regole di permutazione degli operatori $\sigma_x, \sigma_y, \sigma_z$. Ciò è stato fatto dal PAULI ammettendo che le componenti dello spin si comportino a questo riguardo come le componenti di un ordinario momento angolare, le quali soddisfano, come si è dimostrato al § 30, le relazioni di permutazione (125). Poichè queste valgono per momenti misurati nelle ordinarie unità, dovremo in esse sostituire \mathfrak{M}_x con $\frac{h}{4\pi}\sigma_x$, ecc.; troviamo così le relazioni di permutazione seguenti:

$$(235) \quad \begin{aligned} \sigma_y\sigma_z - \sigma_z\sigma_y &= 2i\sigma_x \\ \sigma_z\sigma_x - \sigma_x\sigma_z &= 2i\sigma_y \\ \sigma_x\sigma_y - \sigma_y\sigma_x &= 2i\sigma_z. \end{aligned}$$

(1) Si noti incidentalmente che da ciò segue $\sigma_x^2 + \sigma_y^2 + \sigma_z^2 = 3$, e quindi lo *spin totale* σ (in unità $h/4\pi$), definito da $\sigma^2 = \sigma_x^2 + \sigma_y^2 + \sigma_z^2$, risulta $\sqrt{3}$ e non 1 come si ammette di solito nel modello vettoriale. Ciò dipende dalla inadeguatezza, più volte rilevata, del detto modello: si

Moltiplichiamo ora (a destra) la seconda per σ_z e la terza per σ_y , e sommiamole: si ha

$$\sigma_z \sigma_x \sigma_z - \sigma_y \sigma_x \sigma_y + \sigma_x (\sigma_y^2 - \sigma_z^2) = 2i(\sigma_y \sigma_z + \sigma_z \sigma_y).$$

Sostituendo nei primi due termini per σ_x l'espressione ricavata dalla prima delle (235), e ricordando le (234), si riconosce che tutto il primo membro è uguale a zero, e quindi resta

$$\sigma_y \sigma_z + \sigma_z \sigma_y = 0.$$

Altre due relazioni analoghe a questa si ricaverebbero nello stesso modo: le componenti dello spin sono dunque *anticommutative*. Tenendo poi conto di nuovo delle (235) si può scrivere:

$$(236) \quad \begin{aligned} \sigma_y \sigma_z &= -\sigma_z \sigma_y = i\sigma_x \\ \sigma_z \sigma_x &= -\sigma_x \sigma_z = i\sigma_y \\ \sigma_x \sigma_y &= -\sigma_y \sigma_x = i\sigma_z. \end{aligned}$$

Quando si tiene conto dello spin, una misura delle coordinate x, y, z dell'elettrone non è più una « osservazione massima », poichè non definisce ancora completamente lo stato dell'elettrone: per completarla bisogna aggiungere una osservazione della proiezione dello spin (o, ciò che è lo stesso, del momento magnetico) su una direzione qualunque, che generalmente si sceglie coincidente con l'asse z . La determinazione dello stato si completa dunque con l'osservazione di σ_z (e si noti che non si può aggiungere ancora una osservazione di σ_x e di σ_y , essendo queste incompatibili con quella di σ_z , poichè i relativi operatori non sono permutabili): lo stato sarà dunque ora definito da una funzione $\psi(x, y, z, \sigma_z, t)$ anzichè solo da una $\psi(x, y, z, t)$. Ma poichè la nuova variabile introdotta (detta *variabile di spin*) può avere solo i due valori ± 1 , è spesso opportuno considerarla piuttosto come un *indice* che come un'ordinaria variabile, e cioè scrivere $\psi_1(x, y, z; t)$, $\psi_2(x, y, z; t)$ invece di $\psi(x, y, z, +1; t)$,

osservi infatti che quando nella teoria di UHLENBECK e GOUDSMIT si fa uso non della proiezione dello spin su una data direzione, ma dello spin totale, si giunge a formule che richiedono delle lievi modificazioni per essere messe d'accordo con l'esperienza (v. p. es. bibl., n. 27, p. 198).

$\psi(x, y, z, -1; t)$. L'introduzione della variabile di spin equivale dunque a introdurre due funzioni ψ_1, ψ_2 , in luogo di una sola ψ , ed è questa la caratteristica principale della teoria di Pauli (v. § 25, p. II). Il significato di queste due funzioni è ovviamente il seguente: $|\psi_1|^2 dx dy dz$ rappresenta la probabilità di trovare l'elettrone nell'elemento di volume definito da $x, y, z, x + dx, y + dy, z + dz$, e con $\sigma_z = 1$ (ovvero con $\mu_z = -\mu_0$), mentre $|\psi_2|^2 dx dy dz$ è la probabilità di trovarlo nello stesso elemento di volume, ma con $\sigma_z = -1$ (ossia con $\mu_z = +\mu_0$): usando il linguaggio del modello vettoriale si direbbe che nel primo caso lo spin è *parallelo* all'asse z , nel secondo *antiparallelo*. La parte privilegiata data all'asse z si potrebbe, naturalmente, dare invece all'asse x o all'asse y , o ad un'altra direzione qualunque: bisognerebbe allora usare un'altra coppia di funzioni ψ .

La condizione di normalizzazione è evidentemente:

$$(237) \quad \iiint (|\psi_1|^2 + |\psi_2|^2) dx dy dz = 1.$$

Talvolta conviene (analogamente a quanto si è fatto al § 6 bis) considerare la coppia ψ_1, ψ_2 come una matrice a una sola colonna e a due righe, che si indica simbolicamente con ψ , cioè porre

$$(238) \quad \psi = \begin{vmatrix} \psi_1 \\ \psi_2 \end{vmatrix}.$$

In corrispondenza a questa rappresentazione, è conveniente rappresentare gli operatori $\sigma_x, \sigma_y, \sigma_z$ mediante matrici, osservando che essi hanno solo due autovalori e perciò saranno rappresentati da matrici di due sole righe e colonne: saranno cioè della forma

$$\begin{vmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{vmatrix}$$

dove le k sono costanti, e, per l'hermiticità, deve essere $k_{21} = k_{12}^*$. Siffatti operatori, applicati alla (238), sostituiscono le funzioni ψ_1, ψ_2 con due loro combinazioni lineari, secondo lo schema (conforme alla regola di moltiplicazione delle matrici):

$$(239) \quad \begin{vmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{vmatrix} \begin{vmatrix} \psi_1 \\ \psi_2 \end{vmatrix} = \begin{vmatrix} k_{11}\psi_1 + k_{12}\psi_2 \\ k_{21}\psi_1 + k_{22}\psi_2 \end{vmatrix}.$$

(Poichè in tale processo ψ_1, ψ_2 si comportano come se non dipendessero da x, y, z , si suol dire che un tale operatore « opera solo sulla variabile di spin »). Passiamo alla determinazione effettiva di queste tre matrici, che si indicano con gli stessi simboli $\sigma_x, \sigma_y, \sigma_z$ degli operatori che rappresentano. Osserviamo anzitutto che ψ_1 e ψ_2 , per il significato dato loro più sopra, non sono altro che le autofunzioni dell'operatore σ_z , corrispondenti rispettivamente agli autovalori $+1$ e -1 : deve aversi cioè

$$\begin{aligned} \sigma_z \psi_1 &= \psi_1, & \sigma_z \psi_2 &= -\psi_2, \\ \text{ossia} & & & \\ (240) \quad \sigma_z \begin{vmatrix} \psi_1 \\ \psi_2 \end{vmatrix} &= \begin{vmatrix} \psi_1 \\ -\psi_2 \end{vmatrix}. \end{aligned}$$

Da ciò si deduce, a norma della (239):

$$(241) \quad \sigma_z = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}.$$

Per determinare le matrici σ_x e σ_y , scriviamole dapprima nella forma generica hermitiana

$$\sigma_x = \begin{vmatrix} a_1 & a_3 \\ a_3^* & a_2 \end{vmatrix}, \quad \sigma_y = \begin{vmatrix} b_1 & b_3 \\ b_3^* & b_2 \end{vmatrix},$$

e imponiamo a ciascuna di esse di essere anticommutativa con σ_z : si ha, per la (241),

$$\sigma_x \sigma_z = \begin{vmatrix} a_1 & -a_3 \\ a_3^* & -a_2 \end{vmatrix}, \quad \sigma_z \sigma_x = \begin{vmatrix} -a_1 & a_3 \\ -a_3^* & -a_2 \end{vmatrix},$$

e quindi, perchè sia $\sigma_x \sigma_z + \sigma_z \sigma_x = 0$, deve essere $a_1 = a_2 = 0$. Si ha poi

$$\sigma_x^2 = \begin{vmatrix} a_3 a_3^* & 0 \\ 0 & a_3 a_3^* \end{vmatrix}$$

e, affinchè questa si riduca alla matrice unitaria, deve aversi $a_3 a_3^* = 1$, cioè $a_3 = e^{i\alpha}$, con α reale (arbitrario). Ragionando in modo analogo per σ_y , si conclude che deve aversi

$$\sigma_x = \begin{vmatrix} 0 & e^{i\alpha} \\ e^{-i\alpha} & 0 \end{vmatrix}, \quad \sigma_y = \begin{vmatrix} 0 & e^{i\beta} \\ e^{-i\beta} & 0 \end{vmatrix}.$$

Imponendo poi la condizione $\sigma_x \sigma_y + \sigma_y \sigma_x = 0$, si trova tra α e β il legame $\alpha - \beta = \pi/2 + n\pi$ (con n intero qualunque).

L'arbitrarietà di n e quella che resta in una delle due costanti α, β non ha conseguenze fisiche ⁽¹⁾: assumendo $\alpha = 0$, $\beta = -\pi/2$ si ottiene infine

$$(241') \quad \sigma_x = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad \sigma_y = \begin{vmatrix} 0 & -i \\ i & 0 \end{vmatrix}.$$

Queste, insieme alla (241), danno le cercate espressioni. Si noti che σ_z ha forma diagonale, perchè, per la parte privilegiata che abbiamo conferito all'asse z , le matrici sono riferite allo « schema σ_z »: adottando un altro schema (e quindi un altro significato per ψ_1 e ψ_2) le tre matrici si trasformerebbero come è stato spiegato al § 8.

Per ovvia estensione dei principi del § 22, l'operatore che corrisponde ad una qualsiasi grandezza relativa allo spin si ottiene scrivendo l'espressione classica di questa in funzione delle componenti di spin, e sostituendovi queste con i rispettivi operatori. P. es., la proiezione dello spin su una direzione qualsiasi n , di coseni α, β, γ , sarà rappresentata dall'operatore

$$(242) \quad \sigma_n = \alpha \sigma_x + \beta \sigma_y + \gamma \sigma_z,$$

ossia, nello schema σ_z , dalla matrice:

$$(242') \quad \sigma_n = \begin{vmatrix} \gamma & \alpha - i\beta \\ \alpha + i\beta & -\gamma \end{vmatrix}.$$

che si ottiene immediatamente sostituendo nella (242) le (241) e (241').

Come applicazione, supponiamo che un'osservazione dello spin rispetto a una certa direzione n , di coseni α, β, γ , abbia dato il risultato $+1$, e supponiamo che subito dopo si esegua un'osservazione dello spin rispetto all'asse z : quale è la probabilità di trovare $+1$ e quale è quella di trovare -1 ? Lo stato risultante dalla prima osservazione sarà definito da una $\begin{vmatrix} \psi_1 \\ \psi_2 \end{vmatrix}$ tale che (v. § 22):

$$\sigma_n \begin{vmatrix} \psi_1 \\ \psi_2 \end{vmatrix} = 1 \begin{vmatrix} \psi_1 \\ \psi_2 \end{vmatrix},$$

vale a dire, applicando la (242'),

$$\begin{aligned} \gamma \psi_1 + (\alpha - i\beta) \psi_2 &= \psi_1 \\ (\alpha + i\beta) \psi_1 - \gamma \psi_2 &= \psi_2. \end{aligned}$$

⁽¹⁾ Essa corrisponde all'arbitrarietà della costante θ nell'argomento della ψ , rilevata a pag. 166.

Queste due equazioni omogenee (il cui determinante è nullo in virtù di $\alpha^2 + \beta^2 + \gamma^2 = 1$) danno:

$$\frac{\psi_1}{\psi_2} = \frac{\alpha - i\beta}{1 - \gamma},$$

e quindi il rapporto delle probabilità dei due risultati $+1$ e -1 è:

$$\left| \frac{\psi_1}{\psi_2} \right|^2 = \frac{\alpha^2 + \beta^2}{(1 - \gamma)^2} = \frac{1 + \gamma}{1 - \gamma}.$$

Se, in particolare, la direzione n fosse normale all'asse z , le due probabilità risulterebbero uguali.

Si osservi che tutto ciò non è interpretabile col semplice modello vettoriale, secondo il quale la seconda osservazione darebbe con certezza il risultato $\sigma_n = \gamma$.

L'hamiltoniana di un elettrone dotato di spin in un campo magnetico si scrive per analogia con quella della meccanica classica, la quale, in prima approssimazione ⁽¹⁾, si ottiene semplicemente aggiungendo all'hamiltoniana di un elettrone senza spin, trovata al § 31, l'energia di interazione tra il suo momento e il campo magnetico: si avrà cioè, detta m_0 la massa, $-e$ la carica ⁽²⁾, \mathbf{E} il campo elettrico, \mathbf{H} quello magnetico, V il potenziale scalare e \mathbf{A} quello vettoriale, e detto μ il momento magnetico

$$(243) \quad \mathcal{H} = \frac{1}{2m_0} \sum_r^3 \left(p_r + \frac{e}{c} A_r \right)^2 - eV - \mathbf{H} \times \boldsymbol{\mu}.$$

Sostituendo p_r e μ_r con gli operatori corrispondenti $\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_r}$, $-\mu_0 \sigma_r$, questa espressione si trasforma nell'operatore

$$(244) \quad \mathfrak{H} = \frac{1}{2m_0} \sum_r^3 \left(\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_r} + \frac{e}{c} A_r \right)^2 - eV + \mu_0 \mathbf{H} \times \boldsymbol{\sigma}$$

dove, estendendo in modo ovvio la notazione del prodotto

⁽¹⁾ La formula rigorosa conterrebbe anche dei termini dell'ordine di v/c rispetto agli altri, rappresentanti l'azione del campo elettrico sul magnete in moto.

⁽²⁾ Indicheremo in tutto questo capitolo con e la carica dell'elettrone in valore assoluto, e con m_0 la sua massa di quiete.

interno, si è indicato col simbolo $\mathbf{H} \times \boldsymbol{\sigma}$ l'operatore (o matrice)

$$(245) \quad \mathbf{H} \times \boldsymbol{\sigma} = \sum_r^3 H_r \sigma_r = \begin{vmatrix} H_z & H_x - iH_y \\ H_x + iH_y & -H_z \end{vmatrix}.$$

L'operatore hamiltoniano \mathfrak{H} , così formato, permette poi di scrivere, nel modo solito, l'equazione per la ψ , e cioè

$$(246) \quad \mathfrak{H}\psi = -\frac{\hbar}{2\pi i} \frac{\partial \psi}{\partial t}$$

e, per gli stati stazionari

$$(246') \quad \mathfrak{H}\psi = E\psi.$$

La (246), come pure la (246'), si può considerare formalmente come una equazione nella $\psi(x, y, z, \sigma_z; t)$, ovvero, più esplicitamente, come un sistema di due equazioni nelle due funzioni $\psi_k(x, y, z; t)$ (con $k = 1, 2$): p. es., se si indica con \mathfrak{H}_0 la parte dell'hamiltoniano (244) che non opera sullo spin, cioè se si pone $\mathfrak{H} = \mathfrak{H}_0 + \mu_0 \mathbf{H} \times \boldsymbol{\sigma}$, la (246) si può esplicitare, mediante la (245), nelle due equazioni

$$\begin{aligned} \mathfrak{H}_0 \psi_1 + \mu_0 H_z \psi_1 + \mu_0 (H_x - iH_y) \psi_2 &= -\frac{\hbar}{2\pi i} \frac{\partial \psi_1}{\partial t} \\ \mathfrak{H}_0 \psi_2 + \mu_0 (H_x + iH_y) \psi_1 - \mu_0 H_z \psi_2 &= -\frac{\hbar}{2\pi i} \frac{\partial \psi_2}{\partial t}; \end{aligned}$$

e analogamente per la (11'). Se il campo magnetico è nullo o trascurabile, ciascuna delle due ψ soddisfa l'ordinaria equazione di Schrödinger: perciò, se si tratta di uno stato stazionario non degenere ⁽¹⁾, esse devono differire solo per un coefficiente costante.

Separazione della variabile di spin. — Se il campo magnetico è uniforme, la parte dell'hamiltoniano dipendente dallo spin non contiene le coordinate di posizione, e allora è possibile (per uno stato stazionario) operare la « separazione » della variabile di spin σ_z da x, y, z , cioè porre la $\psi(x, y, z, \sigma_z; t)$ sotto la forma $\Psi_n(x, y, z; t) \varphi_s(\sigma_z)$ dove $\sigma_z = \pm 1$ e quindi $\varphi_s(\sigma_z)$ rappresenta, più che una vera funzione, l'insieme di due c costanti $\varphi_s(1) = \alpha_s$, $\varphi_s(-1) = \beta_s$, ossia la matrice $\begin{vmatrix} \alpha_s \\ \beta_s \end{vmatrix}$. L'in-

⁽¹⁾ Esclusa, s'intende, la degenerazione dovuta allo spin.

dice s è il numero quantico di spin che può assumere, come vedremo, solo due valori, mentre n rappresenta al solito il gruppo di tre numeri quantici orbitali.

Infatti la (246') si scinde allora nelle due

$$(247) \quad \mathfrak{S}_0 \Psi_n = E'_n \Psi_n$$

$$(248) \quad \mu_0 \mathbf{H} \times \mathfrak{T} \varphi_s = E''_s \varphi_s$$

con $E'_n + E''_s = E$. La prima è l'ordinaria equazione di Schrödinger: dunque E'_n rappresenta uno dei livelli energetici della meccanica ondulatoria ordinaria, e Ψ_n è una autofunzione ad esso corrispondente. La seconda (dove E''_s rappresenta l'energia dovuta all'azione del campo magnetico sul momento magnetico di spin) si può esplicitare nelle due equazioni algebriche

$$(249) \quad \begin{cases} \mu_0 H_z \alpha_s + \mu_0 (H_x - iH_y) \beta_s = E''_s \alpha_s \\ \mu_0 (H_x + iH_y) \alpha_s - \mu_0 H_z \beta_s = E''_s \beta_s \end{cases}$$

le quali sono lineari e omogenee in α_s e β_s . Poichè queste non sono entrambe nulle, dovrà essere

$$\begin{vmatrix} \mu_0 H_z - E''_s & \mu_0 (H_x - iH_y) \\ \mu_0 (H_x + iH_y) & -\mu_0 H_z - E''_s \end{vmatrix} = 0,$$

cioè $E''_s{}^2 = \mu_0^2 (H_x^2 + H_y^2 + H_z^2) = \mu_0^2 H^2$: di qui i due valori di E''_s , corrispondenti ai due valori 1,2 dell'indice s ,

$$(250) \quad E''_1 = \mu_0 H, \quad E''_2 = -\mu_0 H.$$

Questo risultato giustifica il successo della teoria modelistica dello spin, in quanto essa postulava che lo spin potesse disporsi o parallelamente o antiparallelamente al campo, corrispondendo ai due casi i due valori (250) dell'energia magnetica.

Se ora supponiamo il campo magnetico diretto secondo l'asse z , e risolviamo il sistema (249) (determinando la costante di normalizzazione in modo che $|\alpha_s|^2 + |\beta_s|^2 = 1$) troviamo:

$$\text{per } E''_1: \quad \alpha_1 = 1, \quad \beta_1 = 0,$$

$$\text{per } E''_2: \quad \alpha_2 = 0, \quad \beta_2 = 1.$$

Nel primo caso si ha dunque $\psi_{n,1} = \begin{vmatrix} \Psi_n \\ 0 \end{vmatrix}$, vale a dire lo spin è diretto con certezza nel verso dell'asse z , nel secondo caso $\psi_{n,2} = \begin{vmatrix} 0 \\ \Psi_n \end{vmatrix}$ e lo spin è diretto con certezza nel verso opposto.

§ 46. - L'ESTENSIONE RELATIVISTICA DELL'EQUAZIONE DI SCHRÖDINGER.

Prima di esporre la teoria di Dirac, mostriamo a quale equazione per la ψ si giungerebbe se, applicando il principio del § 22, si partisse dall'espressione relativistica dell'hamiltoniana (anzichè dall'espressione classica come si è fatto al § 19) e la si trasformasse in operatore mediante la solita sostituzione (S), (S') del § 19.

Consideriamo una particella ⁽¹⁾ di carica ε (in unità elettrostatiche) e massa (di quiete) m_0 , posta in un campo elettrico e magnetico, derivante da un potenziale scalare V e da un potenziale vettoriale \mathbf{A} : l'hamiltoniana relativistica è ⁽²⁾

$$(251) \quad \mathcal{H} = c \sqrt{m_0^2 c^2 + \sum_k \left(p_k - \frac{\varepsilon}{c} A_k \right)^2} + \varepsilon V,$$

dove i momenti p_k sono dati da (v. § 31):

$$(252) \quad p_k = \frac{m_0 v_k}{\sqrt{1 - v^2/c^2}} + \frac{\varepsilon}{c} A_k.$$

L'integrale della forza viva è

$$(253) \quad \mathcal{H} - E = 0;$$

⁽¹⁾ Indichiamo con ε la carica di una particella generica, perchè in questo capitolo riserviamo la lettera e per il valore assoluto della carica elettronica. Nel caso dell'elettrone dunque $\varepsilon = -e$.

⁽²⁾ Difatti, la forza viva è $m_0 c^2 \left(\frac{1}{\sqrt{1 - v^2/c^2}} - 1 \right)$, l'energia intrinseca è $m_0 c^2$, quella elettrostatica εV : l'energia totale è dunque $\frac{m_0 c^2}{\sqrt{1 - v^2/c^2}} + \varepsilon V$. Per esprimerla mediante le p_k , si noti che da (252) si ha $\frac{m_0^2 v^2}{1 - v^2/c^2} = \sum_k \left(p_k - \frac{\varepsilon}{c} A_k \right)^2$, da cui $\frac{m_0^2 c^2}{1 - v^2/c^2} = m_0^2 c^2 + \sum_k \left(p_k - \frac{\varepsilon}{c} A_k \right)^2$: sostituendo nell'espressione dell'energia si ha la (251).

essendo però il primo membro irrazionale nelle p_k , conviene, prima di applicare la sostituzione (S) , (S') , renderlo razionale isolando il radicale ed elevando a quadrato: partiremo dunque, anzichè dalla (253), dalla relazione

$$(254) \quad \frac{1}{c^2} (\varepsilon V - E)^2 - \sum_k \left(p_k - \frac{\varepsilon}{c} A_k \right)^2 - m_0^2 c^2 = 0.$$

Operando la sostituzione (S) , (S') e applicando l'operatore ottenuto alla ψ (indicando inoltre con U , come d'uso, l'energia potenziale εV), si ha la seguente equazione, che dovrebbe rappresentare l'estensione relativistica dell'equazione di Schrödinger

$$(255) \quad \left[\frac{1}{c^2} \left(U + \frac{\hbar}{2\pi i} \frac{\partial}{\partial t} \right)^2 - \sum_k \left(\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_k} - \frac{\varepsilon}{c} A_k \right)^2 - m_0^2 c^2 \right] \psi = 0.$$

Verifichiamo anzitutto che questa equazione, nella approssimazione non relativistica, cioè quando c si può considerare assai grande rispetto alle altre velocità in giuoco, ci riconduce alla teoria di Schrödinger: troveremo però che la ψ che figura nella (255) differisce dalla ψ di Schrödinger (che indicheremo momentaneamente con $\bar{\psi}$) per un fattore di modulo 1, inessenziale nel calcolo di ρ e \mathbf{i} e che corrisponde al fatto che nella energia viene computata anche l'energia intrinseca $m_0 c^2$ dell'elettrone. Infatti, se si pone

$$\psi = \bar{\psi} e^{-\frac{2\pi i}{h} m_0 c^2 t},$$

sostituendo questa espressione nella (255) si ha, con facile calcolo,

$$\left[\frac{1}{c^2} \left(U + \frac{\hbar}{2\pi i} \frac{\partial}{\partial t} \right)^2 - 2m_0 \left(U + \frac{\hbar}{2\pi i} \frac{\partial}{\partial t} \right) - \sum_k \left(\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_k} - \frac{\varepsilon}{c} A_k \right)^2 \right] \bar{\psi} = 0.$$

Se si considera trascurabile il primo termine a causa del fattore $1/c^2$, si ha l'approssimazione non relativistica

$$\frac{1}{2m_0} \sum_k \left(\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_k} - \frac{\varepsilon}{c} A_k \right)^2 \bar{\psi} + U \bar{\psi} = -\frac{\hbar}{2\pi i} \frac{\partial \bar{\psi}}{\partial t},$$

equazione che coincide con la (142') del § 31, salvo le differenze di notazione.

È stato però riconosciuto che, per vari motivi, non si può adottare in generale la (255) come generalizzazione rela-

tivistica dell'equazione di Schrödinger (p. es., secondo la (255) l'integrale di $|\psi|^2$ risulta variabile col tempo, cosicchè non può essere uguagliato a 1). Queste difficoltà però si possono evitare nel caso particolare di un elettrone non soggetto a forze ($V = \mathbf{A} = 0$), nel qual caso l'equazione diviene

$$(256) \quad \Delta \psi - \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \frac{4\pi^2 m_0^2 c^2}{h^2} \psi = 0$$

e costituisce il punto di partenza della teoria di Dirac.

§ 47. — PRINCIPIO DELLA TEORIA DI DIRAC.

L'idea fondamentale che ha condotto alla teoria di Dirac è la seguente: postuliamo, in analogia alla teoria di Schrödinger e a quella di Pauli, che la densità di probabilità P si calcoli (come si era detto al § 25, p. II) mediante un certo numero N , per ora indeterminato, di funzioni $\psi_1, \psi_2, \dots, \psi_N$ (nella teoria di Schrödinger $N = 1$, in quella di Pauli $N = 2$: vedremo in seguito che nella teoria di Dirac bisogna prendere $N = 4$), attraverso la relazione

$$(257) \quad P = \psi_1 \psi_1^* + \psi_2 \psi_2^* + \dots + \psi_N \psi_N^*.$$

Se ciascuna delle ψ fosse vincolata solo da un'equazione differenziale del secondo ordine rispetto al tempo, bisognerebbe assegnare i valori iniziali di ψ e $\frac{\partial \psi}{\partial t}$ per determinare l'evoluzione ulteriore della ψ col tempo: è invece ovvio postulare che debba bastare l'assegnazione della ψ iniziale (in tutto lo spazio) per determinare la P in tutti gli istanti successivi, e quindi ammettere che le N funzioni ψ soddisfino un sistema di N equazioni differenziali del primo ordine rispetto al tempo. Siccome poi in ogni teoria relativistica la variabile t deve essere trattata alla stessa stregua delle coordinate spaziali x_1, x_2, x_3 , ne segue che tali equazioni dovranno essere del primo ordine anche rispetto a x_1, x_2, x_3 . Naturalmente, da equazioni del primo ordine si possono sempre ricavare, con operazioni di derivazione, delle equazioni del secondo ordine, conseguenze necessarie delle prime (ma non viceversa): richiederemo perciò che, nel caso particolare della

assenza di campo elettromagnetico, sia verificata per ciascuna delle ψ l'equazione relativistica (256) come conseguenza delle equazioni del primo ordine che ci accingiamo a stabilire (1). Ci limiteremo dapprima al caso di un elettrone non soggetto a forze.

La più semplice ipotesi che si possa fare sulle equazioni cercate è che siano lineari e a coefficienti costanti (2): come vedremo, si riesce con ciò a soddisfare tutte le condizioni richieste. Supposto poi che le equazioni siano risolubili rispetto alle derivate temporali, il sistema si potrà scrivere nella forma

$$(258) \quad \frac{1}{c} \frac{\partial \psi_\mu}{\partial t} + \sum_{\lambda, k} \alpha_{\mu\lambda}^k \frac{\partial \psi_\lambda}{\partial x_k} + \frac{2\pi i}{h} \sum_{\lambda} \beta_{\mu\lambda} \psi_\lambda = 0$$

(con $\mu = 1, 2, \dots, N$), dove i coefficienti $\alpha_{\mu\lambda}^k$, $\beta_{\mu\lambda}$ sono delle costanti da determinare (la ragione del fattore $2\pi i/h$ messo in evidenza nell'ultima somma apparirà in seguito); inoltre abbiamo introdotto la convenzione, che sarà mantenuta in seguito, di indicare con lettere greche λ, μ, \dots gli indici (assunti valori da 1 a N) che distinguono le varie ψ , e con lettere latine gli indici ($= 1, 2, 3$) che distinguono le tre coordinate spaziali (che sono indicate indifferentemente con x, y, z o x_1, x_2, x_3). La scrittura si semplifica notevolmente introducendo la notazione delle matrici, cioè indicando con α^k e β rispettivamente le 4 matrici a N righe e N colonne, il cui elemento della riga μ e della colonna λ è $\alpha_{\mu\lambda}^k$, $\beta_{\mu\lambda}$, e considerando ψ come il simbolo di una matrice a N righe e a una sola colonna, come si è fatto al § 45 per $N = 2$. Allora le N equazioni (258) si compendiano nella formula

$$(259) \quad \frac{1}{c} \frac{\partial \psi}{\partial t} + \sum_k \alpha^k \frac{\partial \psi}{\partial x_k} + \frac{2\pi i}{h} \beta \psi = 0,$$

la quale, introducendo gli operatori

$$p_k = \frac{h}{2\pi i} \frac{\partial}{\partial x_k}, \quad p_4 = -\frac{1}{c} \frac{h}{2\pi i} \frac{\partial}{\partial t}$$

(1) Analogamente, nella teoria elettromagnetica della luce, l'equazione delle onde (del 2° ordine) è conseguenza delle equazioni di Maxwell (del 1° ordine).

(2) Tale ipotesi si può del resto giustificare con la considerazione che nessun punto dello spazio-tempo deve risultare privilegiato.

si scrive più comodamente

$$(259') \quad [-p_4 + \beta + \sum_k \alpha^k p_k] \psi = 0.$$

Prima di procedere alla determinazione dei coefficienti della (258), cioè delle quattro matrici α^k , β , conviene procurarsi le espressioni della densità elettrica media ϱ e della densità media di corrente \mathbf{j} , che saranno la generalizzazione di quelle già trovate per il caso non relativistico.

§ 48. — DENSITÀ ELETTRICA E DENSITÀ DI CORRENTE.

La densità di probabilità P è data dalla (257), che, introducendo la matrice ψ e inoltre la matrice a una riga e N colonne

$$(260) \quad \bar{\psi} = |\psi_1^*, \psi_2^*, \dots, \psi_N^*|$$

(notazione conforme a quella del § 7) si può scrivere brevemente:

$$(261) \quad P = \bar{\psi} \psi,$$

come si verifica subito, applicando la regola del prodotto di due matrici. Conseguentemente, la densità elettrica media $\varrho = -eP$ sarà data da

$$(262) \quad \varrho = -e \bar{\psi} \psi.$$

Per trovare la densità media di corrente elettrica \mathbf{j} , osserviamo che essa dovrà soddisfare l'equazione « di continuità »

$$(263) \quad \frac{\partial \varrho}{\partial t} + \text{div } \mathbf{j} = 0,$$

la quale, introducendovi l'espressione di ϱ diviene

$$\text{div } \mathbf{j} = e \left(\frac{\partial \bar{\psi}}{\partial t} \psi + \bar{\psi} \frac{\partial \psi}{\partial t} \right).$$

In questa, sostituiamo le derivate di ψ e $\bar{\psi}$ con le loro espressioni ricavate dalla (259) e dalla sua coniugata, che è (designando al solito con $\bar{\alpha}^k$, $\bar{\beta}$ le matrici ottenute da α^k e β cambiando le linee con le colonne, e prendendo il coniugato

gato di ogni elemento) (1)

$$\frac{1}{e} \frac{\partial \bar{\psi}}{\partial t} + \sum_k \frac{\partial \bar{\psi}}{\partial x_k} \bar{\alpha}^k - \frac{2\pi i}{h} \bar{\psi} \bar{\beta} = 0.$$

Sostituendo queste derivate nella espressione di $\text{div } \mathbf{j}$ si ha

$$\text{div } \mathbf{j} = -ce \sum_k \left(\frac{\partial \bar{\psi}}{\partial x_k} \bar{\alpha}^k \psi + \bar{\psi} \alpha^k \frac{\partial \psi}{\partial x_k} \right) - \frac{2\pi i}{h} ce \bar{\psi} (\beta - \bar{\beta}) \psi.$$

Affinchè il secondo membro abbia effettivamente la forma di una divergenza, basta imporre alle matrici α^k , β le condizioni

$$\bar{\alpha}^k = \alpha^k, \quad \bar{\beta} = \beta,$$

le quali significano che dette matrici devono essere hermitiane. La formula diviene allora

$$\text{div } \mathbf{j} = -ce \sum_k \frac{\partial}{\partial x_k} \left(\bar{\psi} \alpha^k \psi \right)$$

e si possono quindi prendere come componenti di \mathbf{j} le espressioni

$$(264) \quad j_k = -ce \bar{\psi} \alpha^k \psi,$$

ovvero, in forma esplicita,

$$(264') \quad j_k = -ce \sum_{\lambda\mu} \psi_\lambda^* \alpha_{\lambda\mu}^k \psi_\mu.$$

§ 49. - LE EQUAZIONI DI DIRAC PER L'ELETTRONE NON SOGGETTO A FORZE.

Determineremo ora i coefficienti delle equazioni (258), ossia le matrici α^k , β , imponendo la condizione che dalle dette equazioni del primo ordine discenda, come conseguenza necessaria, l'equazione relativistica (256) per ciascuna delle N funzioni ψ_λ (ossia, che l'equazione (256) sia soddisfatta anche considerandovi la ψ come matrice). Applichiamo perciò alla (259) l'operatore differenziale seguente

(1) Si noti che, per conservare la validità della regola di moltiplicazione, la matrice ψ va sempre scritta a destra di α^k , β , e la $\bar{\psi}$ a sinistra.

(che è l'unico per il quale si eliminino i termini in $\frac{\partial}{\partial t}$):

$$-\frac{1}{e} \frac{\partial}{\partial t} + \sum_k \alpha^k \frac{\partial}{\partial x_k} + \frac{2\pi i}{h} \beta.$$

Tenendo presente che le matrici α^k , β sono permutabili con i simboli di derivazione, ma non sono da ritenersi, in generale, permutabili tra loro, otteniamo

$$-\frac{1}{e^2} \frac{\partial^2 \psi}{\partial t^2} + \sum_k \alpha^k \alpha^j \frac{\partial^2 \psi}{\partial x_k \partial x_j} + \frac{2\pi i}{h} \sum_k (\beta \alpha^k + \alpha^k \beta) \frac{\partial \psi}{\partial x_k} - \frac{4\pi^2}{h^2} \beta^2 \psi = 0.$$

Affinchè questa si identifichi con la (256), deve essere:

$$\begin{aligned} \alpha^k \alpha^j + \alpha^j \alpha^k &= 0 & (\text{per } j \neq k), \\ (\alpha^k)^2 &= \{1\}, & \beta^2 = m_0^2 c^2 \{1\}, \\ \beta \alpha^k + \alpha^k \beta &= 0. \end{aligned}$$

dove $\{1\}$ designa la matrice unità ad N righe ed N colonne. Introducendo, invece di β , la matrice α^4 definita da

$$(265) \quad \beta = m_0 c \alpha^4,$$

le formule precedenti si compendiano così

$$(266) \quad \begin{cases} \alpha^\lambda \alpha^\mu + \alpha^\mu \alpha^\lambda = 0 & \text{per } \lambda \neq \mu \\ (\alpha^\lambda)^2 = \{1\}, \end{cases}$$

dove gli indici λ e μ , secondo la convenzione già fatta, assumono i valori 1, 2, ... N . Si può anche scrivere, riunendo le (266) in una sola formula,

$$(266') \quad \alpha^\lambda \alpha^\mu + \alpha^\mu \alpha^\lambda = 2\delta_{\lambda\mu} \{1\}.$$

Si tratta dunque di trovare quattro matrici hermitiane che soddisfino queste condizioni. Si può dimostrare che non è possibile trovarle di rango inferiore a 4: è invece possibile (in infiniti modi) trovarle con $N \geq 4$ (1). Volendo dunque limitare quanto più è possibile la complicazione della teoria, assumeremo $N = 4$, cioè rappresenteremo lo stato di un elettrone con 4 funzioni $\psi_1, \psi_2, \psi_3, \psi_4$ di x, y, z, t , ossia con la matrice

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}.$$

(1) E precisamente per tutti i valori di N multipli di 4: tali soluzioni però si possono ricondurre a quella per $N = 4$.

Quanto alle quattro matrici α , si possono prendere le seguenti, che, come si verifica, sono hermitiane e soddisfano la (266')

$$(267) \quad \left\{ \begin{array}{l} \alpha^1 = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix} \\ \alpha^2 = \begin{vmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{vmatrix} \\ \alpha^3 = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{vmatrix} \\ \alpha^4 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} \end{array} \right.$$

L'equazione fondamentale (259'), che, introducendovi per β la (265), diviene

$$(268) \quad [\sum_k \alpha^k p_k - p_4 + m_0 c \alpha^4] \psi = 0,$$

equivale, quando si assumano per le α le espressioni (267), alle quattro equazioni seguenti (equazioni di Dirac per l'elettrone non soggetto a forze):

$$(269) \quad \left\{ \begin{array}{l} (-p_4 + m_0 c) \psi_1 + (p_1 - i p_2) \psi_4 + p_3 \psi_3 = 0, \\ (-p_4 + m_0 c) \psi_2 + (p_1 + i p_2) \psi_3 - p_3 \psi_4 = 0, \\ (-p_4 - m_0 c) \psi_3 + (p_1 - i p_2) \psi_2 + p_3 \psi_1 = 0, \\ (-p_4 - m_0 c) \psi_4 + (p_1 + i p_2) \psi_1 - p_3 \psi_2 = 0, \end{array} \right.$$

che si potrebbero anche scrivere, esplicitando gli operatori,

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \frac{2\pi i m_0 c}{h} \right) \psi_1 + \left(\frac{\partial}{\partial x_1} - i \frac{\partial}{\partial x_2} \right) \psi_4 + \frac{\partial \psi_3}{\partial x_3} = 0,$$

ecc.

Le condizioni (266') si possono soddisfare anche con altre infinite quaterne di matrici hermitiane: si hanno allora altrettante forme diverse delle equazioni di Dirac, cui corrispondono diverse quaterne di funzioni ψ . Si vede subito, anzitutto, che le (266') sono soddisfatte anche dalle matrici α'^k che si ricavano dalle α^k , definite dalle (267), con la trasformazione $\alpha'^k = S \alpha^k S^{-1}$, dove S è una matrice di trasformazione arbitraria (purchè « unitaria »). Allora la (268) diviene

$$[-p_4 + m_0 c S \alpha^4 S^{-1} + \sum_k S \alpha^k S^{-1} p_k] \psi' = 0$$

(indicando con ψ' la nuova quaterna di funzioni ψ): questa equazione è evidentemente soddisfatta prendendo $\psi' = S \psi$. Si noti che, applicando le (262) e (264), partendo da ψ' , si giunge agli stessi valori per ρ e \mathbf{j} come partendo da ψ : la soluzione considerata dunque non è fisicamente diversa dalla precedente. Si può dimostrare che, oltre a queste, non vi sono altre quaterne di matrici (con $N = 4$) soddisfacenti le condizioni volute.

§ 50. — ELETTRONE IN UN CAMPO ELETTROMAGNETICO.

Bisogna ora estendere le equazioni di Dirac al caso di un elettrone posto in un campo elettrico e magnetico, derivante da un potenziale scalare (elettrostatico) V e da un potenziale vettoriale \mathbf{A} . L'estensione si fonda sull'analogia con la meccanica ondulatoria non relativistica, nella quale l'equazione per una particella in un campo magnetico (v. form. (142'), § 31) differisce da quella in assenza di campo per avere, al posto degli operatori $\frac{h}{2\pi i} \frac{\partial}{\partial q_r}$, $-\frac{h}{2\pi i} \frac{\partial}{\partial t}$, gli operatori $\frac{h}{2\pi i} \frac{\partial}{\partial q_r} - \frac{e}{c} A_r$, $-\frac{h}{2\pi i} \frac{\partial}{\partial t} - U$. Tenendo conto che ora la carica dell'elettrone è indicata con $-e$, e che $U = -eV$, si è condotti a sostituire, nella (263), gli operatori p_k e p_4 rispettivamente con

$$(270) \quad \mathfrak{P}_k = \frac{h}{2\pi i} \frac{\partial}{\partial x_k} + \frac{e}{c} A_k, \quad \mathfrak{P}_4 = -\frac{h}{2\pi i} \frac{1}{c} \frac{\partial}{\partial t} + \frac{e}{c} V.$$

L'equazione di Dirac diviene allora

$$(271) \quad [\sum_k \alpha^k \mathfrak{P}_k - \mathfrak{P}_4 + m_0 c \alpha^4] \psi = 0$$

e, assumendo per le α le espressioni (267), si traduce nelle quattro equazioni

$$(272) \quad \left\{ \begin{array}{l} (-\mathfrak{P}_4 + m_0 c) \psi_1 + (\mathfrak{P}_1 - i \mathfrak{P}_2) \psi_4 + \mathfrak{P}_3 \psi_3 = 0 \\ (-\mathfrak{P}_4 + m_0 c) \psi_2 + (\mathfrak{P}_1 + i \mathfrak{P}_2) \psi_3 - \mathfrak{P}_3 \psi_4 = 0 \\ (-\mathfrak{P}_4 - m_0 c) \psi_3 + (\mathfrak{P}_1 - i \mathfrak{P}_2) \psi_2 + \mathfrak{P}_3 \psi_1 = 0 \\ (-\mathfrak{P}_4 - m_0 c) \psi_4 + (\mathfrak{P}_1 + i \mathfrak{P}_2) \psi_1 - \mathfrak{P}_3 \psi_2 = 0. \end{array} \right.$$

La giustificazione di questo postulato sta, naturalmente, nelle conseguenze che se ne deducono: in particolare nella

circostanza che i risultati di queste equazioni coincidono, in prima approssimazione, con quelli della teoria di Pauli e, trascurando lo spin, con quelli della teoria di Schrödinger, e soddisfano inoltre rigorosamente (al contrario delle dette teorie) al principio di relatività. Tutto ciò si vedrà nei §§ successivi. Vogliamo ora far notare che se la (271) si risolve rispetto alla derivata temporale contenuta in \mathfrak{P}_4 , essa si può mettere nella forma usuale:

$$(273) \quad \mathfrak{H}\psi = -\frac{\hbar}{2\pi i} \frac{\partial \psi}{\partial t},$$

ponendo

$$(274) \quad \mathfrak{H} = c \sum_k \alpha^k \mathfrak{P}_k - eV + m_0 c^2 \alpha^4.$$

Questo operatore dunque si può considerare come l'operatore hamiltoniano della teoria di Dirac. Si noti che dalla (273) si ricaverebbe, con lo stesso procedimento usato al § 28, la formula (118) per la derivata di un'osservabile: questa si può dunque ritenere valida anche nella teoria di Dirac, purchè per \mathfrak{H} si intenda l'operatore (274).

Infine, osserveremo che per uno stato stazionario di energia $(^1) W$ si ha

$$\psi_\lambda = u_\lambda e^{-\frac{2\pi i}{\hbar} W t}$$

e la matrice delle quattro u soddisfa (come anche, in questo caso, la ψ) all'equazione ottenuta dalla (271) sostituendovi P_4 con $\frac{1}{c}(W + eV)$.

§ 51. - LA TEORIA DI PAULI COME PRIMA APPROSSIMAZIONE.

Vogliamo ora mostrare come le formule di Dirac, ove si trascurino termini in $1/c^2$, si riducano a quelle della teoria di Pauli. È opportuno a questo scopo introdurre, in luogo

(¹) Indichiamo con W l'energia totale, compresa cioè quella intrinseca richiesta dalla teoria della relatività, e uguale a $m_0 c^2$: tra questa W e la E usata finora vi è dunque la relazione: $W = E + m_0 c^2$.

delle quattro ψ , due coppie di funzioni $\varphi_1, \varphi_2, \chi_1, \chi_2$, legate ad esse dalle seguenti relazioni

$$(275) \quad \begin{cases} \psi_1 = \varphi_1 e^{-\frac{2\pi i}{\hbar} m_0 c^2 t} & \psi_2 = \varphi_2 e^{-\frac{2\pi i}{\hbar} m_0 c^2 t} \\ \psi_3 = \chi_1 e^{-\frac{2\pi i}{\hbar} m_0 c^2 t} & \psi_4 = \chi_2 e^{-\frac{2\pi i}{\hbar} m_0 c^2 t} \end{cases}$$

Applicando gli operatori \mathfrak{P} definiti al § 50 si trovano subito le formule:

$$\mathfrak{P}_k \begin{vmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{vmatrix} = e^{-\frac{2\pi i}{\hbar} m_0 c^2 t} \mathfrak{P}_k \begin{vmatrix} \varphi_1 \\ \varphi_2 \\ \chi_1 \\ \chi_2 \end{vmatrix}, \quad \mathfrak{P}_4 \begin{vmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{vmatrix} = e^{-\frac{2\pi i}{\hbar} m_0 c^2 t} (\mathfrak{P}_4 + m_0 c) \begin{vmatrix} \varphi_1 \\ \varphi_2 \\ \chi_1 \\ \chi_2 \end{vmatrix}.$$

Tenendo conto dell'ultima di queste, si vede che nelle prime due delle equazioni (272) si elimina il termine $m_0 c$ della prima parentesi, mentre nelle altre due tale termine si raddoppia: le equazioni divengono infatti

$$(276) \quad \begin{cases} -\mathfrak{P}_4 \varphi_1 + (\mathfrak{P}_1 - i\mathfrak{P}_2) \chi_2 + \mathfrak{P}_3 \chi_1 = 0 \\ -\mathfrak{P}_4 \varphi_2 + (\mathfrak{P}_1 + i\mathfrak{P}_2) \chi_1 - \mathfrak{P}_3 \chi_2 = 0 \\ (-\mathfrak{P}_4 - 2m_0 c) \chi_1 + (\mathfrak{P}_1 - i\mathfrak{P}_2) \varphi_2 + \mathfrak{P}_3 \varphi_1 = 0 \\ (-\mathfrak{P}_4 - 2m_0 c) \chi_2 + (\mathfrak{P}_1 + i\mathfrak{P}_2) \varphi_1 - \mathfrak{P}_3 \varphi_2 = 0. \end{cases}$$

Introduciamo le matrici a due righe $\varphi = \begin{vmatrix} \varphi_1 \\ \varphi_2 \end{vmatrix}$, $\chi = \begin{vmatrix} \chi_1 \\ \chi_2 \end{vmatrix}$, e inoltre le tre matrici (a due righe e due colonne) $\sigma_x, \sigma_y, \sigma_z$, definite al § 45, e che ora per comodità indicheremo con $\sigma_1, \sigma_2, \sigma_3$: allora le quattro equazioni precedenti si possono riassumere nelle formule:

$$(277) \quad \mathfrak{P}_4 \varphi = \sum_k \mathfrak{P}_k \sigma_k \chi,$$

$$(278) \quad (\mathfrak{P}_4 + 2m_0 c) \chi = \sum_k \mathfrak{P}_k \sigma_k \varphi.$$

Nei casi ordinari (corrispondenti cioè nel modello classico a particelle dotate di velocità piccole rispetto a c , si da potersi usare la meccanica non relativistica), $\mathfrak{P}_4 \chi$ risulta trascurabile rispetto a $2m_0 c \chi$, cosicchè dalla (278) si ricava per χ l'espressione approssimata

$$(278') \quad \chi = \frac{1}{2m_0 c} \sum_k \mathfrak{P}_k \sigma_k \varphi,$$

da cui si vede che le χ sono ordinariamente piccole rispetto alle φ (ossia ψ_3 e ψ_4 sono piccole rispetto a ψ_1 e ψ_2). Sostituendo nella (277) questa espressione di χ si ha

$$(277') \quad \mathfrak{P}_4\varphi = \frac{1}{2m_0c} \sum_{kl} \mathfrak{P}_k \mathfrak{P}_l \sigma_k \sigma_l \varphi.$$

In questa sommatoria doppia, i sei termini in cui $k \neq l$ si possono riunire due a due nel modo seguente. Si considerino p. es. i due termini $(\mathfrak{P}_2 \mathfrak{P}_3 \sigma_2 \sigma_3 + \mathfrak{P}_3 \mathfrak{P}_2 \sigma_3 \sigma_2) \varphi$: in virtù delle (236) essi si possono scrivere $(\mathfrak{P}_2 \mathfrak{P}_3 - \mathfrak{P}_3 \mathfrak{P}_2) i \sigma_1 \varphi$. D'altra parte, dalle (270) si ricava subito che

$$(\mathfrak{P}_2 \mathfrak{P}_3 - \mathfrak{P}_3 \mathfrak{P}_2) \varphi = \frac{\hbar}{2\pi i} \frac{e}{c} \left(\frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3} \right) \varphi = \frac{\hbar}{2\pi i} \frac{e}{c} H_1 \varphi;$$

cosicchè i due termini in questione si riducono a $\frac{\hbar}{2\pi c} H_1 \sigma_1 \varphi$, e tutto il gruppo dei sei termini con $k \neq l$ dà:

$$\frac{\hbar e}{2\pi c} \mathbf{H} \times \boldsymbol{\sigma} \varphi.$$

Quanto ai tre termini in cui $k = l$, essi danno, tenendo presenti le (234),

$$\sum_k \mathfrak{P}_k^2 \varphi.$$

Così, in conclusione, la (277') diviene:

$$\mathfrak{P}_4 \varphi = \left[\frac{\hbar e}{4\pi m_0 c^2} \mathbf{H} \times \boldsymbol{\sigma} + \frac{1}{2m_0 c} \sum_k \mathfrak{P}_k^2 \right] \varphi,$$

o anche, esplicitando \mathfrak{P}_4 e ricordando l'espressione del magnetone di Bohr, $\mu_0 = \frac{\hbar e}{4\pi m c}$,

$$(279) \quad -\frac{\hbar}{2\pi i} \frac{\partial \varphi}{\partial t} = \left[\frac{1}{2m_0} \sum_k \mathfrak{P}_k^2 - eV + \mu_0 \mathbf{H} \times \boldsymbol{\sigma} \right] \varphi.$$

La parentesi quadra al secondo membro di questa equazione si identifica con l'operatore \mathfrak{S} della (244), e quindi questa equazione coincide con quella della teoria di Pauli.

Di qui si vede intanto come le formule di Dirac contengano implicitamente, almeno in prima approssimazione, l'esi-

stenza di un momento magnetico eguale a μ_0 e diretto in senso opposto allo spin: ciò si vedrà in altro modo nel § seguente.

§ 52. — MOMENTO MAGNETICO DELL'ELETTRONE.

Mostriamo ora che l'elettrone di Dirac si comporta (in prima approssimazione) come se avesse un momento magnetico μ_0 , non solo nei riguardi dei livelli energetici, il che si è visto al § precedente, ma anche in quanto genera intorno a sè un campo magnetico medio ⁽¹⁾ eguale al campo generato da un magnete di momento μ_0 . Ci limiteremo, per semplicità, al caso di un elettrone in quiete, immerso eventualmente in un campo elettrico (non in un campo magnetico di origine esterna) ⁽²⁾, e avente lo spin parallelo all'asse z , cosicchè, nell'approssimazione di Pauli, sia $\psi_2 = \varphi_2 = 0$.

Calcoliamo anzitutto le componenti della densità media di corrente \mathbf{j} (da cui dovremo poi ricavare il campo magnetico medio generato dall'elettrone) mediante la formula (264), introducendovi le φ e χ definite nel § precedente mediante le (275): avremo, usando le (267):

$$\begin{aligned} j_1 &= -ce(\varphi_1^* \chi_2 + \chi_2^* \varphi_1) \\ j_2 &= -ice(-\varphi_1^* \chi_2 + \chi_2^* \varphi_1) \\ j_3 &= -ce(\varphi_1^* \chi_1 + \chi_1^* \varphi_1). \end{aligned}$$

Introduciamo ora per le χ la loro espressione approssimata (278'), che diviene nel caso attuale, utilizzando le (241) e (241'),

$$\chi = \frac{\hbar}{4\pi i m_0 c} \left| \begin{array}{c} \frac{\partial \varphi_1}{\partial z} \\ \frac{\partial \varphi_1}{\partial x} + i \frac{\partial \varphi_1}{\partial y} \end{array} \right|.$$

⁽¹⁾ Nel senso spiegato al § 27.

⁽²⁾ Per una trattazione più generale, v. p. es. bibl. n. 26, cap. XIII e XIV. Con la teoria rigorosa si trova (v. BREIT, *Nature* 122 (1928) p. 649) che un elettrone vincolato a un nucleo di carica Ze ha, nello stato fondamentale, un momento magnetico $\frac{1}{3}(1 + 2\sqrt{1 - \alpha^2 Z^2})\mu_0$, dove α è la costante della struttura fina.

Ricordiamo inoltre che la φ_1 soddisfa l'ordinaria equazione di Schrödinger (v. pag. 395) e che questa, per uno stato stazionario di energia E , ammette sempre una soluzione della forma $\varphi_1 = ue^{-\frac{2\pi i}{h}Et}$, con la u reale (v. pag. 173): si ottiene allora

$$j_1 = -c\mu_0 \frac{\partial u^2}{\partial y}, \quad j_2 = c\mu_0 \frac{\partial u^2}{\partial x}, \quad j_3 = 0.$$

Queste formule, introducendo il vettore \mathbf{I} di componenti

$$(280) \quad I_1 = 0 \quad I_2 = 0 \quad I_3 = -\mu_0 u^2,$$

si riassumono nella formula vettoriale

$$(281) \quad \mathbf{j} = c \operatorname{rot} \mathbf{I}.$$

La densità di corrente non risulta dunque nulla, come avverrebbe nella teoria di Schrödinger per una ψ della forma qui considerata; è però nulla, come si verifica subito, la corrente totale attraverso un piano indefinito qualsiasi, il che si può interpretare, modellisticamente, dicendo che la « nuvola » di densità elettrica che equivale, in media, all'elettrone, non si sposta nel suo insieme, ma è agitata da correnti elettriche interne, di cui ora ricercheremo l'effetto magnetico.

Il potenziale vettore, da cui deriva il campo magnetico, si ottiene dalla densità di corrente \mathbf{j} con la nota formula dell'elettromagnetismo

$$(282) \quad \mathbf{A} = \frac{1}{c} \int_S \frac{\mathbf{j}}{r} dS.$$

D'altra parte, si può dimostrare ⁽¹⁾ che un corpo magnetizzato di cui \mathbf{I} sia l'intensità di magnetizzazione, genera un campo magnetico il cui potenziale vettore è dato da

$$\mathbf{A} = \int_S \frac{\operatorname{rot} \mathbf{I}}{r} dS.$$

Confrontando queste due formule e tenendo presente la (281), si conclude che l'effetto magnetico delle correnti in

questione è quello stesso che sarebbe prodotto da una magnetizzazione dello spazio rappresentata dal vettore \mathbf{I} , definito dalle (280). In altre parole, la « nuvola » di densità elettrica si comporta come se fosse magnetizzata, in direzione dell'asse z *negativo* (cioè in senso opposto al momento angolare) con densità $\mu_0 u^2$: il momento magnetico totale, che si ottiene integrando questa densità in tutto lo spazio, risulta μ_0 in virtù della normalizzazione di u .

Se si fosse supposto invece lo spin antiparallelo all'asse z , cioè $\varphi_1 = 0$ e $\varphi_2 = ue^{-\frac{2\pi i}{h}Et}$, si sarebbe giunti a una conclusione analoga, ma il momento magnetico sarebbe risultato diretto nel verso positivo dell'asse z , cioè anche in questo caso opposto allo spin.

Nel caso generale, si trova che la magnetizzazione equivalente è data, nella stessa approssimazione, da

$$(283) \quad \begin{cases} I_1 = -\mu_0(\psi_1^* \psi_2 + \psi_2^* \psi_1) \\ I_2 = i\mu_0(\psi_1^* \psi_2 - \psi_2^* \psi_1) \\ I_3 = -\mu_0(\psi_1^* \psi_1 - \psi_2^* \psi_2). \end{cases}$$

Si potrebbe pensare di rivelare sperimentalmente il momento magnetico dell'elettrone (come si fa per gli atomi) mediante un'esperienza del tipo di quella di Stern e Gerlach. Ciò non è possibile poichè, comunque si disponga l'esperienza, lo sparpagliamento del fascio di elettroni dovuto alla diffrazione è dello stesso ordine di grandezza delle deviazioni dovute alle forze magnetiche, che bisognerebbe rivelare ⁽¹⁾. Questa circostanza ha carattere generale: infatti, come è stato rilevato da BOHR, nell'ordine di approssimazione in cui la meccanica ondulatoria si identifica con la meccanica del punto, lo spin e il momento magnetico dell'elettrone (che figurano entrambi in termini contenenti \hbar) scompaiono, e quindi nessuna esperienza, in cui all'elettrone si possa applicare il modello corpuscolare, permette di mettere in evidenza questi momenti.

Non è invece impossibile immaginare delle esperienze, di tipo non corpuscolare, che permettano di rivelare l'esistenza dello spin: p. es., un fascio di elettroni diffratto da un cristallo in direzione presso a poco ortogonale al fascio incidente dovrebbe presumi-

⁽¹⁾ V. p. es. bibl. n. 26, p. 173.

⁽¹⁾ Per la dimostrazione v. N. F. MOTT, *Proc. Roy. Soc., A.* 124, p. 425, (1929).

bilmente risultare « polarizzato », vale a dire, se fatto ulteriormente diffrangere su un altro cristallo dovrebbe manifestare diversa capacità di diffrazione nei diversi piani passanti per la sua direzione di propagazione (nel modello vettoriale, questo si interpreterebbe pensando che dopo la prima diffrazione gli spin degli elettroni non sono più orientati a caso, ma con prevalenza in uno dei detti piani). Questa esperienza, eseguita in vario modo da diversi sperimentatori, ha dato risultati tuttora soggetti a discussione.

§ 53. — LO SPIN NELLE EQUAZIONI DI DIRAC.

Per mostrare che la (271) contiene implicitamente l'esistenza dello spin e la sua proprietà di quantizzazione (il che risulta anche dal § 51, ma solo in prima approssimazione), consideriamo il caso in cui la forza che agisce sull'elettrone ha momento nullo rispetto all'asse z , come avviene p. es. in un campo centrale. In tal caso abbiamo dimostrato (v. § 30) che, nella teoria di Schrödinger, si mantiene costante il momento dell'impulso rispetto all'asse z , cioè che

$$(284) \quad M_z = xp_y - yp_x$$

è un integrale primo, come nella meccanica ordinaria. Mostriamo ora che, invece, nella teoria di Dirac ciò non si verifica, e che in luogo di M_z si mantiene costante l'osservabile $N_z = M_z + S_z$, dove S_z è un'osservabile i cui autovalori sono $\pm \hbar/4\pi$: ciò significa che il momento totale dell'impulso rispetto all'asse z non è M_z ma N_z e che il termine S_z rappresenta un momento d'impulso intrinseco dell'elettrone, la cui proiezione sull'asse z è sempre $\pm \hbar/4\pi$.

Per metterci nelle condizioni anzidette, supponiamo che sia $\mathbf{A} = 0$ (assenza di campo magnetico), e V simmetrico intorno all'asse z , cioè, se si adottano coordinate polari, e φ è la « longitudine », che V sia indipendente da φ . Calcoliamo allora la derivata di M_z rispetto a t , servendoci della (118) del § 28; avremo (tralasciando la distinzione tra il simbolo di una osservabile e quello del suo operatore):

$$(285) \quad \dot{M}_z = \frac{2\pi i}{\hbar} (\mathcal{H}M_z - M_z\mathcal{H}).$$

Si osservi ora che nel caso attuale l'hamiltoniano si riduce (v. form. (274)) a

$$(286) \quad \mathcal{H} = c\sum_k \alpha^k p_k - eV + m_0 c^2 a^4,$$

e che M_z , come risulta dalla (284), è permutabile con α^1 , α^2 , α^3 , α^4 , con p_z e anche con V (perchè, come si è visto al § 30, in coordinate polari $M_z = \frac{\hbar}{2\pi i} \frac{\partial}{\partial \varphi}$, e V è indipendente da φ): quindi nella (285) resta solo il contributo dei primi due termini della sommatoria delle (286), cioè

$$\dot{M}_z = \frac{2\pi i}{\hbar} [c\alpha^1(p_x M_z - M_z p_x) + c\alpha^2(p_y M_z - M_z p_y)].$$

Ma dalle formule di permutazione e dalla (284) si ricava

$$p_x M_z - M_z p_x = \frac{\hbar}{2\pi i} p_y, \quad p_y M_z - M_z p_y = -\frac{\hbar}{2\pi i} p_x;$$

quindi

$$(287) \quad \dot{M}_z = c(\alpha^1 p_y - \alpha^2 p_x).$$

Come si vede, questa derivata non risulta identicamente nulla, il che significa che M_z non è un integrale primo.

Consideriamo ora l'osservabile S_z il cui operatore è

$$(288) \quad S_z = \frac{\hbar}{4\pi i} \alpha^1 \alpha^2,$$

e calcoliamone la derivata con la formula usata sopra: avremo

$$\dot{S}_z = \frac{1}{2} (\mathcal{H}\alpha^1\alpha^2 - \alpha^1\alpha^2\mathcal{H}).$$

Sostituiamo ora per \mathcal{H} la sua espressione (286), e osserviamo che $\alpha^1\alpha^2$ è permutabile con le p e con V , e che inoltre, come risulta immediatamente dalle (266), si ha

$$\begin{aligned} \alpha^1(\alpha^1\alpha^2) - (\alpha^1\alpha^2)\alpha^1 &= 2\alpha^2 \\ \alpha^2(\alpha^1\alpha^2) - (\alpha^1\alpha^2)\alpha^2 &= -2\alpha^1 \\ \alpha^3(\alpha^1\alpha^2) - (\alpha^1\alpha^2)\alpha^3 &= 0 \\ \alpha^4(\alpha^1\alpha^2) - (\alpha^1\alpha^2)\alpha^4 &= 0. \end{aligned}$$

Si trova così

$$S_z = c(\alpha^2 p_x - \alpha^1 p_y).$$

Confrontando con la (287), si vede che

$$\dot{M}_z + \dot{S}_z = 0$$

e cioè che l'osservabile $N_z = M_z + S_z$ è un integrale primo, come si era annunciato. Analogo ragionamento si potrebbe fare per le componenti y e z : se ne conclude che le proiezioni dello spin sui tre assi sono rappresentate dagli operatori:

$$S_x = \frac{\hbar}{4\pi i} \alpha^2 \alpha^3, \quad S_y = \frac{\hbar}{4\pi i} \alpha^3 \alpha^1, \quad S_z = \frac{\hbar}{4\pi i} \alpha^1 \alpha^2,$$

ovvero, adottando le matrici (267) ed eseguendo i prodotti:

$$(289) \quad S_x = \frac{\hbar}{4\pi} \begin{vmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{vmatrix}, \quad S_y = \frac{\hbar}{4\pi} \begin{vmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{vmatrix}, \quad S_z = \frac{\hbar}{4\pi} \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}.$$

Si noti anzitutto che queste matrici sono hermitiane, e quindi le osservabili che rappresentano sono reali ⁽¹⁾. Inoltre ciascuna di esse ha per autovalori $\pm \hbar/4\pi$: infatti si verifica subito che le tre matrici scritte sopra hanno per quadrato la matrice unità, e quindi che S_x^2, S_y^2, S_z^2 hanno il solo autovalore $(\hbar/4\pi)^2$. Si noti inoltre che S_x, S_y, S_z non sono permutabili tra loro ⁽²⁾: esse soddisfano invece le stesse relazioni di permutazione che nel § 30 si sono trovate per gli ordinari momenti angolari.

Agli operatori di spin si applicano, naturalmente, tutte le considerazioni esposte nel cap. II per gli operatori corrispondenti alle altre osservabili: unica differenza è che gli operatori di spin operano su funzioni di una variabile a soli quattro valori, che è l'indice λ delle quattro $\psi_\lambda(x, y, z)$ (poichè non coinvolgono le altre variabili x, y, z , trattasi di operatori « incompleti »): siffatti operatori non sono dunque altro che sostituzioni lineari su gruppi

⁽¹⁾ La particolare scelta adottata per le matrici α fa sì che S_z risulti diagonale: ciò significa che gli operatori sono rappresentati « nello schema S_z ». Con una trasformazione del tipo descritto alla fine di pag. 428 si potrebbe passare a una rappresentazione in cui invece risulta diagonale S_y , oppure S_x .

⁽²⁾ Quindi le tre componenti dello spin non sono osservabili compatibili: da ciò dipende il fatto che le proprietà dello spin non corrispondono in tutto a quelle di un ordinario girescopio.

di quattro numeri. In corrispondenza a ciò, ciascuno di essi ha solo 4 autovalori e 4 autofunzioni (anzi, i 4 autovalori si riducono a 2 autovalori doppi).

Per chiarire la cosa con un esempio, cerchiamo gli autovalori e le autofunzioni dell'operatore S_z definito da (288), ossia dall'ultima delle (289). Indichiamo con φ^μ , ($\mu = 1, 2, 3, 4$), una generica delle 4 autofunzioni (che si ridurrà a un gruppo di 4 numeri $\varphi_1^\mu, \varphi_2^\mu, \varphi_3^\mu, \varphi_4^\mu$), e con s^μ il corrispondente autovalore: dovrà aversi

$$S_z \varphi^\mu = s^\mu \varphi^\mu,$$

ossia, in forma esplicita:

$$\frac{\hbar}{4\pi} \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} \begin{vmatrix} \varphi_1^\mu \\ \varphi_2^\mu \\ \varphi_3^\mu \\ \varphi_4^\mu \end{vmatrix} = s^\mu \begin{vmatrix} \varphi_1^\mu \\ \varphi_2^\mu \\ \varphi_3^\mu \\ \varphi_4^\mu \end{vmatrix}, \quad \text{ovvero:} \quad \frac{\hbar}{4\pi} \begin{vmatrix} \varphi_1^\mu \\ -\varphi_2^\mu \\ \varphi_3^\mu \\ -\varphi_4^\mu \end{vmatrix} = s^\mu \begin{vmatrix} \varphi_1^\mu \\ \varphi_2^\mu \\ \varphi_3^\mu \\ \varphi_4^\mu \end{vmatrix}$$

il che equivale alle quattro equazioni:

$$\left(s^\mu - \frac{\hbar}{4\pi}\right)\varphi_1^\mu = 0, \quad \left(s^\mu + \frac{\hbar}{4\pi}\right)\varphi_2^\mu = 0, \quad \left(s^\mu - \frac{\hbar}{4\pi}\right)\varphi_3^\mu = 0, \quad \left(s^\mu + \frac{\hbar}{4\pi}\right)\varphi_4^\mu = 0.$$

Dovendosi escludere la soluzione $\varphi_1^\mu = \varphi_2^\mu = \varphi_3^\mu = \varphi_4^\mu = 0$, dovrà aversi o $s^\mu = -\frac{\hbar}{4\pi}$, o $s^\mu = +\frac{\hbar}{4\pi}$: e poichè ognuno di questi autovalori, come vedremo subito, è doppio, lo conteremo per due, e porremo

$$s^1 = s^3 = +\frac{\hbar}{4\pi}, \quad s^2 = s^4 = -\frac{\hbar}{4\pi}.$$

Nel primo caso ($\mu = 1, 3$) le equazioni danno $\varphi_2^\mu = \varphi_4^\mu = 0$, mentre φ_1^μ e φ_3^μ restano arbitrarie (salvo l'ortogonalità e la normalizzazione) e si possono prendere uguali a 1 e a 0 rispettivamente (per $\mu = 1$), o a 0, 1 (per $\mu = 3$): similmente, nel secondo caso ($\mu = 2, 4$), risulterà $\varphi_1^\mu = \varphi_3^\mu = 0$ e $\varphi_2^\mu, \varphi_4^\mu$ si potranno prendere uguali a (1, 0) e (0, 1) rispettivamente. Avremo così le quattro autofunzioni dell'operatore S_z

$$(290) \quad \varphi^1 = \begin{vmatrix} 1 \\ 0 \\ 0 \\ 0 \end{vmatrix}, \quad \varphi^2 = \begin{vmatrix} 0 \\ 1 \\ 0 \\ 0 \end{vmatrix}, \quad \varphi^3 = \begin{vmatrix} 0 \\ 0 \\ 1 \\ 0 \end{vmatrix}, \quad \varphi^4 = \begin{vmatrix} 0 \\ 0 \\ 0 \\ 1 \end{vmatrix}.$$

Serviamoci ora di queste per risolvere il problema: date le ψ , trovare la probabilità che una osservazione dello spin rispetto all'asse z dia il risultato $+\frac{\hbar}{4\pi}$ (o $-\frac{\hbar}{4\pi}$). Applichiamo letteral-

mente il procedimento del § 22 (osservazione), cioè sviluppiamo la matrice ψ mediante le autofunzioni φ , ponendo (la serie si riduce a quattro termini: indichiamo con c_λ i coefficienti indicati nel § 22 con ψ_r^y):

$$\psi = c_1\varphi^1 + c_2\varphi^2 + c_3\varphi^3 + c_4\varphi^4.$$

Si verifica subito usando le (290) che i quattro coefficienti c si identificano rispettivamente con $\psi_1, \psi_2, \psi_3, \psi_4$: applicando la (97') del § 22 troviamo che la probabilità corrispondente all'autovalore s^1 è $P_1 = \iiint |\psi_1|^2 dx dy dz$, e similmente quella corrispondente a s^3 è $P_3 = \iiint |\psi_3|^2 dx dy dz$; essendo i due autovalori coin-

cidenti, la probabilità totale dell'autovalore doppio $+\frac{\hbar}{4\pi}$ è

$$(291) \quad P_+ = \iiint (|\psi_1|^2 + |\psi_3|^2) dx dy dz$$

Similmente si troverebbe che la probabilità dell'autovalore doppio $-\frac{\hbar}{4\pi}$ è

$$(291') \quad P_- = \iiint (|\psi_2|^2 + |\psi_4|^2) dx dy dz.$$

Possiamo dunque dire che, delle quattro ψ , le due con indice dispari corrispondono (nel modello vettoriale) allo spin orientato parallelamente all'asse z , le due con indice pari allo spin antiparallelo all'asse z . Nell'approssimazione non relativistica, come si è visto, ψ_3 e ψ_4 si possono trascurare rispetto a ψ_1 e ψ_2 , e le formule (291), (291') si riducono a quelle che danno il significato delle due ψ di Pauli.

§ 54. - ONDE PIANE.

Come esempio di soluzione rigorosa delle equazioni di Dirac, studiamo il caso particolare di un elettrone non soggetto a forze, e avente un impulso \mathbf{p} esattamente determinato. Detta W l'energia totale (compresa quella intrinseca), cerchiamo di soddisfare le equazioni (269), con una soluzione analoga a quella relativa al caso analogo nella teoria di Schrödinger (v. § 44), cioè con onde piane di frequenza W/\hbar e di vettore di propagazione \mathbf{p}/\hbar : porremo perciò

$$(292) \quad \psi_\mu = a_\mu e^{\frac{2\pi i}{\hbar}(-Wt + \mathbf{p} \times \mathbf{r})} \quad (\mu = 1, 2, 3, 4).$$

Sostituendo nelle (269) si ottiene

$$(293) \quad \begin{cases} (-W/c + m_0c)a_1 + (p_1 - ip_2)a_4 + p_3a_3 = 0 \\ (-W/c + m_0c)a_2 + (p_1 + ip_2)a_3 - p_3a_4 = 0 \\ (-W/c - m_0c)a_3 + (p_1 - ip_2)a_2 + p_3a_1 = 0 \\ (-W/c - m_0c)a_4 + (p_1 + ip_2)a_1 - p_3a_2 = 0. \end{cases}$$

Queste quattro equazioni lineari omogenee nelle quattro costanti a_1, a_2, a_3, a_4 , hanno soluzione non nulla solo se

$$(294) \quad \begin{vmatrix} -W/c + m_0c & 0 & p_3 & p_1 - ip_2 \\ 0 & -W/c + m_0c & p_1 + ip_2 & -p_3 \\ p_3 & p_1 - ip_2 & -W/c - m_0c & 0 \\ p_1 + ip_2 & -p_3 & 0 & -W/c - m_0c \end{vmatrix} = 0.$$

Sviluppando il determinante si trova che esso è uguale a

$$(W^2/c^2 - m_0^2c^2 - p^2)^2,$$

dove è posto $p^2 = p_1^2 + p_2^2 + p_3^2$; perciò la condizione che esso sia nullo equivale a

$$(295) \quad \frac{W}{c} = \pm \sqrt{m_0^2c^2 + p^2}.$$

Questa coincide con la relazione data dalla meccanica relativistica tra energia W e impulso p : prendendo il segno $+$, sviluppando in serie il radicale e trascurando le potenze superiori di p , si può scrivere

$$(295') \quad W = m_0c^2 + \frac{p^2}{2m_0} + \dots$$

Questa espressione dell'energia si riduce, per piccole velocità, alla forza viva della meccanica ordinaria, aumentata dell'energia intrinseca m_0c^2 . Se invece si prende il segno $-$, si ha per W un valore prossimo a $-m_0c^2$, che non ha corrispondente nella meccanica ordinaria: su questi valori anomali (negativi) dell'energia cinetica torneremo nel § 60.

Passiamo ora a determinare le a : poichè, come si può verificare, la (295) rende nulli anche tutti i minori del terzo ordine del determinante, si possono fissare arbitrariamente due delle quattro a , e ricavare le altre due: le (293) hanno quindi due soluzioni linearmente indipendenti, che corrispondono alle due possibili orientazioni dello spin. Prenderemo come soluzioni fondamentali, nel caso $W > 0$:

$$(296) \quad \begin{cases} \text{I) } a_1 = A, & a_2 = 0, & a_3 = A \frac{cp_3}{W + m_0c^2}, & a_4 = A \frac{c(p_1 + ip_2)}{W + m_0c^2} \\ \text{II) } a_1 = 0, & a_2 = A, & a_3 = A \frac{c(p_1 - ip_2)}{W + m_0c^2}, & a_4 = A \frac{-cp_3}{W + m_0c^2}; \end{cases}$$

il modulo della costante A si determinerà con la condizione di normalizzazione (v. § 10, p. II).

Si osservi che i denominatori delle (296) sono dell'ordine di $2m_0c^2$, e quindi, se $p \ll m_0c$ come è d'ordinario, a_3 e a_4 risultano

piccole rispetto ad A , come del resto si è già visto in generale nel § 51. Nell'ordine di approssimazione in cui $|a_3|^2$ e $|a_4|^2$ si ritengono trascurabili rispetto a $|A|^2$, la soluzione I corrisponde al caso in cui lo spin è parallelo all'asse z , la II invece al caso in cui lo spin è antiparallelo all'asse z : la soluzione più generale, che si ottiene combinando linearmente I e II, rappresenta i casi in cui lo spin secondo l'asse z non ha un valore determinato.

Passando ora al caso di $W < 0$, conviene prendere come soluzioni fondamentali

$$(297) \quad \left\{ \begin{array}{l} \text{I) } a_1 = B \frac{cp_3}{W - m_0c^2}, \quad a_2 = B \frac{c(p_1 + ip_2)}{W - m_0c^2}, \quad a_3 = B, \quad a_4 = 0, \\ \text{II) } a_1 = B \frac{c(p_1 - ip_2)}{W - m_0c^2}, \quad a_2 = B \frac{-cp_3}{W - m_0c^2}, \quad a_3 = 0, \quad a_4 = B, \end{array} \right.$$

e la condizione di normalizzazione dà $|B|$.

In questo caso sono a_1 e a_2 piccole rispetto a B (supposto $p \ll m_0c$); e, ritenendole trascurabili, la soluzione I corrisponde allo spin parallelo all'asse z , la II allo spin antiparallelo.

§ 55. - ALTRA FORMA DELLE EQUAZIONI DI DIRAC.

Metteremo ora le equazioni di Dirac in un'altra forma che, trattando simmetricamente la variabile t e le x, y, z , si presta soprattutto per considerazioni di carattere relativistico.

Moltiplichiamo l'equazione di Dirac (271) per $-i\alpha^4$ (a sinistra), e poniamo ($k = 1, 2, 3$):

$$(298) \quad -i\alpha^4\alpha^k = \gamma^k, \quad \alpha^4 = \gamma^4,$$

$$(299) \quad P_k = \Pi_k, \quad iP_4 = \Pi_4.$$

Ciò permette di raccogliere in una unica sommatoria anche il termine con l'indice 4, e l'equazione si scrive così nella forma ⁽¹⁾:

$$(300) \quad \left[\sum_1^4 \gamma^\mu \Pi_\mu - im_0c \right] \psi = 0.$$

⁽¹⁾ Ricordiamo che, in tutto questo capitolo, si indicano con lettere greche gli indici che assumono i valori 1, 2, 3, 4, e con lettere latine quelli che assumono solo i valori 1, 2, 3.

Le matrici γ^μ , definite dalle (298), sono hermitiane al pari delle α^μ , e soddisfano anche esse (come si vede subito) le relazioni:

$$(301) \quad \gamma^\lambda \gamma^\mu + \gamma^\mu \gamma^\lambda = 2\delta_{\lambda\mu} \quad (\lambda, \mu = 1, 2, 3, 4).$$

Gli operatori Π_μ si possono scrivere in forma più simmetrica introducendo in luogo di t la variabile

$$x_4 = ict,$$

e considerando A_1, A_2, A_3, iV come le quattro componenti di un « quadrivettore » Φ nello spazio delle variabili x_1, x_2, x_3, x_4 (spazio di Minkowsky): è noto infatti dalla teoria della relatività che in una trasformazione di Lorentz queste quattro quantità si trasformano appunto come le componenti di un quadrivettore invariante. Porremo dunque

$$(302) \quad A_k = \Phi_k \quad iV = \Phi_4$$

e allora le (299) si raccolgono nell'unica formola

$$(299') \quad \Pi_\mu = \frac{\hbar}{2\pi i} \frac{\partial}{\partial x_\mu} + \frac{e}{c} \Phi_\mu.$$

Occupiamoci ora delle espressioni della densità elettrica (media) ρ e della densità di corrente (media) \mathbf{j} ; conviene porre:

$$(303) \quad i\overline{\psi} \gamma^4 = \psi^+,$$

da cui (moltiplicando a destra per $-i\gamma^4$),

$$(303') \quad \overline{\psi} = -i\psi^+ \gamma^4.$$

Le espressioni (262) e (264) divengono allora rispettivamente

$$(304) \quad \rho = ie\psi^+ \gamma^4 \psi,$$

$$(305) \quad j_k = -ce\psi^+ \gamma^k \psi,$$

il che mostra che le quattro quantità,

$$(306) \quad J_k = j_k, \quad J_4 = ieo,$$

(che, come si vedrà, costituiscono anch'esse le componenti di un quadrivettore invariante, cioè la « tetracorrente ») si esprimono in modo uniforme mediante ψ e ψ^+ , poichè le

(304) e (305) si possono compendiare nella formula:

$$(307) \quad J_{\mu} = -ce\psi + \gamma^{\mu}\psi.$$

Si noti che l'equazione di continuità (263) assume, con le notazioni (306), la forma

$$(308) \quad \sum_{\mu} \frac{\partial J_{\mu}}{\partial x_{\mu}} = 0$$

che è relativisticamente invariante.

Può essere utile scrivere l'equazione cui soddisfa la ψ^{+} . Dall'equazione di Dirac (271) si ha, prendendone la coniugata (e notando che, se a è una matrice hermitiana, la coniugata di $a\psi$ è $\bar{\psi}a$):

$$(309) \quad \sum_k P_k^* \bar{\psi} \alpha^k - P_4^* \bar{\psi} + m_0 c \bar{\psi} \alpha^4 = 0.$$

Introducendovi la (303') e tenendo conto delle (298) si ha

$$\sum P_k^* \psi^{+\gamma^k} + iP_4^* \psi^{+\gamma^4} - im_0 c \psi^{+} = 0.$$

Per raccogliere i primi due termini in un'unica sommatoria, conviene definire gli operatori

$$(310) \quad \Pi_k^{\dagger} = \Pi_k^* = P_k^*, \quad \Pi_4^{\dagger} = -\Pi_4^* = iP_4^*,$$

cioè

$$(310') \quad \Pi_{\mu}^{\dagger} = -\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_{\mu}} + \frac{e}{c} \Phi_{\mu}.$$

Così l'equazione diviene

$$(311) \quad \sum_{\mu} \Pi_{\mu}^{\dagger} \psi^{+\gamma^{\mu}} - im_0 c \psi^{+} = 0.$$

§ 56. - INVARIANZA RELATIVISTICA DELLE EQUAZIONI DI DIRAC.

Questa invarianza va intesa nel modo seguente. Si consideri un secondo sistema di riferimento in moto traslatorio rettilineo e uniforme rispetto al primo, cioè si passi con una trasformazione di Lorentz dalle variabili $x_1, x_2, x_3, x_4 = ict$ a $x'_1, x'_2, x'_3, x'_4 = ict'$: dimostreremo che nel nuovo sistema di riferimento vale l'equazione di Dirac (300) nella stessa forma che nel primo, cioè si può scrivere:

$$(312) \quad \left[\sum_{\mu} \gamma^{\mu} \Pi'_{\mu} - im_0 c \right] \psi' = 0.$$

Ciò significa non già che la ψ' che si ricava da questa sia uguale alla ψ , ma che le quantità ρ', \mathbf{j}' (densità elettrica e densità di corrente nel secondo sistema di riferimento) che si ricavano da ψ' con le formule analoghe alle (262), (264), sono uguali a quelle che si ricavano da P e \mathbf{j} trasformandole con le leggi con cui, nella teoria elettromagnetica ordinaria, si trasformano la densità elettrica e la densità di corrente, vale a dire in modo che le quattro quantità $J_k = j^k, J_4 = ic\rho$ costituiscano le componenti di un quadrivettore invariante nello spazio di Minkowsky.

Per dimostrare quanto abbiamo ora enunciato, consideriamo la trasformazione di Lorentz più generale, ossia la più generale trasformazione ortogonale nello spazio di Minkowsky, espressa dalle formule:

$$(313) \quad x'_{\mu} = \sum_{\nu} a_{\mu\nu} x_{\nu},$$

dove i coefficienti $a_{\mu\nu}$ sono vincolati dalle relazioni

$$(314) \quad \sum_{\mu} a_{\mu\sigma} a_{\mu\tau} = \delta_{\sigma\tau}.$$

Si tenga poi presente che questi coefficienti sono reali, salvo quelli che contengono l'indice 4 una volta sola, che sono immaginari puri (essendo tali x_4 e x'_4). Gli operatori Π_{μ} si trasformano come le componenti di un quadrivettore (come si riconosce subito dalle (299')) cioè secondo le formule

$$(315) \quad \Pi'_{\mu} = \sum_{\nu} a_{\mu\nu} \Pi_{\nu}.$$

Perciò la (312) si può scrivere:

$$(316) \quad \left[\sum_{\mu\nu} a_{\mu\nu} \gamma^{\mu} \Pi_{\nu} - im_0 c \right] \psi' = 0.$$

Tentiamo ora di soddisfare questa equazione ammettendo, salvo verifica *a posteriori*, che le ψ si trasformino linearmente, analogamente al comportamento delle componenti del campo elettrico e magnetico in una trasformazione di Lorentz: porremo dunque

$$(317) \quad \psi'_{\lambda} = \sum_{\mu} S_{\lambda\mu} \psi_{\mu},$$

lasciando i coefficienti $S_{\lambda\mu}$ per ora indeterminati: più bre-

vemente scriveremo, indicando con S la matrice i cui elementi sono $S_{\lambda\mu}$,

$$(318) \quad \psi' = S\psi.$$

Sostituendo nella (316), e moltiplicando a sinistra per S^{-1} , si ha (si noti che S è permutabile con le Π ma non con le γ):

$$[\sum_{\mu\nu} a_{\mu\nu} S^{-1} \gamma^\mu S \Pi_\nu - im_0 c] \psi = 0.$$

Questa equazione sarà soddisfatta se la matrice S è tale che sia

$$(319) \quad \sum_{\mu} a_{\mu\nu} S^{-1} \gamma^\mu S = \gamma^\nu,$$

poichè in tal caso l'equazione si identifica con la (300), che per ipotesi è soddisfatta da ψ : si tratta dunque di dimostrare che per ogni trasformazione di Lorentz esiste una matrice S che soddisfa la (319). Per dimostrarlo, consideriamo dapprima una trasformazione di Lorentz «infinitesima», cioè poniamo:

$$(320) \quad a_{\mu\nu} = \delta_{\mu\nu} + \varepsilon_{\mu\nu},$$

dove le $\varepsilon_{\mu\nu}$ sono infinitesime del 1° ordine: la (314) si traduce nella condizione di emisimmetria:

$$(321) \quad \varepsilon_{\rho\sigma} + \varepsilon_{\sigma\rho} = 0.$$

È ovvio in questo caso ricercare per la S una forma infinitamente vicina alla matrice unità, cioè porre $S_{\mu\nu} = \delta_{\mu\nu} + T_{\mu\nu}$, dove le $T_{\mu\nu}$ sono infinitesime, ossia:

$$(322) \quad S = \{ 1 \} + T.$$

Sarà allora (trascurando gli infinitesimi di ordine superiore)

$$(323) \quad S^{-1} = \{ 1 \} - T$$

e la (319) assume la forma

$$(324) \quad \sum_{\mu} \varepsilon_{\mu\nu} \gamma^\mu = T \gamma^\nu - \gamma^\nu T.$$

Ora, si verifica facilmente, tenendo presenti le (301), che questa è soddisfatta prendendo

$$T = \frac{1}{4} \sum_{\mu\lambda} \varepsilon_{\mu\lambda} \gamma^\mu \gamma^\lambda,$$

ossia prendendo

$$(325) \quad S = \{ 1 \} + \frac{1}{4} \sum_{\mu\lambda} \varepsilon_{\mu\lambda} \gamma^\mu \gamma^\lambda.$$

Dimostrata così l'esistenza della matrice S per una trasformazione infinitesima ne segue subito (poichè le trasformazioni di Lorentz, come è noto, formano gruppo) che tale matrice esiste per una trasformazione di Lorentz qualunque, e si può costruire come prodotto di infinite matrici del tipo (325).

Dimostriamo ora che, costruita la S in tal modo, la ψ^+ si trasforma con la legge:

$$(326) \quad \psi^{+'} = \psi^+ S^{-1}.$$

A tal uopo osserviamo che $\psi^{+'}$ è definita dalla formula analoga alla (303), cioè da

$$(327) \quad \psi^{+'} = \widetilde{\psi}' \gamma^4,$$

formula che si può trasformare, osservando che dalla (318) si ha (introducendo la matrice \widetilde{S} definita, come al solito, da $\widetilde{S}_{\mu\lambda} = S_{\lambda\mu}^*$):

$$(328) \quad \widetilde{\psi}' = \widetilde{\psi} \widetilde{S};$$

sostituendo questa nella (327), e tenendo presente la (303'), la formula che definisce $\psi^{+'}$ diviene:

$$(329) \quad \psi^{+'} = \psi^+ \gamma^4 \widetilde{S} \gamma^4.$$

La (326) resterà dunque dimostrata se faremo vedere che la S relativa a ogni trasformazione di Lorentz gode la proprietà:

$$(330) \quad \gamma^4 \widetilde{S} \gamma^4 = S^{-1},$$

e basterà dimostrare questa formula per una S della forma (325), poichè si verifica subito che, se essa vale per due matrici S_1, S_2 , vale anche per il loro prodotto. Per dimostrare la (330), osserviamo che da (325) si ricava:

$$\begin{aligned} \widetilde{S} &= \{ 1 \} + \frac{1}{4} \sum_{\mu\lambda} \varepsilon_{\mu\lambda}^* \gamma^\mu \gamma^\lambda = \\ &= \{ 1 \} + \frac{1}{4} \sum_{hk} \varepsilon_{hk}^* \widetilde{\gamma}^h \widetilde{\gamma}^k + \frac{1}{4} \sum_k (\varepsilon_{4k}^* \widetilde{\gamma}^4 \gamma^k + \varepsilon_{k4}^* \gamma^k \widetilde{\gamma}^4). \end{aligned}$$

Di qui, tenendo presente che, per matrici hermitiane come sono le γ , si ha $\overline{AB} = \overline{B} \overline{A} = BA$, e che inoltre $\varepsilon_{hk}^* = \varepsilon_{hk}$, $\varepsilon_{4k}^* = -\varepsilon_{4k}$, $\varepsilon_{k4}^* = -\varepsilon_{k4}$, si ottiene:

$$\overline{S} = \{ 1 \} - \frac{1}{4} \sum_{hk} \varepsilon_{hk} \gamma^k \gamma^h - \frac{1}{4} \sum_k \varepsilon_{4k} (\gamma^k \gamma^4 - \gamma^4 \gamma^k).$$

Moltiplicando a destra e a sinistra ambo i membri per γ^4 , e ricordando le (301) si trova:

$$\begin{aligned} \gamma^4 \overline{S} \gamma^4 &= \{ 1 \} - \frac{1}{4} \sum_{hk} \varepsilon_{kh} \gamma^k \gamma^h - \frac{1}{4} \sum_k \varepsilon_{4k} (\gamma^4 \gamma^k - \gamma^k \gamma^4) = \\ &= \{ 1 \} - \frac{1}{4} \sum_{\mu\lambda} \varepsilon_{\mu\lambda} \gamma^\mu \gamma^\lambda = \{ 1 \} - T = S^{-1}, \end{aligned}$$

e così resta dimostrata la (330).

Trovate così le leggi di trasformazione ⁽¹⁾ di ψ e di ψ^+ , si trova subito quella delle componenti della quadricorrente, cioè delle J_μ definite dalla (307): si ha infatti

$$J'_\mu = -c e \psi^+ S^{-1} \gamma^\mu S \psi^+$$

e poichè dalla (319) si ricava

$$S^{-1} \gamma^\mu S = \sum_\nu a_{\mu\nu} \gamma^\nu,$$

si ha

$$(331) \quad J'_\mu = \sum_\nu a_{\mu\nu} J_\nu,$$

il che prova che le J_μ si trasformano come le componenti di un quadrivettore invariante, come volevasi dimostrare.

§ 57. - INVARIANZA ELETTROMAGNETICA DELLE EQUAZIONI DI DIRAC.

Nelle equazioni di Dirac il campo elettromagnetico è rappresentato dai potenziali V, A . Ora, è ben noto che que-

⁽¹⁾ Un gruppo di quattro grandezze che in una trasformazione di Lorentz si trasformano come $\psi_1, \psi_2, \psi_3, \psi_4$ dicesi uno *spinore* (concetto analogo a quello di *tensore*). Tale concetto dal VAN DER WAERDEN e da altri è stato posto a base di una trattazione sistematica analoga al calcolo tensoriale (v. *Göttinger Nachr.*, 1929, p. 100). Le componenti di uno spinore non rappresentano grandezze fisiche direttamente osservabili (il che sarebbe contrario al principio di Relatività) ma hanno interesse in quanto determinano indirettamente i valori di grandezze osservabili.

sti non sono fisicamente determinati in modo unico, perchè allo stesso campo elettromagnetico si possono attribuire (come si verifica subito) anche i potenziali

$$(332) \quad V' = V - \frac{1}{c} \frac{\partial \varphi}{\partial t}, \quad A' = A + \text{grad } \varphi,$$

dove $\varphi(x, y, z, t)$ è una funzione arbitraria. Se si utilizza il quadripotenziale Φ definito dalla (302), si può dire che alle

Φ_μ si possono sostituire le $\Phi'_\mu = \Phi_\mu + \frac{\partial \varphi}{\partial x_\mu}$. Ora, poichè

tale sostituzione non corrisponde a nessun cambiamento fisico, bisogna che le equazioni di Dirac siano insensibili a questo mutamento, ossia, più precisamente, che per esso non si alterino i risultati osservabili da esse forniti, come p. es. i valori medi della densità elettrica e della densità di corrente, cioè le J_μ ; (la ψ invece può cambiare, poichè non ha significato fisico diretto). Questa proprietà, che ora verificheremo, dicesi «invarianza elettromagnetica» (dai tedeschi è detta *Eichinvarianz*).

Scriviamo l'equazione di Dirac nella forma (300), sostituendovi le Φ_μ con le Φ'_μ , e la ψ con una nuova funzione ψ' ; avremo, esplicitando gli operatori mediante la (299')

$$\left[\sum_\mu \gamma^\mu \left(\frac{\hbar}{2\pi i} \frac{\partial}{\partial x_\mu} + \frac{e}{c} \Phi_\mu + \frac{e}{c} \frac{\partial \varphi}{\partial x_\mu} \right) - im_0 c \right] \psi' = 0.$$

Ora, questa può essere soddisfatta prendendo

$$(333) \quad \psi' = \psi e^{-\frac{2\pi i}{\hbar} \frac{e}{c} \varphi},$$

poichè con questa sostituzione essa si riduce alla (300) come si verifica immediatamente. A questa ψ' corrisponde, a norma della (303),

$$\psi'^+ = \overline{\psi'} \gamma^4 = \psi^+ e^{\frac{\hbar}{2\pi i} \frac{e}{c} \varphi},$$

e formando, con queste espressioni di ψ' e ψ'^+ , le J'_μ mediante la formula (307), si vede che gli esponenti si elidono e la funzione arbitraria φ scompare.

Più in generale: secondo il principio generale della meccanica quantistica (§ 22) la probabilità di un determinato

risultato G_n nella misura di un'osservabile G , quando il sistema è nello stato ψ , è data, come si è visto, da $|\psi \times \varphi_n|^2$, dove φ_n non dipende dalla ψ ma solo dall'osservabile G e dal suo autovalore G_n : sostituendo ψ con ψ' il prodotto $\psi \times \varphi_n$ si altera solo per un fattore di modulo 1, e quindi il quadrato del suo modulo non varia. La probabilità cercata risulta dunque la stessa.

§ 58. — ELETTRONE IN UN CAMPO CENTRALE.

Come applicazione della teoria di Dirac tratteremo ora il problema di un elettrone soggetto solo a una forza elettrostatica centrale, di legge qualunque: le formule così trovate ci serviranno poi (specializzando la legge di questa forza col supporla coulombiana) a trattare, nel § seguente, il problema dei sistemi idrogenoidi. Si noterà che i risultati ottenuti corrispondono qualitativamente a quelli che si ottengono con la teoria modellistica dello spin, cioè usando il modello di Bohr e Sommerfeld, completato con l'introduzione del momento magnetico dell'elettrone considerato come un magnete ordinario, soggetto però alla regola di quantizzazione spaziale (v. § 62, p. II): però la teoria di Dirac, oltre ad essere razionalmente più coerente, fornisce risultati lievemente diversi e in migliore accordo con l'esperienza.

Indicando al solito con $U = -eV(r)$ l'energia potenziale del campo centrale in cui si trova l'elettrone, consideriamo uno stato stazionario di energia W : le equazioni di Dirac (272) assumono la forma

$$(334) \quad \begin{cases} \frac{2\pi i - W + U + m_0 c^2}{h} \psi_1 + \left(\frac{\partial}{\partial x_1} - i \frac{\partial}{\partial x_2} \right) \psi_4 + \frac{\partial}{\partial x_3} \psi_3 = 0 \\ \frac{2\pi i - W + U + m_0 c^2}{h} \psi_2 + \left(\frac{\partial}{\partial x_1} + i \frac{\partial}{\partial x_2} \right) \psi_3 - \frac{\partial}{\partial x_3} \psi_4 = 0 \\ \frac{2\pi i - W + U - m_0 c^2}{h} \psi_3 + \left(\frac{\partial}{\partial x_1} - i \frac{\partial}{\partial x_2} \right) \psi_2 + \frac{\partial}{\partial x_3} \psi_1 = 0 \\ \frac{2\pi i - W + U - m_0 c^2}{h} \psi_4 + \left(\frac{\partial}{\partial x_1} + i \frac{\partial}{\partial x_2} \right) \psi_1 - \frac{\partial}{\partial x_3} \psi_2 = 0. \end{cases}$$

Poichè, come si è visto al § 46, p. II, l'equazione di Schrödinger corrispondente a questo problema ammette soluzioni

del tipo $\psi = Y_l^m f(r)$, dove Y_l^m è una funzione sferica di superficie (i cui indici l e m rappresentano rispettivamente il quanto azimutale e quello magnetico dell'elettrone), tentiamo di soddisfare analogamente le (334) prendendo ciascuna delle quattro ψ della forma:

$$(335) \quad \psi_\mu = Z(\theta, \varphi) f_\mu(r),$$

dove la funzione sferica è stata indicata ⁽¹⁾ con Z anzichè con Y_l^m , perchè, per semplificare le formule, conviene qui (seguendo DARWIN) normalizzarla in modo diverso da quello seguito al § 46, p. II: precisamente prenderemo

$$(336) \quad Z_l^m = \sqrt{\frac{4\pi}{2l+1}} (l+m)! (l-m)! Y_{lm}.$$

Si può allora verificare facilmente, utilizzando le formule di passaggio dalle coordinate cartesiane x_1, x_2, x_3 alle coordinate polari r, θ, φ , che per le derivate di espressioni del tipo $f Z_l^m$ (dove $f = f(r)$ è una qualunque funzione di r) valgono le formule seguenti:

$$(337) \quad \left\{ \begin{aligned} \left(\frac{\partial}{\partial x_1} + i \frac{\partial}{\partial x_2} \right) f Z_l^m &= \frac{1}{2l+1} \left[\left(\frac{df}{dr} - \frac{l}{r} f \right) Z_{l+1}^{m+1} - (l-m)(l-m-1) \left(\frac{df}{dr} + \frac{l+1}{r} f \right) Z_{l-1}^{m+1} \right] \\ \left(\frac{\partial}{\partial x_1} - i \frac{\partial}{\partial x_2} \right) f Z_l^m &= \frac{1}{2l+1} \left[- \left(\frac{df}{dr} - \frac{l}{r} f \right) Z_{l+1}^{m-1} + (l+m)(l+m-1) \left(\frac{df}{dr} + \frac{l+1}{r} f \right) Z_{l-1}^{m-1} \right] \\ \frac{\partial}{\partial x_3} f Z_l^m &= \frac{1}{2l+1} \left[\left(\frac{df}{dr} - \frac{l}{r} f \right) Z_{l+1}^m + (l+m)(l-m) \left(\frac{df}{dr} + \frac{l+1}{r} f \right) Z_{l-1}^m \right]. \end{aligned} \right.$$

Con l'aiuto di queste formule si verifica senza difficoltà che, se si prendono le quattro ψ della forma:

$$(338) \quad \begin{cases} \psi_1 = ia_1 F_+(r) Z_{l+1}^m & \psi_2 = ia_2 G_+(r) Z_{l+1}^{m+1} \\ \psi_3 = a_3 G_+(r) Z_l^m & \psi_4 = a_4 G_+(r) Z_l^{m+1} \end{cases}$$

(dove le a sono costanti, e F_+, G_+ sono due funzioni di r , per ora indeterminate) e si sostituiscono nelle (334), basta poi imporre alle costanti a le condizioni

$$\frac{a_2}{a_1} = 1 \quad \frac{a_3}{a_4} = -\frac{l+m+1}{l-m}$$

⁽¹⁾ Tralasciamo, per semplicità di notazione, di scrivere gli indici della funzione sferica Z : si noti che essi saranno in genere diversi per le quattro ψ , come si vedrà più avanti.

perchè le funzioni sferiche si eliminino dalle equazioni, e queste si riducano a due sole (poichè la prima e la seconda diventano equivalenti, e così la terza e la quarta), e precisamente alle seguenti

$$\begin{aligned} -\frac{2\pi}{h} \frac{-W+U+m_0c^2}{c} a_1 F_+ - \frac{a_4}{l-m} \left(\frac{dG_+}{dr} - \frac{l}{r} G_+ \right) &= 0 \\ -\frac{2\pi}{h} \frac{-W+U-m_0c^2}{c} a_1 G_+ + a_1 \left(\frac{dF_+}{dr} + \frac{l+2}{r} F_+ \right) &= 0. \end{aligned}$$

Le costanti a_1 e a_4 restano arbitrarie, e le prenderemo uguali rispettivamente a 1 e a $-(l-m)$, cosicchè sarà:

$$(339) \quad a_1 = 1, \quad a_2 = 1, \quad a_3 = l+m+1, \quad a_4 = -(l-m)$$

e le precedenti equazioni divengono

$$(340) \quad \begin{cases} -\frac{2\pi}{h} \frac{-W+U+m_0c^2}{c} F_+ + \frac{dG_+}{dr} - \frac{l}{r} G_+ = 0 \\ \frac{2\pi}{h} \frac{-W+U-m_0c^2}{c} G_+ + \frac{dF_+}{dr} + \frac{l+2}{r} F_+ = 0, \end{cases}$$

Si noti che una soluzione della forma qui considerata può esistere solo per m compreso tra $-(l+1)$ ed l (estremi inclusi), altrimenti vi figurerebbero dei simboli di funzioni sferiche con l'indice superiore più grande (in valore assoluto) dell'inferiore, simboli a cui non abbiamo dato significato. Quanto a l , esso può assumere i valori 0, 1, 2,...

Un secondo modo di soddisfare le (334) consiste nel prendere le ψ della forma

$$(341) \quad \begin{cases} \psi_1 = ib_1 F_-(r) Z_l^{m-1} & \psi_2 = ib_2 F_-(r) Z_l^{m+1} \\ \psi_3 = b_3 G_-(r) Z_l^m & \psi_4 = b_4 G_-(r) Z_l^{m+1}. \end{cases}$$

Sostituendo nelle (334) e procedendo come poc'anzi, si trova che, se si prende

$$(342) \quad b_1 = (l+m), \quad b_2 = -(l-m-1), \quad b_3 = 1, \quad b_4 = 1,$$

esse si riducono alle due seguenti equazioni nelle funzioni $F(r)$, $G(r)$:

$$(343) \quad \begin{cases} -\frac{2\pi}{h} \frac{-W+U+m_0c^2}{c} F_- + \frac{dG_-}{dr} + \frac{l+1}{r} G_- = 0 \\ \frac{2\pi}{h} \frac{-W+U-m_0c^2}{c} G_- + \frac{dF_-}{dr} - \frac{l-1}{r} F_- = 0. \end{cases}$$

Si noti che queste equazioni differiscono dalle (340) cui soddisfano F_+ e G_+ semplicemente per la sostituzione di l con $-(l+1)$. Una soluzione di questa forma può esistere solo per m compresa fra $-l$ ed $l-1$, estremi inclusi: inoltre, essa manca se $l=0$, poichè non esistono funzioni sferiche con l'indice inferiore negativo.

Per interpretare i due tipi di soluzione così trovati, ricordiamo (v. § 53) che il momento angolare totale rispetto all'asse z corrisponde all'operatore:

$$\mathfrak{M}_z = \mathfrak{M}_z + \mathfrak{S}_z = \frac{\hbar}{2\pi i} \frac{\partial}{\partial \varphi} + \frac{\hbar}{4\pi} \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}.$$

Applichiamo ora questo operatore alle ψ della forma (338) o della forma (341), osservando che

$$\frac{\partial}{\partial \varphi} Z_l^m = im Z_l^m;$$

troviamo, sia nell'un caso che nell'altro

$$\mathfrak{M}_z \psi = \left(m + \frac{1}{2} \right) \frac{\hbar}{2\pi} \psi,$$

il che mostra che tutte le soluzioni qui considerate rappresentano stati in cui il momento angolare rispetto all'asse z ha un valore definito, e precisamente il valore $\left(m + \frac{1}{2} \right) \frac{\hbar}{2\pi}$. Tenendo poi presenti i limiti entro cui può variare m , si vede che N_z può variare, nel caso (338), da $-\left(l + \frac{1}{2} \right) \frac{\hbar}{2\pi}$ a $\left(l + \frac{1}{2} \right) \frac{\hbar}{2\pi}$ e, nel caso (341), da $-\left(l - \frac{1}{2} \right) \frac{\hbar}{2\pi}$ a $\left(l - \frac{1}{2} \right) \frac{\hbar}{2\pi}$. Questi risultati sono identici a quelli che si ricavano dalla teoria modellistica, interpretando N_z come la proiezione sull'asse z del momento angolare totale e prendendo questo uguale a $j \frac{\hbar}{2\pi}$, dove j è il quanto interno (v. § 62, p. II), cioè $j = l + 1/2$ nel caso (338) e $j = l - 1/2$ nel caso (341): la prima soluzione corrisponde dunque, nell'interpretazione modellistica, al caso dello spin parallelo al momento orbitale, la seconda al caso dello spin antiparallelo.

Nel caso poi di $l = 0$ manca, come si è detto, la soluzione (341), vale a dire N_z può avere solo il valore $\frac{1}{2} \frac{h}{2\pi}$ (ossia j solo il valore 1/2) come, del resto, risulta immediatamente anche dal modello.

§ 59. — TEORIA DEI SISTEMI IDROGENOIDI (STRUTTURA FINA).

Applichiamo ora i risultati del § precedente al caso di un sistema idrogenoide, cioè specializziamo la funzione U prendendo

$$(344) \quad U = -\frac{Ze^2}{r}.$$

Trattiamo dapprima il caso della soluzione (338) cioè di $j = l + 1/2$ e cerchiamo gli autovalori (per il parametro W) e le autofunzioni delle equazioni (340), limitatamente al caso in cui l'energia $E = W - m_0c^2$ è negativa (caso che corrisponde alle orbite ellittiche). Introduciamo le notazioni:

$$(345) \quad A^2 = \frac{2\pi}{h} \left(m_0c + \frac{W}{c} \right), \quad B^2 = \frac{2\pi}{h} \left(m_0c - \frac{W}{c} \right),$$

e inoltre introduciamo la costante della struttura fina $\alpha = \frac{2\pi e^2}{ch}$: le (340) allora si scrivono

$$(346) \quad \begin{cases} \left(-B^2 + \frac{\alpha Z}{r} \right) F_+ + \frac{dG_+}{dr} - \frac{l}{r} G_+ = 0, \\ \left(-A^2 - \frac{\alpha Z}{r} \right) G_+ + \frac{dF_+}{dr} + \frac{l+2}{r} F_+ = 0. \end{cases}$$

Si osservi che per r tendente all'infinito queste equazioni tendono alla forma

$$-B^2 F_+ + \frac{dG_+}{dr} = 0, \quad -A^2 G_+ + \frac{dF_+}{dr} = 0,$$

da cui si ricava

$$\frac{d^2 F_+}{dr^2} - A^2 B^2 F_+ = 0, \quad \frac{d^2 G_+}{dr^2} - A^2 B^2 G_+ = 0.$$

Quindi le soluzioni F_+ e G_+ delle (346) avranno per espressione asintotica $e^{\pm AB r}$: poichè si cercano soluzioni tendenti

a zero per $r \rightarrow \infty$, si dovrà scartare il segno $+$: si è così condotti a ricercare soluzioni della forma

$$(347) \quad \begin{cases} F_+ = e^{-ABr} (a_0 r^\gamma + a_1 r^{\gamma+1} + \dots) \\ G_+ = e^{-ABr} (b_0 r^\gamma + b_1 r^{\gamma+1} + \dots). \end{cases}$$

L'altro punto eventualmente singolare è $r = 0$: ivi le espressioni precedenti sono regolari se $\gamma \geq 0$, mentre se $\gamma < 0$ divergono infinite dell'ordine di r^γ : tale singolarità è accettabile purchè sia $|\gamma| < 1$ (v. § 28, p. II).

Sostituendo queste espressioni nelle (346) e annullando intanto i coefficienti di $r^{\gamma-1}$, si trova

$$\begin{aligned} Z\alpha a_0 + (\gamma - l)b_0 &= 0, \\ (\gamma + l + 2)a_0 - Z\alpha b_0 &= 0, \end{aligned}$$

e, annullando il determinante dei coefficienti di queste due equazioni lineari in a_0, b_0 , si trova per γ l'equazione

$$(\gamma - l)(\gamma + l + 2) + Z^2 \alpha^2 = 0,$$

che dà (scartando la soluzione col segno $-$ che darebbe $\gamma < -1$)

$$(348) \quad \gamma = -1 + \sqrt{(l+1)^2 - Z^2 \alpha^2}.$$

Annullando poi il coefficiente di $r^{\gamma+s}$ ($s = 0, 1, 2, \dots$) si trova:

$$(349) \quad \begin{cases} -B(Ba_s - Ab_s) + Z\alpha a_{s+1} + (\gamma + s - l + 1)b_{s+1} = 0, \\ A(Ba_s - Ab_s) - Z\alpha b_{s+1} + (\gamma + s + l + 3)a_{s+1} = 0, \end{cases}$$

da cui, moltiplicando la prima per A e la seconda per B e sommando,

$$a_{s+1}[AZ\alpha + B(\gamma + s + l + 3)] + b_{s+1}[A(\gamma + s - l + 1) - BZ\alpha] = 0.$$

Convienne allora introdurre, in luogo di a_{s+1} e b_{s+1} , una sola costante c_{s+1} , ponendo

$$\begin{aligned} a_{s+1} &= c_{s+1}[A(\gamma + s - l + 1) - BZ\alpha], \\ b_{s+1} &= -c_{s+1}[AZ\alpha + B(\gamma + s + l + 3)]. \end{aligned}$$

Sostituendo queste espressioni, e le analoghe per a_s e b_s , nelle (349), si trova per le c_s la formula ricorrente

$$(350) \quad c_{s+1} = \frac{-Z\alpha(A^2 - B^2) - 2AB(\gamma + s + 1)}{Z\alpha^2 + (\gamma + s + l + 3)(\gamma + s - l + 1)} c_s.$$

Le funzioni (347) risultano certamente nulle all'infinito se le serie si riducono a polinomi: detto n' il grado di questi, dovrà essere a tal uopo $c_{n'} \neq 0$ e $c_{n'+1} = 0$, e quindi, per la (350),

$$\frac{Z\alpha(B^2 - A^2)}{2AB} = \gamma + n' + 1,$$

ossia, per la (348),

$$Z\alpha \frac{B^2 - A^2}{2AB} = \sqrt{(l+1)^2 - Z^2c^2} + n'.$$

Sostituendo per A e B le espressioni (345) e risolvendo rispetto a W si trova

$$(351) \quad W = m_0c^2 \left[1 + \frac{Z^2\alpha^2}{(\sqrt{(l+1)^2 - Z^2c^2} + n')^2} \right]^{-1/2}$$

e, ricordando che per la soluzione di cui ci occupiamo si ha $j = l + 1/2$,

$$(352) \quad W = m_0c^2 \left[1 + \frac{Z^2\alpha^2}{(\sqrt{(j+1/2)^2 - Z^2c^2} + n')^2} \right]^{-1/2}.$$

Passando ora a considerare la soluzione (341), corrispondente a $j = l - 1/2$, non occorre rifare il calcolo, bastando ricordare che le equazioni (343) differiscono dalle (340) solo per la sostituzione di l con $-(l+1)$. Operando quindi la stessa sostituzione nella formula (351), si trova:

$$W = m_0c^2 \left[1 + \frac{Z^2\alpha^2}{(\sqrt{l^2 - Z^2c^2} + n')^2} \right]^{-1/2},$$

ma introducendo poi $j + 1/2$ in luogo di l si ricade nella formula (352). Questa esprime dunque i livelli energetici in entrambi i casi. Sviluppiamola in serie di potenze di α , fino ai termini in α^4 inclusi (approssimazione più che sufficiente per i confronti con l'esperienza) e togliamo il termine m_0c^2 che rappresenta l'energia intrinseca: si trova allora per l'energia non intrinseca E l'espressione seguente, dove si è posto $n = j + 1/2 + n'$ e si è tenuto presente che $m_0c^2\alpha^2 = 2hcR$ (R = costante di Rydberg):

$$(353) \quad E_{nj} = -\frac{Z^2Rhc}{n^2} \left[1 - \frac{Z^2\alpha^2}{n^2} \left(\frac{3}{4} - \frac{n}{j+1/2} \right) \right].$$

Se si trascura il termine in α^2 si ritrova la ben nota espressione dei termini balmeriani, data dalla teoria di Schrödinger (v. § 48, p. II): si riconosce così che l' n testè introdotto (il quale risulta intero) si deve identificare col quanto totale. Se invece si tiene conto del termine correttivo in α^2 , si vede che, fissato il valore di n , in luogo dell'unico livello energetico della teoria di Schrödinger, si ricavano da questa formula n livelli energetici diversi, corrispondenti agli n valori che può assumere j (da $1/2$ a $n - 1/2$): tali livelli risultano però assai vicini tra loro, perchè j figura solo nel termine in α^2 . Questa scissione è dovuta all'effetto combinato della correzione relativistica e dello spin, contenuti globalmente nelle formule di Dirac. Alla scissione dei livelli corrisponde naturalmente una scissione delle righe spettrali, che ne costituisce la « struttura fina » (v. § 60, p. II): la riga emessa nel passaggio da un livello di quanto totale n a uno di quanto totale n' deve essere costituita di nn' componenti, di cui però alcune sono escluse dal principio di selezione del quanto interno j (v. § 64 *d*, p. II).

La formula testè trovata per la struttura fina degli spettri idrogenoidi differisce dalla (341) del § 60, p. II (che risultava dall'applicazione della correzione relativistica alla teoria di Bohr e Sommerfeld) solo per avere $j + 1/2$ in luogo di k : ma poichè tanto $j + 1/2$ che k assumono la stessa serie di valori (da 1 ad n), le due formule conducono esattamente agli stessi livelli energetici (sebbene coordinati a numeri quantici diversi) e quindi alla stessa struttura delle righe, la quale coincide assai bene con quella che si osserva sperimentalmente. Diverso è però il risultato delle due teorie nei riguardi dell'intensità, che si può calcolare, nel caso dell'antica teoria, mediante il principio di corrispondenza, e, nel caso della nuova mediante la teoria di Dirac dell'irraggiamento: noi tralasciamo questi calcoli ⁽¹⁾, limitandoci a dire che l'esperienza è in favore della teoria di Dirac. Per un confronto dettagliato fra la teoria sopra esposta e le osservazioni della struttura fina dello spettro dell'idrogeno e di He^+

(1) Vedasi p. es. bibl. n. I, p. 447.

(il quale ultimo, come si è detto, si presta meglio allo studio sperimentale) rimandiamo alla bibl., n. 1, p. 316: qui ci limitiamo a riportare, come esempio, un confronto grafico relativo alla riga 4686 di He^+ , corrispondente al passaggio da $n = 4$ a $n = 3$ (fig. 47): la curva rappresenta l'intensità in funzione della frequenza ricavata dalle osservazioni spettrofotometriche della riga: i segmenti verticali indicano le posizioni delle componenti previste dalla teoria, e la loro lunghezza misura, in unità arbitrarie, l'intensità calcolata (il confronto riguarda solo i rapporti delle intensità delle varie componenti e le loro mutue distanze).

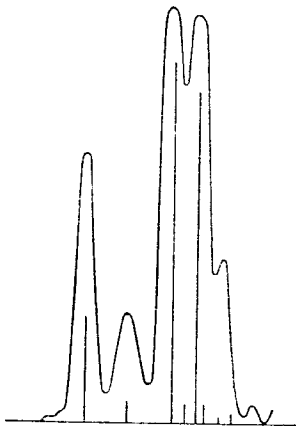


Fig. 47

§ 60. — STATI A ENERGIA CINETICA NEGATIVA E TEORIA DEL POSITRONE.

L'equazione di Dirac ammette (come si è visto al § 54 per il caso particolare di onde piane e $U = 0$) accanto ad ogni soluzione rappresentante uno stato stazionario con energia cinetica \mathfrak{E} positiva ⁽¹⁾, anche una soluzione analoga con energia cinetica negativa. Ciò deriva dal fatto che, nella meccanica relativistica, l'energia cinetica \mathfrak{E} di una particella è legata al suo impulso p dalla relazione

$$(353) \quad \frac{\mathfrak{E}^2}{c^2} = m_0^2 c^2 + p^2,$$

che è quadratica in \mathfrak{E} , mentre la corrispondente relazione della meccanica classica è lineare: perciò anche nella meccanica relativistica non quantistica si presenta la possibi-

⁽¹⁾ Per brevità di locuzione, comprendiamo nell'energia cinetica anche l'energia intrinseca $m_0 c^2$, e la indichiamo con \mathfrak{E} per distinguerla dall'energia cinetica ordinaria o « forza viva » T : si ha $\mathfrak{E} = m_0 c^2 + T$. L'energia totale $\mathfrak{E} + U$ sarà poi indicata, al solito, con W . Nel caso del § 54 si era assunto $U = 0$ e perciò $W = \mathfrak{E}$.

lità di soluzioni con energia cinetica negativa. Precisamente si ricava dalla (353) (trascurando di scrivere i termini d'ordine superiore al secondo in v/c),

$$\mathfrak{E} = \pm \left(m_0 c^2 + \frac{p^2}{2m_0} + \dots \right).$$

Come si vede, \mathfrak{E} può variare entro due intervalli (da $-\infty$ a $-m_0 c^2$ e da $m_0 c^2$ a $+\infty$) separati tra loro: perciò, poichè nella meccanica *non* quantistica l'energia cinetica può variare solo con continuità, essa non può passare da un intervallo all'altro, e così, se il moto si inizia con \mathfrak{E} positivo, stati ad energia cinetica negativa non vengono mai in considerazione. Nella meccanica quantistica invece questo ragionamento non vale, perchè non si può più affermare che \mathfrak{E} vari con continuità: vi sono anzi dei casi in cui vi è una probabilità finita, o anche la certezza, che l'elettrone passi da uno stato con \mathfrak{E} positiva ($> m_0 c^2$) a uno con \mathfrak{E} negativa ($< -m_0 c^2$). Uno tipico di questi casi è stato indicato da O. KLEIN ⁽¹⁾, ed è quello di un elettrone con energia totale $W (> 0)$ lanciato contro un gradino di potenziale come quello della fig. 25, di altezza $U_0 > W + m_0 c^2$: allora, se la ψ a sinistra del gradino è rappresentata dalla (292), a destra essa è rappresentata da onde analoghe in cui \mathbf{p} è sostituito da un \mathbf{p}' reale e tale che

$$\frac{W - U_0}{c} = -\sqrt{m_0^2 c^2 + p'^2}.$$

Le onde quindi proseguono oltre il gradino (in forma sinusoidale, non esponenziale come nella teoria di Schrödinger, v. § 40, p. II), dove l'energia cinetica è $\mathfrak{E} = W - U_0$ ed è quindi negativa. Ciò significa che l'elettrone ha una certa probabilità, che può essere assai notevole, di superare il gradino passando contemporaneamente a uno stato con energia cinetica negativa ⁽²⁾. Del resto, anche senza l'intervento di forze esterne come quelle considerate dal Klein, si pos-

⁽¹⁾ *Zs. f. Phys.*, 53 (1929), 157.

⁽²⁾ La ragione analitica di questo fatto sta nella circostanza che le onde piane a energia cinetica positiva non costituiscono, da sè sole, un sistema *completo* di funzioni ortogonali: per avere un sistema completo si devono aggiungere le onde a energia cinetica negativa. Perciò,

sono avere transizioni a stati di energia cinetica negativa: risulta anzi dalla teoria della radiazione di Dirac che un elettrone, appena posto in uno stato a energia cinetica positiva, compirebbe subito spontaneamente una tale transizione, irradiando la differenza di energia.

Le proprietà di un elettrone con energia cinetica negativa dovrebbero essere assai singolari: esso in un campo elettrico e magnetico acquisterebbe una accelerazione diretta in senso inverso a quella di un elettrone ordinario (e in questo si comporterebbe come un elettrone positivo) ma, cosa più singolare, la sua energia andrebbe decrescendo col crescere della velocità.

Per dimostrare che un elettrone negativo di energia cinetica $\bar{\mathfrak{E}} (< 0)$ si muove come si muoverebbe, nello stesso campo, un elettrone positivo di energia cinetica positiva $-\bar{\mathfrak{E}}$, indichiamo con ψ un'autofunzione (matrice) del primo, relativa a uno stato stazionario di energia cinetica negativa $\bar{\mathfrak{E}}$: essa soddisferà l'equazione di Dirac

$$(354) \quad (\Sigma \alpha^k P_k - P_4 + m_0 c \alpha^4) \psi = 0,$$

con

$$(355) \quad P_k = \frac{\hbar}{2\pi i} \frac{\partial}{\partial x_k} + \frac{e}{c} A_k, \quad P_4 = \frac{1}{c} (W + eV) = \frac{1}{c} \bar{\mathfrak{E}}.$$

D'altra parte, a un elettrone positivo di energia cinetica $-\bar{\mathfrak{E}}$ corrisponderebbe un'autofunzione ψ' soddisfacente l'equazione

$$(356) \quad (-\Sigma \alpha^k P_k^* + P_4^* + m_0 c \alpha^4) \psi' = 0,$$

ottenuta dalla precedente cambiando e in $-e$ e $\bar{\mathfrak{E}}$ in $-\bar{\mathfrak{E}}$.

Ora mostriamo che la (356) è soddisfatta prendendo

$$(357) \quad \psi' = S \psi^*, \quad \text{con } S = \alpha^4 \alpha^2.$$

Infatti, la matrice S così definita ha la proprietà seguente:

$$(358) \quad \alpha^k S = S \alpha^{k*}, \quad \alpha^4 S = -S \alpha^{4*},$$

il che si verifica immediatamente, applicando le (266) e osservando che, per le (267), le coniugate delle matrici α sono uguali

data arbitrariamente una ψ iniziale, non è possibile in generale svilupparla in integrale di Fourier senza far intervenire anche termini con energia cinetica negativa.

alle matrici stesse, tranne α^{2*} che è uguale a $-\alpha^2$. Sostituendo dunque la (357) nella (356), e tenendo conto delle (358), il primo membro diviene

$$-S(\Sigma \alpha^{k*} P_k^* - P_4^* + m_0 c \alpha^{4*}) \psi^*,$$

cioè è uguale al coniugato del primo membro della (354) moltiplicato a sinistra per $-S$: esso è dunque nullo in virtù della (354), e la (356) è soddisfatta. Poichè la (357) vale per qualunque autofunzione, varrà anche per una somma di autofunzioni, cosicchè, se ψ rappresenta un pacchetto d'onde corrispondente a un elettrone negativo di energia cinetica negativa (il che significa che ψ è formata almeno prevalentemente con autofunzioni corrispondenti a valori negativi di $\bar{\mathfrak{E}}$), allora $\psi' = \alpha^4 \alpha^2 \psi^*$ rappresenta un pacchetto d'onde corrispondente a un elettrone positivo di energia cinetica positiva. E poichè, come si verifica subito, $|\psi'|^2 = |\psi|^2$, i due pacchetti coincidono e si muovono nello stesso modo.

Non sarebbe però lecito identificare senz'altro un elettrone negativo a energia cinetica negativa con un positrone: basta osservare che la sua energia cinetica, in funzione dell'impulso p , è espressa in prima approssimazione da

$$(359) \quad \bar{\mathfrak{E}} = -m_0 c^2 - \frac{1}{2m_0} p^2$$

e quindi decresce col crescere dell'impulso, ossia della velocità, il che non corrisponde certamente alle proprietà di nessuna delle particelle sperimentalmente osservate.

L'esistenza degli stati a energia cinetica negativa, e la impossibilità di prescindere da essi, sembrò costituire, fin da principio, una delle più gravi obiezioni alla teoria di Dirac. Per uscire da questa difficoltà, lo stesso DIRAC propose ⁽¹⁾ una ingegnosa interpretazione di questi stati, la quale conduce inoltre a prevedere l'esistenza di un elettrone positivo, e dà il modo di calcolarne le proprietà: tale teoria, sebbene non priva di punti ancora deboli e oscuri, ha acquistato molto credito, specialmente dopo la scoperta sperimentale del positrone. Ecco in breve di che si tratta.

DIRAC fa l'ipotesi che uno spazio privo di cariche elettriche non debba pensarsi come uno spazio in cui non vi

⁽¹⁾ *Proc. Roy. Soc.*, A. **126**, (1931) 360; *id.*, A. **133**, (1931) 60. V. anche bibl. n. 6.

sono elettroni, ma che invece vi siano in esso infiniti elettroni a energia negativa, e precisamente uno per ogni stato stazionario. Questa particolare distribuzione di elettroni, che chiameremo « normale », è stabile in virtù del principio di Pauli che vieta ad un elettrone di occupare uno stato già occupato da un altro elettrone. La distribuzione « normale » non deve dar luogo ad alcun campo elettrico, e per questo bisogna fare l'ulteriore ipotesi che nell'equazione $\text{div } E = -4\pi\rho$ si debba sempre intendere per ρ non già la carica dovuta a tutti gli elettroni presenti (che sarebbe infinita) ma solo la differenza tra questa carica e quella dovuta alla distribuzione « normale ». Se ora modifichiamo la distribuzione normale, facendo passare uno degli elettroni da uno degli stati a energia negativa (rappresentato, p. es., dalla autofunzione $\psi_{n'}$) a uno di energia positiva (la cui autofunzione sia $\psi_{n''}$), la nuova distribuzione ottenuta differirà da quella normale per due rispetti: presenza di un elettrone nello stato n'' , e assenza di un elettrone nello stato n' . Il primo di questi fatti darà luogo a tutte le ordinarie manifestazioni di un ordinario elettrone nello stato n'' , mentre la « lacuna » lasciata negli stati di energia negativa si manifesterà come un elettrone a carica positiva, o *positrone*, nello stato n' . Più generalmente un positrone in uno stato generico sarà rappresentato *sottraendo* dalla distribuzione normale una $\sum_n c_n \psi_n$ (la somma essendo estesa a tutti o a parte degli stati di energia negativa), vale a dire, sottraendo dalla distribuzione normale un pacchetto d'onde a energia cinetica negativa, il quale, come si è visto, si muove nello stesso modo di un elettrone di carica $+e$. Interpretando così un positrone non già come un pacchetto d'onde *esistente*, ma come uno *mancante*, si elimina la difficoltà, cui si è accennato sopra, relativa all'energia: difatti, se aumenta la velocità dei positroni, diminuisce l'energia cinetica mancante dal sistema, ossia, in definitiva, la sua energia cinetica totale aumenta, come deve essere ⁽¹⁾.

⁽¹⁾ Nel lavoro citato il DIRAC interpretò le « lacune » come protoni, essendo allora sconosciuta l'esistenza dei positroni: restava però ine-

Da questa teoria risulta evidentemente che deve essere possibile la creazione di una coppia elettrone-positrone, e questo fenomeno sarebbe in certo modo analogo alla ordinaria eccitazione di un atomo, poichè consisterebbe nel portare un elettrone da uno stato di energia cinetica negativa (dove esso non si manifesta in alcun modo) a uno stato di energia cinetica positiva: si noti però che essendo l'energia del primo stato inferiore a $-m_0c^2$, e quella del secondo superiore a m_0c^2 , l'energia necessaria per la creazione della coppia dovrebbe superare, sia pure di poco, $2m_0c^2$, vale a dire circa un milione di volt. Questa energia può essere fornita da raggi X o γ , purchè di frequenza ν tale che $h\nu > 2m_0c^2$ (cioè $\bar{\nu} > 8,3 \cdot 10^9 \text{ cm}^{-1}$): in tal caso il fotone può « materializzarsi », cioè trasformarsi in una coppia di elettroni, uno negativo e uno positivo, aventi complessivamente la forza viva $h\nu - 2m_0c^2$. Si può anzi calcolare, sulle basi della teoria di Dirac, la probabilità perchè avvenga questa *materializzazione della radiazione* e si trova che il fenomeno è praticamente possibile solo quando il fotone attraversa il campo elettrico di un nucleo.

Queste previsioni teoriche si accordano bene con numerosi risultati sperimentali, dovuti principalmente a I. CURIE ed F. JOLIOT, CHADWICK, BLACKETT e OCCHIALINI, ANDERSON e NEDDERMEYER, MEITNER e PHILIPP, ecc. ⁽¹⁾. Da queste esperienze risulta, p. es., che se si lanciano i raggi γ del torio C'' (il cui fotone ha un'energia di $2,65 \cdot 10^6$ volt) contro uno schermo di piombo in una camera di Wilson, dal piombo escono elettroni positivi e negativi (questi ultimi assai più abbondanti), di cui si può determinare il segno e la velocità facendone deflettere le traiettorie da un campo magnetico: si trova che, mentre la forza viva degli elettroni negativi può raggiungere anche l'intera ener-

splicabile che la massa delle particelle positive fosse tanto diversa da quella degli elettroni. Con la scoperta sperimentale del positrone (1932), le « lacune » di DIRAC trovarono la loro più diretta e naturale interpretazione.

⁽¹⁾ Per indicazioni bibliografiche più complete e per maggiori particolari su questo argomento e sulla annichilazione degli elettroni, rimandiamo al n. 26 e al n. 31 della bibl., nonchè all'opuscolo di I. CURIE ed F. JOLIOT, *L'électron positif* (Parigi, Hermann, 1934).

gia del fotone, quella dei positivi è abbastanza nettamente limitata a non più di $1,6 \cdot 10^6$ volt. Ciò si interpreta così: gli elettroni negativi sono dovuti, oltrechè alla materializzazione dei fotoni, anche all'ordinario effetto fotoelettrico (nel quale, come è noto, quasi tutta l'energia del fotone può trasformarsi in forza viva), mentre i positivi sono dovuti esclusivamente alla materializzazione: poichè questo fenomeno assorbe un'energia di circa un milione di volt, la coppia formata ha complessivamente una forza viva di $2,6 - 1 = 1,6$ milioni di volt (di cui poi una parte viene perduta nell'attraversare il piombo): perciò ciascuna delle due particelle non può avere forza viva superiore a questo limite. In qualche raro caso, poi, si è potuta osservare la formazione di una coppia di elettroni (+ e -) in seno al gas della camera di Wilson anzichè nel piombo: in tal caso le due particelle non vengono sensibilmente rallentate, e quindi la somma delle loro forze vive misurate deve risultare uguale ad $h\nu - 2m_0c^2$, ciò che in generale si verifica abbastanza bene.

Il fenomeno inverso del precedente è la caduta di un elettrone da uno stato di energia positiva ad uno dei pochi stati liberi di energia negativa, vale a dire la ricombinazione di un elettrone negativo con uno positivo, i quali scompaiono liberando, sotto forma di radiazione, una quantità di energia uguale a $2m_0c^2$ (più la loro eventuale forza viva, in generale trascurabile). Questo fenomeno, detto *annichilazione*, può avvenire essenzialmente in due modi: *a*) un positrone incontra un elettrone libero e si ricombina con esso, emettendo due fotoni in direzioni opposte (non uno solo, poichè i due elettroni non potrebbero fornire a questo l'impulso necessario): ciascuno dei due fotoni ha quindi l'energia m_0c^2 , cioè circa mezzo milione di volt (trascurando la forza viva dei due elettroni); *b*) un positrone si ricombina con un elettrone fortemente legato a un nucleo: in tal caso è possibile l'emissione di un solo fotone, di energia circa uguale a $2m_0c^2$, cioè di frequenza doppia di quello del caso (*a*), e l'impulso di rinculo viene comunicato al nucleo. La probabilità del processo (*b*) risulta, da un calcolo di FERMI e UHLENBECK, assai minore di quella del processo (*a*). Esistono anche altre possibilità teoriche di annichilazione, ma ci limiteremo a considerare

queste due, di cui la prima è confermata da notevoli esperienze di diversi autori⁽¹⁾: tra queste citeremo soltanto quelle di J. THIBAUD e di F. JOLIOT.

Questi autori hanno indipendentemente osservato che se si lanciano dei positroni contro un pezzo di piombo, o di alluminio, o di platino, dal metallo urtato esce una radiazione il cui coefficiente di assorbimento nel piombo corrisponde a quello di raggi γ di circa mezzo milione di volt (precisamente, i limiti assegnati da JOLIOT sono da 425.000 a 645.000 volt). Inoltre, una misura (necessariamente imprecisa) dell'intensità di questa radiazione in rapporto al numero dei positroni incidenti ha permesso a JOLIOT di concludere che per ogni positrone incidente vengono emessi da 1,6 a 3 fotoni, con buona probabilità per il valore 2 previsto dalla teoria.

Il fenomeno della annichilazione rende ragione del fatto che gli elettroni positivi sono tanto più rari ad osservarsi dei negativi, e che essi non intervengono nella conduzione elettrica, nel fenomeno termoionico, ecc. Infatti, un positrone in seno alla materia (sia pure a un gas rarefatto) si trova in presenza di un gran numero di elettroni negativi ed ha una notevole tendenza a ricombinarsi con uno di essi: la probabilità di ricombinazione è stata calcolata dal DIRAC, il quale ha trovato che essa aumenta, tendendo ad un valore massimo, col diminuire della velocità del positrone, cosicchè il positrone può sussistere in pratica solo se è animato da notevole velocità. La vita media di un positrone lento nell'acqua risulterebbe p. es. di $3,5 \cdot 10^{-10}$ sec.

⁽¹⁾ CHAO, *Proc. Nat. Acad.*, 16 (1930), 431; *Phys. Rev.* 36 (1930), 1519; *Proc. Roy. Soc. A.* 135 (1931), 206; TARRANT, *ibid. A.* 136 (1932), 662; GRAY e TARRANT, *ibid.* 143 (1934), 681; J. THIBAUD, *C. R.* 197 (1933), 1629; F. JOLIOT, *C. R.* 197 (1933), 1622 e *Journ de Phys.* V (1934), 299.

CAPITOLO VI.

SISTEMI CON PARTICELLE UGUALI

§ 61. — UGUAGLIANZA TRA PARTICELLE ELEMENTARI.

La rappresentazione intuitiva degli elettroni, dei protoni, dei fotoni, ecc. come punti materiali è inesatta non solo perchè attribuisce ad essi delle proprietà cinematiche non corrispondenti a realtà (come si è ampiamente illustrato nella parte I) ma anche perchè attribuisce ad essi una « individualità » che in realtà non hanno. Per comprendere meglio che cosa si voglia intendere con ciò, si pensi a un sistema contenente vari elettroni (p. es., un atomo): se ci lasciamo guidare dal modello corpuscolare, e pensiamo gli elettroni come minuscole palline, possiamo sempre attribuire un significato all'operazione di scambiare tra loro due di queste palline, anche se esse sono uguali (nel senso ordinario di questa parola). Ma se si analizza operativamente il significato di questo scambio, si vede che esso è fondato sulla possibilità di attribuire una individualità alla pallina « segnanbola » in qualche modo (sia pure così lievemente da non alterarne le proprietà essenziali), oppure seguendola ininterrottamente durante l'operazione di scambio. Ma per gli elettroni (e per qualunque altra particella elementare) è *concettualmente impossibile* sia segnarli, sia seguirli con continuità: dunque il modello corpuscolare è inadeguato, anche per questa ragione, a rappresentarli. La proprietà della « individualità » non corrisponde in tal caso ad alcuna realtà fisica, e l'unica ragione per cui siamo abituati ad attribuirle alle particelle elementari, è che essa è, nella nostra mente e

nel linguaggio comune, indissolubilmente legata (al pari dei concetti di posizione, velocità, traiettoria, ecc.) alla parola « particella » con cui convenzionalmente designamo gli elettroni, i protoni, ecc. Ora, si è visto che, per quanto riguarda i concetti cinematici, la inadeguatezza del linguaggio corpuscolare si può compensare introducendo il principio di indeterminazione: similmente, per ciò che riguarda la questione della individualità, è possibile ancora conservare il linguaggio corpuscolare, ma con l'espressa convenzione che due particelle elementari della stessa specie sono da considerarsi *uguali* in un senso assai più stretto di quello dato ordinariamente a questa parola, e cioè nel senso che l'espressione « scambio di due qualunque di esse » è priva di significato, con che si sottrae al concetto di particella la caratteristica della individualità ⁽¹⁾. È in questo senso speciale che deve intendersi l'*uguaglianza* delle particelle, di cui ci occupiamo in questo capitolo. Come si vedrà, questa proprietà delle particelle elementari, che non ha riscontro tra i corpi ordinari, conduce a delle conseguenze singolari, che non possono essere interpretate intuitivamente col modello dei corpuscoli, ma che in molti casi sono suscettibili di verifica sperimentale.

Considereremo dapprima i sistemi con due sole particelle uguali (come è, p. es., l'atomo di elio); poi estenderemo sommarariamente i ragionamenti a sistemi con quante si vogliono particelle uguali.

§ 62. — STATI DI UN SISTEMA DI DUE PARTICELLE UGUALI.
SIMMETRIA E ANTISIMMETRIA.

Si abbia un sistema costituito di due particelle uguali (immerse in un campo assegnato) e indichiamo brevemente

⁽¹⁾ Per chiarire meglio la cosa con un altro esempio di enti privi di individualità, nel senso sopra specificato, si pensi a una corda vibrante in cui sono eccitati contemporaneamente due sistemi di onde stazionarie, uno di frequenza ν_1 , ampiezza A_1 , fase φ_1 , l'altro di frequenza ν_2 , ampiezza A_2 , fase φ_2 . La frase « scambio del sistema di onde (ν_1, A_1, φ_1) col sistema (ν_2, A_2, φ_2) » è completamente priva di significato. Allo stesso modo, se in un atomo vi sono due elettroni, uno in uno stato quantico definito da certi numeri quantici n_1, l_1, m_1 , l'altro in uno stato definito da n_2, l_2, m_2 , non ha alcun significato il dire che si passa il primo elettrone nel secondo stato, e il secondo nel primo.

con q^1 , p^1 l'insieme delle coordinate e dei momenti di una di esse, con q^2 , p^2 quelli dell'altra, includendo nelle q anche la variabile di spin, che designeremo con σ , e potrà essere, p. es., la σ_z del capitolo precedente (quindi, p. es., q^1 sta per il gruppo x_1, y_1, z_1, σ_1). Qualunque osservabile del sistema deve avere un'espressione $F(q^1, p^1, q^2, p^2)$ *simmetrica* rispetto ai due gruppi di variabili, cioè tale che scambiando le q^1, p^1 con le corrispondenti q^2, p^2 l'espressione risulti la stessa: altrimenti lo scambio delle due particelle avrebbe un significato fisico. Indicheremo la simmetria della funzione F scrivendo brevemente

$$F(1, 2) = F(2, 1),$$

invece di

$$F(q^1, p^1, q^2, p^2) = F(q^2, p^2, q^1, p^1).$$

Sarà quindi simmetrico anche l'operatore corrispondente a una generica osservabile F . In particolare, sarà simmetrica l'espressione dell'energia, e quindi l'operatore hamiltoniano \mathfrak{H} : ne segue che se $\psi_n(1, 2)$ è una autofunzione appartenente all'autovalore E_n , cioè soddisfa l'equazione

$$(360) \quad \mathfrak{H}\psi_n(1, 2) = E_n\psi_n(1, 2),$$

anche $\psi_n(2, 1)$ è una autofunzione appartenente allo stesso autovalore, perchè questa equazione è ancora soddisfatta se nella ψ_n si scambiano le q^1 con le q^2 .

Di qui si traggono notevoli conseguenze. Supponiamo dapprima che E_n sia un autovalore semplice: in tal caso $\psi_n(2, 1)$ non può essere essenzialmente distinta da $\psi_n(1, 2)$, vale a dire deve risultare

$$\psi_n(2, 1) = c\psi_n(1, 2),$$

dove c è una costante: se poi in questa relazione si scambiano le q^1 con le q^2 , e si moltiplicano membro a membro queste due equazioni, si trova $c^2 = 1$, ossia $c = \pm 1$. Dunque, un'autofunzione appartenente a un autovalore semplice, o gode la proprietà

$$\psi_n(2, 1) = \psi_n(1, 2),$$

vale a dire è *simmetrica*, oppure l'altra

$$\psi_n(2, 1) = -\psi_n(1, 2),$$

nel qual caso si dice *antisimmetrica*.

Esaminiamo ora il caso della degenerazione, cioè supponiamo che E_n sia un autovalore multiplo d'ordine p , e sia

$$(361) \quad \psi_n^{(1)}(1, 2), \psi_n^{(2)}(1, 2), \dots, \psi_n^{(p)}(1, 2)$$

un sistema fondamentale di autofunzioni ad esso appartenenti, ortogonali tra loro (v. § 6, p. II). È noto che ad esso si può sostituire un qualsivoglia altro sistema ottenuto dal precedente con una sostituzione ortogonale: mostreremo ora che si può scegliere la sostituzione in modo che questo sistema risulti composto solo di funzioni simmetriche e antisimmetriche (nonchè ortogonali tra loro).

Infatti, consideriamo, p. es., la prima delle autofunzioni (361) e scambiamo in essa le q^1 con le q^2 : otterremo una nuova autofunzione $\psi_n^{(1)}(2, 1)$ appartenente allo stesso autovalore. Essa potrà coincidere (a meno di un fattore costante) con la stessa $\psi_n^{(1)}(1, 2)$, e allora si dimostra, come si è visto poc'anzi, che questa è già simmetrica o antisimmetrica; oppure potrà coincidere con un'altra delle (361), p. es. $\psi_n^{(r)}(1, 2)$, e in tal caso sostituiamo alla coppia di funzioni $\psi_n^{(1)}(1, 2)$, $\psi_n^{(r)}(1, 2)$ le due funzioni (evidentemente non nulle)

$$(362) \quad \frac{1}{2} \left[\psi_n^{(1)}(1, 2) \pm \psi_n^{(r)}(2, 1) \right],$$

che sono pure linearmente indipendenti tra loro e con le altre, e inoltre sono l'una simmetrica e l'altra antisimmetrica; o infine potrà $\psi_n^{(1)}(2, 1)$ non coincidere con nessuna delle (361), e allora potremo sostituire a $\psi_n^{(1)}(1, 2)$ una qualunque delle due funzioni (362), che sono entrambe indipendenti dalle altre. Applicando successivamente questo processo a tutte le autofunzioni (361), si riesce a sostituire queste con altrettante autofunzioni indipendenti, alcune simmetriche e altre antisimmetriche. Resta da ottenere l'ortogonalità. Consideriamo il gruppo delle autofunzioni simmetriche: su di esse si può sempre operare una sostituzione lineare tale da sostituirle con altrettante autofunzioni indipendenti ortogonali tra loro (v. p. es. § 6, p. II) le quali risulteranno evidentemente simmetriche; similmente, le autofunzioni antisimmetriche si possono sostituire con altrettante loro com-

binazioni lineari indipendenti ortogonali tra loro (ed, evidentemente, antisimmetriche). Si osservi infine che una qualsiasi funzione simmetrica ψ_s ed una antisimmetrica ψ_a sono necessariamente ortogonali tra loro, poichè l'integrale

$$\psi_s \times \psi_a = \int \psi_s^*(1, 2) \psi_a(1, 2) dq^1 dq^2$$

non deve alterarsi scambiando le variabili q^1 con le q^2 e, d'altra parte, questo scambio muta solo di segno l'integrando: dunque l'integrale è nullo. Si riconosce così che le p autofunzioni, parte simmetriche e parte antisimmetriche, che abbiamo sostituito alle (361), sono tutte ortogonali tra loro.

Se per ogni eventuale autovalore multiplo il sistema fondamentale di autofunzioni viene scelto nel modo descritto, si ottiene un sistema *completo* di autofunzioni ortogonali, il quale non contiene nessuna autofunzione che non sia simmetrica o antisimmetrica.

Passiamo ora a considerare in generale gli stati possibili (anche non stazionari) per il sistema. Ammetteremo, come è fisicamente plausibile, non solo che la densità di probabilità $P = \psi\psi^*$ debba essere simmetrica rispetto allo scambio delle due particelle, ma anche che se $\psi(1, 2)$ rappresenta uno stato *possibile* del sistema, anche $\psi(2, 1)$ [rappresenti uno stato possibile: si può allora dimostrare che $\psi(1, 2)$ deve necessariamente essere o simmetrica o antisimmetrica: una ψ che non fosse tale (come si può ottenere combinando linearmente delle autofunzioni simmetriche con delle antisimmetriche), sebbene soddisfi l'equazione temporale di Schrödinger (o di Dirac), non rappresenta uno stato fisicamente possibile del sistema.

Infatti, sia $\psi(1, 2)$ non simmetrica nè antisimmetrica: se essa rappresentasse uno stato possibile, sarebbe altrettanto di $\psi(2, 1)$ e quindi delle loro combinazioni, simmetrica e antisimmetrica (non nulle): $\psi_s = \frac{1}{2} [\psi(1, 2) + \psi(2, 1)]$, $\psi_a = \frac{1}{2} [\psi(1, 2) - \psi(2, 1)]$. Ma allora anche una qualunque altra combinazione del tipo $\varphi = \psi_s + e^{i\delta}\psi_a$ (dove δ è una costante arbitraria) rappresenterebbe uno stato possibile, e ciò contraddice la prima delle nostre ipotesi, perchè la densità di probabilità di questo stato:

$\varrho = \varphi\varphi^* = \psi_s\psi_s^* + e^{i\delta}\psi_a\psi_s^* + e^{-i\delta}\psi_a^*\psi_s + \psi_a\psi_a^*$, non risulta in generale simmetrica: infatti, scambiando le variabili q^1 con le q^2 la ϱ si muta in $\varrho' = \psi_s\psi_s^* - e^{i\delta}\psi_a\psi_s^* - e^{-i\delta}\psi_a^*\psi_s + \psi_a\psi_a^*$: la differenza $\varrho - \varrho' = 2(e^{i\delta}\psi_a\psi_s^* + e^{-i\delta}\psi_a^*\psi_s)$, in generale, data l'arbitrarietà di δ , non risulta nulla. Dunque, se fosse possibile lo stato $\psi(1, 2)$, sarebbero anche possibili degli stati in cui ϱ non è simmetrica, il che è da escludersi.

Resta così dimostrato che gli stati *a priori* possibili per il sistema si dividono in due classi che chiameremo degli *stati simmetrici* e degli *stati antisimmetrici*, a seconda che la ψ è simmetrica o antisimmetrica (si sottintende rispetto allo scambio di x_1, y_1, z_1, σ_1 con x_2, y_2, z_2, σ_2).

Dimostreremo ora che il sistema *non può mai passare da uno stato simmetrico ad uno stato antisimmetrico o viceversa*, e che, per conseguenza, nell'evoluzione di un dato sistema possono presentarsi stati di una sola delle due classi (come si dirà più avanti, per gli elettroni e i protoni sono gli stati antisimmetrici i soli che si presentano). Infatti, dall'equazione fondamentale:

$$(363) \quad \mathfrak{H}\psi = -\frac{\hbar}{2\pi i} \frac{d\psi}{dt},$$

si ricava che l'incremento della ψ nel tempo dt è

$$(364) \quad d\psi = -\frac{2\pi i}{\hbar} \mathfrak{H}\psi dt:$$

ora, essendo \mathfrak{H} un operatore simmetrico, se ψ in un certo istante t è simmetrica (o antisimmetrica) tale risulta anche $d\psi$ e quindi $\psi + d\psi$: dunque la ψ al tempo $t + dt$ ha lo stesso carattere di simmetria o antisimmetria che ha la ψ al tempo t .

§ 63. — CENNO SULLA ESTENSIONE AD N PARTICELLE UGUALI.

Se il sistema consta di N particelle uguali, e si indicano con q^i, p^i le coordinate e i momenti della particella i -esima (incluse le variabili di spin), è ovvio che ogni grandezza avente significato fisico, e in particolare la densità di probabilità $|\psi|^2$, deve risultare simmetrica rispetto a tutte le

particelle, cioè deve restare inalterata permutando tra loro in modo qualunque gli indici che caratterizzano le variabili. Dal fatto che, in particolare, l'hamiltoniana \mathcal{H} gode questa proprietà, si deduce, con lo stesso procedimento usato nel caso di due particelle, che se in una autofunzione $\psi_n(q^1, q^2 \dots q^N)$ si permutano in modo qualunque gli indici delle variabili, si ottiene una autofunzione appartenente allo stesso autovalore. Se questo è semplice, tale autofunzione deve coincidere con la primitiva a meno di un fattore costante che (lo si dimostra come nel caso precedente) non può essere che ± 1 . Dimostreremo anzi che questo fattore o è $+1$ per tutte le permutazioni (e allora la funzione dicesi *simmetrica*) oppure è $+1$ per le permutazioni pari ⁽¹⁾ e -1 per le dispari (e allora la funzione dicesi *antisimmetrica*).

Infatti, si consideri dapprima una *trasposizione* (r, s) cioè lo scambio di due soli indici, r, s , e si indichi con C_{rs} il fattore ($= \pm 1$) per cui resta moltiplicata la funzione in seguito a questo scambio: si osservi poi che la trasposizione (r, s) può anche ottenersi mediante l'applicazione successiva delle trasposizioni (r, t) , (t, s) , (r, t) , dove t è un altro indice qualunque: si avrà dunque tra i corrispondenti fattori C la relazione $C_{rs} = C_{rt}C_{ts}C_{rt} = C_{ts}C_{rt}^2 = C_{ts}$. Ma se $C_{rs} = C_{ts}$ qualunque siano gli indici r, s, t , ciò significa che tutte le C sono uguali tra loro, e quindi: o sono tutte uguali a $+1$, nel qual caso qualunque trasposizione, e quindi anche qualunque permutazione lascia inalterata la funzione, la quale dunque è simmetrica, ovvero sono tutte uguali a -1 , e allora un numero pari di trasposizioni la lascia inalterata, e un numero dispari la cambia di segno, ossia si tratta di una funzione antisimmetrica. Si noti che in questo ragionamento non si è fatto uso della equazione cui soddisfa ψ_n , e quindi esso vale non solo per una autofunzione, ma per qualunque funzione delle coordinate che, di fronte a ogni trasposizione degli indici di queste, resti moltiplicata per ± 1 .

Se l'autovalore è multiplo, permutando gli indici in ψ_n questa si muta in un'altra autofunzione, in generale indipendente dalla primitiva: si può dimostrare però che per ogni

⁽¹⁾ Una permutazione è pari o dispari a seconda che può essere ottenuta con un numero pari o dispari di trasposizioni, cioè di scambi di due soli elementi tra loro.

autovalore multiplo d'ordine p esiste un sistema di p autofunzioni (indipendenti e ortogonali) di cui una è simmetrica, una antisimmetrica, mentre le altre hanno un comportamento più complicato. Queste però non ci interessano perchè, come vedremo, non rappresentano stati fisicamente possibili ⁽¹⁾.

Consideriamo ora uno stato generico, anche non stazionario, rappresentato da una certa ψ . Se si ammette (come nel caso di due particelle) che, scambiando tra loro le coordinate di due qualunque delle particelle, la ψ si muti in un'altra rappresentante uno stato fisicamente possibile, si può applicare, per ciascuno di questi scambi o trasposizioni, il ragionamento del § precedente, e si giunge alla conclusione che la ψ deve essere o simmetrica o antisimmetrica rispetto a ogni trasposizione, e quindi (poichè, come si è visto poco anzi, una funzione non può essere simmetrica per alcune trasposizioni e antisimmetrica per altre) la ψ deve essere o *simmetrica (rispetto a ogni permutazione degli indici)* o *antisimmetrica*: corrispondentemente, si parlerà di *stati simmetrici* e di *stati antisimmetrici*, e non sono possibili stati che non rientrino in una di queste due categorie.

Si estende poi immediatamente al caso di più particelle il ragionamento fatto a pag. 471, per dimostrare che il sistema non può passare in alcun modo da uno stato simmetrico a uno antisimmetrico o viceversa.

§ 64. — PRINCIPIO DI PAULI E PRINCIPIO DI ANTISIMMETRIA.

Nella teoria dell'atomo di Bohr e Sommerfeld a ciascun elettrone, come si sa, vengono attribuiti tre numeri quantici — il *quanto totale* n_i , quello *azimutale* l_i , quello *magnetico* m_i — che ne individuano completamente l'orbita: se poi a questi si aggiunge, conformemente all'ipotesi dell'elettrone rotante, il numero quantico di *spin* s_i (susceptibile solo di due

⁽¹⁾ Lo studio generale delle proprietà di simmetria delle autofunzioni di N particelle è stato fatto coi metodi della teoria dei gruppi da E. WIGNER, *Zs. f. Phys.*, 40 (1927), 883.

valori, corrispondenti alle due orientazioni opposte che può avere lo spin), lo stato quantico di ciascuno degli elettroni dell'atomo resta individuato da una *quaterna* di numeri quantici. Questi numeri intervengono in tutte le leggi che determinano il comportamento dell'elettrone in questione (p. es., nelle regole di selezione) e quindi sono strettamente legati ai fenomeni chimici e spettroscopici per gli elettroni più esterni (elettroni di valenza e ottici) e all'emissione e all'assorbimento dei raggi X per i più interni: da ciò la possibilità di controllare sperimentalmente i loro valori.

Veramente, nel § 56, p. II, abbiamo definito i numeri quantici solo per un unico elettrone soggetto a un campo centrale: in un atomo con più elettroni le mutue azioni tra questi complicano il problema. In prima approssimazione si può, come si è visto al § 59, p. II, schematizzare l'azione esercitata sull'elettrone *i*-esimo dai rimanenti come una forza centrale da aggiungersi a quella prodotta dal nucleo, e allora si ricade nel caso precedente: ma anche indipendentemente da ogni approssimazione è possibile sempre definire per ciascun elettrone la sua quaterna di numeri quantici n_i, l_i, m_i, s_i , nel modo seguente. Supponiamo che l'atomo venga posto in un campo magnetico il cui valore, dapprima nullo, cresca lentamente fino a che le forze da esso prodotte sugli elettroni divengano preponderanti rispetto a tutte le loro azioni reciproche, e anche rispetto all'azione del campo magnetico orbitale sugli spin: si può dimostrare che, se l'atomo era originariamente in uno stato quantico, continua ad esserlo anche nel campo magnetico forte. Ma in questo campo ciascun elettrone può essere quantizzato separatamente e quindi gli si possono attribuire nel modo noto i suoi quattro numeri quantici: poichè il sistema è passato con continuità dal campo magnetico nullo al campo fortissimo, mentre i numeri quantici non possono variare con continuità, così si dovranno attribuire a ciascun elettrone, nell'atomo non soggetto al campo, gli stessi numeri quantici che gli spettano in un campo magnetico sufficientemente intenso.

Ora, è stato rilevato empiricamente (per la prima volta dal PAULI, studiando gli spettri dei metalli alcalini ⁽¹⁾) che *in un atomo non vi sono mai due elettroni aventi la stessa qua-*

terna di numeri quantici ⁽¹⁾. Questo fatto non si deduce da nessuna altra legge della teoria dei quanti, ma poichè si verifica costantemente in tutti gli atomi (e anche in sistemi più complessi come le molecole) è stato assunto come un principio fondamentale della fisica atomica, col nome di « *principio di Pauli* » o « *di esclusione* », e si è dimostrato fecondissimo di conseguenze in pieno accordo con l'esperienza ⁽²⁾. Si può dire che tutta la interpretazione teorica degli spettri è fondata su tale principio, e ne costituisce una imponente conferma: la sua esposizione forma oggetto di altro volume di questo trattato. In questo volume, avremo occasione di citare una sola di tali applicazioni (v. § 67).

Il principio di Pauli ha fornito la chiave per l'interpretazione di una delle più fondamentali leggi naturali, e cioè del sistema periodico di Mendelèiev. Per una esposizione dettagliata di questa interpretazione rimandiamo al volume sulla spettroscopia: qui ci limitiamo a un breve cenno schematico e semplificato. Se non valesse il principio di Pauli ciascun elettrone tenderebbe a portarsi nell'orbita di energia minore, che è (almeno per gli elementi più leggeri, ai quali qui ci limiteremo) $n = 1, l = 0, m = 0$: perciò tutti gli elettroni dell'atomo nello stato fondamentale avrebbero questi numeri quantici (ossia, in linguaggio spettroscopico, sarebbero nello stato $1s$) e le proprietà ottiche e chimiche varierebbero progressivamente da un atomo al successivo, corrispondentemente all'accrescersi del numero di questi elettroni: in particolare, il termine fondamentale dello spettro sarebbe sempre $1s$. Tutto ciò è in assoluto contrasto con la realtà. Invece, per il principio di Pauli, nella detta orbita non vi possono essere più di 2 elettroni: perciò nell'atomo di litio (numero atomico $Z = 3$, quindi 3 elettroni), due

⁽¹⁾ Poichè il quarto numero quantico, s , può assumere solo due valori, se si fissano i valori dei tre numeri quantici n, l, m (detti « orbitali ») vi possono essere al più *due* elettroni aventi questa terna di numeri quantici orbitali (o, come si dice brevemente, questa « orbita »). In questa forma fu enunciato per la prima volta dal PAULI il suo principio. I due elettroni aventi la stessa orbita devono poi differire per il numero quantico s , ossia devono avere gli spin in direzioni opposte.

⁽²⁾ Si ammette il principio di Pauli per qualunque sistema contenente più elettroni: p. es. la « statistica di Fermi », valida per un gas di elettroni, e in particolare per gli elettroni di conduzione dei metalli, è fondata su tale principio. Si ha poi ragione di ritenere che il principio di Pauli valga anche per i protoni e per alcuni nuclei; per altri invece (p. es. le particelle α) esso non vale.

⁽¹⁾ *Zs. f. Phys.*, **31**, 765 (1925).

si trovano nella prima orbita ma il terzo deve trovarsi nell'orbita immediatamente successiva (in ordine di energia), che è $n = 2, l = 0, m = 0$: questo elettrone è quindi legato al nucleo meno strettamente degli altri, e funge perciò da elettrone di valenza (per cui il Li è monovalente) e ottico (quindi il termine fondamentale è $2s$). Nel Be ($Z = 4$) un secondo elettrone viene a occupare la seconda orbita (dove la valenza due, e il termine fondamentale $2s$): nel B ($Z = 5$) le prime due orbite essendo completamente occupate, il quinto elettrone va in un'orbita $n = 2, l = 1$, il che spiega che il termine fondamentale del B è $2p$. Così, aggiungendo via via un elettrone nello stato di energia più bassa rimasto libero compatibilmente col principio di Pauli (e, naturalmente, aumentando ogni volta di una unità la carica del nucleo) si determina lo stato quantico dei vari elettroni per i successivi elementi: in particolare lo stato quantico dell'ultimo elettrone aggiunto definisce il termine fondamentale dello spettro. Le previsioni così ottenute sono in mirabile accordo con le proprietà dei vari elementi sperimentalmente osservate. L'esistenza dei cosiddetti « periodi » si spiega col fatto che gli elettroni successivamente aggiunti occupano (di regola) prima le orbite con $l = 0$, poi con $l = 1$, poi con $l = 2, \dots$ fino ad $n - 1$, che è il massimo l compatibile con un dato n : dopo di che si passa alle orbite con $n + 1$ cominciando di nuovo da $l = 0, l = 1$ ecc. E poichè da un facile calcolo risulta che fissato n vi sono n^2 orbite, cioè $2n^2$ posti, si avrà un primo periodo di $2 \cdot 1^2 = 2$ elementi (H ed He), un secondo di $2 \cdot 2^2 = 8$ e così via.

Per maggiori particolari si veda p. es. bibl. n. 27.

Nell'enunciare il principio di Pauli, ci siamo finora serviti del linguaggio della teoria di Bohr e Sommerfeld. Ma nella meccanica ondulatoria quel principio ha trovato, per opera di DIRAC, una formulazione ancora più semplice e più generale, poichè si applica anche agli stati non stazionari. Essa dice che: *in un sistema contenente più elettroni, sono possibili solo gli stati antisimmetrici* (nel senso spiegato a pag. 471). Anche la meccanica ondulatoria non fornisce nessuna dimostrazione di questo principio (che chiameremo di *antisimmetria*) il quale pertanto deve considerarsi come un postulato: essa però dimostra (come si è visto) che se un sistema soddisfa in un certo istante a questo principio, seguita a soddisfarvi sempre.

Faremo ora vedere che dal principio di antisimmetria di Dirac discende come conseguenza necessaria il principio di esclusione di Pauli. Infatti, supponiamo che il sistema sia in uno stato stazionario, e poniamolo in un campo magnetico la cui intensità, dapprima debolissima, faremo poi crescere fino a un valore tale da poter trascurare le interazioni fra i diversi elettroni, come si è detto a pag. 474: sia allora $\psi_{n_i}(q^i)$ la autofunzione dell' i -esimo elettrone (n_i sta per il gruppo dei quattro numeri n_i, l_i, m_i, s_i , e q^i per x_i, y_i, z_i, σ_i): essendo i vari elettroni dinamicamente indipendenti, l'equazione del sistema complessivo ammette come autofunzione il prodotto delle autofunzioni corrispondenti ai singoli elettroni che lo compongono (v. § 20), cioè $\psi_{n_1}(q^1) \psi_{n_2}(q^2) \dots \psi_{n_N}(q^N)$. Ma, essendo l'equazione simmetrica rispetto alle N quaterne di coordinate, essa sarà soddisfatta (in corrispondenza al medesimo autovalore) anche da ogni funzione che sia ottenuta dalla precedente permutando in modo qualunque le q^i , cioè da

$$(365) \quad \psi_{n_1}(q^{r_1}) \psi_{n_2}(q^{r_2}) \dots \psi_{n_N}(q^{r_N})$$

dove r_1, r_2, \dots, r_N rappresenta una qualunque permutazione dei numeri $1, 2, \dots, N$. La soluzione generale sarà una combinazione lineare di tutte quelle così ottenute. Di queste combinazioni ve ne è una simmetrica, che è la somma di tutti i $N!$ prodotti del tipo (365):

$$(366) \quad \psi_s = \sum \psi_{n_1}(q^{r_1}) \psi_{n_2}(q^{r_2}) \dots \psi_{n_N}(q^{r_N})$$

(la sommatoria essendo estesa a tutte le permutazioni r_1, r_2, \dots, r_N), ed una antisimmetrica, la quale si può scrivere sotto forma di determinante nel modo seguente:

$$(367) \quad \psi_a = \begin{vmatrix} \psi_{n_1}(q^1) & \psi_{n_1}(q^2) & \dots & \psi_{n_1}(q^N) \\ \psi_{n_2}(q^1) & \psi_{n_2}(q^2) & \dots & \psi_{n_2}(q^N) \\ \dots & \dots & \dots & \dots \\ \psi_{n_N}(q^1) & \psi_{n_N}(q^2) & \dots & \psi_{n_N}(q^N) \end{vmatrix};$$

(infatti, lo sviluppo di questo determinante non è altro che la somma di tutte le espressioni del tipo (365), precedute dal

segno + o — secondo che la permutazione r_1, r_2, \dots, r_N è pari o dispari; l'espressione è poi evidentemente antisimmetrica, perchè lo scambio di due qualunque delle q , equivalendo allo scambio di due colonne del determinante, non fa che cambiare il segno di questo). Se ora ammettiamo il principio di antisimmetria, dovremo adottare ψ_a come unica possibile fra tutte le anzidette autofunzioni. Ciò posto, se vi fossero nel sistema due o più elettroni con gli stessi numeri quantici, vi sarebbero nel determinante due o più righe uguali e perciò la ψ_a sarebbe identicamente nulla: dunque il principio di antisimmetria porta ad escludere la possibilità di più elettroni nello stesso stato quantico e quindi contiene in sé il principio di Pauli.

Reciprocamente, il principio di antisimmetria è conseguenza necessaria del principio di Pauli, purchè questo si intenda in un senso alquanto più restrittivo di quello esposto precedentemente: cioè purchè si ammetta non solo che non esistano nel sistema due elettroni nello stesso stato quantico, ma che non sia possibile portare il sistema in uno stato siffatto, qualunque siano le forze a cui lo si assoggetta (purchè, naturalmente, agenti in modo simmetrico su tutti gli elettroni). E infatti, se la ψ , anzichè antisimmetrica, fosse simmetrica (abbiamo già visto al § 63 che non possono darsi altri casi), sarebbe possibile con una conveniente hamiltoniana (simmetrica), farla evolvere in modo che dopo un certo tempo si identifichi con la autofunzione simmetrica ψ_s data dalla (366) nella quale si siano fissati comunque gli indici n_1, n_2, \dots, n_N , anche prendendone quanti si vogliano uguali tra loro.

Il principio di antisimmetria dunque non solo contiene in sé il principio di Pauli, ma rappresenta la condizione necessaria per assicurarne la validità permanente in un sistema, qualunque siano le circostanze fisiche a cui questo è assoggettato. In questo senso, esso si può considerare equivalente al principio di Pauli, e difatti spesso viene chiamato anch'esso con questo nome.

§ 65. — CALCOLO APPROSSIMATO DEGLI STATI DI UN SISTEMA DI DUE PARTICELLE UGUALI.

In un sistema costituito di due particelle, ciascuna di queste si può considerare soggetta: *a*) a forze indipendenti dalla presenza della seconda particella; *b*) a forze (elettriche e magnetiche) dovute all'altra particella. In molti casi queste ultime forze si possono considerare piccole rispetto alle prime, così da poterle trascurare in una approssimazione preliminare (approssimazione d'ordine zero), e introdurre poi, col metodo delle perturbazioni, in una prima approssimazione e, occorrendo, in approssimazioni successive. Con questo metodo, come diremo, si possono trattare l'atomo di elio e la molecola di idrogeno, che sono le due applicazioni più interessanti di questa teoria ⁽¹⁾.

Poniamoci dapprima nell'approssimazione zero, cioè trascuriamo l'interazione tra le due particelle. Possiamo allora considerare ciascuna di esse come se appartenesse a un sistema distinto (prescindere cioè dalla sovrapposizione spaziale dei due sistemi) e, in uno stato stazionario, attribuire alla prima particella una autofunzione $\psi_{n_1}(1)$, alla seconda una $\psi_{n_2}(2)$ (dove n_1, n_2 , come al solito, rappresentano due gruppi di quattro numeri quantici, ed 1, 2, due gruppi di quattro coordinate): queste autofunzioni soddisfano le due equazioni

$$(368) \quad \mathfrak{H}_1^0 \psi_{n_1} = E_{n_1} \psi_{n_1}, \quad \mathfrak{H}_2^0 \psi_{n_2} = E_{n_2} \psi_{n_2},$$

dove si è indicato con \mathfrak{H}_1^0 l'hamiltoniano relativo alla prima particella, con \mathfrak{H}_2^0 quello della seconda (trascurando la loro interazione): questi operatori naturalmente hanno la stessa forma, e il secondo differisce dal primo solo per la materiale sostituzione delle lettere x_1, y_1, z_1, σ_1 con x_2, y_2, z_2, σ_2 .

⁽¹⁾ Per ottenere una precisione maggiore, si può anche includere nell'approssimazione zero una parte della forza di interazione, schematizzandola come un campo centrale agente su ciascuna delle particelle e uguale, presso a poco, al valor medio della forza esercitata su di essa dall'altra (*azione di schermo*, v. § 59, p. II): la perturbazione si riduce allora alla sola differenza tra l'interazione vera e questa. Ciò non altera nulla alla sostanza di quanto segue.

Quindi $\psi_{n_1 s_1}$ e $\psi_{n_2 s_2}$ si ottengono, in sostanza, risolvendo un medesimo problema di autovalori: basta scrivere l'equazione seguente (che riproduce le (368), ma senza alcun indice):

$$(369) \quad \mathfrak{H}^0 \psi(x, \sigma) = E \psi(x, \sigma),$$

poi prenderne una particolare autofunzione ψ_{n_1} e porre in essa l'indice 1 alle variabili, e un'altra particolare autofunzione ψ_{n_2} (che può eventualmente coincidere con la precedente), e porre in essa l'indice 2 alle variabili: E_{n_1} , E_{n_2} sono i corrispondenti autovalori, eventualmente uguali. Se ora consideriamo le due particelle come appartenenti a un unico sistema (sempre però trascurando le forze reciproche) sappiamo (v. § 20) che a questo spetta l'operatore $\mathfrak{H}_1^0 + \mathfrak{H}_2^0$, e che il suo stato, quando la prima particella è nello stato (n_1) e la seconda nello stato (n_2), è espresso dal prodotto delle due rispettive autofunzioni, cioè dalla seguente ψ :

$$(370) \quad \psi_1^0 = \psi_{n_1}(q^1) \psi_{n_2}(q^2);$$

L'energia corrispondente è

$$(371) \quad E^0 = E_{n_1} + E_{n_2}.$$

Ora, per l'osservazione fondamentale fatta al § 62, se nella funzione (371) si scambiano le variabili q^1 con le q^2 si ottiene ancora un'autofunzione del sistema, appartenente

allo stesso autovalore: indicandola con ψ_2^0 avremo

$$(372) \quad \psi_2^0 = \psi_{n_1}(q^2) \psi_{n_2}(q^1).$$

Supponiamo ora che il gruppo dei quattro indici n_1 non sia identico al gruppo n_2 , vale a dire che le due particelle siano in stati diversi: allora ψ_2^0 risulta diversa da ψ_1^0 , e quindi il livello energetico (371) risulta doppio (indipendentemente da altre eventuali degenerazioni, da cui prescindiamo). Dunque: *nell'approssimazione zero si ha una particolare degenerazione derivante dall'identità delle due particelle, in virtù della quale tutti i livelli energetici risultano doppi, tranne quelli corrispondenti a stati quantici uguali delle due particelle.* Questa degenerazione si chiama, per una ragione

che verrà spiegata tra breve, *degenerazione di scambio o di risonanza.*

Passiamo ora alla prima approssimazione, tenendo conto dell'azione reciproca tra le due particelle, che introdurremo come perturbazione, seguendo i metodi del cap. IV. L'hamiltoniano del sistema si scriverà allora sotto la forma

$$(373) \quad \mathfrak{H}_1^0 + \mathfrak{H}_2^0 + \mathfrak{Q},$$

dove l'operatore \mathfrak{Q} corrisponde all'energia mutua delle due particelle, e quindi contiene in modo simmetrico le variabili q^1 e le q^2 (p. es., se l'azione mutua si riduce alla repulsione elettrostatica, sarà $\mathfrak{Q} = e^2/r_{12}$, dove

$$r_{12} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2};$$

se invece si tien conto anche delle azioni magnetiche, \mathfrak{Q} conterrà anche le variabili di spin σ_1 e σ_2 , ma sempre in modo simmetrico). A causa della degenerazione di scambio, si dovranno applicare le formule trovate al § 39, e quindi si dovrà anzitutto costruire la « matrice di perturbazione » i cui elementi sono

$$(374) \quad \begin{cases} L_{11} = \sum_{s_1, s_2} \int \psi_1^0 * \mathfrak{Q} \psi_1^0 dS, & L_{12} = \sum_{s_1, s_2} \int \psi_1^0 * \mathfrak{Q} \psi_2^0 dS \\ L_{21} = \sum_{s_1, s_2} \int \psi_2^0 * \mathfrak{Q} \psi_1^0 dS, & L_{22} = \sum_{s_1, s_2} \int \psi_2^0 * \mathfrak{Q} \psi_2^0 dS, \end{cases}$$

dove ogni integrale è sestuplo, e dS sta per $dx_1 dy_1 dz_1 dx_2 dy_2 dz_2$, mentre le sommatorie sono doppie, e sostituiscono l'integrazione rispetto alle due variabili discontinue s_1 , s_2 (numeri quantici di spin), ciascuna delle quali assume solo due valori. Si noti ora che in ciascuna di queste espressioni è naturalmente lecito cambiare la designazione delle variabili d'integrazione e degli indici di sommatoria senza alterare il risultato: ebbene, se si scambiano le variabili q^1 con le q^2 , si muta ψ_1^0 in ψ_2^0 e viceversa, quindi L_{11} in L_{22} , L_{12} in L_{21} e viceversa: dunque

$$(375) \quad L_{11} = L_{22} \quad L_{12} = L_{21}.$$

In prima approssimazione, la perturbazione ε' del valore dell'energia si trova, come si è visto al § 39, risolvendo l'equazione

$$(376) \quad \begin{vmatrix} L_{11} - \varepsilon' & L_{12} \\ L_{21} & L_{22} - \varepsilon' \end{vmatrix} = 0,$$

la quale, in virtù delle (375), fornisce per ε' i due valori seguenti (la ragione della notazione ε'_s , ε'_a apparirà tra breve):

$$(377) \quad \varepsilon'_s = L_{11} + L_{12}, \quad \varepsilon'_a = L_{11} - L_{12},$$

e quindi il livello (371) si scinde nei due livelli

$$(378) \quad E_s = E_{n_1} + E_{n_2} + L_{11} + L_{12}$$

$$(378') \quad E_a = E_{n_1} + E_{n_2} + L_{11} - L_{12}.$$

Ricerchiamo ora le autofunzioni di approssimazione zero corrispondenti a questi autovalori: esse sono date (v. § 39) da:

$$(379) \quad \begin{cases} \psi_s^0 = c_{11}^0 \psi_1^0 + c_{12}^0 \psi_2^0, \\ \psi_a^0 = c_{21}^0 \psi_1^0 + c_{22}^0 \psi_2^0. \end{cases}$$

dove i coefficienti c^0 sono ricavati dalle (185'), che nel caso attuale si scrivono, prendendo $k = 1$ (per $k = 2$ si avrebbe un sistema equivalente):

$$\begin{aligned} c_{11}^0(L_{11} - \varepsilon'_s) + c_{12}^0 L_{12} &= 0 \\ c_{21}^0(L_{11} - \varepsilon'_a) + c_{22}^0 L_{12} &= 0, \end{aligned}$$

ossia, tenuto conto delle (375) e (377), e supponendo ⁽¹⁾ che sia $L_{12} \neq 0$,

$$c_{11}^0 = c_{12}^0 \quad c_{21}^0 = -c_{22}^0.$$

Normalizzando ψ_s e ψ_a si trova che il modulo di questi coefficienti deve essere $1/\sqrt{2}$, cosicchè si può scrivere

$$(380) \quad \psi_s^0 = \frac{1}{\sqrt{2}} (\psi_1^0 + \psi_2^0), \quad \psi_a^0 = \frac{1}{\sqrt{2}} (\psi_1^0 - \psi_2^0)$$

ossia

$$(381) \quad \psi_s^0 = \frac{1}{\sqrt{2}} [\psi_{n_1}(q^1)\psi_{n_2}(q^2) + \psi_{n_1}(q^2)\psi_{n_2}(q^1)]$$

$$(381') \quad \psi_a^0 = \frac{1}{\sqrt{2}} [\psi_{n_1}(q^1)\psi_{n_2}(q^2) - \psi_{n_1}(q^2)\psi_{n_2}(q^1)]$$

⁽¹⁾ Il caso $L_{12} = 0$ si verifica quando le due nuvole di probabilità rappresentate dalle funzioni ψ_{n_1} , ψ_{n_2} non hanno punti in comune: in tal caso a ciascuno dei due elettroni è assegnata una regione separata dello spazio ed è come se ciascuno avesse la sua individualità: non vi è dunque luogo al fenomeno di scambio.

Come si vede, la prima di queste autofunzioni è simmetrica, la seconda antisimmetrica: ciò era prevedibile per quanto è stato detto al § 62. Nel caso che le particelle siano elettroni o protoni il principio di Pauli porta ad escludere la prima di esse, e resta solo la ψ_a^0 a rappresentare uno stato possibile del sistema: quindi la (378') da sola rappresenta tutti i livelli energetici fisicamente possibili.

Il caso, finora escluso, che le due particelle abbiano gli stessi numeri quantici ($n_1 = n_2$), non dà luogo a degenerazione: si ha quindi in tal caso una unica autofunzione, che è simmetrica. Quando si tratta di particelle che obbediscono al principio di Pauli, questi stati sono evidentemente da escludersi.

Accenneremo ora brevemente a una considerazione che, sebbene abbia valore puramente formale, spiega il nome di fenomeni di scambio o di risonanza, dato generalmente ai fenomeni caratteristici dei sistemi con particelle uguali ⁽¹⁾. Supponiamo perciò che le due particelle in questione, pur essendo dinamicamente uguali, siano in qualche modo distinguibili tra loro, mediante un «segno» che non ne alteri le proprietà meccaniche (supposizione evidentemente fittizia e priva di significato fisico). In questa ipotesi è possibile distinguere tra il caso in cui la particella 1 è nello stato n_1 , e la 2 nello stato n_2 e quello in cui esse sono scambiate, cioè tra i due stati del sistema rappresentati, in approssimazione zero, da ψ_1^0 e ψ_2^0 . Sarà allora ammissibile, oltre a ψ_s e a ψ_a , qualunque loro combinazione lineare

$$(382) \quad \psi = c_s \psi_s + c_a \psi_a,$$

con c_s e c_a costanti. Poichè alla ψ_s corrisponde in prima approssimazione l'autovalore E_s e a ψ_a l'autovalore E_a , esse potranno scriversi sotto la forma:

$$\psi_s = u_s e^{-\frac{2\pi i}{\hbar} E_s t}, \quad \psi_a = u_a e^{-\frac{2\pi i}{\hbar} E_a t},$$

dove u_s e u_a sono funzioni delle coordinate, ma non di t , ed E_s e E_a sono dati da (378) e (378'). Sostituendo in (382), e ponendo

$$(383) \quad C_s = c_s e^{-\frac{2\pi i}{\hbar} (L_{11} + L_{12}) t}, \quad C_a = c_a e^{-\frac{2\pi i}{\hbar} (L_{11} - L_{12}) t},$$

⁽¹⁾ V. W. HEISENBERG, *ZS. f. Phys.*, **38**, (1926) p. 411.

si ha

$$(384) \quad \psi = (C_s u_s + C_a u_a) e^{-\frac{2\pi i}{h} E^0 t}.$$

Si noti che C_s e C_a non sono costanti come c_s e c_a , ma variano periodicamente e lentamente: la loro frequenza $(L_{11} \pm L_{12})/h$ è infatti piccola, rispetto alla frequenza E^0/h del fattore principale di ψ , come la perturbazione è piccola rispetto all'energia. Se la soluzione (384) si considera per un tempo breve rispetto al periodo di variazione di C_s e C_a , così da poter riguardare queste come costanti, essa si può approssimativamente identificare con una combinazione lineare

$$(385) \quad \psi = C_s \psi_s^0 + C_a \psi_a^0$$

delle soluzioni di approssimazione zero (380): in altre parole, la nostra ψ si può considerare una combinazione lineare, a coefficienti lentamente variabili, di ψ_s^0 e ψ_a^0 , quindi anche di ψ_1^0 e ψ_2^0 : si ha infatti, sostituendo le (380) nella (384),

$$\psi = \frac{1}{\sqrt{2}}(C_s + C_a)\psi_1^0 + \frac{1}{\sqrt{2}}(C_s - C_a)\psi_2^0.$$

Supponiamo ora che, al tempo 0, si sia constatato che la particella 1 è nello stato n_1 e la 2 nello stato n_2 , vale a dire, che la ψ è rappresentata (approssimativamente) da ψ_1^0 : allora dalla formula precedente e dalle (383), ponendovi $t = 0$ e $\psi = \psi_1^0$, si avrà $c_s = c_a = 1/\sqrt{2}$ e quindi la ψ per un tempo t qualunque potrà scriversi

$$(386) \quad \psi = e^{-\frac{2\pi i}{h} L_{11} t} \left[\psi_1^0 \cos \frac{2\pi}{h} L_{12} t + i \psi_2^0 \sin \frac{2\pi}{h} L_{12} t \right].$$

Ora, il quadrato del modulo del coefficiente di ψ_1^0 , cioè $\cos^2 \frac{2\pi}{h} L_{12} t$, rappresenta la probabilità di trovare al tempo t il sistema nello stato ψ_1^0 , cioè la particella 1 nello stato n_1 e la 2 nello stato n_2 , mentre analogamente $\sin^2 \frac{2\pi}{h} L_{12} t$ rappresenta la probabilità di trovare le due particelle scambiate. Come si vede, al tempo $\tau/2$, con $\tau = h/2L_{12}$, le due particelle si trovano con certezza scambiate, al tempo τ sono di nuovo come al tempo 0, e così di seguito: negli istanti intermedi, lo stato del sistema è tale che non è definito quale particella è nello stato n_1 e quale nello stato n_2 . Il periodo τ con cui oscillano le due probabilità è tanto più lungo, quanto minore è l'accoppiamento tra le due particelle, misurato dall'integrale L_{12} , che chiamasi « integrale di scambio ».

Questo curioso fenomeno di scambio è analogo, sotto l'aspetto analitico, allo scambio periodico di energia che si verifica tra due oscillatori di ugual frequenza e leggermente « accoppiati »: p. es., tra due pendoli della stessa lunghezza, appesi a uno stesso filo orizzontale non troppo teso: questa analogia meccanica ha suggerito il nome di « risonanza ».

§ 66. — SEPARAZIONE DELLE COORDINATE DI SPIN DA QUELLE DI POSIZIONE.

In molti casi, in un sistema con due elettroni le forze dovute agli spin sono trascurabili in prima approssimazione. Formalmente, ciò significa che le variabili di spin non intervengono nè in \mathfrak{S}^0 , nè nella parte principale del termine di interazione \mathfrak{L} , ma solo in un termine di ordine superiore che nella prima approssimazione si trascura. Quando sono realizzate queste condizioni, si dice che tra i due elettroni vi è l'*accoppiamento di Russell e Saunders*.

La prima di queste ipotesi consente (come si è visto al § 45 nell'approssimazione di Pauli) di trattare il problema imperturbato separando la variabile di spin σ da quelle di posizione, cioè di scrivere una generica soluzione della (369) nella forma

$$(387) \quad \psi_{n_i s_i} = \Psi_{n_i}(x_i) \varphi_{s_i}(\sigma_i)$$

dove $i = 1, 2$, ed n_i rappresenta, ora, solo il gruppo dei tre numeri quantici *orbitali* dell'elettrone i -esimo, mentre il numero quantico di spin, s_i , è ora considerato a parte: similmente x_i sta per le tre coordinate di posizione x_i, y_i, z_i . Il fattore Ψ_{n_i} soddisfa l'equazione di Schrödinger

$$(388) \quad \mathfrak{S}^0 \Psi_{n_i} = E_{n_i} \Psi_{n_i},$$

mentre l'autofunzione di spin φ_{s_i} si riduce, in sostanza, a un gruppo di due costanti α_{s_i} e β_{s_i} (corrispondenti rispettivamente ai due valori ± 1 della variabile σ_i) cosicchè si può scrivere, indicando con s un indice che può assumere i due valori ± 1 ,

$$\varphi_s(\sigma) = \begin{vmatrix} \alpha_s \\ \beta_s \end{vmatrix}.$$

Si noti che, poichè s può assumere solo due valori, esistono solo due « funzioni φ », ossia coppie (α_s, β_s) : supporremo tali « funzioni » ortogonali e normalizzate, vale a dire:

$$(389) \quad \sum_{\sigma} \varphi_{s_1}(\sigma) \varphi_{s_2}^*(\sigma) = \alpha_{s_1} \alpha_{s_2}^* + \beta_{s_1} \beta_{s_2}^* = \delta_{s_1 s_2}.$$

Ciò premesso, le due autofunzioni di approssimazione zero, simmetrica e antisimmetrica, (381), (381'), si scrivono ora, tenuto conto della (387):

$$\begin{aligned} \psi_s^0 &= \frac{1}{\sqrt{2}} [\Psi_{n_1}(x_1) \Psi_{n_2}(x_2) \varphi_{s_1}(\sigma_1) \varphi_{s_2}(\sigma_2) + \Psi_{n_1}(x_2) \Psi_{n_2}(x_1) \varphi_{s_1}(\sigma_2) \varphi_{s_2}(\sigma_1)] \\ \psi_a^0 &= \frac{1}{\sqrt{2}} [\Psi_{n_1}(x_1) \Psi_{n_2}(x_2) \varphi_{s_1}(\sigma_1) \varphi_{s_2}(\sigma_2) - \Psi_{n_1}(x_2) \Psi_{n_2}(x_1) \varphi_{s_1}(\sigma_2) \varphi_{s_2}(\sigma_1)]. \end{aligned}$$

Ora formiamo con le autofunzioni posizionali Ψ le seguenti combinazioni, simmetrica la prima e antisimmetrica la seconda:

$$(390) \quad \begin{cases} \mathfrak{S}(x_1, x_2) = \frac{1}{2} [\Psi_{n_1}(x_1) \Psi_{n_2}(x_2) + \Psi_{n_1}(x_2) \Psi_{n_2}(x_1)] \\ \mathfrak{A}(x_1, x_2) = \frac{1}{2} [\Psi_{n_1}(x_1) \Psi_{n_2}(x_2) - \Psi_{n_1}(x_2) \Psi_{n_2}(x_1)] \end{cases},$$

e similmente con le autofunzioni di spin:

$$(391) \quad \begin{cases} \mathfrak{S}(\sigma_1, \sigma_2) = \frac{1}{2} [\varphi_{s_1}(\sigma_1) \varphi_{s_2}(\sigma_2) + \varphi_{s_1}(\sigma_2) \varphi_{s_2}(\sigma_1)] \\ \mathfrak{A}(\sigma_1, \sigma_2) = \frac{1}{2} [\varphi_{s_1}(\sigma_1) \varphi_{s_2}(\sigma_2) - \varphi_{s_1}(\sigma_2) \varphi_{s_2}(\sigma_1)]. \end{cases}$$

Allora ψ_s^0 e ψ_a^0 si possono esprimere mediante queste nuove combinazioni, e divengono:

$$(392) \quad \psi_s^0 = \sqrt{2} (\mathfrak{S}\mathfrak{S} + \mathfrak{A}\mathfrak{A})$$

$$(392') \quad \psi_a^0 = \sqrt{2} (\mathfrak{S}\mathfrak{A} + \mathfrak{A}\mathfrak{S}).$$

Si osservi ora che, poichè trascuriamo le forze dovute agli spin, gli autovalori risultano indipendenti dai numeri quantici di spin s_1, s_2 , dipendendo solo dai rimanenti numeri quantici n_1, n_2 : da ciò deriva una ulteriore degenerazione che ora vogliamo esaminare.

Fissati n_1 ed n_2 , vi sono per s_1 ed s_2 le seguenti quattro possibilità:

$$(393) \quad \begin{cases} 1) & s_1 = +1, & s_2 = +1 \\ 2) & s_1 = -1, & s_2 = -1 \\ 3) & s_1 = +1, & s_2 = -1 \\ 4) & s_1 = -1, & s_2 = +1, \end{cases}$$

cui corrispondono altrettante autofunzioni simmetriche del tipo (392), che indicheremo rispettivamente con $\psi_s^1, \psi_s^2, \psi_s^3, \psi_s^4$, e altrettante antisimmetriche del tipo (392'), che chiameremo $\psi_a^1, \psi_a^2, \psi_a^3, \psi_a^4$: esse corrispondono tutte, in approssimazione zero, all'autovalore $E_{n_1} + E_{n_2}$.

Quando però si introduce la perturbazione \mathfrak{Q} , tali autofunzioni devono essere sostituite (v. § 39) con altrettante combinazioni lineari opportunamente scelte, che indicheremo con $\bar{\psi}_s^i, \bar{\psi}_a^i$ ($i = 1, 2, 3, 4$), di cui le $\bar{\psi}_s^i$ saranno formate con le sole ψ_s , le $\bar{\psi}_a^i$ con le sole ψ_a , perchè, come si è detto al § 62, non è fisicamente ammissibile una combinazione lineare di una ψ_s con una ψ_a . Anzi, poichè parliamo di elettroni, che obbediscono al principio di Pauli, dovremo escludere le autofunzioni simmetriche $\bar{\psi}_s^i$, e quindi ci occuperemo d'ora in avanti solo delle $\bar{\psi}_a^i$. Esse sono date dalle formule

$$\bar{\psi}_a^i = \sum_j c^{ij} \psi_a^j,$$

dove i coefficienti c^{ij} sono ottenuti (v. § 39) mediante i quattro sistemi di equazioni lineari:

$$(394) \quad \begin{cases} (L^{11} - \varepsilon^i) c^{i1} + L^{12} c^{i2} + L^{13} c^{i3} + L^{14} c^{i4} = 0, \\ L^{21} c^{i1} + (L^{22} - \varepsilon^i) c^{i2} + L^{23} c^{i3} + L^{24} c^{i4} = 0, \\ L^{31} c^{i1} + L^{32} c^{i2} + (L^{33} - \varepsilon^i) c^{i3} + L^{34} c^{i4} = 0, \\ L^{41} c^{i1} + L^{42} c^{i2} + L^{43} c^{i3} + (L^{44} - \varepsilon^i) c^{i4} = 0, \end{cases}$$

e gli elementi L^{jl} della matrice di perturbazione sono dati da

$$(395) \quad L^{jl} = \sum_{\sigma_1 \sigma_2} \int \psi_a^j * \mathfrak{Q} \psi_a^l dS;$$

ε^i è una delle radici ottenute eguagliando a zero il determinante dei coefficienti, e l'autofunzione $\bar{\psi}_a^i$ appartiene all'autovalore

$$E_{n_1} + E_{n_2} + \varepsilon^i.$$

Per calcolare le L^i , calcoliamo, mediante le (391), le autofunzioni di spin, corrispondenti alle quattro coppie di valori (393) per s_1 ed s_2 ; otteniamo:

$$(396) \quad \begin{cases} \varrho_1 = \varphi_1(\sigma_1)\varphi_1(\sigma_2) & \varrho_2 = \varphi_{-1}(\sigma_1)\varphi_{-1}(\sigma_2) \\ \varrho_3 = \varrho_2 = \frac{1}{2}[\varphi_1(\sigma_1)\varphi_{-1}(\sigma_2) + \varphi_1(\sigma_2)\varphi_{-1}(\sigma_1)] \\ \alpha_1 = 0 & \alpha_2 = 0 \\ \alpha_3 = -\alpha_4 = \frac{1}{2}[\varphi_1(\sigma_1)\varphi_{-1}(\sigma_2) - \varphi_1(\sigma_2)\varphi_{-1}(\sigma_1)]. \end{cases}$$

Tenendo conto della (392') e badando che

$$(397) \quad \int \mathfrak{S}^* \mathfrak{Q} \mathfrak{Q} dS = \int \mathfrak{Q}^* \mathfrak{Q} \mathfrak{S} dS = 0$$

(a causa della antisimmetria delle funzioni integrande), la (395) si scrive:

$$(398) \quad L^i = 2I_S \sum_{\sigma_1 \sigma_2} \varphi^* \varrho_i + 2I_A \sum_{\sigma_1 \sigma_2} \varphi^* \varrho_i,$$

dove si è posto

$$(399) \quad I_S = \int \mathfrak{S}^* \mathfrak{S} dS, \quad I_A = \int \mathfrak{Q}^* \mathfrak{Q} dS.$$

Le due sommatorie doppie si calcolano, per le varie coppie (j, l), utilizzando le (391) e la (389), e si trova così in definitiva per la matrice delle L^i il risultato seguente:

$$(399) \quad \begin{pmatrix} 2I_A & 0 & 0 & 0 \\ 0 & 2I_A & 0 & 0 \\ 0 & 0 & (I_A + I_S) & (I_A - I_S) \\ 0 & 0 & (I_A - I_S) & (I_A + I_S) \end{pmatrix}.$$

L'equazione secolare che dà le ε^i si può dunque scrivere:

$$(2I_A - \varepsilon)^2 \begin{vmatrix} I_A + I_S - \varepsilon & I_A - I_S \\ I_A - I_S & I_A + I_S - \varepsilon \end{vmatrix} = 0$$

e le sue radici sono:

$$(400) \quad \varepsilon^1 = \varepsilon^2 = \varepsilon^3 = 2I_A, \quad \varepsilon^4 = 2I_S.$$

Sostituendo nei sistemi (394) le (399) e le (400) si trova che, per $i = 1, 2, 3$, le prime due equazioni sono identicamente soddisfatte e le altre due equivalgono (supponendo,

come faremo ⁽¹⁾, $I_S \neq I_A$) a $c^{i3} = c^{i4}$, mentre c^{i1}, c^{i2}, c^{i3} restano arbitrarie: per $i = 4$ si trova invece dalle prime due equazioni $c^{41} = c^{42} = 0$, e dalle altre due $c^{43} = -c^{44}$. In conclusione si ha:

$$\begin{aligned} \text{per } i = 1, 2, 3, \quad \bar{\psi}_a^i &= c^{i1} \psi_a^1 + c^{i2} \psi_a^2 + c^{i3} (\psi_a^3 + \psi_a^4); \\ \text{per } i = 4, \quad \bar{\psi}_a^4 &= c^{43} (\psi_a^3 - \psi_a^4); \end{aligned}$$

i coefficienti c contenuti in queste formule restando arbitrari, salvo le condizioni di normalizzazione. Tenendo conto della (392') e delle (396), e ponendo

$$\bar{\varrho}_i = c^{i1} \varrho_1 + c^{i2} \varrho_2 + 2c^{i3} \varrho_3,$$

$$\bar{\alpha} = 2c^{43} \alpha_3,$$

le formule precedenti si scrivono:

$$(401) \quad \begin{cases} \bar{\psi}_a^i = \sqrt{2} \mathfrak{Q} \bar{\varrho}_i & (i = 1, 2, 3) \\ \bar{\psi}_a^4 = \sqrt{2} \mathfrak{S} \bar{\alpha}. \end{cases}$$

(Si noti che, nel caso $n_1 = n_2$, si ha $\mathfrak{Q} = 0$ e quindi mancano gli stati corrispondenti ad $i = 1, 2, 3$).

Le tre prime autofunzioni corrispondono (nella nostra approssimazione) all'autovalore

$$(402) \quad E_{n_1} + E_{n_2} + 2I_A,$$

mentre la quarta corrisponde all'autovalore

$$(402') \quad E_{n_1} + E_{n_2} + 2I_S.$$

Il livello primitivo $E_{n_1} + E_{n_2}$ si scinde dunque, in generale, in un livello triplo ⁽²⁾ (402), e in un livello semplice (402'): per $n_1 = n_2$ manca il livello triplo. Come si vede

⁽¹⁾ Il caso $I_S = I_A$ corrisponde al caso in cui l'integrale di scambio è nullo; v. nota al paragrafo precedente.

⁽²⁾ Se si tien conto delle forze dovute agli spin, questo livello triplo si scinde a sua volta in tre livelli semplici, eccettuato il caso che l'autofunzione posizionale \mathfrak{Q} rappresenti uno stato a quanto azimutale nullo (« stato S »), nel qual caso l'energia di perturbazione dovuta agli spin risulta nulla e i tre livelli del tripletto rimangono coincidenti.

dalle (401), le autofunzioni del tripletto si ottengono moltiplicando una autofunzione di Schrödinger antisimmetrica per un'autofunzione di spin simmetrica, mentre la ψ del singoletto è il prodotto di un'autofunzione di Schrödinger simmetrica per una di spin antisimmetrica. La separazione tra il livello semplice e il triplo risulta espressa da $2(I_S - I_A)$, ossia, come si trova facilmente tenendo presenti le (399) e (390), da

$$(402'') \quad \int \Psi_{n_1}^*(x_1) \Psi_{n_2}^*(x_2) \mathfrak{E} \Psi_{n_1}(x_2) \Psi_{n_2}(x_1) dS.$$

Questo non è altro che l'integrale di scambio definito al § precedente, fatto però tenendo conto solo delle autofunzioni in x, y, z , senza i fattori di spin.

Questo risultato trova un'espressione suggestiva nel modello vettoriale. In questo modello, a un'autofunzione di spin simmetrica si fanno corrispondere due spin orientati nello stesso senso, ossia « paralleli », mentre un'autofunzione di spin antisimmetrica corrisponde a spin orientati in senso opposto o « antiparalleli » ⁽¹⁾. Si può allora dire che, quando vale l'accoppiamento di Russell e Saunders, è lecito usare l'ordinaria teoria di Schrödinger anzichè quella di Dirac, ma aggiungendovi la seguente limitazione (che esprime il principio di Pauli): che l'autofunzione *nelle sole coordinate*

⁽¹⁾ Infatti, si consideri l'operatore che rappresenta il quadrato dello spin totale, e che è:

$$\sigma^2 = (\sigma_x^{(1)} + \sigma_x^{(2)})^2 + (\sigma_y^{(1)} + \sigma_y^{(2)})^2 + (\sigma_z^{(1)} + \sigma_z^{(2)})^2,$$

dove $\sigma_x^{(1)}, \sigma_y^{(1)}, \sigma_z^{(1)}$ sono gli operatori corrispondenti alle componenti dello spin del primo elettrone (formati a norma del § 45) e $\sigma_x^{(2)}, \sigma_y^{(2)}, \sigma_z^{(2)}$ sono quelli del secondo. Applicando σ^2 a una funzione (delle variabili di spin) antisimmetrica α o ad una simmetrica β , si trova rispettivamente

$$\sigma^2 \alpha = 0 \quad \sigma^2 \beta = 8\beta.$$

Ciò significa, per il principio generale della meccanica quantistica, che una misura dello spin totale σ darebbe nel primo caso 0, nel secondo $\sqrt{8}$. (Il modello vettoriale darebbe: per spin antiparalleli $\sigma = 0$, per spin paralleli $\sigma = 2$; la divergenza numerica di questo ultimo valore da $\sqrt{8}$ dipende dalla già notata imperfezione del modello vettoriale: v. p. es. § 57, p. II).

posizionali sia antisimmetrica se gli spin sono paralleli, e sia simmetrica se sono antiparalleli: nel primo caso si ha un livello triplo (che il modello vettoriale interpreta mediante le tre orientazioni che può assumere lo spin totale rispetto al campo magnetico orbitale), mentre nel secondo caso (essendo lo spin risultante nullo) non si può avere che un livello semplice. Se poi i due elettroni sono « nella stessa orbita », cioè se i loro numeri quantici orbitali sono uguali (simbolicamente: $n_1 = n_2$), i loro spin non possono disporsi paralleli, perchè in tal caso i due elettroni avrebbero tutti e quattro i numeri quantici uguali, il che è vietato dal principio di Pauli: per $n_1 = n_2$ manca dunque il tripletto, e si ha solo il singoletto.

Un tratto caratteristico della suesposta teoria quantistica dello « scambio » è che la differenza di energia tra il livello semplice e quello triplo corrispondente risulta, come si è visto, anche trascurando le azioni magnetiche sugli spin, mentre, secondo il modello vettoriale, i due stati differirebbero solo per l'orientazione reciproca degli spin, e quindi la loro differenza di energia sarebbe dovuta solo alle forze reciproche tra i loro momenti magnetici. La ragione è che, nella teoria precedente, gli spin, sebbene non intervengano nella espressione dell'hamiltoniana, intervengono indirettamente nell'applicazione del principio di Pauli, poichè obbligano a scegliere l'autofunzione (nelle coordinate posizionali) simmetrica, o quella antisimmetrica, secondo che gli spin sono antiparalleli o paralleli. Si può dimostrare ⁽¹⁾ che ciò equivale formalmente ad ammettere una fortissima energia reciproca tra gli spin, proporzionale al coseno del loro angolo: si giustifica così il successo che questa ipotesi modellistica ha avuto nella interpretazione degli spettri di multipletti, prima della meccanica quantistica.

Se si svolgessero i calcoli di questo § prendendo le funzioni *simmetriche* $\bar{\psi}_s$ invece delle $\bar{\psi}_a$, si troverebbe, con procedimento identico, che, fissati n_1 ed n_2 , vi è un tripletto di autofunzioni della forma $\psi_s^i = \sqrt{2} \bar{\mathfrak{S}}_i$ ($i = 1, 2, 3$) e una autofunzione singola della forma $\bar{\psi}_s^4 = \sqrt{2} \bar{\mathfrak{A}}_4$: vale a dire, i livelli tripli corrisponderebbero in questa ipotesi agli spin paralleli, i semplici agli spin antiparalleli: per $n_1 = n_2$, essendo $\mathfrak{A} = 0$, mancherebbe lo

⁽¹⁾ V. bibl. n. 6, p. 226.

stato semplice anzichè quello triplo. Come si dirà al § seguente, i risultati sperimentali sullo spettro dell'elio e analoghi contraddicono questa ipotesi, confermando invece pienamente il principio di antisimmetria.

Le considerazioni precedenti ci hanno condotti a dividere gli stati antisimmetrici (rispetto allo scambio di tutte le coordinate) in due classi, a seconda che sono *simmetrici* o *antisimmetrici* rispetto allo scambio delle sole coordinate posizionali: i termini spettrali della prima classe costituiscono il « sistema dei tripletti », quelli della seconda il « sistema dei singoletti ». Dimosteremo ora un teorema analogo a quello dimostrato nel § 62, e cioè che i passaggi tra le due classi di stati sono vietati, ossia che i due sistemi di termini spettrali non si combinano tra loro. Mentre però il teorema del § 62 è rigorosamente vero, quello che dimosteremo ora vale soltanto nell'approssimazione di questo paragrafo, cioè trascurando le forze di spin: quindi tali passaggi non si devono ritenere come assolutamente impossibili, ma soltanto come molto improbabili. E difatti, come si dirà più oltre a proposito dello spettro dell'elio, qualcuna delle righe spettrali ad essi corrispondenti è stata osservata, ma con intensità assai debole. La dimostrazione è del tutto analoga a quella del § 62: poichè l'operatore \mathfrak{S} , nell'approssimazione adottata, non coinvolge le variabili di spin, esso è necessariamente simmetrico rispetto alle variabili posizionali: quindi dalla (364) risulta che se $\psi(t)$ è simmetrica (o antisimmetrica) rispetto alle coordinate posizionali, tale viene ad essere anche $d\psi$ e quindi $\psi(t + dt)$.

§ 67. — SPETTRO DELL'ATOMO DI ELIO E ANALOGHI.

I risultati precedenti trovano una delle più importanti applicazioni nel caso dell'atomo di elio. Già molto prima che sorgesse la meccanica quantistica si era osservato che le righe spettrali dell'elio (non ionizzato) si possono ripartire in due classi, per ciascuna delle quali separatamente era stato trovato un sistema di termini (v. fig. 48). Ciò aveva fatto pensare che si trattasse di due specie distinte di atomi, a

cui fu dato il nome di *parelio* e *ortoelio* ⁽¹⁾. Le righe del parelio sono tutte semplici, quelle dell'ortoelio invece, esaminate con mezzi di alto potere risolutivo, rivelano generalmente

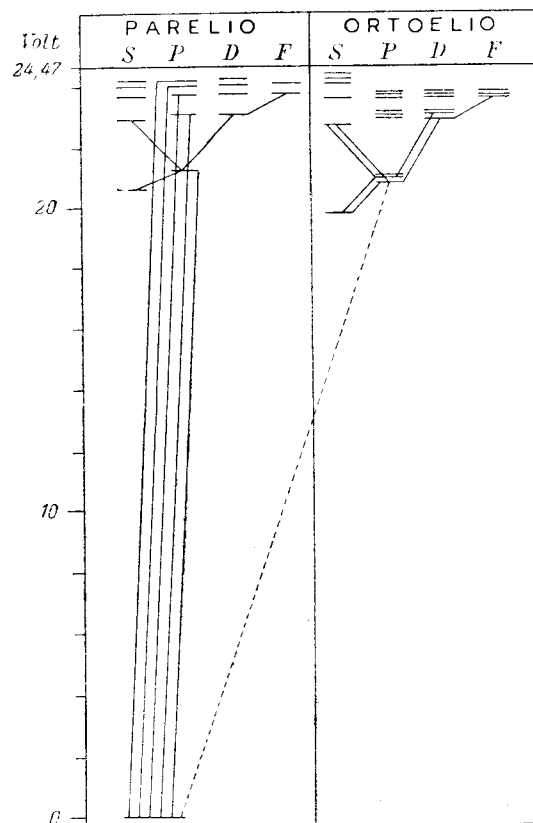


Fig. 48

una struttura fina, che permette di attribuirle a un sistema di tripletti ⁽²⁾. Il livello fondamentale (assai più basso degli

⁽¹⁾ La ragione di questi nomi sta nel fatto che si credeva che le orbite dei due elettroni fossero complanari nel parelio e ortogonali nell'ortoelio: i nomi sono rimasti anche oggi, benchè il modello (che non ha mai condotto a risultati quantitativamente soddisfacenti) sia stato abbandonato.

⁽²⁾ Ciò non significa che i termini dell'ortoelio siano tutti tripli, ma che essi sono tripli ad eccezione dei termini *S* che sono sempre sem-

altri) appartiene al sistema dei singoletti e corrisponde a uno stato $1s$: manca il corrispondente termine nel sistema dei tripletti. È stata poi osservata una riga di combinazione, cioè corrispondente al passaggio da un livello dei tripletti a uno dei singoletti (v. linea tratteggiata nella fig. 48), ma la sua intensità è assai tenue, il che significa che siffatti passaggi hanno probabilità molto piccola.

Questi risultati sperimentali sono in eccellente accordo con la teoria esposta nei §§ precedenti: gli atomi del parelio sono quelli a spin antiparalleli (cioè quelli la cui ψ è antisimmetrica nelle coordinate posizionali), gli atomi dell'ortoelio sono quelli a spin paralleli (o a ψ simmetrica nelle coordinate posizionali). Nello stato fondamentale, come è naturale, entrambi gli elettroni sono nello stato più basso che è $1s$: essi devono quindi avere gli spin antiparalleli per obbedire al principio di Pauli, e perciò tale stato appartiene ai singoletti e non ha corrispondente nei tripletti (è questa la conferma, annunciata nel § precedente, del postulato che le ψ che si realizzano sono quelle antisimmetriche e non le simmetriche). Negli stati eccitati, uno degli elettroni continua a occupare l'orbita $1s$, mentre l'altro passa a orbite superiori: perciò i loro spin possono disporsi tanto paralleli quanto antiparalleli.

Accenneremo ora al confronto quantitativo fra teoria ed esperienza, che costituisce uno dei più brillanti successi della meccanica quantistica, tanto più che tutti i precedenti tentativi di interpretazione teorica sulle basi della teoria di Sommerfeld o di altre erano completamente falliti. La teoria della meccanica quantistica contiene in sé gli elementi per calcolare non solo il valore numerico dei singoli termini spettrali, ma anche tutte le altre costanti fisiche dell'atomo di elio, come suscettività diamagnetica, costante dielettrica,

plici (si veda a tal proposito il volume sulla spettroscopia, o anche il n. 27 della bibl.). Aggiungeremo che i tripletti dell'ortoelio sono molto serrati, così che molti non si possono risolvere, o si risolvono solo parzialmente, apparendo come doppietti: la loro natura di tripletti è però confermata dagli spettri analoghi (Li^+ , Be^{++} , B^{+++} , C^{++++}) in cui le separazioni sono molto maggiori (come di regola).

costante di Van der Waals, punto di ebollizione, ecc.: tali calcoli sono però assai laboriosi, dovendosi necessariamente adoperare metodi di approssimazioni successive, e hanno dato origine a numerosi lavori di carattere matematico ⁽¹⁾: ci limiteremo a citare qualche esempio dei risultati. La quantità che è stata calcolata con maggior cura è il valore numerico del termine fondamentale (σ , in altri termini, l'energia di ionizzazione): si è trovato

$$198.308 \text{ cm.}^{-1}$$

contro il risultato sperimentale:

$$198.298 \pm 8 \text{ cm.}^{-1}$$

lo scarto relativo è di 1 : 20.000. L'energia di ionizzazione risulta di circa 24,47 volt. Per gli stati eccitati il calcolo numerico risulta più difficile e non è stato spinto ad una approssimazione così alta, ma è sempre in soddisfacente accordo con l'esperienza. Sull'effetto di un campo elettrico o magnetico ci limiteremo a citare i risultati seguenti:

	calcolata	osservata
Suscettività magnetica	$-1,87 \cdot 10^{-6}$	$-1,88 \cdot 10^{-6}$
Costante dielettrica	1,000.071	1,000.074 .

Spettri analoghi a quello dell'elio sono forniti, come è naturale, dagli ioni Li^+ , Be^{++} , B^{+++} , C^{++++} , i quali hanno due elettroni come l'atomo di He , e differiscono da questo solo per il nucleo: le formule dell' He quindi si applicano anche ad essi, cambiando i valori numerici. L'accordo con l'esperienza è anche in questi casi eccellente, come mostra la tabella seguente, che dà i valori del termine fondamentale in cm^{-1} :

	calcolato	osservato
Li^+	609.985	610.090 \pm 100
Be^{++}	1.241.222	1.241.350 \pm 200
B^{+++}	2.091.770	2.092.000 \pm 300
C^{++++}	3.161.770	3.161.900 \pm 800

⁽¹⁾ Un'esposizione d'insieme di questi lavori, con le relative indicazioni bibliografiche, si trova nel n. 1 della bibl., da cui sono tolti i dati numerici che riportiamo.

Infine, aggiungeremo che gli atomi della seconda colonna del sistema periodico, o alcalino-terrosi, presentano pure una certa analogia con l'atomo di elio, in quanto si possono schematizzare col criterio del § 59, p. II, considerando i due elettroni di valenza come soggetti all'attrazione del nucleo diminuita dall'azione « di schermo » dei rimanenti $Z - 2$ elettroni (i quali formano strati completi, e quindi possono con discreta approssimazione essere sostituiti da una carica spaziale statica: v. nota a pag. 267). Il problema differisce così da quello dell'elio perchè il campo in cui si muovono i due elettroni non è newtoniano. Naturalmente la teoria non può in questo caso essere spinta a previsioni quantitative molto esatte, ma le sue previsioni qualitative sono confermate dall'esperienza: si ha infatti anche in questi casi un sistema di singoletti e uno di tripletti (questi ultimi, con separazioni assai maggiori che nell'elio): le combinazioni tra singoletti e tripletti sono però in questo caso assai meno improbabili, come prevede la teoria e conferma l'osservazione. Ma per questa, come per le altre applicazioni spettroscopiche della meccanica quantistica, rimandiamo al volume sulla spettroscopia.

§ 68. — CENNO SU ALTRE APPLICAZIONI.

La teoria, esposta in questo capitolo, dei sistemi con particelle identiche — nella quale è caratteristico il concetto quantistico di *risonanza* o *scambio* spiegato al § 65 — è stata applicata con successo anche allo studio degli urti di un elettrone lento contro un atomo di idrogeno o di elio ⁽¹⁾, nel quale fenomeno lo scambio tra l'elettrone urtante e quelli dell'atomo urtato ha una influenza considerevole sulla distribuzione angolare degli elettroni dopo l'urto, influenza che si rivela nelle curve sperimentali. Un'altra applicazione importante è la spiegazione, dovuta ad HEISENBERG ⁽²⁾, del-

⁽¹⁾ V. MOTT e MASSEY, *The theory of atomic collisions*, Oxford, 1933; oppure: G. WENTZEL, *Wellenmechanik der Stoss und Strahlungsvorgänge*, in GEIGER e SCHEEL, *Handbuch der Physik*, Bd. XXIV, 2^a ediz., Berlino, 1933.

⁽²⁾ *ZS. f. Phys.*, 49, (1928), p. 619.

l'origine del cosiddetto « campo molecolare » di Weiss, che è necessario a spiegare il ferromagnetismo, ma di cui non si era finora trovata una soddisfacente giustificazione. Importantissima poi, tra le applicazioni del concetto di scambio, è la teoria della molecola di idrogeno, H_2 , dovuta ad HEITLER e LONDON, che per la prima volta hanno dato una soddisfacente spiegazione della forza che tiene insieme i due atomi di idrogeno, la quale è dovuta essenzialmente al citato fenomeno di scambio o risonanza tra i due elettroni, e chiamasi perciò « forza di risonanza »: per l'esposizione di questa teoria rinviamo al volume « Molecole e cristalli » del presente trattato. Qui ci limitiamo a dire che siffatte « forze di risonanza » hanno in generale una parte essenziale nelle azioni tra atomi vicini, costituenti una molecola o facenti parte di un reticolo cristallino, e che la teoria fisica della valenza, inaugurata da HEITLER e LONDON col caso della molecola H_2 , è stata estesa, almeno qualitativamente, a molecole più complesse, con notevole successo, e sembra suscettibile di ulteriori sviluppi.

BIBLIOGRAFIA ⁽¹⁾

A) Trattati di meccanica quantistica o ondulatoria.

1. BETHE H. - *Quantenmechanik der Ein- und Zwei-Elektronenprobleme*, in « Handbuch der Physik », di H. Geiger e K. Scheel, vol. XXIV, parte I, cap. 3 (Berlino, 1933, 2^a ediz.). Da integrare col n. 14.
2. BLOCH E. - *L'ancienne et la nouvelle théorie des quanta* (Parigi, 1930).
3. CONDON E. U. e MORSE P. M. - *Quantum mechanics* (New York e Londra, 1929).
4. DE BROGLIE L. - *La théorie de la quantification dans la nouvelle mécanique* (Parigi, 1932).
5. DE BROGLIE L. - *Introduction à l'étude de la mécanique ondulatoire* (Parigi, 1930).
6. DIRAC P. A. M. - *The principles of quantum mechanics* (Oxford, 1935).
- 6 bis. DIRAC P. A. M. - *Les principes de la mécanique quantique* (Parigi, 1931).
7. FRENKEL J. - *Wave mechanics, elementary theory* (Oxford, 1932).
8. FRENKEL J. - *Wave mechanics, advanced general theory* (Oxford, 1934).
9. HAAS A. - *Materiewellen und Quantenmechanik* (Lipsia, 1934).
10. HEISENBERG W. - *Die physikalischen Prinzipien der Quantentheorie* (Lipsia, 1930).
- 10 bis. HEISENBERG W. - *Les principes physiques de la théorie des quanta* (Parigi, 1932).
11. MARCH A. - *Die Grundlagen der Quantenmechanik* (Lipsia, 1931).
12. MOTT N. F. - *An outline of wave mechanics* (Cambridge, 1930).

⁽¹⁾ Ci limitiamo a indicare qualcuna delle numerose opere riguardanti gli argomenti trattati in questo volume. Indicazioni bibliografiche più ampie si troveranno nelle opere n. 2, 8, 20.

13. v. NEUMANN J. - *Mathematische Grundlagen der Quantenmechanik* (Lipsia 1931).
14. PAULI W. - *Die allgemeinen Prinzipien der Wellenmechanik*, in «*Handbuch der Physik*» di H. Geiger e K. Scheel, vol. XXIV, parte I, cap. 2 (Berlino, 1933, 2ª ediz.). Da integrare col n. 1.
15. PERSICO E. - *Lezioni di meccanica ondulatoria* (Padova, 1935).
16. RUARK A. E. e UREY H. C. - *Atoms, molecules and quanta* (New York e Londra, 1930).
17. SCHRÖDINGER E. - *Abhandlungen zur Wellenmechanik* (Lipsia, 1928).
- 17 bis. SCHRÖDINGER E. - *Mémoires sur la mécanique ondulatoire* (Parigi, 1933).
18. SOMMERFELD A. - *Wellenmechanischer Ergänzungsband* (Braunschweig, 1929).
20. WEIL H. - *Gruppentheorie und Quantenmechanik* (Lipsia, 1931).

B) Opere da consultare per argomenti speciali.

21. BOHR N. - *La théorie atomique et la description des phénomènes* (Parigi, 1932).
22. BRILLOUIN L. - *Notions de mécanique ondulatoire; les méthodes d'approximation* (Parigi, 1932).
23. BRUNETTI R. - *L'atomo e le sue radiazioni* (Bologna, 1932).
24. CASTELFRANCHI G. - *Fisica Moderna* (Milano, 1931).
25. COURANT R. e HILBERT D. - *Methoden der Mathematischen Physik* (Berlino, 1931).
26. DE BROGLIE L. - *L'électron magnétique* (Parigi, 1934).
27. FERMI E. - *Introduzione alla fisica atomica* (Bologna, 1928).
28. FRANK J. e JORDAN P. - *Anregung von Quantensprüngen durch Stöße* (Lipsia, 1926).
29. GAMOW G. - *Der Bau des Atomkerns und die Radioaktivität* (Lipsia, 1932).
30. MARK H. e WIERL R. - *Die experimentellen und theoretischen Grundlagen der Elektronenbeugung* (Berlino, 1931).
31. MILLIKAN R. A. - *Electrons (+ and -), protons, photons, neutrons and cosmic rays* (Cambridge, 1935).
32. PERSICO E. - *Ottica* (Milano, 1932).
33. PINCHERLE S. e AMALDI U. - *Le operazioni distributive e le loro applicazioni all'analisi* (Bologna, 1901).
34. RIEMANN-WEBER - *Partielle Differentialgleichungen der Mathematischen Physik herausgegeben von R. v. Mises* (Braunschweig, 1925).
36. TONELLI L. - *Serie trigonometriche* (Bologna, 1928).

INDICE ANALITICO

(I numeri romani indicano la parte e quelli arabi il paragrafo)

- | | |
|--|--|
| <p>Accoppiamento di RUSSELL e SAUNDERS, III, 66.</p> <p>Ampiezza di probabilità, II, 25.</p> <p>Annichilazione, III, 60.</p> <p>Antisimmetria, principio di —, III, 64.</p> <p>Autodifferenziale, II, 10.</p> <p>Autofunzioni, II, 2, 17, 49; III, 10.</p> <p>Autofunzioni normalizzate, II, 4, 17; III, 2.</p> <p>Autofunzioni ortogonali, II, 5, 17; III, 2.</p> <p>Autovalori, II, 2, 17, 28, 48; III, 10, 35.</p> <p>AVOGADRO, numero di —, I, 1.</p> <p>BALMER, serie di —, I, 14.</p> <p>Barriera di potenziale, II, 40.</p> <p>BOHR, magnetone di —, I, 25; II, 61, 62.</p> <p>BOHR, teoria di —, I, 16.</p> <p>BOTHE e GEIGER, esperienza di —, I, 11.</p> <p>Caso puro, III, 27.</p> <p>Coefficiente di riflessione, II, 37.</p> <p>Coefficiente di trasmissione, II, 37, 40.</p> <p>COMPTON e DEBYE, teoria di —, I, 10.</p> | <p>COMPTON, effetto —, I, 9.</p> <p>COMPTON e SIMON, esperienza di —, I, 12.</p> <p>Condizione di normalizzazione, II, 27.</p> <p>Condizioni di SOMMERFELD, I, 16; II, 52, 56, 58.</p> <p>Corpo nero, I, 7.</p> <p>Costante della struttura fina, II, 60.</p> <p>Costante di PLANCK, I, 7; II, 52.</p> <p>DAVISSON e GERMER, esperienze di —, I, 29, 30, 31.</p> <p>DE BROGLIE, onde di —, I, 29.</p> <p>Degenerazione, II, 6, 28; III, 39.</p> <p>Degenerazione di scambio, III, 65.</p> <p>Densità di probabilità, II, 19, 25.</p> <p>DIRAC, funzione di —, III, 14.</p> <p>DIRAC, teoria di —, II, 32; III, 47, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60.</p> <p>Effetto COMPTON, I, 9.</p> <p>Effetto fotoelettrico, I, 2, 8.</p> <p>Effetto termoionico, I, 2.</p> <p>Effetto ZEEMAN, I, 2, 25.</p> <p>EHRENFEST, teorema di —, III, 28.</p> <p>Elettroni di valenza, II, 59.</p> <p>Elettrone ottico, I, 25; II, 59.</p> |
|--|--|

- Elettrone rotante, I, 25; II, 62.
 Energia di risonanza, I, 18.
 Equazione autoaggiunta, II, 17.
 Equazione dei moti armonici, II, 8.
 Equazione del raggio, II, 12.
 Equazione unidimensionale di
 SCHRÖDINGER, II, 33.
 Equazione temporale di SCHRÖ-
 DINGER, II, 30; III, 19, 20, 41.
 Esperienza di BOTHE e GEIGER,
 I, 11.
 Esperienza di COMPTON e SIMON,
 I, 12.
 Esperienze di DAVISSON e GERMER,
 I, 29, 30, 31.
 Esperienze di FRANCK e HERTZ,
 I, 19.
 Esperienze di PONTE, I, 33.
 Esperienze di RUPP, I, 32.
 Esperienze di THOMSON, I, 32.

 Fattore di normalizzazione, II, 4.
 Fenomeni elementari, III, 27.
 FERMAT, principio di —, II, 26.
 Fluorescenza indiretta, I, 24.
 Formula di RAYLEIGH e JEANS,
 I, 7.
 Formula di RITZ, I, 15.
 Formula di RYDBERG, I, 15.
 Fotone, I, 8; II, 18.
 FOURIER, sviluppo di —, II, 9, 11,
 12; III, 2.
 FRANCK e HERTZ, esperienze di —,
 I, 19, 20.
 Funzioni associate di LEGENDRE,
 II, 46.
 Funzione di DIRAC, III, 14.

 Hamiltoniana, II, 52; III, 19.
 HERMITE, polinomi di —, II, 39.

 Indice di rifrazione delle onde di
 DE BROGLIE, II, 26.
 Integrali fondamentali, II, 1.
 Integrali primi, III, 29.

 Ipotesi della struttura corpusco-
 lare della materia, I, 1.
 Ipotesi di PLANCK, I, 7.
 Ipotesi di RUTHERFORD, I, 3.
 Ipotesi di THOMSON, I, 3.
 Ipotesi di UHLENBECK e GOUD-
 SMIT, I, 25; III, 44, 45.
 Isotopi, I, 4.

 Legge di STOKES, I, 23.
 Livelli energetici, I, 16; II, 28, 56.
 Livelli multipli, II, 60.
 LYMAN, serie di —, I, 14.

 Magnetone di BOHR, I, 25; II, 61,
 62.
 Matrici, III, 5, 6, 6-bis, 8, 33, 40.
 Matrici continue, III, 13.
 Matrici diagonali, III, 12.
 Matrici di trasformazione, III, 7.
 Matrici hermitiane, III, 9.
 Meccanica ondulatoria, I, 28.
 Metodo della separazione delle va-
 riabili, II, 17.
 Metodo delle matrici, I, 27.
 Metodo di SCHRÖDINGER, II, 27,
 57; III, 19, 20, 46.
 Metodo di WENTZEL e BRILLOUIN,
 II, 51.
 Miscuglio, III, 27.
 Momenti angolari, III, 30.
 Moti armonici, equazione dei —,
 II, 8.

 Nodi delle autofunzioni, II, 7.
 Numeri quantici, I, 25; II, 52.
 Numero atomico, I, 3.
 Numero di schermo, II, 59.
 Numero d'onde, II, 12.
 Numero di AVOGADRO, I, 1.
 Numero di RYDBERG, I, 14, 16;
 II, 58.

 Onde di DE BROGLIE, I, 29.
 Operatori hermitiani, III, 9.
 Operatori incompleti, III, 10.

- Operatori lineari, III, 3, 4, 5, 8, 22.
 Operatori permutabili, III, 11.
 Ortoelio, III, 67.
 Oscillatore armonico, II, 39, 54,
 64; III, 36.
 Osservabili, III, 15, 16, 17, 22, 25.
 Osservazioni massime, III, 18.

 Pacchetto d'onde, II, 15, 25.
 Parelio, III, 67.
 Particelle, varie specie di —, I, 4.
 PASCHEN, serie di —, I, 14.
 PAULI, principio di —, III, 64.
 PAULI, teoria di —, II, 25; III,
 47, 51.
 Perturbazioni, III, 37, 38, 39, 40,
 41, 42, 43.
 PICKERING, serie di —, I, 14.
 PLANCK, costante di —, I, 7; II,
 52.
 PLANCK, ipotesi di —, I, 7.
 Polinomi di HERMITE, II, 39.
 Polinomi di LAGUERRE, II, 49.
 Polinomi di LEGENDRE, II, 46.
 PONTE, esperienze di —, I, 33.
 Positrone, III, 60.
 Potenziale di ionizzazione, I, 18, 19.
 Potenziale di risonanza, I, 18, 19,
 20.
 Principio di antisimmetria, III, 64.
 Principio di combinazione, I, 15.
 Principio di corrispondenza, II, 32,
 63.
 Principio di FERMAT, II, 26.
 Principio di indeterminazione, I,
 27; II, 21, 22, 44; III, 27.
 Principio di PAULI, III, 64.
 Principio di selezione, II, 64.
 Probabilità di transizione, III, 42.

 Quanti di luce, I, 8.
 Quantizzazione spaziale, II, 56.
 Quanto azimutale, II, 59.
 Quanto interno, I, 25; II, 62.
 Quanto magnetico, II, 59.

 Quanto radiale, II, 56.
 Quanto totale, II, 59.

 RAYLEIGH e JEANS, formula di —
 I, 7.
 Regole di selezione, II, 32, 50,
 60, 64.
 Relazione di indeterminazione, II,
 21.
 Relazioni di permutazione, III, 34.
 Resto atomico, I, 25; II, 59.
 Risonanza, III, 43, 65, 68.
 Risonanza ottica, I, 22.
 RITZ, formula di —, I, 15.
 RYDBERG, formula di —, I, 15.
 RYDBERG, numero di —, I, 14, 16.
 Rotatore, II, 55, 64.
 RUPP, esperienze di —, I, 32.
 RUSSELL e SAUNDERS, accoppia-
 mento di —, III, 66.
 RUTHERFORD, ipotesi di —, I, 3.

 SCHRÖDINGER, equazione unidimen-
 sionale di —, II, 33.
 SCHRÖDINGER, equazione temporale
 di —, II, 30; III, 19, 20, 41.
 SCHRÖDINGER, metodo di —, II,
 27, 57; III, 19, 20, 46.
 Separazione delle variabili, metodo
 della —, II, 17.
 Serie di BALMER, I, 14.
 Serie di LYMAN, I, 14.
 Serie di PASCHEN, I, 14.
 Serie di PICKERING, I, 14.
 Singolarità essenziale, II, 16.
 Singolarità fuchsiana, II, 16.
 Sistemi idrogenoidi, I, 16; II, 48,
 56, 60; III, 21, 59.
 SOMMERFELD, condizioni di —, I,
 16; II, 52, 56, 58.
 Soglia fotoelettrica, I, 8.
 Sostituzione ortogonale, II, 6.
 Spazio funzionale, III, 1.
 Spazio hilbertiano, III, 2.
 Spin, I, 25; II, 62; III, 44, 45,
 50, 51, 52, 53, 54.

- Spostamento radiattivo, I, 5.
 Stati antisimmetrici, III, 62, 63.
 Stati quantici, II, 27; III, 18.
 Stati simmetrici, III, 62, 63.
 Stati stazionari, II, 27, 32; III, 18, 38.
 STOKES, legge di —, I, 23.
 Struttura corpuscolare della materia, ipotesi della —, I, 1.
 Struttura fina, I, 14; II, 60; III, 39, 59.
 Struttura fina, costante della —, II, 60.
 Struttura iperfina, I, 14.
 Sviluppo di FOURIER, II, 9, 11, 12; III, 2.

 Teorema di EHRENFEST, III, 28.
 Teoria dell'emissione dei raggi α , II, 43.
 Teoria delle trasformazioni, I, 35.
 Teoria di BOHR, I, 16.
 Teoria di COMPTON e DEBYE, I, 10.
- Teoria di DIRAC, II, 32; III, 47, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60.
 Teoria di PAULI, II, 25; III, 47, 51.
 THOMSON, esperienze di —, I, 32.
 THOMSON, ipotesi di —, I, 3.

 UHLENBECK e GOUDSMIT, ipotesi di —, I, 25; III, 44, 45.
 Urto di seconda specie, I, 21.

 Varie specie di particelle, I, 4.
 Velocità di fase, II, 14.
 Velocità di gruppo, II, 14.
 Vettore di propagazione, II, 15.
 Vettore di stato, III, 19.

 WENTZEL e BRILLOUIN, metodo di —, II, 51.

 ZEEMAN, effetto —, I, 2, 25.

INDICE DEL VOLUME

PARTE I.

SVILUPPO STORICO E BASI SPERIMENTALI

CAP. I. — <i>Il modello dell'atomo</i>	Pag. 5
1. Il concetto di atomo nella chimica e nella fisica	» 5
2. L'elettrone	» 8
3. I corpuscoli positivi. L'atomo di RUTHERFORD	» 10
4. La struttura del nucleo	» 14
5. Legge dello spostamento radioattivo. Spiegazione dell'isotopia	» 16
CAP. II. — <i>I quanti di energia e i quanti di luce</i>	» 19
6. Inapplicabilità delle leggi classiche agli atomi	» 19
7. Lo spettro del corpo nero	» 20
8. L'effetto fotoelettrico e i « quanti di luce »	» 23
9. L'effetto COMPTON	» 27
10. Teoria quantistica dell'effetto COMPTON	» 28
11. Esperienza di BOTHE e GEIGER	» 31
12. Esperienza di COMPTON e SIMON	» 31
13. Impossibilità di una teoria letteralmente corpuscolare della luce	» 33
CAP. III. — <i>I livelli energetici</i>	» 37
14. Lo spettro dell'idrogeno e dei sistemi idrogenoidi	» 37
15. Serie spettrali. Termini spettroscopici	» 41
16. Cenno sulla teoria di BOHR dell'atomo di idrogeno	» 43
17. Stati quantici e livelli energetici	» 48
18. Eccitazione e ionizzazione per urto elettronico	» 50
19. Esperienze di FRANCK e HERTZ	» 54
20. Controllo ottico della eccitazione per urto	» 57
21. Urto di seconda specie	» 59
22. Risonanza ottica	» 59
23. Fluorescenza	» 61
24. Fluorescenza sensibilizzata	» 62

25. Cenno sullo sviluppo della spettroscopia. L'elettrone rotante	Pag. 63
26. La crisi delle teorie modellistiche	» 66
CAP. IV. — <i>La meccanica quantistica</i> » 69	
27. La meccanica quantistica di HEISENBERG (metodo delle matrici)	» 69
28. La meccanica ondulatoria	» 70
29. La diffrazione degli elettroni e degli atomi	» 73
30. Esperienze eseguite col metodo di LAUE	» 75
31. Esperienze eseguite col metodo di BRAGG	» 79
32. Esperienze eseguite col metodo di DEBYE e SCHERRER	» 81
33. Dipendenza della lunghezza d'onda dalla velocità	» 81
34. Assetto logico della meccanica atomica. La definizione operativa dei concetti fisici	» 84
35. Il metodo degli operatori e la teoria di DIRAC dell'elettrone	» 88

PARTE II.

MECCANICA ONDULATORIA DI UNA PARTICELLA

CAP. I. — <i>Introduzione matematica</i> Pag. 91	
1. Generalità sulle equazioni lineari omogenee del secondo ordine	» 91
2. Equazioni con un parametro. Autovalori ed autofunzioni	» 94
3. Forma autoaggiunta di un'equazione	» 96
4. Normalizzazione delle autofunzioni	» 97
5. Ortogonalità delle autofunzioni	» 98
6. Degenerazione (autovalori multipli)	» 99
7. Nodi delle autofunzioni	» 101
8. L'equazione dei moti armonici	» 102
9. Sviluppo di una funzione in serie di autofunzioni. Serie di FOURIER	» 105
10. Caso di un intervallo infinito. Spettro continuo di autovalori	» 107
11. Sviluppo in integrale di FOURIER	» 113
12. Interpretazione ondulatoria dello sviluppo di FOURIER	» 114
13. Lunghezza di un gruppo d'onde e larghezza della riga spettrale corrispondente	» 118
14. Velocità di gruppo	» 122
15. Onde nello spazio. Pacchetti d'onde	» 124
16. Cenno sulle equazioni con singolarità	» 128
17. Cenno sulle equazioni a derivate parziali	» 130

CAP. II. — <i>Impostazione probabilistica dei problemi atomici</i> . Pag. 133	
18. Analisi del concetto di fotone	» 133
19. Significato probabilistico dell'ottica ondulatoria	» 138
20. Principio di sovrapposizione dell'ottica	» 140
21. Principio di indeterminazione per i fotoni	» 142
22. Principio di indeterminazione per le particelle materiali	» 145
23. Misura delle coordinate di una particella	» 148
24. Misura dell'impulso o della velocità di una particella	» 153
25. Impostazione probabilistica della meccanica di una particella	» 155
26. Moto di una particella e moto di un pacchetto di onde	» 159
27. L'equazione di SCHRÖDINGER	» 163
28. I livelli energetici come autovalori dell'equazione di SCHRÖDINGER	» 166
29. Principio di sovrapposizione della meccanica ondulatoria	» 167
30. Equazione temporale di SCHRÖDINGER	» 169
31. Densità di corrente	» 170
32. Campo elettromagnetico e radiazione	» 172
CAP. III. — <i>Problemi unidimensionali</i> » 175	
33. Caratteristiche dei problemi unidimensionali	» 175
34. Discussione qualitativa dei problemi unidimensionali	» 176
35. Particella di energia determinata e non soggetta a forze	» 178
36. Particella di energia indeterminata (gruppo d'onde)	» 180
37. Gradino di potenziale	» 185
38. Particella su un segmento	» 189
39. Oscillatore armonico	» 192
40. Barriera di potenziale	» 198
41. Cenno sulla teoria dell'emissione elettronica a freddo	» 204
42. Particella tra due barriere di potenziale	» 206
43. Cenno sulla teoria dell'emissione dei raggi α	» 208
CAP. IV. — <i>Problemi a tre dimensioni</i> » 211	
44. Particella non soggetta a forze	» 211
45. Particella in una scatola parallelepipedica	» 215
46. Forze centrali (parte generale)	» 216
46. L'equazione radiale nel problema delle forze centrali	» 223
48. Teoria dei sistemi idrogenoidi. Gli autovalori	» 225
49. Le autofunzioni dei sistemi idrogenoidi per $E < 0$	» 230
50. Regole di selezione	» 236
CAP. V. — <i>La teoria di Bohr e Sommerfeld</i> » 239	
51. Il metodo di WENTZEL e BRILLOUIN	» 239
52. Le condizioni di SOMMERFELD	» 245

53. Cenzo sulla scelta del sistema di coordinate	Pag. 250
54. Oscillatore armonico	» 251
55. Rotatore e atomo di BOHR	» 252
56. Sistemi idrogenoidi	» 254
57. Confronto con la teoria di SCHRÖDINGER	» 261
58. Movimento del nucleo	» 264
59. Atomi con più elettroni	» 266
60. Cenzo sulla teoria relativistica dei sistemi idrogenoidi	» 271
61. Il magnetone di BOHR	» 273
62. L'elettrone rotante e il quanto interno	» 276
63. Principio di corrispondenza	» 280
64. Principio di selezione	» 284

PARTE III.

METODI GENERALI DELLA MECCANICA QUANTISTICA

CAP. I. - <i>Introduzione matematica</i>	Pag. 291
1. Lo spazio funzionale	» 291
2. Lo spazio hilbertiano	» 293
3. Operatori lineari	» 297
4. Algebra degli operatori lineari	» 299
5. Rappresentazione di un operatore lineare mediante una matrice	» 303
6. Algebra delle matrici	» 305
6 bis. Rappresentazione di una funzione mediante una matrice di una sola colonna	» 307
7. Cambiamenti di assi nello spazio hilbertiano. Matrici di trasformazione	» 308
8. Rappresentazione di uno stesso operatore in diversi sistemi di riferimento	» 311
9. Operatori e matrici hermitiani	» 312
10. Assi principali di un operatore	» 315
11. Teorema fondamentale sugli operatori permutabili	» 319
12. Matrici diagonali	» 321
13. Matrici a indici continui	» 323
14. Funzione di DIRAC	» 326
CAP. II. - <i>Teoria generale della meccanica quantistica</i>	» 329
15. Concetto di « osservabile »	» 329
16. Compatibilità delle osservabili	» 331
17. Osservabili funzioni di altre	» 332
18. Osservazioni massime. Stato di un sistema	» 335
19. Interpretazione del metodo di SCHRÖDINGER nello spazio hilbertiano	» 337

20. Estensione a un sistema di N particelle distinte	Pag. 339
21. Il problema dei due corpi	» 344
22. Principio generale della meccanica quantistica. Corrispondenza tra osservabili e operatori	» 347
23. Cenzo sull'uso di coordinate curvilinee	» 353
24. Alcune conseguenze del principio generale della meccanica quantistica	» 354
25. Criterio di compatibilità delle osservabili	» 357
26. Relazioni di permutazione	» 358
27. Valori medi	» 361
28. Derivate rispetto al tempo. La meccanica classica come caso limite	» 363
29. Integrali primi	» 367
30. Momenti angolari e loro operatori	» 368
31. Forze magnetiche	» 372
32. Il determinismo e la meccanica quantistica	» 374
CAP. III. - <i>Metodo delle matrici</i>	» 379
33. Cenni generali	» 379
34. Relazioni di permutazione	» 382
35. Ricerca degli autovalori col metodo delle matrici	» 383
36. Applicazione all'oscillatore	» 383
CAP. IV. - <i>Teoria delle perturbazioni</i>	» 389
37. Generalità	» 389
38. Perturbazione degli stati stazionari (caso non degenere)	» 390
39. Perturbazione degli stati stazionari (caso degenere e quasi-degenere)	» 394
40. Teoria delle perturbazioni col metodo delle matrici	» 403
41. Perturbazione di stati non stazionari. Perturbazioni dipendenti dal tempo	» 405
42. Probabilità di transizione	» 407
43. Perturbazione di tipo sinusoidale. Risonanza	» 409
CAPITOLO V. - <i>Relatività e spin</i>	» 411
44. Considerazioni generali	» 411
45. Fondamenti della teoria non relativistica dello spin (PAULI)	» 412
46. L'estensione relativistica dell'equazione di SCHRÖDINGER	» 421
47. Principio della teoria di DIRAC	» 423
48. Densità elettrica e densità di corrente	» 425
49. Le equazioni di DIRAC per l'elettrone non soggetto a forze	» 426
50. Elettrone in un campo elettromagnetico	» 429
51. La teoria di PAULI come prima approssimazione	» 430
52. Momento magnetico dell'elettrone	» 433

53. Lo spin nelle equazioni di DIRAC	Pag. 436
54. Onde piane	» 440
55. Altra forma delle equazioni di DIRAC	» 442
56. Invarianza relativistica delle equazioni di DIRAC	» 444
57. Invarianza elettromagnetica delle equazioni di DIRAC	» 448
58. Elettrone in un campo centrale	» 450
59. Teoria dei sistemi idrogenoidi (struttura fina)	» 454
60. Stati a energia cinetica negativa e teoria del positrone	» 458
CAP. VI. - Sistemi con particelle uguali	» 466
61. Uguaglianza tra particelle elementari	» 466
62. Stati di un sistema di due particelle uguali. Simmetria e antisimmetria	» 467
63. Cenno sulla estensione ad N particelle uguali	» 471
64. Principio di PAULI e principio di antisimmetria	» 473
65. Calcolo approssimato degli stati di un sistema di due particelle uguali	» 479
66. Separazione delle coordinate di spin da quelle di posi- zione	» 485
67. Spettro dell'atomo di elio e analoghi	» 492
68. Cenno su altre applicazioni	» 496
BIBLIOGRAFIA	» 499
INDICE ANALITICO	» 501