PAUL R. CHERNOFF AND JERROLD E. MARSDEN

SOME REMARKS ON HAMILTONIAN SYSTEMS AND QUANTUM MECHANICS*

1. INTRODUCTION

These notes contain some remarks on the general structure of a class of physical systems called Hamiltonian, and on quantum mechanical systems in particular. A complete treatment with emphasis on technical details is found in [6]; our goal here is to point out some unifying structures and special properties of physical systems that may be of interest to this conference. Thus many of our remarks are deliberately brief and sometimes vague. Most of the results are known in the literature (cf. [18, 23]) although perhaps from a different point of view.

We shall begin in Section 2 with the general features that a physical system admitting a probabilistic interpretation should have. The distinguishing features of classical and quantum mechanical systems are pointed out. In Section 3 the C*-algebra approach to quantum mechanics, as delineated by Segal, is reviewed. Then in Section 4 we study the dynamics of classical and quantum mechanics – we endeavor to show that both systems are Hamiltonian, when the latter condition is interpreted from the modern point of view of symplectic manifolds (see [2]). In Section 5 we briefly describe some other classes of Hamiltonian systems: specifically hydrodynamics and general relativity. Finally in Sections 6 and 7 we mention a few problems connected with hidden variables and the theory of measurement.

2. BASIC PROPERTIES OF PHYSICAL SYSTEMS

A physical system consts of two collections of objects, denoted $\mathscr{S}$ and $\mathscr{O}$ – called *states* and *observables* respectively, together with a mapping

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

Additionally, there is usually a Hamiltonian structure described below in Section 4.

Harper and Hooker (eds.), Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science, Vol. III, 35–53.

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

From the general point of view, the above seem to be minimal features that any physical system should possess. They have been axiomatized and studied by several authors; cf. Mackey [16]. The set $\mathscr{S}$ describes the basic mathematical structure we are dealing with, while $\mu_{A,\psi}$ provides the physical interpretation.

Normally there is some additional structure as well, namely the dynamics. The dynamics tells us how the system changes from time s to a (later) time t. Thus we are provided with a family of mappings

$$U_{t,s}: \mathscr{S} \to \mathscr{S},$$

called evolution operators.[1]

Causality, in the sense that a state at a given time uniquely determines the state at any other time[2] forces us to postulate the *flow property*:

$$\begin{cases} U_{t_2,t_1} = U_{t_2,s} \circ U_{s,t_1}; \\ U_{t,t} = \text{identity}. \end{cases}$$

If the dynamics or the 'law of motion' is idependent of time, then $U_{t,s}$ depends only on $t-s$; i.e. we can write $U_{t,s} = U_{t-s}$. Then U_t satisfies:

$$U_{t+s} = U_t \circ U_s.$$

One also calls U_t a semigroup or flow (nonlinear in general).

The flow $U_t: \mathscr{S} \to \mathscr{S}$ is determined by a law of motion: let $x \in \mathscr{S}$, write $x(t) = U_t(x)$ and set $A(x) = dx/dt \,|_{t=0}$. The flow property gives us the required law: $(d/dt)\, x(t) = A(x(t))$. We need enough structure on $\mathscr{S}$ for this to make sense; i.e. that $\mathscr{S}$ be a differentiable manifold.

The dynamics U_t determines and is usually determined by the operator A, called the *generator* of U_t. The differential equation $dx/dt = A(x)$ becomes the Schrödinger equation in quantum mechanics or Hamilton's canonical equations (or Newton's second law, if you wish) for classical mechanics.

Already at this stage the technical problems are enormous. For example,

one wishes to know just what operators A generate flows. This is an active area of research, especially in the nonlinear case; cf. [5].

It is in connection with the dynamics that the Hamiltonian structure enters. However, let us first consider some other general features of classical and quantum mechanics. We begin with classical mechanics from the modern point of view [2].

2.1. *Classical Mechanics*

A *classical mechanical system* consists of a finite dimensional manifold P called the *phase space*; a *symplectic structure* ω on P, that is, a closed ($d\omega=0$) differential 2-form of maximal rank; and a *Hamiltonian* or energy function $H:P\to\mathbb{R}$. From ω we can construct a measure, called Liouville measure: $\mu=\omega\wedge\ldots\wedge\omega$; from ω and H we can construct Hamilton's equations X_H, a vector field on P. X_H is determined by the relation $\iota_{X_H}\omega=dH$ where ι_X denotes the interior product and d the differential. Integrating the vector field X_H (i.e. solving a system of ordinary differential equations) yields the dynamics, $U_t:P\to P$.

Considerations from statistical mechanics lead to the following interpretation:

(a) States: $\mathscr{S}$ consists of probability measures ν on P

(b) Observables: $\mathscr{O}$ consists of measureable realvalued functions $A:P\to\mathbb{R}$

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

The states are measures rather than points of P to allow for the possibility that we may only have a statistical knowledge of the 'exact' state. Note that $\mathscr{S}$ is a convex set. Its extreme points are called *pure states* (see [7] for a detailed discussion of convex sets and extreme points; see [22] for details on the connections with statistical mechanics).

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 $\mu_{\nu,A}$ on $\mathbb{R}$ 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 $\mathscr{H}$ denote the Hilbert space of

all functions $\psi : P \to C$, square integrable with respect to Liouville measure. Each $\psi \in \mathscr{H}$ determines a probability measure $\nu_\psi = |\psi|^2 \, \mu$ if $\|\psi\| = 1$. If A is an observable, its expected value is

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

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

The dynamics $U_t : P \to P$ on phase space P induces in a natural way, and (under certain conditions) is induced by, a dynamics on $\mathscr{S}$ and on $\mathscr{H}$.

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

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

Because the phase group is so large, neither the inner products $\langle \psi, \varphi \rangle$ nor their squares $|\langle \psi, \varphi \rangle|^2$ can have any physical meaning. This is related to the *absence* of *coherence phenomena* in classical mechanics.

The dynamics U_t on P is Hamiltonian; that is, U_t preserves the 2-form ω. In other words U_t consists of canonical transformations. In particular U_t preserves Liouville measure. Thus the induced dynamics on $\mathscr{H}$ is unitary. Hence the dynamics is consistent with the statistical interpretation: probability is conserved.

2.2. *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 ($\psi \in \mathscr{S}$ is *dispersion-free* when $\mu_{A,\psi}$ is a point measure for each $A \in \mathscr{O}$).

In classical mechanics, each state $\nu \in \mathscr{S}$ was a 'mixture' of pure states, reflecting our ignorance of the true state. Increasing our knowledge will 'reduce' ν to a measure with smaller variance.

In quantum mechanics there are 'irreducible' statistical phenomena even when the system is in a pure state. This is clearly illustrated by experiments with beams of plane-polarized coherent light, or even with single

photons. In such experiments, states can be described by unit vectors $\psi \in \mathbb{R}^2$ giving the direction of polarization. The probability that a φ wave will pass through a ψ filter is observed to be $|\langle \varphi, \psi \rangle|^2$. Furthermore, the emerging wave is ψ-polarized. But a little thought shows that the φ-polarized state is not a statistical mixture of other polarized states.

This sort of experimental fact leads one to consider the states as being the unit rays in a Hilbert space $\mathscr{H}$[3]. (These are the pure states; mixed states corresponding to ν's above are introduced later.) Thus, letting $\mathscr{S}_0$ denote the rays in $\mathscr{H}$ ($\mathscr{S}_0$ is called projective Hilbert space), we have a map $\mathscr{H} \to \mathscr{S}_0$, again many-to-one. This time the *phase group* is the circle group $\{e^{i\alpha}, \alpha \in \mathbb{R}\}$. The reason $\mathscr{S}_0$ is chosen this way is that one imagines general *elementary selective measurements* wherein $|\langle \psi, \varphi \rangle|^2$, for each $\psi, \varphi \in \mathscr{H}$, $\|\psi\| = \|\varphi\| = 1$, is the object with physical meaning – it represents the probability that we will find φ in state ψ or, if you like, the 'transition probability' for going from φ to ψ.

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

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

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

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

Since μ is a probability measure we must have:

(1) $$P_\phi^A = 0, \qquad P_{\mathbb{R}}^A = I$$

(2) $$P^A_{\bigcup_{i=1}^{\infty} E_i} = \sum_{i=1}^{\infty} P_{E_i}^A$$

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

$$P_{E \cup F}^A = P_{E \setminus F}^A + P_{F \setminus E}^A + P_{E \cap F}^A$$

$$P_E^A = P_{E \setminus F}^A + P_{E \cap F}^A$$

and

$$P_F^A = P_{F\setminus E}^A + P_{E\cap F}^A$$

Hence

$$P_E^A P_F^A = P_{E\cap F}^A = P_F^A P_E^A; \quad \text{i.e. the } P_E^A\text{'s commute.}$$

The spectral theorem [4] now tells us that there is a unique self-adjoint operator, also denoted A, such that $A = \int_{-\infty}^{\infty} \lambda \, dP_\lambda^A$; $\{P_\lambda^A\}$ 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.[5]

Notice that the expected value of A in a state ψ is

$$\mathscr{E}(A) = \int_{-\infty}^{\infty} \lambda \, d\mu_{A,\psi}(\lambda) = \int_{-\infty}^{\infty} \lambda \, d\langle P_\lambda^A \psi, \psi\rangle = \langle A\psi, \psi\rangle .$$

Thus a state φ yields a mapping $F \mapsto \langle P_F \varphi, \varphi\rangle$ of subspaces in $\mathscr{H}$ 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 $\mathscr{H}$. It is a famous theorem of Gleason [6] that such a measure is given by $F \mapsto \text{trace } (P_F D)$ where D is a positive operator of trace one on $\mathscr{H}$, called a *density matrix*. (Here we must add the condition that $\mathscr{H}$ have dimension at least three.)

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

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

It is not hard to see that the pure states (extreme points of $\mathscr{S}$) are identifiable with unit vectors in $\mathscr{H}$, modulo the phase group – what we previously called $\mathscr{S}_0$.

We also postulate dynamics $U_t : \mathscr{S} \to \mathscr{S}$ on $\mathscr{S}$, and assume that U_t consists of convex automorphisms. It is a theorem (going back to Bargmann

and Wigner) that U_t is naturally induced by a one-parameter unitary group V_t on $\mathscr{H}$. The generator of V_t is H, the Hamiltonian. Conversely if H is self adjoint, it determines V_t by Stone's theorem: $V_t = e^{iHt}$ (see [28]).

Actually this dynamics is Hamiltonian in the same sense that classical dynamics is, as discussed in 2.1 above. We shall see this in Section 4 below.

3. The C^*-algebra approach to quantum mechanics

There are many ways of generalizing the examples of physical systems given in Sections 2.1, 2.2 above. One of these, taken by von Neumann and Segal, 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. Another type of generalization is the 'quantum logic' point of view described in other lectures and in [26].

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 version of this formulation proceeds as follows. Let $\mathfrak{A}$ be a C^*-algebra [7], (for example all bounded operators on Hilbert space). The observables [8] are the self-adjoint elements of $\mathfrak{A}$.

The *states* are the normalized positive linear functionals on $\mathfrak{A}$. They are automatically continuous.[9] We are to think of states in the same way as before. If $\mathscr{E}$ is a state, $\mathscr{E}(A)$ is the expectation of A in the state $\mathscr{E}$.

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

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

In face $\mathscr{H}$, ψ, π are unique up to unitary equivalence. See Lanford [15] for details.

In this way, we can construct our probability measure $\mu_{A,\mathscr{E}}$. Thus we have a general example of a physical system – consisting of $\mathscr{S}$, $\mathscr{O}$ and the

map $\mu_{A,\mathscr{E}}$ just constructed – which includes both classical and quantum systems as special cases.

The above construction is similar to Gleason's theorem in that it characterizes states. The Gelfand-Naimark-Segal construction essentially enables one to recover the Hilbert space formalism from the abstract C^*-algebra formalism. Often it is convenient to stick with the general formalism (e.g. see Section 6 below). For example, one can characterize pure states $\mathscr{E}$ as those for which $\pi_{\mathscr{E}}$ is irreducible.

Several other ideas carry over also. For example, a general form of the uncertainty principle is valid: for observables A, $B\in\mathfrak{A}$, and a state $\mathscr{E}$,

$$\sigma(A,\mathscr{E})\,\sigma(B,\mathscr{E})\geqslant\tfrac{1}{2}\mathscr{E}(C);\qquad C=i(AB-BA)$$

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

Proof. Let $[X, Y]=\mathscr{E}(XY^*)$. This is an inner product on $\mathfrak{A}$, so obeys the Schwarz inequality. Note that it is enough to prove the inequality in case $\mathscr{E}(A)=0$, $\mathscr{E}(B)=0$ for we can replace A, B by $A-\mathscr{E}(A)\,I$, and $B-\mathscr{E}(B)\,I$. Then

$$\begin{aligned}\mathscr{E}(C)&=i[\mathscr{E}(AB)-\mathscr{E}(BA)]\\&=2\,\mathrm{Im}\,[A,B]\\&\leqslant 2[A,A]^{1/2}[B,B]^{1/2}\end{aligned}$$

so $\frac{1}{2}\mathscr{E}(C)\leqslant\sigma(A,\mathscr{E})\,\sigma(B,\mathscr{E})$. Q.E.D.

4. The Hamiltonian structure

A general Hamiltonian system consists of a manifold P, possibly infinite dimensional, together with (a) a (weakly) nondegenerate, closed two form ω on P (i.e. ω is an alternating bilinear form on each tangent space T_xP of P, $\mathrm{d}\omega=0$, and for $x\in P$, $\omega_x(v,w)=0$ for all $v\in T_xP$ implies $w=0$) and (b) a Hamiltonian function $H:P\to\mathbb{R}$.

Then P, H, ω determine, in nice cases, a vector field X_H by the relation $i_{X_H}\omega=\mathrm{d}H$; i.e.,

$$\omega(X_H(x),v)=\mathrm{d}H_x\cdot v\quad\text{for all}\quad x\in P\quad\text{and}\quad v\in T_xP.$$

As in Section 2 we can require that X_H be the generator of a flow U_t on P.

Let us enrich our definition of a *physical system* as follows. It consists of

(a) a convex set of states $\mathscr{S}$
(b) a set $\mathscr{O}$ of observables
(c) a map $(\psi, A) \mapsto \mu_{A,\psi}$ from $\mathscr{S} \times \mathscr{O}$ to (Borel probability measures on $\mathbb{R}$)
(d) a flow of convex automorphism of $\mathscr{S}$, $U_t: \mathscr{S} \to \mathscr{S}$; this induces a flow $U_t: P \to P$, where P = extreme points of $\mathscr{S}$ = pure states.

We also assume:

(e) The flow on P arises from a Hamiltonian system.

The fact that we deal with $\mathscr{S}$ rather than P is due to our ignorance about the precise initial state. This stochastic aspect of physical models is fundamental. Moreover, there are other related questions. Usually we don't know the dynamics exactly either, so only *stable* states are of physical interest. Here stability means that the time evolution should not be affected much by a perturbation of either the initial condition or the vector field itself. (The latter property is called structural stability.) Such questions are not simple and much remains to be discovered; cf. [1].

We have already explained the Hamiltonian structure for classical mechanics. It remains to discuss it for quantum mechanics. (cf. [17]).

Thus let $\mathscr{H}$ be complex Hilbert space and let P be the corresponding projective Hilbert space. Let A be a self-adjoint operator on $\mathscr{H}$. We claim that for a suitable symplectic form ω on $\mathscr{H}$, iA is Hamiltonian. By 'quotienting' by the phase group, it can be shown that this symplectic structure is inherited by P.

The symplectic form will be given by an antisymmetric bilinear form $\omega: \mathscr{H} \times \mathscr{H} \to \mathbb{R}$, namely

$$\omega(\psi, \varphi) = \operatorname{Im}\langle \psi, \varphi \rangle .$$

The energy is

$$H(\psi) = \tfrac{1}{2}\langle A\psi, \psi \rangle .$$

The Hamiltonian condition, namely

$$\omega(iA\psi, \varphi) = \mathrm{d}H_\psi \cdot \varphi$$

is verified as follows:

$$\begin{aligned}
&\operatorname{Im}\langle iA\psi, \varphi\rangle \\
&\quad = \operatorname{Re}\langle A\psi, \varphi\rangle \\
&\quad = \tfrac{1}{2}[\langle A\psi, \varphi\rangle + \overline{\langle A\psi, \varphi\rangle}] \\
&\quad = \tfrac{1}{2}[\langle A\psi, \varphi\rangle + \langle \varphi, A\psi\rangle] \\
&\quad = \tfrac{1}{2}[\langle A\psi, \varphi\rangle + \langle A\varphi, \psi\rangle] \\
&\quad = \mathrm{d}H_\psi \cdot \varphi .
\end{aligned}$$

Thus we conclude that the usual Hilbert space formalism of quantum mechanics satisfies all the requirements (a)–(e) listed above.

The structure of some of the basic observables in quantum mechanics can be derived by the elegant Mackey-Wightman analysis and other group-theoretic arguments. In particular, that the position and momentum operators have to be x^j and $i\,(\partial/\partial x^j)$ follows from very fundamental and non-controversial hypotheses. See [26], [6], and [13]. Note that the usual Heisenberg uncertainty principle follows, as was explained in Section 3.

5. Some other examples of Hamiltonian systems

Several interesting physical systems can be put into the above Hamiltonian framework. They are all classical in the sense that $\mathscr{S}$ consists of the probability measures and $\mathscr{O}$ the real-valued functions on a given phase space P.

Examples are:

(a) wave equations such as

$$\frac{\partial^2\varphi}{\partial t^2} = \Delta\varphi - m^2\varphi - \alpha\varphi^p, \quad p \text{ an integer} \geqslant 2, \quad \alpha \in \mathbb{R}, \quad m \geqslant 0.$$

(b) Maxwell's equations

(c) hydrodynamics

(d) conservative continuum mechanics

(e) elasticity

(f) Einstein's equations of general relativity.

Of course many of these examples are interrelated. For details, see [6] and [18]. We shall make a few brief remarks concerning (c) and (f).

5.1. *The Motion of a Perfect Fluid*

Let v be the velocity field of a perfect (incompressible, homogeneous, non-viscous) fluid in a region D of $\mathbb{R}^3$. Euler's equations assert:

$$\text{(E)} \qquad \begin{cases} \dfrac{\partial v}{\partial t} + (v \cdot \nabla)\, v = -\operatorname{grad} p \\ \operatorname{div} v = 0 \\ v \text{ parallel to } \partial D \end{cases}$$

These equations are not in Hamiltonian form. But they can be so put as follows. Let $\eta(t, x) \in D$, $x \in D$, $t \in \mathbb{R}$ be defined by

$$\frac{\mathrm{d}}{\mathrm{d}t}\eta(t, x) = v(t, \eta(t, x))$$
$$\eta(0, x) = x.$$

Then $t \to \eta(t, x)$ is the trajectory followed by the particle which was initially at x. Let $\eta_t(x) = \eta(t, x)$. Then $\eta_t : D \to D$ is invertible (a diffeomorphism) and is volume-preserving. Let $\mathscr{D}_\mu$ denote all volume-preserving diffeomorphisms on D.

In 1967, V. Arnold showed that v satisfies (E) if and only if η_t is a geodesic on the infinite dimensional manifold $\mathscr{D}_\mu$. (The metric on $\mathscr{D}_\mu$ is canonically associated with the kinetic energy of the fluid.) Thus we immediately have the required Hamiltonian structure, since geodesics on a manifold M are well-known to arise from a Hamiltonian system on TM. See for example [2]. Thus $P = T\mathscr{D}_\mu$.

This point of view has turned out to be very useful technically for proving existence and representation theorems for the solutions of the equations of ydrodynamics. It may be of importance in problems of quantization as well.

5.2. *The Equations of General Relativity*

Consider a Lorentz manifold V and associated metric ${}^{(4)}g$. Outside regions of matter, Einstein's equations assert that V is Ricci-flat: $R_{\alpha\beta} = 0$. Pick a space-like hypersurface M in V and some orthogonal coordinate t; then, near M, V looks like $M \times$ (an interval in $\mathbb{R}$). Hence ${}^{(4)}g$ yields a curve $t \mapsto g_t$ of positive-definite Riemannian metrics on M. Let $\mathfrak{M}$ denote the set of all such metrics on M. Then one can show that $R_{\alpha\beta} = 0$ if and only if $t \mapsto g_t$ is a geodesic (with a potential term) in the space $\mathfrak{M}$.

Of course we can choose different M's and t-coordinates. In this way we get different Hamiltonian representations in $\mathfrak{M}$, but they are all equivalent.[11]

The phase space is thus $P=T\mathfrak{M}$. The symplectic structure is that associated with an indefinite, but non-degenerate metric on $\mathfrak{M}$ called the deWitt metric:

$$\mathscr{I}_g(h, k) = \int_M \{(\operatorname{tr} h)(\operatorname{tr} k) - \langle h, k\rangle\}\, d\mu_g$$

where tr is the trace, $\langle h, k\rangle = h^{ij}k_{ij}$ and $\mu_g = \sqrt{\det g_{ij}}\, dx^1 \wedge dx^2 \wedge dx^3$.

6. A HIDDEN VARIABLES THEOREM

The orthodox physical interpretation of quantum mechanics has discomforted many physicists, notably including Planck, Einstein, de Broglie, and Schrödinger; cf. [11]. 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 [27] 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 $\mathfrak{A}$. If $A\in\mathfrak{A}$ is an observable and ρ is a state, the dispersion of A in the state ρ is given by $\sigma^2(A, \rho)=\rho(A^2)-\rho(A)^2=\rho((A-\rho(A)I)^2)$. We shall say that ρ is a *dispersion-free state* provided that $\sigma^2(A, \rho)=0$ for every observable $A\in\mathfrak{A}$. 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 ρ 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.) Mathe-

matically, the hypothesis states that every state ρ is of the form

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

where each ρ_ω is a dispersion-free state and μ is a probability measure on some space Ω. The coordinate $\omega \in \Omega$ represents, of course, the indeterminate 'hidden variables'.

THEOREM. (Segal [24]). *A C^*-algebra $\mathfrak{A}$ admits hidden variables in the above sense only if $\mathfrak{A}$ is abelian. (The corresponding physical system is then 'classical'.)*

Proof. The first step is to show that a dispersion-free state ρ_ω is multiplicative. Note that the bilinear form $\langle\langle A, B\rangle\rangle = \rho_\omega(AB^*)$ is a Hermitian inner product on $\mathfrak{A}$. ($\langle\langle A, A\rangle\rangle = \rho_\omega(AA^*)$ is $\geqslant 0$ by hypothesis. From this it follows easily that $\rho_\omega(C^*) = \overline{\rho_\omega(C)}$ for any $C \in \mathfrak{A}$. In particular we have $\langle\langle B, A\rangle\rangle = \rho_\omega(BA^*) = \rho_\omega((AB^*)^*) = \rho_\omega(AB^*) = \langle\langle A, B\rangle\rangle$.) Hence, by the Schwarz inequality,

$$|\rho_\omega(AB)| \leqslant \rho_\omega(AA^*)^{1/2}\, \rho_\omega(B^*B)^{1/2}$$

for all $A, B \in \mathfrak{A}$. 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 $\mathfrak{A}$ 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 $\mathfrak{A}$ is dispersion-free it follows that $\rho_\omega(AB) = \rho_\omega(BA)$.

But if A admits hidden variables, it follows immediately from (1) that every state ρ satisfies $\rho(AB) = \rho(BA)$. Since there are enough states to distinguish the members of $\mathfrak{A}$ (e.g. states of the form $A \mapsto \langle A\psi, \psi\rangle$), it follows that $AB = BA$. Thus $\mathfrak{A}$ 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 [25].) The states of $\mathfrak{A}$ are simply the probability measures on X, which are convex superpositions of the δ-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 [13], Mackey has proposed the consideration of 'ε-dispersion-free' states. A state ρ is called ε-dispersion-free if for every *projection* $E \in \mathfrak{A}$ we have $\sigma^2(E, P) < \varepsilon$. 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 ρ_ω are ε-dispersion-free. If $\mathfrak{A}$ admits quasi-hidden variables and ρ is a *pure state* of $\mathfrak{A}$, then it is easy to see that ρ is ε-dispersion free for every ε. Then by the argument above ρ must be multiplicative on the algebra generated by the projections in $\mathfrak{A}$. This will be all of $\mathfrak{A}$ in many interesting cases – in particular, if $\mathfrak{A}$ is a von Neumann algebra (i.e. closed in the strong operation topology). But then, because the pure states separate elements of $\mathfrak{A}$, it follows as before that $\mathfrak{A}$ 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. But see [29].)

The essential point of the argument given above was the non-existence in general of a large supply of linear functionals on $\mathfrak{A}$ which carry squares to squares. A much deeper analysis has been carried out by Kochen and Specker [14]; cf. also Bell [3]. 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. Finally, we mention some recent experimental work in this area, centering around 'Bell's inequality'; the outcome argues against the hidden variable hypothesis. See [8, 12].

7. The measurement process

Let us now discuss the process of measurement in some detail, following von Neumann [27]. (A clear summary of von Neumann's ideas may be found in the book of Nelson [20]; see also Jauch [13] and de Broglie [9].)

Various solutions of the problems of measurement have been proposed; cf. [4] and J. Bub's lecture in this conference. 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 $\mathscr{H}$, and those of a second system S' correspond to the rays of $\mathscr{H}'$, then the pure states of the compound system consisting of S and S' correspond to the unit rays of the tensor product[12] $\mathscr{H}\otimes\mathscr{H}'$. (The tensor product of Hilbert spaces $\mathscr{H}$ and $\mathscr{H}'$ is by definition the completion of their algebraic tensor product with respect to the following inner product:

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

For example, $L^2(\mathbb{R}^3)\otimes L^2(\mathbb{R}^3) = L^2(\mathbb{R}^6)$. If $\{e_i\}$ and $\{f_j\}$ are orthonormal bases of $\mathscr{H}$ and $\mathscr{H}'$ respectively, then $\{e_i\otimes f_j\}_{i,j=1}$ is an orthonormal basis of $\mathscr{H}\otimes\mathscr{H}'$.) An observable A of S corresponds to the operator $A\otimes I$ on $\mathscr{H}\otimes\mathscr{H}'$; 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 $\mathscr{H}\otimes\mathscr{H}'$ is a limit of operators of the form $\sum (A_i\otimes I)\cdot(I\otimes B_i)$. A state ρ 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 ρ_S will in general be a mixture even if ρ is pure. Thus, if ρ is given by the vector $\sum \varphi_i\otimes\varphi_i'$, with $\{\varphi_i\}$, $\{\varphi_i'\}$ orthogonal systems in $\mathscr{H}$ and $\mathscr{H}'$, we have

$$\rho_s(A) = \sum \|\varphi_i'\|^2 \langle A\varphi_i, \varphi_i\rangle ,$$

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

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 φ_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 $\mathscr{H}'$. 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 θ, 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 θ_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 θ_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 φ_j. This is the famous (or notorious) 'reduction of the wave packet'.[13]

We now venture to make some philosophical remarks. It is important to realize that 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't 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 φ_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' 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.

Department of Mathematics,
University of California at Berkeley, San Francisko

NOTES

* Presented by J. Marsden to this conference. Partially supported by NSF Grant GP-15735.

[1] This appears to require a universal Newtonian time, thereby excluding relativistic effects. In fact, as we shall see, general relativity can be included in this formalism.

[2] In other words, the state contains complete information of the system, and the dynamical laws are followed exactly. There is a serious philosophical point here which is further discussed in Section 4 below.

[3] We take $\mathscr{H}$ 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; cf. [6] and [16] and related references.

[4] See for example [28]. Of course 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 $\mu_{A,\psi}$; cf. [21].

[5] For example it is not clear how to measure (position)+(momentum)$=q+p$ in the laboratory.

[6] See Mackey [16] for further discussion. Gleason's theorem is proved in [26].

[7] See Dixmier [10] for their theory. Simmons [25] contains a very readable account of elementary facts.

[8] Unbounded operators like x, p are included via their spectral projections.

[9] See for example Lanford [15], p. 160. Real linearity is the main content of 'linear', for complex linearity is a convention by which the state extends from the self-adjoint operators to all of $\mathfrak{A}$. Countably additive states are called 'normal'.

[10] Whenever one has a symplectic manifold P and a symmetry group G acting on P once can construct another symplectic manifold, called the reduced phase space. Namely, let $P_\mu = \psi^{-1}(\mu)/G_\mu$ where $\psi: P \to \mathfrak{G}^*$ is a 'moment' or energy function for the action, $\mathfrak{G}^*$ is the dual of the Lie algebra and G_μ is the isotropy group of the action of G on $\mathfrak{G}^*$. In the quantum mechanical case $G = S^1$ is the quantum mechanical phase group and $\psi: \mathscr{H} \to \mathbb{R}$, $\varphi \to \frac{1}{2}\langle \varphi, \varphi \rangle$. See Marsden-Weinstein [19] for details.

[11] One can pass to a suitable reduced phase as one can do in quantum mechanics by dividing out the phase group. See [19].

[12] The tensor product $\mathscr{H} \otimes \mathscr{H}'$ 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.

[13] 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 observer's mind into S'. Nevertheless, apparently one cannot find a mathematical device (within the framework of orthodox quantum mechanics) to yield the reduction of pure states. This is the fundamental problem in interpreting the foundations of quantum mechanics.

BIBLIOGRAPHY

[1] Abraham, R., *Introduction to Morphology*. Quatrième Rencontre entre mathématiciens et physiciens, Vol. 4, Fasc. 1 Publ. du Département de Mathématiques, Université de Lyon, France, 1972.

[2] Abraham, R. and Marsden, J., *Foundations of Mechanics* W. A. Benjamin, 1967.

[3] Bell, J. S., 'On the Problem of Hidden Variables in Quantum Mechanics', *Rev. Mod. Phys.* **38** (1966), 447–452.

[4] Bohm, D. and Bub, J., 'A Proposed Solution of the Measurement Problem in Quantum Mechanics by a Hidden Variable Theory', *Rev. Mod. Phys.* **38** (1966), 453–469, 470–475.

[5] Carroll, R., *Abstract Methods in Partial Differential Equations*, Harper and Row, 1969.

[6] Chernoff, P. and Marsden, J., 'Hamiltonian Systems and Quantum Mechanics', in preparation.

[7] Choquet, G., *Lectures on Analysis*, W. A. Benjamin, 1969.

[8] Clauser, J. F., Horne, M. A., Shimony, A., and Holt, R. A., 'Proposed Experiment to Test Local Hidden Variable Theories', *Phys. Rev. Letters* **23** (1969), 880–884.

[9] DeBroglie, L., *The Current Interpretation of Wave Mechanics*, Elsevier, 1964.

[10] Dixmier, J., *Les C* algèbres et leurs représentations*, Gauthier-Villars, 1964.

[11] Einstein, A., Podolsky, B., and Rosen, N., 'Can Quantum-Mechanical Description of Physical Reality be Considered Complete?', *Phys. Rev.* **47** (1935), 777–780.

[12] Freedman, S. J. and Clauser, J. F., 'Experimental Test of Local Hidden-Variable Theories', *Phys. Rev. Letters* **28** (1972), 938–941.

[13] Jauch, J. M., *Foundations of Quantum Mechanics*, Addison-Wesley, 1968.

[14] Kochen, S. and Specker, E. P., 'The Problem of Hidden Variables in Quantum Mechanics', *Jour. Math. and Mech.* **17** (1967), 59–88.

[15] Lanford, O. E., 'Selected Topics in Functional Analysis', in *Statistical Mechanics and Quantum Field Theory*, (ed. by C. DeWitt and R. Stora), Gordon and Breach, 1972.
[16] Mackey, G., *Mathematical Foundations of Quantum Mechanics*, W. A. Benjamin, 1963.
[17] Marsden, J., 'Hamiltonian One Parameter Groups', *Arch. Rat. Mech. and An.* **28** (1968), 362–396.
[18] Marsden, J., Ebin, D., and Fischer, A., 'Diffeomorphism Groups, Hydrodynamics and Relativity', *Proc. of 13th Biennial Seminar of Canadian Math. Congress*, (ed. by J. R. Vanstone), Montreal 1972.
[19] Marsden, J. and Weinstein, A., 'Reduction of Symplectic Manifolds with Symmetry', *Reports on Math. Phys.*, to appear.
[20] Nelson, E., *Dynamical Theories of Brownian Motion*, Princeton University Press, 1967.
[21] Robinson, M. C., 'A Thought Experiment Violating Heisenberg's Uncertainty Principle', *Can. Jour. Phys.* **47** (1969), 963–967.
[22] Ruelle, D., *Statistical Mechanics – Rigorous Results*, W. A. Benjamin, 1969.
[23] Segal, I., *Mathematical Problems of Relativistic Physics*, Amer. Math. Soc., 1963.
[24] Segal, I., 'A Mathematical Approach to Elementary Particles and Their Fields', University of Chicago Lecture Notes, 1955.
[25] Simmons, G. F., *Introduction to Topology and Modern Analysis*, McGraw-Hill, 1963.
[26] Varadarajan, V. S., *Geometry of Quantum Theory*, vols. I, II, Van Nostrand, 1968.
[27] Von Neumann, J., *Mathematical Foundations of Quantum Mechanics*, Princeton Univ. Press, 1955.
[28] Yosida, K., *Functional Analysis*, Springer-Verlag, 1965.
[29] Deliyannis, P. C., 'Generalized Hidden Variables Theorem', *Jour. of Math. Physics* **12** (1971), 1013–1017.