

**I FONDAMENTI DELLA TEORIA QUANTISTICA
SECONDO VON NEUMANN**

1. INTRODUZIONE

La Teoria Quantistica fu elaborata agli inizi del secolo XX, per spiegare fenomeni fisici, quali lo spettro del corpo nero e dell'atomo di idrogeno, che avevano resistito ad ogni tentativo di spiegazione da parte della Fisica Classica. La nuova teoria ha sconvolto la fisica sin dalle sue fondamenta, più di quanto non fece la teoria della relatività, elaborata da Albert Einstein negli stessi anni.

Una dettagliata rassegna dei lavori che condussero alla formulazione della meccanica quantistica si può trovare in [1]. Essa è il risultato del lavoro di scienziati come Max Planck, Albert Einstein, Max Born, Ernest Rutherford, Louis De Broglie, Niels Bohr, Wolfgang Pauli, Erwin Schroedinger, Werner Heisenberg, Paul Dirac.

La nuova teoria si formò attraverso parecchi stadi evolutivi, ed è in grado oggi di spiegare e descrivere con una straordinaria potenza predittiva i fenomeni fisici che avvengono a livello atomico, sub-atomico, nucleare, sub-nucleare, quanto quelli a livelli macroscopici.

La teoria Quantistica è caratterizzata dalla profonda astrazione della rappresentazione matematica dei suoi concetti, astrazione ancor più evidente se messa a confronto con la fisica classica. Ad esempio, la meccanica classica rappresenta la posizione di una particella mediante un vettore $\mathbf{x} = (x, y, z) \in \mathbf{R}^3$, le cui componenti sono le proiezioni del punto occupato dalla particella sui tre assi del sistema di riferimento. Per descrivere una di queste componenti secondo la meccanica quantistica nella formulazione di Heisenberg, occorre introdurre una matrice a infinite righe e infinite colonne, i cui elementi sono funzioni del tempo che devono rispettare delle precise condizioni prescritte dalla teoria.

Questa astrazione è sempre risultata incomprensibile, non solo allo studente che per la prima volta si cimenta con la disciplina, ma anche agli uomini stessi che l'hanno creata.

R. Feynman scrive [2] “c’era un tempo in cui sui giornali si leggeva che solo dodici persone al mondo comprendevano la teoria della relatività [...]. Invece credo di poter dire con sicurezza che nessuno ancora comprende la meccanica quantistica [...]. Se ci riuscite, cercate di non chiedervi “ma come può essere così?” Perché entrerete in un vicolo cieco da cui nessuno è ancora uscito”.

Più formulazioni della teoria quantistica si svilupparono durante la sua genesi. Le due principali sono la “meccanica ondulatoria” di Erwin Schroedinger e la “meccanica delle matrici” di Werner Heisenberg. Benchè in apparenza tali formulazioni appaiono nettamente diverse, fu possibile dimostrarne la completa equivalenza.

In ogni caso, la prima formulazione della teoria quantistica, rigorosa dal punto di vista matematico e logicamente consistente è opera di J. Von Neumann (vedi [3]).

In questo capitolo daremo una formulazione della teoria quantistica secondo Von Neumann in una forma più moderna, conveniente per gli scopi del presente lavoro. In particolare studieremo le nozioni di *osservabili* $1 - 0$, *stati* e *stati puri*, che saranno utilizzate nel seguito della tesi.

2. TEORIA QUANTISTICA GENERALE

2.1. Base concettuale

I concetti fisici di base della teoria di Von Neuman [3] sono quello di *osservabile* e quello di *valore di aspettazione*.

- Per osservabile si intende una qualunque grandezza fisica misurabile sul sistema fisico, che può assumere, fissata l’unità di misura, valori nel campo dei numeri reali. Per una particella esempi di osservabili sono le coordinate della posizione, la quantità di moto,

il momento angolare lungo una direzione, l'energia, etc. L'insieme di tutte le osservabili viene indicato con \mathcal{O} .

- Un valore di aspettazione (*funzioni R*, nella nomenclatura di Von Neumann) è una funzione

$$v : \mathcal{O}_v \rightarrow \mathbf{R}, \quad \text{dove } \mathcal{O}_v \subseteq \mathcal{O}$$

che ad ogni osservabile $\mathcal{A} \in \mathcal{O}$ associa un numero reale $v(\mathcal{A})$ che rappresenta il *valore di aspettazione* della grandezza \mathcal{A} . Pertanto esso va riferito ad un ensemble statistico di sistemi fisici. Se \mathcal{A} viene misurata su un numero N di sistemi di un ensemble, si otterranno N risultati $a_1, a_2, a_3, \dots, a_N$. Il valore medio

$$\langle \mathcal{A} \rangle = \frac{\sum_{i=1}^N a_i}{N}$$

tende a $v(\mathcal{A})$ se $N \rightarrow \infty$.

Il concetto di osservabile induce le seguenti considerazioni.

Osservazione 2.1. Sia \mathcal{A} un'osservabile e $\tilde{\sigma}_A$ l'insieme dei suoi possibili valori (detto spettro fisico) e sia

$$f : \tilde{\sigma}_A \rightarrow \mathbf{R}$$

una funzione reale. Misurando \mathcal{A} e applicando f al risultato a di \mathcal{A} , si ottiene un'altra osservabile che ha $f(a)$ come risultato (Vedi Fig.1).

Questa nuova osservabile viene indicata con $f(\mathcal{A})$. Essenzialmente, \mathcal{A} e $f(\mathcal{A})$ misurano la stessa grandezza, con una scala diversa, se f è iniettiva.

Osservazione 2.2. Se $\{v_1, v_2, \dots\}$ è un insieme numerabile di funzioni R , per ogni sequenza μ_1, μ_2, \dots di numeri reali non negativi tali che $\sum_k \mu_k = 1$ esisterà un altro valore d'aspettazione v come *combinazione convessa*, definita su $\mathcal{O}_v = \cap_k \mathcal{O}_{v_k}$, $v = \sum_i \mu_i v_i$. Tale v si

riferisce all'ensemble di N sistemi che si ottiene prendendo un numero $N(v_i) = \mu_i N$ di sistemi da ogni ensemble corrispondente a v_i , e mettendo insieme tutte le frazioni. $v = \sum_i \mu_i v_i$ rappresenta quindi una miscela statistica degli ensembles descritti dalle v_i .

2.2. Il sistema di assiomi

Gli assiomi della meccanica quantistica secondo Von Neumann possono essere enunciati come segue:

Assioma I. *Ad ogni sistema fisico può essere associato uno spazio di Hilbert complesso e separabile \mathcal{H} . Ad ogni osservabile corrisponde uno ed uno solo operatore autoaggiunto A di \mathcal{H} . La corrispondenza è bijectiva.*

Per introdurre l'assioma successivo occorre premettere il teorema della rappresentazione spettrale per operatori autoaggiunti.

Teorema 2.1. *(Rappresentazione spettrale). Per ogni operatore autoaggiunto A con spettro $\sigma(A)$ e dominio di definizione D_A denso in \mathcal{H} , esiste un'unica famiglia di proiettori ortogonali $\{E_\lambda^A | \lambda \in \mathbf{R}\}$ (risoluzione dell'identità di A) non decrescente in λ , continua a destra rispetto alla norma dei vettori, con $E^A(-\infty) = 0$ e $E^A(\infty) = I$, tale che $\forall g \in D_A$*

$$\langle Ag|g\rangle = \int_{-\infty}^{+\infty} \lambda d\langle E_\lambda^A g|g\rangle$$

(l'integrale va inteso come l'integrale di Lebesgue-Stieltjes $\int \lambda d\alpha$ rispetto alla funzione $\alpha(\lambda) = \langle E_\lambda^A g|g\rangle$). Più brevemente scriveremo

$$A = \int_{\sigma(A)} \lambda dE_\lambda^A.$$

In particolare:

- Se E^A non è crescente nel punto λ_0 , cioè se esiste $\delta > 0$ tale che $E_\lambda^A = E_{\lambda_0}^A, \quad \forall \lambda \in (\lambda_0 - \delta, \lambda_0 + \delta)$, allora $\lambda_0 \notin \sigma(A)$, ovvero λ_0 appartiene al risolvente $\rho(A)$.
- Se E^A è crescente e continua in λ_0 , allora λ_0 appartiene allo spettro continuo.
- Se E^A è crescente e non continua in λ_0 , allora λ_0 appartiene allo spettro puntuale $\sigma_P(A)$, cioè è un autovalore.

Per una trattazione matematica si veda, ad esempio, [4], [5]. Possiamo ora formulare gli altri assiomi.

Assioma II. Siano \mathcal{A} ed f un'osservabile e una funzione reale (di variabile reale) misurabile; se ad \mathcal{A} corrisponde l'operatore A , allora ad $f(\mathcal{A})$ corrisponde l'operatore $f(A) = \int f(\lambda)dE_\lambda^A$, dove E_λ^A rappresenta la risoluzione dell'identità di A .

Assioma III. Siano $\mathcal{A}, \mathcal{S}, \dots$ osservabili a cui corrispondono gli operatori A, S, \dots . Allora esiste un'altra osservabile $\mathcal{A} + \mathcal{S} + \dots$ a cui corrisponde l'operatore $A+S+\dots$.

Assioma IV. Se \mathcal{A} è non negativa, nel senso che i suoi possibili valori sono non negativi, allora $v(A) \geq 0$, per ogni valore d'aspettazione v tale che $\mathcal{A} \in \mathcal{O}_v$.

Assioma V. Per ogni valore d'aspettazione v , se a, b, \dots sono numeri reali e $\mathcal{A}, \mathcal{B}, \dots \in \mathcal{O}_v$, allora $v(a\mathcal{A} + b\mathcal{B} + \dots) = av(\mathcal{A}) + bv(\mathcal{B}) + \dots$.

Gli assiomi della teoria di Von Neumann (in particolare l'assioma I) dimostrano la profonda astrazione della rappresentazione matematica dei concetti fisici nella teoria quantistica, a cui abbiamo accennato. Gli assiomi II-V affermano essenzialmente che la struttura algebrica dell'insieme degli operatori di \mathcal{H} riflette la struttura algebrica naturale delle grandezze che esse rappresentano.

Esempio 2.1. L'osservabile \mathcal{Z} che ha sempre risultato 0 deve corrispondere all'operatore $Z = \mathbf{0}$. Infatti, se \mathcal{A} è una qualunque osservabile rappresentata dall'operatore A , abbiamo $\mathcal{Z} = f(\mathcal{A})$ prendendo $f(\lambda) = 0$ per ogni λ . Dall'assioma II segue allora

$$Z = f(A) = \int f(\lambda)dE_\lambda = \int 0dE_\lambda = \mathbf{0}.$$

In maniera analoga, si ottiene che l'osservabile \mathcal{I} che ha sempre risultato 1 deve corrispondere all'operatore $I = \mathbf{1}$. Ovviamente $\mathcal{Z}, \mathcal{I} \in \mathcal{O}_v$, per ogni valore d'aspettazione v .

3. COMPATIBILITÀ TRA OSSERVABILI

Nella Fisica Quantistica non sempre è possibile misurare insieme, cioè sullo stesso campione individuale del sistema fisico, due osservabili \mathcal{A} e \mathcal{B} . Nella Fisica Classica limitazioni di questo tipo. Per esempio in meccanica classica è possibile misurare contemporaneamente la posizione e la quantità di moto di una particella, mentre secondo la teoria quantistica non può esistere un apparato di misura che permette di misurarle entrambe sullo stesso esemplare del sistema fisico. In teoria quantistica si introduce allora il concetto di *compatibilità*.

Definizione 3.1. *Due osservabili \mathcal{A} ed \mathcal{B} sono compatibili se sono simultaneamente misurabili.*

Per mostrare l'esistenza di coppie di osservabili quantistiche incompatibili, osserviamo che affinché due osservabili siano contemporaneamente misurabili deve esistere una procedura sperimentale che permette di ottenere i valori di entrambe: ad ogni risultato di questo esperimento, rappresentato dal numero reale λ , deve corrispondere una ben precisa coppia $(f(\lambda), g(\lambda))$ di risultati di \mathcal{A} e \mathcal{B} . Allora, in accordo

con l'osservazione 2.1, questo esperimento definisce una osservabile \mathcal{T} con risultati λ , tali che $\mathcal{A} = f(\mathcal{T})$ e $\mathcal{B} = g(\mathcal{T})$. Queste considerazioni permettono di caratterizzare la compatibilità: due osservabili \mathcal{A} e \mathcal{B} sono compatibili se e soltanto se sono funzioni di una terza osservabile \mathcal{T} .

Dall'assioma II e dalla definizione 3.1 segue che due osservabili sono simultaneamente misurabili se i corrispondenti operatori A e B sono funzioni di uno stesso operatore T . La teoria degli operatori lineari in spazi di Hilbert (v. ad esempio [5]) afferma che due operatori autoaggiunti sono funzioni di uno stesso operatore se e solo se commutano. Si ottiene quindi il seguente risultato:

Teorema 3.1. *Due osservabili \mathcal{A} e \mathcal{B} sono compatibili se e solo se i corrispondenti operatori commutano: $[A, B] = \mathbf{0}$.*

Esempio 3.1. (i) Lo spazio di Hilbert per descrivere una particella puntiforme è lo spazio $L_2(\mathbf{R}^3)$ delle funzioni di \mathbf{R}^3 integrabili in modulo al quadrato (nelle formulazioni correnti viene sottinteso che lo spazio di Hilbert è formato in realtà dalle classi di equivalenza di funzioni uguali quasi ovunque, integrabili al quadrato). La coordinata x_i , $i=1,2,3$, della posizione è rappresentata dall'operatore

$$Q_i : D_{Q_i} \rightarrow L_2(\mathbf{R}^3)$$

$$\psi \rightarrow Q_i\psi, (Q_i\psi)(\mathbf{x}) = x_i\psi(\mathbf{x})$$

dove il dominio di definizione D_{Q_i} che rende Q_i autoaggiunto è

$$D_{Q_i} = \{\psi \in L_2(\mathbf{R}^3) : x_i\psi \in L_2(\mathbf{R}^3)\}.$$

Poichè $[Q_i, Q_j] = \mathbf{0} \quad \forall i, j$, dal teorema 3.2 deduciamo che due diverse coordinate della posizione della particella sono sempre misurabili insieme.

(ii) La quantità di moto lungo la direzione x_i è invece rappresentata dall'operatore

$$P_i : D_{P_i} \rightarrow L_2(\mathbf{R}^3)$$

$$(P_i, \psi)(x) = -i\hbar \frac{\partial \psi}{\partial x_i}$$

dove \hbar è la costante di Planck diviso 2π . Poichè $[Q_i, P_j] = i\hbar\delta_{ij}$ dal teorema 3.2 otteniamo che una componente della quantità di moto non può essere misurata insieme alla stessa componente della posizione. Può essere invece misurata insieme ad una diversa componente della posizione.

4. STATI QUANTISTICI E OPERATORI DENSITÀ

L'assioma I stabilisce che le osservabili di un sistema fisico sono rappresentate nel formalismo matetico della teria da operatori autoaggiunti. La rappresentazione matematica dei *valori di aspettazione* avviene attraverso i cosiddetti *operatori densità*, che sono operatori lineari

$$\rho : \mathcal{H} \rightarrow \mathcal{H},$$

positivi, cioè $\langle \psi | \rho \psi \rangle \geq 0 \quad \forall \psi \in \mathcal{H}$, tali che $Tr(\rho) = 1$.

Von Neumann, infatti, ha dimostrato il seguente teorema [3]:

Teorema 4.1. *Per ogni valore di aspettazione v , esiste un operatore ρ , autoaggiunto e positivo, con $Tr(\rho) = 1$, tale che per ogni osservabile $\mathcal{A} \in \mathcal{O}_v$ si ha:*

$$v(\mathcal{A}) = Tr(\rho \mathcal{A}) \tag{4.1}$$

(\mathcal{A} é l'operatore corrispondente ad \mathcal{A}).

Dimostrazione. Utilizzeremo la notazione di Dirac, dove $|l\rangle$ indica un vettore di norma 1 e $|l\rangle\langle l|$ il proiettore sul sottospazio unidimensionale

generato da $|l\rangle$. Dato un sistema ortonormale completo (sonc)

$$(|1\rangle, |2\rangle, \dots, |n\rangle, \dots) \subseteq D_A$$

possiamo scrivere; nella notazione di Dirac,

$$A = \sum_{n,m} |n\rangle a_{n,m} \langle m|, \quad \text{con } a_{nm} = \langle n|A|m\rangle. \quad (4.2)$$

L'operatore A , essendo l'operatore corrispondente all'osservabile \mathcal{A} , è autoaggiunto, e quindi:

$$Re(a_{nm}) = Re(a_{mn}), \quad Im(a_{nm}) = -Im(a_{mn}), \quad (4.3)$$

dove Re indica la parte reale e Im quella immaginaria. Dalle (4.2) e (4.3) segue:

$$\begin{aligned} A &= \sum_{n,m} |n\rangle a_{nm} \langle m| = & (4.4) \\ &= \sum_n |n\rangle \langle n| a_{nn} + \sum_{\substack{n,m \\ n \neq m}} |n\rangle \langle m| a_{nm} = \\ &= \sum_n |n\rangle \langle n| a_{nn} + \sum_{\substack{n,m \\ n > m}} |n\rangle \langle m| a_{nm} + \sum_{\substack{n,m \\ n < m}} |n\rangle \langle m| a_{nm} = \\ &= \sum_n |n\rangle \langle n| a_{nn} + \sum_{\substack{n,m \\ n > m}} |n\rangle \langle m| a_{nm} + \sum_{\substack{n,m \\ n > m}} |m\rangle \langle n| a_{mn} = \\ &= \sum_n |n\rangle \langle n| a_{nn} + \sum_{\substack{n,m \\ n > m}} |n\rangle \langle m| (Re(a_{nm}) + iIm(a_{nm})) + \\ &+ \sum_{\substack{n,m \\ n > m}} |m\rangle \langle n| (Re(a_{mn}) + iIm(a_{mn})) = \\ &= \sum_n |n\rangle \langle n| a_{nn} + \sum_{\substack{n,m \\ n > m}} |n\rangle \langle m| (Re(a_{nm}) + iIm(a_{nm})) + \\ &+ \sum_{\substack{n,m \\ n > m}} |m\rangle \langle n| (Re(a_{nm}) - iIm(a_{nm})) = \end{aligned}$$

$$\begin{aligned}
&= \sum_n |n \rangle \langle n| a_{nn} + \sum_{\substack{n,m \\ n>m}} (|n \rangle \langle m| + |m \rangle \langle n|) \operatorname{Re}(a_{nm}) + \\
&+ \sum_{\substack{n,m \\ n>m}} i(|m \rangle \langle n| - |n \rangle \langle m|) \operatorname{Im}(a_{nm}).
\end{aligned}$$

Gli operatori:

$$D^{(n,n)} = |n \rangle \langle n|$$

$$B^{(n,m)} = |n \rangle \langle m| + |m \rangle \langle n|$$

$$C^{(n,m)} = i(|n \rangle \langle m| - |m \rangle \langle n|)$$

sono autoaggiunti; se indichiamo le osservabili corrispondenti rispettivamente, secondo l'Assioma I, con: $\mathcal{D}^{(n,n)}$, $\mathcal{B}^{(n,m)}$ e $\mathcal{C}^{(n,m)}$, possiamo scrivere:

$$\begin{aligned}
\mathcal{A} &= \sum_n \mathcal{D}^{(n,n)} a_{nn} + \\
&+ \sum_{\substack{n,m \\ n>m}} \mathcal{B}^{(n,m)} \operatorname{Re}(a_{nm}) + \\
&+ \sum_{\substack{n,m \\ n>m}} \mathcal{C}^{(n,m)} \operatorname{Im}(a_{nm}),
\end{aligned}$$

da cui segue, per l'Assioma V,

$$\begin{aligned}
v(\mathcal{A}) &= \sum_n v(\mathcal{D}^{(n,n)}) a_{nn} + \\
&+ \sum_{\substack{n,m \\ n>m}} v(\mathcal{B}^{(n,m)}) \operatorname{Re}(a_{nm}) + \\
&+ \sum_{\substack{n,m \\ n>m}} v(\mathcal{C}^{(n,m)}) \operatorname{Im}(a_{nm}).
\end{aligned}$$

Se poniamo:

$$\begin{aligned}
 \rho_{nn} &= v(\mathcal{D}^{(n,n)}) \\
 \rho_{nm} &= \frac{1}{2}v(\mathcal{B}^{(n,m)}) + \frac{1}{2}iv(\mathcal{C}^{(n,m)}) \quad n > m \\
 \rho_{mn} &= \frac{1}{2}v(\mathcal{B}^{(n,m)}) - \frac{1}{2}iv(\mathcal{C}^{(n,m)}) \quad n > m
 \end{aligned} \tag{4.5}$$

sommando e sottraendo le ultime due otteniamo:

$$\begin{aligned}
 v(\mathcal{B}^{(n,m)}) &= \rho_{nm} + \rho_{mn} \\
 v(\mathcal{C}^{(n,m)}) &= i(\rho_{mn} - \rho_{nm})
 \end{aligned}$$

per cui:

$$\begin{aligned}
 v(\mathcal{A}) &= \sum_n \rho_{nn}a_{nn} + \sum_{\substack{n,m \\ n>m}} (\rho_{nm} + \rho_{mn})Re(a_{nm}) + \\
 &\quad + \sum_{\substack{n,m \\ n>m}} i(\rho_{mn} - \rho_{nm})Im(a_{nm}) = \\
 &= \sum_n \rho_{nn}a_{nn} + \sum_{\substack{n,m \\ n>m}} \rho_{nm}(Re(a_{nm}) - iIm(a_{nm})) + \\
 &\quad + \sum_{\substack{n,m \\ n>m}} \rho_{mn}(Re(a_{nm}) + iIm(a_{nm})) = \\
 &= \sum_n \rho_{nn}a_{nn} + \sum_{\substack{n,m \\ n>m}} \rho_{nm}(Re(a_{mn}) + iIm(a_{mn})) + \\
 &\quad + \sum_{\substack{n,m \\ n>m}} \rho_{mn}(Re(a_{nm}) + iIm(a_{nm})) = \\
 &= \sum_n \rho_{nn}a_{nn} + \sum_{\substack{n,m \\ n>m}} \rho_{nm}a_{mn} + \\
 &\quad + \sum_{\substack{n,m \\ n>m}} \rho_{mn}a_{nm} =
 \end{aligned}$$

$$\begin{aligned}
&= \sum_n \rho_{nn} a_{nn} + \sum_{\substack{n,m \\ n>m}} \rho_{nm} a_{mn} + \\
&\quad + \sum_{\substack{n,m \\ n<m}} \rho_{nm} a_{mn} = \\
&= \sum_n \rho_{nn} a_{nn} + \sum_{\substack{n,m \\ n \neq m}} \rho_{nm} a_{mn} = \sum_{n,m} \rho_{nm} a_{mn}.
\end{aligned}$$

Se definiamo l'operatore $\hat{\rho}$ con

$$\langle n | \hat{\rho} | m \rangle = \rho_{nm}$$

allora

$$v(\mathcal{A}) = \text{Tr}(\hat{\rho}A), \quad (4.6)$$

con $\hat{\rho}$ operatore autoaggiunto indipendente da \mathcal{A} .

Ma $v(\mathcal{I}) = 1 = \text{Tr}(\hat{\rho} \cdot \mathbf{1})$, quindi $\hat{\rho}$ deve avere traccia unitaria. Ora proviamo che $\hat{\rho}$ deve essere definito positivo, ossia

$$\langle \phi | \hat{\rho} | \phi \rangle \geq 0 \quad \forall |\phi\rangle.$$

Sia dunque $|\phi\rangle \in \mathcal{H}$ con norma unitaria e sia $\mathcal{A}_{|\phi\rangle}$ l'osservabile corrispondente a $\hat{P}_{|\phi\rangle} = |\phi\rangle\langle\phi|$, allora, $\mathcal{A}_{|\phi\rangle}^2$ è l'osservabile corrispondente a $\hat{P}_{|\phi\rangle}^2$. $\mathcal{A}_{|\phi\rangle}^2$ è quindi una quantità non negativa, pertanto anche il suo valore d'aspettazione deve essere non negativo:

$$v(\mathcal{A}_{|\phi\rangle}^2) = \text{Tr}(\hat{\rho}\hat{P}_{|\phi\rangle}^2) \geq \mathbf{0}, \quad (4.7)$$

e poichè $\hat{P}_{|\phi\rangle}^2 = \hat{P}_{|\phi\rangle}$ si ha

$$\text{Tr}(\hat{\rho}\hat{P}_{|\phi\rangle}) \geq \mathbf{0};$$

ne segue che:

$$\begin{aligned}
0 \leq Tr(\hat{\rho}\hat{P}_{|\phi\rangle}) &= \sum_n \langle n|\hat{\rho}\hat{P}_{|\phi\rangle}|n\rangle \\
&= \sum_n \langle n|\hat{\rho}(\langle\phi|n\rangle)|\phi\rangle \\
&= \sum_n \langle\phi|n\rangle\langle n|\hat{\rho}|\phi\rangle \\
&= \langle\phi|\hat{\rho}|\phi\rangle.
\end{aligned}$$

Infine proviamo che $\hat{\rho}$ dipende solo dal valore di aspettazione cui si riferisce, e non dalla base utilizzata per determinarlo.

Se $\hat{\rho}$ ed $\hat{\rho}'$ sono due operatori statistici corrispondenti allo stesso valore di aspettazione v , e siano, rispettivamente $\{|n\rangle\}$ e $\{|\mu\rangle\}$ i sonc utilizzati per determinarli. Allora qualunque sia A autoaggiunto, si ha

$$Tr(\hat{\rho}A) = Tr(\hat{\rho}'A).$$

Per $A = \hat{P}_{|\phi\rangle}$, $\langle\phi|\phi\rangle = 1$, abbiamo:

$$\begin{aligned}
Tr(\hat{\rho}\hat{P}_{|\phi\rangle}) &= \sum_n \langle n|\hat{\rho}\hat{P}_{|\phi\rangle}|n\rangle \\
&= \sum_n \langle n|\hat{\rho}(\langle\phi|n\rangle)|\phi\rangle \\
&= \sum_n \langle\phi|n\rangle\langle n|\hat{\rho}|\phi\rangle \\
&= \langle\phi|\hat{\rho}|\phi\rangle.
\end{aligned}$$

$$\begin{aligned}
Tr(\hat{\rho}'\hat{P}_{|\phi\rangle}) &= \sum_n \langle\mu|\hat{\rho}'\hat{P}_{|\phi\rangle}|\mu\rangle \\
&= \sum_n \langle\mu|\hat{\rho}'(\langle\phi|\mu\rangle)|\phi\rangle \\
&= \sum_n \langle\phi|\mu\rangle\langle\mu|\hat{\rho}'|\phi\rangle \\
&= \langle\phi|\hat{\rho}'|\phi\rangle.
\end{aligned}$$

Quindi:

$$\langle \phi | \hat{\rho} | \phi \rangle = \langle \phi | \hat{\rho}' | \phi \rangle$$

e questo per ogni $|\phi\rangle$ con $\langle \phi | \phi \rangle = 1$, allora:

$$\hat{\rho} = \hat{\rho}',$$

cosicchè $\hat{\rho}$ non può dipendere dalla base scelta.

Questo permette di rappresentare nel formalismo matematico il concetto fisico di valore d'aspettazione. L'operatore densità ρ che rappresenta un valore d'aspettazione v viene identificato come lo *stato quantistico* del sistema.

5. OSSERVABILI 1-0 E PROIETTORI ORTOGONALI

Sia \mathcal{T} una osservabile che può assumere solo due valori: 0 oppure 1. Le osservabili che hanno questa caratteristica vengono chiamate osservabili 1 – 0. Mostriamo adesso che tali osservabili sono caratterizzate dall'essere rappresentate da proiettori ortogonali.

Infatti, per tali osservabili è possibile identificare il valore d'aspettazione con la probabilità di accadimento del valore 1: se indichiamo con $p(1)$ e $p(0)$ le probabilità dei risultati 1 e 0, rispettivamente, dovremmo avere:

$$v(\mathcal{T}) = 0 \cdot p(0) + 1 \cdot p(1) = p(1). \quad (5.8)$$

Allora dal teorema 4.1 otteniamo:

$$p(1) = v(\mathcal{T}) = \text{Tr}(\rho \cdot T). \quad (5.9)$$

Data una qualunque osservabile \mathcal{A} , fissiamo un intervallo $\Delta = (a, b] \in \mathcal{B}(\mathbf{R})$, e consideriamo il funzionale caratteristico di Δ ,

$$\chi_{\Delta} : \mathbf{R} \rightarrow \mathbf{R}, \quad \chi_{\Delta}(x) = \begin{cases} 1, & \text{se } x \in \Delta \\ 0, & \text{se } x \notin \Delta. \end{cases}$$

Allora $\chi_\Delta(\mathcal{A})$ è un'osservabile $1 - 0$ che assume il valore 1 se $a \in \Delta$, e il valore 0 se $a \notin \Delta$.

Determiniamo l'operatore autoaggiunto T_Δ corrispondente all'osservabile $1 - 0$ $T_\Delta = \chi_\Delta(\mathcal{A})$; usando l'Assioma I e il Teorema 2.1,

$$T_\Delta = \int \chi_\Delta(\lambda) dE_\lambda^A = \lim_{\varepsilon \rightarrow 0^+} \int_{a+\varepsilon}^{b+\varepsilon} dE_\lambda = E_b^A - E_a^A.$$

T_Δ è un proiettore. Infatti, siccome E_λ^A è non decrescente in λ

$$\mu \leq \lambda \quad \text{implica} \quad E_\mu^A E_\lambda^A = E_\lambda^A E_\mu^A = E_\mu^A,$$

allora

$$\begin{aligned} T_\Delta^2 &= (E_b^A - E_a^A)(E_b^A - E_a^A) \\ &= E_b^{A^2} - E_b^A E_a^A - E_a^A E_b^A + E_a^{A^2} \\ &= E_b^A - 2E_a^A + E_a^A = E_b^A - E_a^A = T_\Delta. \end{aligned}$$

Allora, per la (5.8), la probabilità $P(\mathcal{A}, \Delta)$ che una misurazione di \mathcal{A} produca un risultato $a \in \Delta$ è

$$P(\mathcal{A}, \Delta) = \text{Tr}(\rho[E_b^A - E_a^A]). \quad (5.10)$$

Un numero reale a è un risultato possibile per l'osservabile \mathcal{A} se e soltanto se per ogni $\delta > 0$ la probabilità $P(\mathcal{A}, (a - \delta, a + \delta])$ che una misurazione di \mathcal{A} produca un risultato nell'intorno $(a - \delta, a + \delta]$ di a è diversa da zero.

Chiameremo spettro fisico di \mathcal{A} l'insieme chiuso $\hat{\sigma}(\mathcal{A})$ dei suoi risultati possibili (il suo complementare è aperto).

Allora dal Teorema (2.1) e dalla (5.10) concludiamo che

Teorema 5.1. *Un numero reale a è un risultato possibile di \mathcal{A} se e soltanto se $a \in \sigma(\mathcal{A})$.*

In altre parole lo spettro fisico di \mathcal{A} coincide con lo spettro matematico dell'operatore A .

Sia \mathcal{T} un'osservabile $1 - 0$, e T l'operatore autoaggiunto che la rappresenta. Dal Teorema (5.1) si ha che $\sigma(T) = \{0, 1\}$.

Un operatore autoaggiunto ha lo spettro costituito dai soli autovalori 0 e 1 se e soltanto se $P^2 = P$, cioè se e soltanto se è un proiettore ortogonale. Pertanto

Teorema 5.2. *I proiettori ortogonali corrispondono biunivocamente alle osservabili 1-0.*

6. STATI PURI

In questo paragrafo studieremo l'importante classe degli "stati puri".

Un operatore densità ρ è detto *puro* se

$$\rho = \lambda\rho_1 + (1 - \lambda)\rho_2 \quad \text{con} \quad \lambda^2 < \lambda \quad \text{implica} \quad \rho_1 = \rho_2 = \rho.$$

Dal punto di vista fisico, la purezza di ρ significa che l'ensemble statistico che esso descrive non può essere ottenuto come miscela statistica di ensembles descritti da differenti funzioni R (Osserv. 2.2).

Un operatore densità ha sempre spettro puntuale puro $\sigma(\rho) = (\lambda_1, \lambda_2, \dots)$, con $\lambda_i \geq 0$ (V. [6]). Siano M_1, M_2, \dots, M_n , gli autospazi corrispondenti agli autovalori $\lambda_1, \lambda_2, \dots, \lambda_n, \dots$. Indichiamo con $P_1, P_2, \dots, P_n, \dots$ i proiettori che proiettano su tali autospazi.

Allora

$$\rho = \sum_i \lambda_i P_i.$$

Proposizione 6.1. Se $\lambda_i > 0$, allora

$$\dim M_i < +\infty.$$

Dimostrazione. Se $\dim M_{i_0} = +\infty$ e $\lambda_{i_0} > 0$, poichè $\dim M_i = Tr(P_i)$ si avrebbe:

$$Tr(\rho) = \sum_i \lambda_i Tr(P_i) \geq \lambda_{i_0} Tr(P_{i_0}) = \infty$$

mentre $Tr(\rho) = 1$.

Se ρ è un operatore densità puro, allora ovviamente, $P_i = P_j \quad \forall i, j$ tali che $\lambda_i \neq 0, \lambda_j \neq 0$.

Pertanto, se ρ è puro,

$$\rho = kP$$

dove P è un proiettore di rango finito, per la Proposizione 6.1, e

$$k = \frac{1}{Tr(P)}$$

dove $Tr(P)$ è la dimensione dell'immagine M di P e pertanto è un numero naturale:

$$Tr(P) = n.$$

Ora sia u_1, u_2, \dots, u_n una base ortonormale di M . Allora

$$P = \sum_{i=1}^n |u_i \rangle \langle u_i|$$

e di conseguenza

$$\rho = \sum_{i=1}^n \frac{1}{n} |u_i \rangle \langle u_i|.$$

Pertanto, se ρ è puro, $|u_1 \rangle \langle u_1| = |u_2 \rangle \langle u_2| = \dots = |u_n \rangle \langle u_n|$. Allora ogni operatore di densità puro è un proiettore di rango 1, quindi esiste un vettore $\psi \in \mathcal{H}$, $\|\psi\| = 1$ tale che

$$\rho = |\psi \rangle \langle \psi|.$$

Per gli stati puri $\rho = |\psi\rangle\langle\psi|$ il valore di aspettazione di un'osservabile \mathcal{A} rappresentata dall'operatore A si può esprimere nel seguente modo:

$$v(A) = \langle\psi|A\psi\rangle. \quad (6.1)$$

Infatti, sia $\{u_n\}$ una base ortonormale di \mathcal{H} tale che $u_1 = \psi$. Allora, dal teorema 4.1 si ha:

$$\begin{aligned} v(\mathcal{A}) &= Tr(\rho A) = \sum_k \langle u_k | \rho A u_k \rangle \\ &= \langle u_1 | \rho A u_1 \rangle + \sum_{k>1} \langle u_k | \rho A u_k \rangle \\ &= \langle \psi | \psi \rangle \langle \psi | A \psi \rangle + \sum_{k>1} \langle u_k | u_1 \rangle \langle u_1 | A u_k \rangle = \langle \psi | A \psi \rangle. \end{aligned}$$

- [1] B. Van der Waerden, *Sources of quantum mechanics*, North-Holland Publishing co., Amsterdam, 1967
- [2] R. Feynman, *La legge fisica*, Universale scientifica Boringhieri, Torino, 1971
- [3] J. von Neumann, *Mathematical foundations of quantum mechanics*, Princeton University Press, Princeton 1955
- [4] E. Kreyszig, *Introductory functional analysis with applications*, John Wiley e Sons, New York 1978
- [5] F. Riesz, B. Sz. Nagy, *Lecons D'analyse fonctionnelle*, Gauthier-Villars, Paris 1972.

Chapter 1

Formulation and Interpretation of Quantum Theory

Since the birth of Quantum Mechanics, several formulations of the theory were developed; the main ones are the *wave mechanics*, proposed by Erwin Schrödinger [10, 11, 12], and the *matrix mechanics* of Werner Heisenberg [13]; whereas the starting point of this last is a critical analysis of the old theories, Schrödinger's formulation originates from the works of L. de Broglie on matter waves [14]. As pointed out by Schrödinger himself [11], although apparently different, these are two equivalent formulations of a more general theory whose formalism is due to Dirac [15]. However, the more rigorous and elegant formulation of Quantum Theory is due to Von Neumann [16]; the basic idea is to found the theory on a system of axioms, so that every valid assertion must be proved as consequence of these assumptions. A deep abstraction arises because it is formulated by means of the Hilbert space theory. We will see, in fact, that measurable physical magnitudes are represented by self-adjoint operators on a suitable Hilbert space. A strict link exists between Quantum

Mechanics and the spectral theory; a particularly important role is played by the spectral theorem for self-adjoint operators, according to which self-adjoint operators can be decomposed in terms of projections operators.

In spite of its deep abstraction, this theory turns out to agree with empirical results. It is introduced here in a modern and suitable version.

1.1 Von Neumann Theory

The basic concepts of Von Neumann theory are essentially two:

Observables. By observable we mean any physical magnitude measurable, by means of a concrete apparatus, on individual specimens of the physical system under investigation, and having real numbers as outcomes; position, momentum along a direction, energy are examples of observables. Let \mathcal{O} be the set of all observables.

Expectation Value of observables or R-functions. By expectation value we mean a function $Ev : \mathcal{O} \rightarrow \mathbb{R}$ assigning a numerical value $Ev(\mathcal{R})$ to every observable \mathcal{R} , to be interpreted as the expectation value of the observable \mathcal{R} : thus any expectation value Ev must be referred to a statistical ensemble of physical systems. The following physical interpretation of the notion of expectation value is adopted: let N be an ensemble of individual physical system corresponding to an expectation value Ev ; the measurement of every observable \mathcal{R} on the sub-ensemble N_1 , made up of n_1 physical systems, provides n_1 outcomes a_1, a_2, \dots, a_{n_1} , whose mean value (statistical concept) $\langle \mathcal{R} \rangle_{N_1} = \sum_{i=1}^{n_1} a_i / n_1$ converges to the expectation value $Ev(\mathcal{R})$ (probabilistic concept) as $n_1 \rightarrow \infty$.

For a physical system, more than one expectation value exists: different expectation values correspond to different ways of preparing - in a laboratory or by a natural process - ensembles of physical systems. As noticed by Ballentine [17], the ensembles at issue are different from those used in classical statistical mechanics, where an ensemble is a virtual set of copies of the physical system used for calculations whose results are interpreted as a measurement on a single system. In Quantum Theory, the ensemble is the set of all systems which have undergone the same preparation technique, generally by interaction with a suitable apparatus; these physical systems are similar in their properties, but not in all of them: thus, in general, the theory does not make predictions about a single measurement, but rather it establishes the probability of each possible outcome; then, in general, its prediction can be verified by repeating the preparation and the measurement many times, and then constructing the statistical distribution of the results.

Further assumptions, induced by the interpretation of the concepts of observable and expectation value, are the following:

- Given an observable \mathcal{R} , we can introduce the set $\tilde{\sigma}_{\mathcal{R}}$ of its possible outcomes (said physical spectrum); in correspondence with each function $f : \tilde{\sigma}_{\mathcal{R}} \rightarrow \mathbb{R}$, there is another observable, denoted by $f(\mathcal{R})$, whose outcomes are obtained by applying the function f to the outcomes of \mathcal{R} ; we notice that, if f is injective, \mathcal{R} and $f(\mathcal{R})$ measure the same magnitude, by using two different scales.
- The set of \mathcal{R} -functions is endowed with a convex operation: if $\{Ev_1, Ev_2, \dots\}$ is a countable set of \mathcal{R} -functions, then $\sum_i \alpha_i Ev_i$, where $\alpha_i \geq 0$ and $\sum_i \alpha_i = 1$, is a \mathcal{R} -function too, corresponding to the ensemble obtained by making a mixture with statistical weights α_i of the

statistical ensembles of physical systems corresponding to the Ev_i .

Von Neumann theory is based on the following axioms:

Axiom 1. *Given a physical system, a complex and separable Hilbert space \mathcal{H} can be associated to it such that to every observable \mathcal{R} there corresponds one, and only one, self-adjoint operator R of \mathcal{H} . The correspondence is bijective.*

Axiom 2. *Let \mathcal{R} be an observable and $f : \tilde{\sigma}_{\mathcal{R}} \rightarrow \mathbb{R}$ a measurable function; if \mathcal{R} is represented by the operator R , then $f(\mathcal{R})$ is represented by the operator $f(R) = \int_{\mathbb{R}} f(\lambda) dE^R(\lambda)$, where $E^R(\lambda)$ denote the spectral resolution of R .*

Axiom 3. *Let $\mathcal{R}, \mathcal{S}, \dots$, be observables represented by the operators R, S, \dots , then there exists another observable, denoted by $\mathcal{R} + \mathcal{S} + \dots$, which is represented by the operator $R + S + \dots$.*

Axiom 4. *If \mathcal{R} is a non-negative quantity, then $Ev(\mathcal{R}) \geq 0$.*

Axiom 5. *For every \mathcal{R} -function Ev , $Ev(a\mathcal{R} + b\mathcal{S} + \dots) = aEv(\mathcal{R}) + bEv(\mathcal{S}) + \dots$, where $\mathcal{R}, \mathcal{S}, \dots$ are observables and a, b, \dots are numbers.*

Axiom 1 states that observables of a physical system are represented, in mathematical framework, by self-adjoint operators. Axioms 2-5 establish a strong link between the algebraic structure of the set of observables and that of the self-adjoint operators of \mathcal{H} .

1.2 Spectral Representation

Because of the correspondence between observables and self-adjoint operators, it is useful to investigate the properties of such operators on a Hilbert space.

According to the spectral representation theorem [16], each self-adjoint operator can be analyzed in term of more simple operators (projection operators):

Theorem 1 (Spectral representation). *For every self-adjoint operator R with spectrum $\sigma(R)$ there exists a unique family of projections $\{E^R(\lambda)|\lambda \in \sigma(R)\}$ (spectral resolution), increasing in λ with $E^R(-\infty) = 0$ and $E^R(\infty) = 1$, such that $R = \int_{-\infty}^{+\infty} \lambda dE^R(\lambda)$.*

The integral is to be understood in the sense that $\forall \psi \in D_R$,

$$\langle \psi | R\psi \rangle = \int_{-\infty}^{+\infty} \lambda d\langle \psi | E^R(\lambda)\psi \rangle$$

according to the Lebesgue-Stieltjes integral. In particular:

- if E^R is non-increasing in λ_0 , i.e. if $\delta > 0$ exists such that $E_{\lambda_0}^R = E_{\lambda_0 \pm \delta}^R \quad \forall \lambda \in (\lambda_0 - \delta, \lambda_0 + \delta)$, then $\lambda_0 \notin \sigma(R)$ or, equivalently, $\lambda_0 \in \rho(R)$ ($\rho(R)$ is the resolvent set);
- if E^R is increasing and continuous in λ_0 , then λ_0 is in the continuous spectrum;
- if E^R is increasing and not continuous in λ_0 , then λ_0 is in the point spectrum, i.e. λ_0 is an eigenvalue;

1.3 States and density operators

The mathematical representation of the expectation values makes use of the *density operators*:

Definition 1. *A linear operator $\rho : \mathcal{H} \rightarrow \mathcal{H}$ is said to be a density operator if the following three conditions hold:*

1. ρ is self-adjoint
2. ρ is positive, i.e. $\langle \psi | \rho \psi \rangle \geq 0 \quad \forall \psi \in \mathcal{H}$
3. $Tr(\rho) = 1$

The following theorem characterizes density operators:

Theorem 2. *Every density operator has a purely discrete spectrum. Conversely, if P_1, P_2, \dots are projection operators of rank 1 and if $\alpha_1, \alpha_2, \dots$ are positive numbers such that $\sum_i \alpha_i = 1$, the operator $\rho = \sum_i \alpha_i P_i$ exists and defines a density operator.*

The theorem in this form is in [18](p. 87) (see also [16](p. 328)); it ensures that eigenvalues suffice in order to construct the identity resolution of a density operator. Von Neumann proved the following theorem

Theorem 3. *For every \mathcal{R} -function Ev satisfying axioms 0-4 there is a density operator ρ , such that $Ev(\mathcal{R}) = Tr(\rho R)$.*

Notice that, taking into account theorem 2, the following equality holds

$$Ev(\mathcal{R}) = \sum_i \alpha_i Tr(|\psi_i\rangle\langle\psi_i|R) = \sum_i \alpha_i \langle\psi_i|R\psi_i\rangle.$$

The operator ρ is identified as the *state* of the system; as pointed out in [17, 19], whereas in classical physics the concept of state is used to refer to the coordinates and momenta of an individual system, a quantum state represents an ensemble of similarly prepared systems; it is the mathematical representation of the distribution of the results of a certain state preparation procedure.

The set of all state operators form a convex set. From axioms 0-4, the spectral representation theorem 1 and theorem 3, it follows that:

Proposition 1. *For every \mathcal{R} -function Ev there exists ρ such that*

$$Ev(\mathcal{R}) = \int_{\sigma(\mathcal{R})} \lambda Tr(\rho dE^{\mathcal{R}}(\lambda))$$

As a consequence, $Tr(\rho dE^{\mathcal{R}}(\lambda))$ has to be interpreted as the probability that the outcome of \mathcal{R} is a value in the interval $(\lambda, \lambda + d\lambda)$.

1.4 0-1 observables and projection operators

We call *0-1 observable* any observable \mathcal{P} having 0 and 1 as possible outcomes. For a 0-1 observable \mathcal{P} the expectation value $Ev(\mathcal{P})$ must be interpreted as the probability of occurrence of the outcome 1 for \mathcal{P} ; indeed, if $p(0)$ and $p(1)$ denote the probabilities of occurrence of 0 and 1, then

$$Ev(\mathcal{P}) = 0 \cdot p(0) + 1 \cdot p(1) = p(1).$$

Fixed an observable \mathcal{A} , a 0-1 observable can be constructed in correspondence with every interval. Let \mathcal{A} be any observable and $\Delta = (a, b]$ an interval; let us consider the characteristic function

$$\chi_{\Delta} : \mathbb{R} \rightarrow \mathbb{R}, \quad \chi_{\Delta} = \begin{cases} 1, & \text{if } x \in \Delta \\ 0, & \text{if } x \notin \Delta. \end{cases}$$

Then $\chi_{\Delta}(\mathcal{A})$ is a 0-1 observable, assuming value 1 if $a \in \Delta$ and value 0 otherwise, represented by the self-adjoint operator

$$T_{\Delta} = \int \chi_{\Delta}(\lambda) dE_{\lambda}^{\mathcal{A}} = \lim_{\epsilon \rightarrow 0^+} \int_{a+\epsilon}^{b+\epsilon} dE_{\lambda}^{\mathcal{A}} = E_b^{\mathcal{A}} - E_a^{\mathcal{A}}.$$

T_{Δ} is a projection operator.

Such a kind of observable provides a characterization of the spectrum of an observable. Let \mathcal{A} be an observable, represented by the self-adjoint operator

A ; the probability $p(A, \Delta)$ that a measurement of \mathcal{A} provides a result $a \in \Delta$ is given by

$$p(A, \Delta) = \text{Tr}(\rho[E_b^A - E_a^A]). \quad (1.1)$$

A real number a is a possible outcome for the observable \mathcal{A} if and only if, fixed $\delta \geq 0$, the probability $p(\mathcal{A}, (a - \delta, a + \delta])$ that a measurement of \mathcal{A} gives a result in the neighborhood $(a - \delta, a + \delta]$ of a is not zero. Then, it follows that

Proposition 2. *The real number a is a possible result for the measurement of \mathcal{A} if and only if $a \in \sigma(A)$.*

Then, by theorem 1 and (1.1) it follows that the physical spectrum of \mathcal{A} coincides with the mathematical spectrum of the operator A .

Let \mathcal{P} be a 0-1-observable and P the associated self-adjoint operator; then proposition 1 implies that $\sigma(P) = \{0, 1\}$. A self-adjoint operator P has such a spectrum if and only if $P^2 = P$, i.e. if P is a projection operator; in general, the following theorem holds:

Proposition 3. *If P is the self-adjoint operator associated to a 0-1 observable \mathcal{P} , then $\sigma(P) = \{0, 1\}$, that is to say P is a projection operator. The converse holds too.*

We notice that 0-1 observables can be interpreted as properties that the system may or may not possess: if the outcome of a measurement is 1 than the system possesses the property, otherwise it does not.

1.5 Pure states

Within the set of all states operator we distinguish the important class of *pure states*.

Definition 2. A density operator is pure if

$$\rho = \lambda\rho_1 + (1 - \lambda)\rho_2, \text{ where } \lambda^2 < \lambda \quad \text{imply} \quad \rho_1 = \rho_2 = \rho$$

A general state which is not pure is called *mixed state*. From a physical point of view, if ρ is pure, then the statistical ensemble described by ρ cannot be obtained as statistical mixture of ensembles described by different \mathcal{R} -functions.

By definition, a pure state has the form $\rho = |\psi\rangle\langle\psi|$; the expectation value of the observable \mathcal{R} , in this pure state, is

$$Ev(\mathcal{R}) = Tr(|\psi\rangle\langle\psi|R) = \langle\psi|R\psi\rangle.$$

We observe that the state vector is not unique; indeed, any vector of the form $e^{i\alpha}|\psi\rangle$, with $\alpha \in \mathbb{R}$, is physically equivalent.

1.6 Simultaneous measurability

We have seen that the algebraic structure of the set of observables is mirrored in the algebraic structure of self-adjoint operators, which is not commutative; then, whereas in Classical Mechanics all the observables can be measured simultaneously on a physical system, this is not the case in Quantum Mechanics. Such a limitation is the object of the following theorem:

Theorem 4. Two observables, \mathcal{A} and \mathcal{B} are measurable together if and only if the corresponding self-adjoint operators, A and B , commute i.e. $[A,B] = AB - BA = \mathbf{0}$.

This theorem proved by Von Neumann ([16], p. 223) at first for operators A and B with pure discrete spectra, is extended to the more general case by means of the following theorems ([16], p. 170).

Theorem 5. *If two self-adjoint operators are commutative, then each function $f(A)$ of A commute with each function $g(B)$ of B .*

Since the hypothesis is always satisfied for $A = B$, two functions of the same operator always commute; the converse is also valid:

Theorem 6. *A and B are two commuting self-adjoint operators if and only if there exist a self-adjoint operator C and two functions, f and g , such that $A = f(C)$ and $B = g(C)$.*

(for a proof of theorem 6 see [20]). Hence the following interpretation can be given to simultaneous measurability: observables \mathcal{A} and \mathcal{B} are measurable together if and only if there exist an observable \mathcal{C} and functions, f and g , such that $\mathcal{A} = f(\mathcal{C})$ and $\mathcal{B} = g(\mathcal{C})$, i.e. two observables \mathcal{A} and \mathcal{B} are simultaneously measurable if there is an arrangement, say \mathcal{C} , which measures both \mathcal{A} and \mathcal{B} on the same system.

The theorem can be extended to the case of more than two observables: “without this condition [the commutativity of operator R_1, R_2, \dots], nothing can be said regarding the results of simultaneous measurement of $\mathcal{R}_1, \mathcal{R}_2, \dots$, since simultaneous measurements of these quantities are then in general not possible” ([16], p. 230).

Since we restrict ourselves to the case of projection operators, a characterization can be given to simultaneous measurement by means of the following theorem:

Theorem 7. *Two projection operators A and B are commuting if and only if there exist three mutually orthogonal projection operators A_1 , B_1 and K , such that $A = A_1 + K$ and $B = B_1 + K$.*

Then, the following definition can be given:

Definition 3. Observables \mathcal{A} and \mathcal{B} , corresponding to projection operators A and B such that $[A,B] = 0$, are said compatible.

Examples. Position and momentum are examples of observables; the operator representing the x_i coordinate ($i = 1,2,3$) of position is the multiplication operator

$$(Q_i\psi)\mathbf{x} = x_i\psi(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^3, \quad \forall \psi \in D_{Q_i}$$

where $D_{Q_i} = \{\psi \in L_2(\mathbb{R}^3) : x_i\psi \in L_2(\mathbb{R}^3)\}$. The operator representing the p_i coordinate ($i = 1,2,3$) of momentum is the derivation operator

$$(P_i\psi)\mathbf{x} = -i\frac{\partial\psi(\mathbf{x})}{\partial x_i} \quad \forall \mathbf{x} \in \mathbb{R}^3, \quad \forall \psi \in D_{P_i}$$

where $D_{P_i} = \{\psi \in L_2(\mathbb{R}^3) : x_i\psi \text{ absolutely continuous and } \frac{\partial\psi}{\partial x_i} \in L_2(\mathbb{R}^3)\}$.

Now we ask whether position and momentum can be measured together or not on the same physical system; this is possible if and only if the respective operators commute; straightforward calculations show that

- if $i = j$, then $[Q_i, P_i] = i\mathbf{1}$
- if $i \neq j$, then $[Q_i, P_j] = \mathbf{0}$

Hence we may conclude that different components of position and momentum can be measured together, whereas same components cannot.

1.7 Uncertainty relations

In previous section we have dealt with the question of measuring simultaneously two or more observables. Now we are interested in observables, \mathcal{A} and \mathcal{B} , represented by non-commuting operators, A and B .

In spite of the hypothesis that A and B do not commute, states ψ may exist in which both quantities have defined values: these are, for instance, the common eigenvectors of both A and B (however, they cannot form a complete orthogonal set since, in this case, A and B would commute). In such a case, indeed, if $\psi_1, \psi_2 \dots$, are the set of common eigenvectors, we may extend it to a complete orthonormal set by the addition of an orthonormal set $\phi_1, \phi_2 \dots$; now, let us introduce the self-adjoint operator $T = \sum_n \lambda_n |\psi_n\rangle\langle\psi_n| + \sum_n \mu_n |\phi_n\rangle\langle\phi_n|$ where $\lambda_1, \lambda_2 \dots, \mu_1, \mu_2 \dots$ are non-degenerate eigenvalues of T , pairwise different to each other. The measurement of the associated observable \mathcal{T} produces one of the eigenvalues; if a λ_n is found, then we also know the values of \mathcal{A} and \mathcal{B} ; on the other hand, if a μ_n is the result, nothing can be said about the values of \mathcal{A} and \mathcal{B} .

If no common eigenvector exists for A and B , let $\langle A \rangle_\psi = \langle \psi | A \psi \rangle$ and $\langle B \rangle_\psi = \langle \psi | B \psi \rangle$ be the expectation values of \mathcal{A} and \mathcal{B} respectively; let us define the operators $(\Delta A)^2 = (A - \langle A \rangle_\psi)^2$ and $(\Delta B)^2 = (B - \langle B \rangle_\psi)^2$ and evaluate the variances $(\sigma_A)^2 = \langle \psi | (A - \langle A \rangle_\psi)^2 \psi \rangle$ and $(\sigma_B)^2 = \langle \psi | (B - \langle B \rangle_\psi)^2 \psi \rangle$. Straightforward calculations shows that $(\sigma_A)^2 = \|A\psi\|^2 - \langle A\psi, \psi \rangle^2$ and similarly $(\sigma_B)^2 = \|B\psi\|^2 - \langle B\psi, \psi \rangle^2$; (σ_A) and (σ_B) cannot be jointly zero, since in this case ψ would be a common eigenvector of A and B . But one of the variances, say (σ_A) , can be made arbitrary small by choosing a suitable state vector ψ .

Heisenberg [21] discovered a relation preventing both variances from becoming arbitrary small

$$\sigma_A \sigma_B \geq \frac{1}{4} |\langle \psi | [A, B] \psi \rangle|^2 \quad (1.2)$$

known as *uncertainty relation*. For a proof see [22]; in [16] (p. 233) the same result is established for position and momentum operators (or more generally

for canonically conjugate operators) which satisfy $[\sigma_P\sigma_Q \geq \frac{\hbar}{2}]$.

As pointed out by [23], if conditions

1. A and B do not commute,
2. neither A nor B is bounded,
3. A and B have no point spectrum, or, if either of them has nonempty point spectrum, its eigenvectors lie outside the domain of the other.

are all fulfilled, this inequality states that for certain pairs of observables the right-hand side of (1.2) majorizes some positive number, independently of the choice of ψ ; otherwise, if at least one condition is not fulfilled, it is always possible to exhibit a state vector ψ which makes the right-hand side arbitrary small or null.

Result (1.2) can be interpreted as follows: if A and B do not commute, observables \mathcal{A} and \mathcal{B} cannot be measured together on the same physical system; however two disjoint sub-ensembles can be extracted from the ensemble corresponding with ψ , one for measuring \mathcal{A} and the other for measuring \mathcal{B} . The uncertainty relation states that there exists a lower bound for the product $\sigma_A\sigma_B$ of the variances outcoming from these measurements.

We emphasize that σ_A and σ_B are not errors of measurement; as pointed out by Ballentine [19], the error (or preferably the resolutions) δA and δB of the measuring instruments, are unrelated to σ_A and σ_B , except for the practical requirement that if δA is larger than σ_A then it is not possible to determine σ_B in the experiment, and so it is not possible to test (1.2).