## 7. Quantum Mechanical Systems.*

In this lecture we shall describe a few aspects of quantum mechanics. Obviously we cannot be exhaustive here, but we will try to mention a number of important foundational points. (For further elucidation, see von Neumann [2], Mackey [1, 2], Jauch [1], Varadarajan [1], Chernoff-Marsden [1]).

In order to clarify the differences between classical and quantum mechanics, it is convenient to adopt a probabilistic point of view and think in terms of statistical mechanics rather than particle mechanics. We begin with some general considerations.

### Basic Properties of Physical Systems.

A physical system consists of two collections of objects, denoted $S$ and $\mathcal{O}$ -- called states and observables respectively -- together with a mapping

$$S \times \mathcal{O} \to (\text{Borel probability measures on the real line } R)$$
$$(\psi, A) \mapsto \mu_{A,\psi} \quad .$$

Additionally, there is usually a Hamiltonian structure described below.

Elements $\psi \in S$ describe the state of the system at some instant and elements $A \in \mathcal{O}$ represent "observable quantities"; when $A$ is measured and the system is in state $\psi$, $\mu_{A,\psi}$ represents the probability distribution for the observed values of $A$. Thus if $E \subset R$, $\mu_{A,\psi}(E) \in R$ is the probability that we will measure the value of $A$

*This lecture was prepared in collaboration with P. Chernoff.

to lie in the set $E$ if the system is known to be in state $\psi$.

Normally there is also some dynamics; i.e. a flow or evolution operator $U_t : \mathbb{S} \to \mathbb{S}$.

The set $\mathbb{S}$ is usually a convex set and $U_t$ consists of convex automorphisms. The set $P$ of extreme points -- called the pure states, is usually symplectic and (for conservative systems) the flow $F_t$ on $P$ is Hamiltonian. (The flows $F_t$, $U_t$ determine one another.)

Statistical Mechanics.

Consider now the following example: let $P$, $\omega$ be a symplectic manifold, say finite dimensional and define the states and observables by:

(a) States; $\mathbb{S}$ consists of probability measures $\nu$ on $P$.

(b) Observables; $\mathbb{O}$ consists of real valued functions $A : P \to R$.

(c) The map $\mathbb{S} \times \mathbb{O} \to$ (Borel measures on $R$) is given by $\mu_{\nu,A}(E) = \nu(A^{-1}(E))$, where $E \subset R$.

The states are measures rather than points of $P$ to allow for the fact that we may only have a statistical knowledge of the "exact" state.

It is easy to see that the pure states are point measures, so are in one-to-one correspondence with points of $P$ itself. Note that every observable $A$ is sharp in a pure state; i.e. the corresponding

measure on $R$ is a point measure. In other words there is no dispersion when measuring any observable in a pure state.

Around 1930, B. O. Koopman noted that the above picture can be expressed in Hilbert space language. Let $\mathcal{H}$ denote the Hilbert space of all square integrable functions $\psi : P \to C$, with respect to Liouville measure. Each $\psi \in \mathcal{H}$ determines a probability measure $\nu_\psi = |\psi|^2 \mu$ if $||\psi|| = 1$. If $A$ is an observable, its expected value is

$$\mathcal{E}(A) = \int_P A|\psi|^2 d\mu = \langle A\psi, \psi \rangle$$

where $A$ is regarded as a (self adjoint) multiplication operator on $\mathcal{H}$.

The dynamics $F_t : P \to P$ on phase space $P$ induces in a natural way, and is induced by (under certain conditions) a dynamics on $S$ and on $\mathcal{H}$, namely $U_t \nu = F^*_{-t} \nu$ and $U_t \psi = \psi \circ F_{-t}$.

Consider the map $\psi \mapsto \nu_\psi$ of $\mathcal{H}$ to $S$. It is many-to-one. In fact $\nu_\psi = \nu_{\psi'}$ if $\psi' = e^{i\alpha}\psi$ where $\alpha : P \to R$. These phase transformations $\psi \mapsto e^{i\alpha}\psi$ form the <u>phase group</u> of classical mechanics.

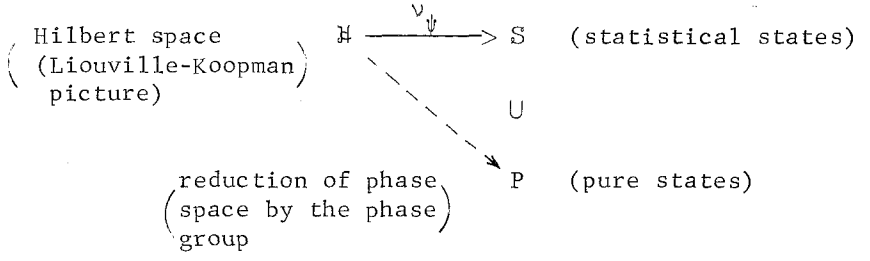
It is not hard to see that an operator $A$ on $\mathcal{H}$ is a multiplication operator iff it commutes with all phase transformations. Classical observables are those $A$'s which are self-adjoint; i.e. real valued.

Since only the measures have physical meaning, we see that any quantity of physical meaning must be invariant under the phase group.

It follows that the inner products $\langle\psi, \varphi\rangle$ and their squares $|\langle\psi, \varphi\rangle|^2$ can have no physical meaning. One says that there is "no coherence" in classical mechanics. (This is because $|\langle e^{i\alpha}\psi, e^{i\beta}\varphi\rangle| \neq |\langle\varphi, \psi\rangle|^2$ in general.)

We can think of $\mathcal{H}$ as a symplectic manifold with the usual symplectic form: $\omega = \mathrm{Im}\langle\ \rangle$. The dynamics induced on $\mathcal{H}$ is unitary and thus symplectic; i.e. it is Hamiltonian (see lecture 2). Thus the dynamics on $\mathcal{H}$ is consistent with the statistical interpretation.

We can regard the phase group $G$ as a symmetry group of $\mathcal{H}$ by $\psi \mapsto e^{i\alpha}\psi$. Indeed, as explained in lecture 6 we can form the reduced phase space; we just get back $P$ and the old dynamics on $P$. We have the following picture




Quantum Mechanics.

Quantum mechanics differs from classical mechanics in that the phase group is much smaller; interference and coherence -- typical wave phenomena -- now play a fundamental role. Furthermore, all predictions are necessarily statistical in that there are no dispersion free states ($\psi \in S$ is dispersion free when $\mu_{A,\psi}$ is a point measure for each $A \in \mathcal{A}$).

In classical mechanics, each state $\nu \in S$ was a "mixture" of pure states. We use $\nu$ because of ignorance as to the true state. Increasing our knowledge will "reduce" $\nu$ to a measure with smaller variance.

In quantum mechanics, states are not always reducible into statistical states of mixtures. This is clearly illustrated by experiments with polarized beams of coherent light (even with single photons). In such an experiment, states can be described by unit vectors $\psi \in R^2$ giving the direction of polarization. The probability that a $\varphi$ wave passes through a $\psi$ filter is observed to be $|\langle\varphi, \psi\rangle|^2$. A little thought shows that no such polarized state $\varphi$ can be realized as a statistical mixture of other polarized states.*

These sorts of experimental facts lead one to consider the states as forming a Hilbert space† $\mathcal{H}$ and the states as being the unit rays in $\mathcal{H}$. (These are the pure states; mixed states corresponding to $\nu$'s above are introduced below.) Thus, letting $P$ denote the rays in $\mathcal{H}$ ($P$ is called projective Hilbert space), we have a map $\mathcal{H} \to P$, again many to one. This time the phase group is the circle

---

* Furthermore, the experiment is not reproducible in the sense that no matter how carefully $\varphi$ is prepared, there is uncertainty in the outcome (unless the probability is 0 or 1). Such an uncertainty seems to be fundamental.

† We take $\mathcal{H}$ to be complex but it is not a priori clear why it shouldn't be real. There are good reasons for the complex structure related to the Hamiltonian structure; (see lecture 2 and references in Jauch [1]).

group $\{e^{i\alpha};\ \alpha \in R\}$ . The reason $P$ is chosen this way is that one imagines general elementary selective measurements wherein $|\langle\psi, \varphi\rangle|^2$ , for each $\psi$ , $\varphi \in \mathcal{H}$ , $\|\psi\| = \|\varphi\| = 1$ is the object with physical meaning -- it represents the probability that we will find $\varphi$ in state $\psi$ or if you like, the "transition probability" for going from $\varphi$ to $\psi$ .

More generally, we can imagine a general selection measurement. let $F \subset \mathcal{H}$ be a (closed) subspace and $\varphi \in \mathcal{H}$ . The probability of transition from $\varphi$ to $F$ is $\langle P_F\varphi, \varphi\rangle$ where $P_F$ is the orthogonal projection onto $F$ .

Just as in the case of statistical mechanics we observe that $P$ is the reduction of $\mathcal{H}$ by the phase group (this was noted in lecture 6).

Once the above view is accepted, then as Mackey has shown, the rest of the picture of what $\mathcal{S}$ , $\mathcal{O}$ and $\mu_{A,\psi}$ have to be is pretty much forced upon us. This goes as follows.

Consider an observable $A$ . For each $E \subset R$ we have $\mu_{A,\psi}(E)$ measuring a probability of observing $A$ to lie in $E$ if the state is $\psi$ . The previous discussion suggests there should be a projection operator $P_E^A$ on $\mathcal{H}$ such that

$$\mu_{A,\psi}(E) = \langle P_E^A\psi, \psi\rangle .$$

Since $\mu$ is a probability measure we must have:

$$P^A_{\emptyset} = 0 \ , \quad P^A_R = I \tag{1}$$

and

$$P^A_{\bigcup_{i=1}^{\infty} E_i} = \sum_{i=1}^{\infty} P^A_{E_i} \tag{2}$$

if $E_i$ are disjoint. It follows that the $P^A_{E_i}$ are mutually orthogonal. We also must have by (2),

$$P^A_{E\cup F} = P^A_{E\setminus F} + P^A_{F\setminus E} + P^A_{E\cap F}$$

$$P^A_E = P^A_{E\setminus F} + P^A_{E\cap F}$$

and $$P^A_F = P^A_{F\setminus E} + P^A_{E\cap F} \ .$$

Hence $P^A_E P^A_F = P^A_{E\cap F} = P^A_F P^A_E$ ; i.e. the $P^A_E$ 's commute.

The spectral theorem (see, e.g. Yosida [1]) now tells us that there is a unique self adjoint operator, also denoted $A$ , such that $A = \int_{-\infty}^{\infty} \lambda dP^A_\lambda$ ; $\{P^A_\lambda\}$ is the spectral measure of $A$ . Conversely any self adjoint operator $A$ yields a spectral measure and hence defines $\mu_{A,\psi}$ .

Thus, to every observable there is a self adjoint operator $A$ , but it is not clear that every self adjoint operator is physically realizable. (For example it is not clear how to measure (position) plus (momentum) = $q + p$ in the laboratory.)

Of course it is well known that a self adjoint operator (like the position operator) need not have any square integrable

eigenfunctions. What is asserted to be of physical relevance is the probability measure $\mu_{A,\psi}$, which is always well defined. Of course, one must avoid trivial "paradoxes" in quantum mechanics which arise from an inadequate understanding of the spectral theorem, or by ascribing more physical meaning (e.g. individual trajectories) to the theory than that given by the $\mu_{A,\psi}$.

Notice that the expected value of $A$ in a state $\psi$ is $\mathcal{E}(A) = \int_{-\infty}^{\infty} \lambda \, d\mu_{A,\psi}(\lambda) = \int_{-\infty}^{\infty} \lambda \, d\langle P_\lambda^A \psi, \psi\rangle = \langle A\psi, \psi\rangle$.

Thus a state $\varphi$ yields a mapping $F \mapsto \langle P_F \varphi, \varphi\rangle$ of subspaces in $\mathcal{H}$ to $[0, 1]$ describing a transition probability. It is a "probability measure" based on the closed subspaces.

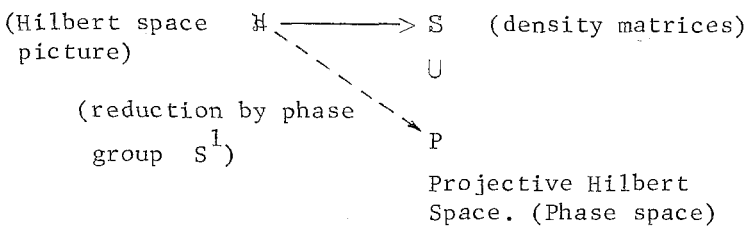
We can generalize the notion of state so as to allow for the possibility of mixed states (with the same statistical interpretation as in the classical case) by just considering a general "measure" defined on the closed subspaces of $\mathcal{H}$. It is a famous theorem of Gleason (see Varadarajan [1] for a proof) that such a state is given by $F \mapsto \text{trace } (P_F D)$ where $D$ is a positive operator of trace one on $\mathcal{H}$, called a <u>density matrix</u>.

Thus quantum mechanics is specified as follows: we are given a complex Hilbert space $\mathcal{H}$ and set

$$\begin{cases} \mathcal{S} = \text{all density matrices, a convex set} \\ \mathcal{O} = \text{self adjoint operators on } \mathcal{H} \\ \mu_A(E) = \text{trace } (P_E^A D) \text{ , } P_E^A \text{ the spectral projections of } A . \end{cases}$$

It is not hard to see that the pure states (extreme points of $\mathbb{S}$) are identifiable with unit vectors in $\mathbb{H}$, modulo the phase group -- what we previously called $P$.

Thus we again get this picture:




Bargmann-Wigner Theorem.

In the case of statistical mechanics we saw that the flow on $P$ naturally induced one on $\mathbb{H}$. For quantum mechanics this is not so obvious, and was considered by Bargmann-Wigner (see Varadarajan [1] for proofs and details). Since $P$ are the extreme points of $\mathbb{S}$, we can just as well work with $\mathbb{S}$ as $P$.

Theorem. Let $U_t$ be a flow by convex automorphisms on $\mathbb{S}$. Then $U_t$ is induced by a one parameter unitary group $V_t$ on $\mathbb{H}$, unique up to phase factors.

The result is conceptually important because $\mathbb{H}$ is a mathematical construct for analytical convenience. Only $P$ should be directly physically relevant.

Note. A convex automorphism of $\mathbb{S}$ can be implemented by either a unitary or anti-unitary operator (Wigner), but for one parameter groups

the later case is excluded.

One can go on to see when actions of groups lift from $P$ (or $S$) to $\mathcal{H}$. Then there is a cohomology condition needed on the group (see Simms [1], Chichilnisky [1] for more information).

We also remark that $S$ is, in the classical case, a simplex. This means that each state in $S$ can be uniquely represented in terms of the extreme points (see Choquet [1] for general information on this point).

Miscellany.

At this point one might ask: what do we choose for $\mathcal{H}$? The answer actually is that $\mathcal{H}$ should be $L_2$ of the configuration (rather than phase) space. The reason for this is best seen through an analysis given by Mackey and Wightman. The result is that if one makes entirely reasonable hypotheses on what the position and momentum operators ought to be, then their structure and that of the Hilbert space is determined. One finds that if the classical phase space is $P = T^*M$, then the quantum mechanical Hilbert space is $\mathcal{H} = L_2(M, C)$ and the quantum operators corresponding to a position observable $f$ (a function $f : M \to R$) and a momentum observable $P(X)$ ($P(X) \cdot \alpha = \alpha(X)$, $X$ a vector field on $M$ -- see lecture 6) are:

$$Q_f = \text{multiplication by } f$$

$$\text{and } P_X = iX \text{ as a differential operator .}$$

The associations $f \mapsto Q_f$, $P(X) \mapsto P_X$ are often called the Dirac

Quantization Rules. They preserve bracket operations.

This suggests that corresponding to the classical system $K + V$ on $TM$ is the quantum system with energy operator $-\Delta + V$ on $\mathcal{H}$. (Some problems related to this are discussed in the next lecture.)

Exact quantum procedures are not so simple. In fact an old theorem of Groenwald and Van Hove asserts that there is no map possible from all classical observables to quantum observables that preserves the bracket operations. However much work is currently being done on some geometric aspects of this problem (see Souriau [1]).

Another fundamental question is the reverse problem: in what sense is classical mechanics a limit of quantum mechanics (as $h$, Planck's constant $\to 0$) ? This has been investigated by many people, but the deepest analysis seems to be due to Maslov (see Arnold [2]). This problem is discussed further in the Appendix.

The C* Algebra Approach to Quantum Mechanics.

There are many ways of generalizing the examples of physical systems given in the first part of the lecture. One of these, taken by von Neumann, is to regard the set of observables as an algebra. This is mathematically convenient although it may not correspond exactly with physical reality for as mentioned above, the sum of two observables need not be observable. Other ways of generalization are the "quantum logic" point of view described in Varadarajan [1] and Mackey [1].

In the classical case the algebra is the algebra of functions

on phase space -- a commutative algebra. The quantum case is distinguished by having a non-commutative algebra. Indeed any C* algebra which is commutative must be isomorphic to a space of continuous functions and so, is in this sense, classical.

Segal's formulation of this point of view proceeds as follows. Let $\mathcal{A}$ be a C* algebra; i.e. a Banach space which is also an algebra and has a conjugation (or adjoint) operation * satisfying certain simple axioms. For example one can think of an algebra of bounded operators on a Hilbert space (unbounded operators are included via their spectral projections). Simmons [1] contains a very readable account of the elementary properties of C* algebras.

Take the observables to be the self adjoint elements of $\mathcal{A}$.

The states are the normalized positive linear functionals on $\mathcal{A}$. (It is easy to see that they are automatically continuous.) We are to think of states in the same way as before. If $\mathcal{E}$ is a state, $\mathcal{E}(A)$ is the expectation of $A$ in the state $\mathcal{E}$.

Of central importance is the Gelfand-Naimark-Segal construction: Let $\mathcal{A}$ be a C* algebra and $\mathcal{E}$ a state of $\mathcal{A}$. Then there is a Hilbert space $\mathcal{H}$, a unit (cyclic) vector $\psi \in \mathcal{H}$ and a *-representation $\pi_{\mathcal{E}} : \mathcal{A} \to \mathcal{L}(\mathcal{H})$ (the bounded operators on $\mathcal{H}$) such that

$$\mathcal{E}(A) = \langle \pi_{\mathcal{E}}(A)\psi, \psi \rangle \quad \text{for all} \quad A \in \mathcal{A}.$$

In fact $\mathcal{H}$, $\psi$, $\pi$ are unique up to unitary equivalence. See Lanford [1] for details.

In this way, we can construct our probability measure $\mu_{A,\mathcal{E}}$. Thus we have a general example of a physical system consisting of $S$, $\mathcal{O}$ and the map $\mu_{A,\mathcal{E}}$ just constructed which includes both classical and quantum systems as special cases.

There is no canonical Hilbert space, but one can be constructed for each $\mathcal{E}$. We can still form $P$, the extreme points of $S$, but in general $P$ won't be a symplectic manifold. (It is in the examples previously constructed however.)

The above Gelfand-Naimark-Segal construction is similar to Gleason's theorem in that it delineates states. It essentially enables one to recover the Hilbert space formalism from the abstract C* algebra formalism. However, often it is convenient to stick with the general C* algebra point of view. For example, one can characterize pure states $\mathcal{E}$ as those for which $\pi_{\mathcal{E}}$ is irreducible.

Several other ideas from the Hilbert space approach carry over. For example the general form of the uncertainty principle is valid: for observables $A$, $B \in \mathcal{A}$, and a state $\mathcal{E}$,

$$\sigma(A, \mathcal{E})\sigma(B, \mathcal{E}) \geq \tfrac{1}{2}\mathcal{E}(C) \quad , \quad C = i(AB - BA)$$

where $\sigma(A, \mathcal{E})$ is the variance of the probability distribution $\mu_{A,\mathcal{E}}$: $\sigma(A, \mathcal{E})^2 = \mathcal{E}(A^2) - (\mathcal{E}(A))^2 = \mathcal{E}((A - \mathcal{E}(A)I)^2)$.

Proof. Let $[X, Y] = \mathcal{E}(XY^*)$. This is an inner product on $\mathcal{A}$ so obeys the Schwartz inequality. Note that it is enough to prove the inequality

in case $\mathcal{E}(A) = 0$ , $\mathcal{E}(B) = 0$ for we can replace $A$ , $B$ by $A - \mathcal{E}(A)I$ , and $B - \mathcal{E}(B)I$ . Then

$$\begin{aligned} \mathcal{E}(C) &= i[\mathcal{E}(AB) - \mathcal{E}(BA)] \\ &= 2 \operatorname{Im} [A, B] \\ &\leq 2 [A, A]^{\frac{1}{2}}[B, B]^{\frac{1}{2}} \end{aligned}$$

$$\text{so} \quad \tfrac{1}{2}\mathcal{E}(C) \leq \sigma(A, \mathcal{E})\sigma(B, \mathcal{E}) \ , \ Q \ E \ D \ .$$

A Hidden Variables Theorem.

The orthodox interpretation of quantum mechanics presented above has discomforted many physicists, notably including Planck, Einstein, de Broglie, and Schrodinger (see for example De Broglie [1] and Einstein-Podolsky-Rosen [1]). It is hard to escape the feeling that a statistical theory must be, in some sense, an incomplete description of reality. One might hope that the probabilistic aspects of the theory are really due, as in the case of classical statistical mechanics, to some sort of averaging over an enormous number of "hidden variables"; in a perfect description of a state, in which these hidden parameters would have well-determined values, all the observables would be sharp. However, von Neumann [2] has given a proof that the results of quantum mechanics are not compatible with a reasonably formulated hidden variable hypothesis. We shall outline an argument along von Neumann's lines, but in the more general setting of Segal's C*-algebra formulation of quantum theory.

Let the observables of a given physical system be represented by the self-adjoint elements of a C* algebra $\mathcal{A}$ . If $A \in \mathcal{A}$ is an

observable and $\rho$ is a state, the dispersion of $A$ in the state $\rho$ is given by $\sigma^2(A, \rho) = \rho(A^2) - \rho((A - \rho(A)I^2)$ . We shall say that $\rho$ is a dispersion-free state provided that $\sigma^2(A, \rho) = 0$ for every observable $A \in \mathcal{A}$ . The results of experiment show that the states of quantum systems prepared in the laboratory are not dispersion-free. The hidden-variable hypothesis is that the physical state $\rho$ owes its dispersion to the fact that it is a statistical ensemble of ideal dispersion-free states. (The latter need not be physically realizable -- just as one cannot really prepare a classical gas with precisely determined positions and velocities for each of its molecules.) Mathematically, the hypothesis states that every state $\rho$ is of the form

$$\rho(A) = \int_{\Omega} \rho_{\omega}(A) d\mu(\omega) \tag{1}$$

where each $\rho_{\omega}$ is a dispersion-free state and $\mu$ is a probability measure on some space $\Omega$ . The coordinate $\omega \in \Omega$ represents, of course, the indeterminate "hidden variables".

Theorem. (See Segal [4].) A C* algebra $\mathcal{A}$ admits hidden variables in the above sense only if $\mathcal{A}$ is abelian. (The corresponding physical system is then "classical".)

Proof. (Chernoff) The first step is to show that a dispersion-free state $\rho_{\omega}$ is multiplicative. Note that the bilinear form $\langle\langle A, B\rangle\rangle = \rho_{\omega}(AB^*)$ is a Hermitian inner product on $\mathcal{A}$ . ($\langle\langle A, A\rangle\rangle = \rho_{\omega}(AA^*)$ is $\geq 0$ by hypothesis. From this it follows

easily that $\rho_\omega(C^*) = \overline{\rho_\omega(C)}$ for any $C \in \mathcal{A}$. In particular we have $\langle\langle B, A\rangle\rangle = \rho_\omega((AB^*)^*) = \overline{\rho_\omega(AB^*)} = \overline{\langle\langle A, B\rangle\rangle}$.) Hence, by the Schwarz inequality,

$$|\rho_\omega(AB)| \leq \rho_\omega(AA^*)^{\frac{1}{2}}\rho_\omega(B^*B)^{\frac{1}{2}}$$

for all $A, B \in \mathcal{A}$. From this we see that if $\rho_\omega(AA^*) = 0$ then $\rho_\omega(AB) = 0$ for all $B$. Suppose that $A$ is self-adjoint. Then, since $\mathcal{A}$ is dispersion-free, $\rho_\omega((A - \rho_\omega(A)I)^2) = 0$. Therefore, for every $B$, $\rho_\omega((A - \rho_\omega(A))B) = 0$. That is, $\rho_\omega(AB) = \rho_\omega(A)\rho_\omega(B)$. This holds as well for non-self-adjoint $A$ by linearity. In particular, if $\mathcal{A}$ is dispersion-free it follows that $\rho_\omega(AB) = \rho_\omega(BA)$.

But if $\mathcal{A}$ admits hidden variables, it follows immediately from (1) that every state $\rho$ satisfies $\rho(AB) = \rho(BA)$. Since there are enough states to distinguish the members of $\mathcal{A}$, (e.g. states of the form $A \mapsto \langle A\psi, \psi\rangle$) it follows that $AB = BA$. Thus $\mathcal{A}$ is abelian. □

Remark. Conversely, a well-known theorem of Gelfand and Naimark states that every abelian C* algebra is isomorphic to $C(X)$, the set of continuous functions on some compact set $X$. (Many accounts of this result are available; a very readable one is in Simmons [1].) The states of $\mathcal{A}$ are simply the probability measures on $X$, which are convex superpositions of the $\delta$-measures at the points of $X$; the latter are, of course, precisely the dispersion-free states.

We can also dispose of a less stringent notion of hidden variables. According to Jauch [11], Mackey has proposed the

consideration of "$\varepsilon$-dispersion-free" states. A state $\rho$ is called $\varepsilon$-dispersion-free if for every projection $E \in \mathcal{A}$ we have $\rho^2(E, \rho) < \varepsilon$. A system is said to admit "quasi-hidden variables" if for all $\varepsilon > 0$, every state can be represented as $\int \rho_\omega d\mu(\omega)$ where all the states $\rho_\omega$ are $\varepsilon$-dispersion-free. (In fact one can say "if for some $\varepsilon > 0$ sufficiently small" but this leads to a harder theorem). If $\mathcal{A}$ admits quasi-hidden variables and $\rho$ is a pure state of $\mathcal{A}$, then it is easy to see that $\rho$ is $\varepsilon$-dispersion free for every $\varepsilon$. Then by the argument above $\rho$ must be multiplicative on the algebra generated by the projections in $\mathcal{A}$. This will be all of $\mathcal{A}$ in many interesting cases -- in particular, if $\mathcal{A}$ is a von Neumann algebra (i.e. closed in the strong operation topology). But then, because the pure states separate elements of $\mathcal{A}$, it follows as before that $\mathcal{A}$ is abelian. (We must hasten to add that Jauch and Mackey were considering these questions in the context of lattices of "questions" which are more general than the projection lattices which we have discussed; so from the foundational point of view the notion of quasi-hidden variables has raised problems which our simple argument cannot handle.)

The essential point of the argument given above was the non-existence in general of a large supply of linear functionals on $\mathcal{A}$ which carry squares to squares. A much deeper analysis has been carried out by Kochen and Specker [11], cf. also Bell [1]. They have faced squarely the fact, which we have mentioned, that it is really not physically reasonable for the sum of non-commuting observables always to be an observable. Drastically reducing the algebraic operations

which they allow, they nevertheless reach the same results; their functionals are required to be linear only on commuting observables. We shall not go into the details of their arguments, for which we refer the reader to their paper, which also includes an interesting discussion of the entire problem of hidden variables and various attempts to introduce them. Some recent work on this subject centering around Bells' inequality has been done. The results again are against hidden variables theories. See Clauser et al [1] and Freedman-Clauser [1].

The Measurement Process.

Let us now discuss the process of measurement in some detail, following von Neumann [1]. (A clear summary of von Neumann's ideas may be found in the book of Nelson [2]; see also Jauch [1] and de Broglie [1].)

Various solutions of the problems of measurement have been proposed; cf. Bohm and Bub [1]. However it is not yet clear that the problems have been solved. The measurement of an observable involves the interaction of a "physical system" with an "observing apparatus", so we should first describe the mathematical treatment of such composite systems.

If the pure states of a system $S$ correspond to the unit rays of $\mathcal{H}$, and those of a second system $S'$ correspond to the rays of $\mathcal{H}'$, then the pure states of the compound system consisting of $S$ and $S'$ correspond to the unit rays of the tensor product* $\mathcal{H} \otimes \mathcal{H}'$.

* The tensor product $\mathcal{H} \otimes \mathcal{H}'$ is the direct product in the category of Hilbert spaces, just as the cartesian product is in the category of manifolds (if $P$ and $P'$ are phase spaces for isolated systems $P \times P'$ is the phase space for the interacting system). A pure state in a composite quantum system is much more complicated than an ordered pair of pure states of the subsystems. This fact seems related to many, if not all, of the so-called "paradoxes" of quantum theory.

The tensor product of Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ is by definition the completion of their algebraic tensor product with respect to the following inner product:

$$\langle \sum_i \varphi_i \otimes \varphi_i', \sum_j \psi_j \otimes \psi_j' \rangle = \sum_{i,j} \langle \varphi_i, \psi_j \rangle \langle \varphi_i', \psi_j' \rangle .$$

For example, $L^2(R^3) \otimes L^2(R^3) = L^2(R^6)$. If $\{e_i\}$ and $\{f_j\}$ are orthonormal bases of $\mathcal{H}$ and $\mathcal{H}'$ respectively, then $\{e_i \otimes f_j\}_{i,j=1}^{\infty}$ is an orthonormal basis of $\mathcal{H} \otimes \mathcal{H}'$. An observable $A$ of $S$ corresponds to the operator $A \otimes I$ on $\mathcal{H} \otimes \mathcal{H}'$; similarly the observable $B$ of $S'$ corresponds to $I \otimes B$. It can be shown that every observable of the composite system is a function of observables of the above sort, in the sense that every bounded operator on $\mathcal{H} \otimes \mathcal{H}'$ is a limit of operators of the form $\Sigma(A_i \otimes I)\cdot(I \otimes B_i)$. A state $\rho$ of the compound system determines a state of $S$ by the relation

$$\rho_S(A) = \rho(A \otimes I) .$$

It is important to note that $\rho_S$ will in general be a mixture even if $\rho$ is pure. Thus, if $\rho$ is given by the vector $\Sigma \varphi_1 \otimes \varphi_i'$, with $\{\varphi_i\}$, $\{\varphi_i'\}$ orthogonal systems in $\mathcal{H}$ and $\mathcal{H}'$, we have

$$\rho_S(A) = \Sigma \|\varphi_i'\|^2 \langle A\varphi_i, \varphi_i \rangle ,$$

so that $\rho_S$ is given by the density matrix $\Sigma \|\varphi_i'\|^2 P_{\varphi_i}$.

Now let $S$ be a physical system which we wish to study.

Suppose that we wish to measure an observable $A$ of $S$. For simplicity let us assume that $A$ has a pure point spectrum, with eigenvectors $\varphi_1, \varphi_2, \ldots$. To measure $A$ it is necessary to allow the system $S$ to interact with an apparatus $S'$. A suitable apparatus for measuring $A$ will have the property that, if the system $S$ is initially in the state $\varphi_i$, after the interaction the composite system of $S$ and $S'$ will be in the state $\varphi_i \otimes \theta_i$, where $\{\theta_i\}$ is a sequence of orthonormal vectors in $\mathcal{H}'$. The interaction, of course, is governed by the Schrodinger equation for the composite system. Hence, if the initial state of $S$ is given by $\psi = \sum_1^\infty c_i\varphi_i$, the final state of $S + S'$ will be $\theta = \sum_1^\infty c_i\varphi_i \otimes \theta_i$ by linearity. Now if $B$ is an observable of $S'$, then after the interaction the expected value of $B$ will be

$$\langle (I \otimes B)\theta, \theta\rangle = \sum_{i=1}^\infty |c_i|^2 \langle B\theta_i, \theta_i\rangle ,$$

so that, although $S + S'$ is in the pure state $\theta$, $S'$ is in the mixed state $\sum_{i=1}^\infty |c_i|^2 P_{\theta_i}$. Similarly, $S$ is in the mixed state $\sum_{i=1}^\infty |c_i|^2 P_{\varphi_i}$.

Now the apparatus is supposed to be of a macroscopic nature, its orthogonal states $\theta_i$ represent, say, different counter readings. After the interaction the observer "looks" at the apparatus. Through his faculty of introspection he realizes that the apparatus is in a definite state, say $\theta_j$. (This occurs with probability $|c_j|^2$.)

Once this act of consciousness has taken place it is no longer true that the state of $S + S'$ is $\sum_{i=1}^{\infty} c_i \varphi_i \times \theta_i$ ; it must be $\varphi_j \times \theta_j$ . One then says that the system has been found to be in the state $\varphi_j$ . This is the famous (or notorious) "reduction of the wave packet".*

We now venture to make some philosophical remarks. It is important to realize that an analogous "reduction" takes place in a classical statistical mechanical system when new information is gained. This is never regarded as a difficulty, because the classical probability packet is always viewed as a mere reflection of the observer's ignorance of the objective underlying state of the system. This is a perfectly consistent interpretation. Why can not the same interpretation serve in the quantum mechanical case?

As long as we are concerned only with a single observable (or with a commuting family of observables) it is perfectly possible to view the quantum system classically. That is, one can interpret the reduction from the mixture to the state $\varphi_j$ as a reduction of classical type. But the existence of incompatible observables in quantum mechanics forces this interpretation to break down. Indeed, the entire point of the negative results concerning "hidden variables"

---

* Of course, "looking at the apparatus" involves interaction with some further apparatus ultimately with the consciousness of the observer. But one can lump all that into $S$ and the observers mind into $S'$ . Nevertheless, apparently one cannot find a mathematical device to yield the reduction of pure states. This is the fundamental problem in interpreting the foundations of quantum mechanics.

is that there is no "objective underlying state" of the system!

Perhaps the quantum probability distributions can be interpreted as reflecting our partial knowledge, as long as we do not insist that there be an objective entity of which we have partial knowledge. This seems reminiscent of the problem of the golden mountain in the sentence "The golden mountain does not exist". If one asks "what does not exist?" and answers "the golden mountain", one is implying that the golden mountain is in fact an entity with some sort of "existence". Some philosophers tried to rescue the situation by stating that the golden mountain "subsists" -- that is, has enough of a shadowy sort of existence to serve as the subject of a sentence. Now Bertrand Russell has observed that the real solution of the problem is to recognize that the original sentence is implicitly quantified, and actually should be regarded as saying "for every $x$ it is false that $x$ is both golden and mountainous". In the absence of new physical discoveries, it seems not impossible that the same sort of purely grammatical trick may be the ultimate solution of the quantum measurement problem.