International Journal of Theoretical Physics, Vol. 12, No. 1 (1975), pp. 11-24

Macroscopic Variables

H. KRIPS

History and Philosophy of Science Department, Melbourne University

Received: 15 *June* 1974

Abstract

The problem of describing macroscopic variables in quantum theory, is discussed. It is suggested that the 'eigenspaces' of macroscopic variables be hyperspheres rather than closed linear subspaces. This is combined with the usual suggestion that macroscopic variables are nearly diagonal in the energy representation. The Schrödinger paradox is resolved in terms of this discussion.

1. *Introduction*

What is said in this paper to some extent hinges on an axiom system for Q.M. which has been presented elsewhere (Krips, 1974). To the extent that we only appeal to the non-controversial theorems of that system, however, our views on m-variables (macro-variables) can be explained without rehearsing the arguments in Krips (1974). We shall now give a short summary of the relevant results from the above which will be assumed here.

With any vector ψ and variable A is associated a set of probabilities $P[A, a_i; \psi]$—which turns out to be the probability that A is measured to have the value a_i if A is measured in S at t under the conditions that S at t has the pure state ψ. All vectors are assumed to have unit norm, so that $\mathbf{P}[\psi]$, the projector into ψ is $|\psi\rangle\langle\psi|$. Also associated with system S at time t is a set of vectors and associated weights $\{p_\alpha, \psi_\alpha\}, p_\alpha > 0, \Sigma\, p_\alpha = 1$, so that there is probability p_α that S at t is in ψ_α out of the set $\{\psi_\alpha\}$. Finally we associate with S at t a set of probabilities $\{P[A, a_i; S, t]\}$ defined by:

(i) $P[A, a_i; S, t] = \Sigma_\alpha p_\alpha P[A, a_i; \psi_\alpha]$.

$P[A, a_i; S, t]$ turns out to be the probability that A is measured to have value a_i in S at t if A is measured in S at t. If S at t is associated with $\{p_\alpha, \psi\}$ then we define $\mathbf{W}(S, t)$, the density operator for S at t, as $\Sigma\, p_\alpha \mathbf{P}[\psi_\alpha]$. We have the theorems/axioms:

(ii) If $\mathbf{W}(S, t) = \Sigma\, p_\alpha \mathbf{P}[\psi_\alpha]$ where each $p_\alpha > 0$ and $\{\mathbf{P}[\psi_\alpha]\}$ are linearly independent, then there is probability p_α that S at t is in ψ_α, for each α, out of the set $\{\psi_\alpha\}$.†

We introduce a set of vectors $\{\psi_{id}\}$ associated with A, where each ψ_{id} is a vector of A for value a_i. We define‡ the space of vectors spanned by $\{\psi_{id}\}_i$ to be the a_i-eigenspace of A. We have that

(iii) If S at t is associated with $\{p_i, \psi_i\}$, where ψ_i is in the a_i-eigenspace of A, for each i, then A has the value a_i in S at t if and only if S at t is in ψ_i out of the set $\{\psi_i\}$.

(iv) $\mathbf{W}(S, t) = \mathrm{Tr}_M \mathbf{W}(S + M, t)$.

(v) $P[A, a_i; \psi] = \Sigma_d |\langle \psi_{id}, \psi \rangle|^2$.

From (v) it is easily shown that the 'exceptation value' of A in S at t—which is defined as $\Sigma\, P[A, i; S, t]a_i$—is just Tr $\mathbf{W}(S, t)\mathbf{A}$, where $\mathbf{A} = \Sigma\, a_i \mathbf{P}[\psi_{id}]$.

A distinction emphasised in Krips (1974), is that to say S at t is in the *pure* state ψ is to say that S at t is associated with $\{1, \psi\}$ or that $\mathbf{W}(S, t) = \mathbf{P}[\psi]$, which is much stronger than just saying that S at t is in ψ (or has state-vector ψ), because the latter is consistent with $\mathbf{W}(S, t)$ taking any form $\Sigma\, p_\alpha \mathbf{P}[\psi_\alpha]$ where at least one of the $\{\psi_\alpha\}$ is ψ. Finally we note that the one numbering sequence will be used through all sections—except that within each proof a separate sequence is started.

The suggestions we shall make in this paper are rather tentative ones, to the extent that the whole question of the nature of macroscopic variables is a rather open one. The way to answer this question is to identify macroscopic variables with those quantum theoretical constructs which satisfy classical equations, to some approximation, and under expected circumstances. (This program is essentially the one put forward by van Kampen (1962) and by Ludwig (1954)). For example, Ehrenfest's theorem (Schiff, 1955, p. 25) suggests that the macroscopic position of a particle be identified as the expectation value of the particle's (quantum theoretical) position. We also require that if a particle has a definite macroscopic position then there is small probability associated with any other (quantum theoretical) positions.§ And finally, we require that no position be available to a particle, from a macroscopic point of view, which is not available to it from a quantum theoretical point of view.

Therefore, we arrive at the conclusion:

(vi) If S at t is in the pure state ψ then the macroscopic variable $\bar{A}$ has value $\bar{a}$ if and only if

† The reason for restricting $\{\mathbf{P}[\psi_\alpha]\}$ to be linearly independent is that, if they are not, then the density operator may be decomposed into various sums of the $\{\mathbf{P}[\psi_\alpha]\}$ each with *different* coefficients—thus giving a non-unique probability. This restriction, in any case, is much weaker than the usually accepted restriction to orthogonal $\{\psi_\alpha\}$. I note also the rule that the events that S at t is in ψ_α and $\psi_{\alpha'}$ respectively, out of the set $\{\psi_\alpha\}$, are mutually exclusive.

‡ We use the abbreviation here that '$\{X_{\alpha\beta}\}_\alpha$' abbreviates the set of all $X_{\alpha\beta}$ for given α and varying β.

§ Van Kampen uses the different condition that the expectation value of the *dispersion* of position is small. Under suitable assumptions those two conditions are equivalent.

(a) the corresponding quantum theoretical variable A has expectation value approximately equal to $\bar{a}$, i.e.

$$|(\Sigma P[A, a_i; \psi])a_i - \bar{a}| \leqslant \delta, \qquad \text{for } \delta \text{ small}$$

(b) $\bar{a}$ is one of the values $\{a_i\}$ of A;
(c) if $\bar{a} = a_i$, then

$$\sum_{\substack{i' \\ i' \neq i}} P[A, a_{i'}; \psi] \leqslant \delta, \qquad \text{for } \delta \text{ small}$$

A second example of macroscopic variables arises in quantum statistical thermostatics. A system's temperature is a function of the mean of its internal energy; but it is only defined when the system is in equilibrium. It is well known, however, that for an equilibrium system, with a given average energy and number of components, the internal energy dispersion is small (Huang, 1963, pp. 159, 189). This in turn suggests that for the temperature variable, (a), (b) and (c) are satisfied.

The statement (vi) does imply that there is an intrinsic vagueness in the concept of a macroscopic variable—to the extent that the number δ is only defined as being 'small'. This is not a disadvantage however, because the difference between micro and macro is only supposed to be one of degree. Later, an upper bound on δ will be suggested.

The macroscopic variables may be formally introduced as follows:

Definition 1$'$. If ψ' and ψ are vectors in H, then $\psi' \approx \psi$ to order δ if and only if $|\text{Tr}\, \mathbf{P}[\psi - \psi']| \leqslant \delta$ where $\delta \geqslant 0$.

Axiom I$'$. In any system S in which the m-variable $\bar{A}$ exists it is associated with the (ordinary quantum theoretical) variable A, and with a set $\{C_i\}$ of clusters of vectors in H, the Hilbert space of S. The vectors in cluster C_i are called 'the vectors of $\bar{A}$ for value a_i'. ψ is in C_i if and only if $\psi \approx \psi'$ to order δ for some ψ' in the a_i-eigenspace of A in H, where δ is small and positive.

Comment. The significance of the C_i associated with $\bar{A}$ will be given later by Axiom II$'$, which implies that if S at t is in the pure state ψ, ψ in C_i, then $\bar{A}$ has value a_i.

The use of the strong norm in Definition 1$'$, is necessary for later theorems (cf. Theorem 3$'$). Note that $|\text{Tr}\, \mathbf{P}[\psi - \psi']| < \delta$ has the geometrical interpretation that the tip of ψ' is located within a hypersphere of radius δ, and with the tip of ψ as centre. Thus the cluster C_i of vectors is *not* a closed linear manifold—rather it is a union of hyperspheres of vectors centred on the vectors in the a_i-eigenspace of A. Therefore we picture the m-variable $\bar{A}$ associated with A, as having the same values as A, and having a hypersphere of vectors around each eigenvector of A, of radius δ.

Theorem 1$'$. If $\{\psi_i'\}$ is any complete orthonormal set and $\psi \approx \psi_i'$ to order δ, for some i, then

$$\sum_{\substack{i' \\ i' \neq i}} |\langle \psi_{i'}', \psi \rangle|^2 < \delta \qquad \text{and} \qquad |\langle \psi_i', \psi \rangle|^2 \approx 1 \qquad \text{to order } \delta$$

Proof.

$$\begin{aligned}
\mathrm{Tr}\,\mathbf{P}[\psi - \psi_i'] &= \sum_{i'} \langle \psi_{i'}', \psi - \psi_i' \rangle \langle \psi - \psi_i', \psi_{i'}' \rangle \\
&= \sum_{i'} |\langle \psi_{i'}', \psi - \psi_i' \rangle|^2 \\
&= \sum_{i'} |\langle \psi_{i'}', \psi \rangle - \langle \psi_{i'}', \psi_i' \rangle|^2 \\
&= \sum_{i'} |\langle \psi_{i'}', \psi \rangle - \delta_{i'i}|^2
\end{aligned}$$

Hence, by Definition 1$'$, Theorem 1$'$ follows.

Comment. It then follows that I$'$ is a formalisation of (vi), since if ψ is in C_i, then from the comment to I$'$ and Theorem 1$'$, we immediately get the conditions (a), (b) and (c).

In introducing *m*-variables, like $\bar{A}$, into the axiom scheme we have not yet introduced a probability $P[\bar{A}, a_i; S, t]$; we shall now discuss this.

As mentioned above, $P[\bar{A}, a_i; S, t]$ plays the role of the probability of measuring the variable A to have value a_i. In our view, however, no such quantity exists for *m*-variables–and hence there is no place for a $P[\bar{A}, a_i; S, t]$. This claim appears startling–surely we do measure *m*-variables? Our answer to this is 'No'–we suggest that when we appear to be measuring the *m*-variable $\bar{A}$, we are in fact measuring A, and when we appear to refer to a probability of measuring $\bar{A}$ to have value a_i, we are in fact referring to the probability of measuring A to have value a_i–at least to within an approximation. Furthermore, we wish to suggest that we do not even determine $P[A, a_i; S, t]$ as an *estimate* of $P[\bar{A}, a_i; S, t]$–at best we determine a time-averaged value of $P[A, a_i; S, t]$, averaged over the error in measuring time-coordinates at the macro-level. This will be of significance later.

It is, perhaps, appropriate here to give an example.

Consider the Stern-Gerlach experiment (Ludwig, 1954). After passage through the magnetic field, the electron is in a mixture of macroscopically distinct states. They are macroscopically distinct because they have widely divergent average values of position (and small dispersions). Nevertheless, we do not make a measurement to distinguish the various components of the mixture. All we measure is the *actual* position (or coarse-grained position) of the electron. Since the components of the mixture do overlap in configuration space (albeit with small probability), the measured position probabilities are only taken to approximate (albeit very accurately) the probabilities attached to measuring the various components of the mixture.†

2. *Theorems*

Definition 2′. Two clusters of vectors C_1 and C_2 are approximately orthogonal to order δ if and only if, for any ψ_1 in C_1 and ψ_2 in C_2, there exist vectors ψ'_1 and ψ'_2 such that $\langle \psi'_1, \psi'_2 \rangle = 0$ and $\psi_1 \approx \psi'_1$ to order δ and $\psi_2 \approx \psi'_2$ to order δ.

Theorem 2′. The $\{C_i\}$ for any given $\bar{A}$ are approximately orthogonal to order δ.

Proof. Obvious from I′ and Definition 2′.

We now introduce one more definition and a theorem, which become of importance shortly.

Definition 3′. The set of clusters $\{C_i\}$ is linearly independent if and only if any set of vectors, containing only members of the $\{C_i\}$ and at most one member of any one $\{C_i\}$, is linearly independent.

Theorem 3′. Any finite set of clusters, which are approximately orthogonal to order δ, are linearly independent, for δ small enough, but $\delta > 0$.

Proof. Let $\{C_i\}$ be a finite set of N clusters, approximately orthogonal to order δ. Let ψ_i be an arbitrary vector in C_i, for each i. Then by I′, there exists an orthonormal set of vectors $\{\psi'_i\}$ for which $\psi'_i \approx \psi_i$ to order δ, for each i.

By Theorem 1′

(i) $$\sum_{\substack{i' \\ i' \neq i}} |\langle \psi_{i'}, \psi'_i \rangle|^2 \leqslant \delta \qquad \text{for any } i' \neq i$$

and

(ii) $$-|\langle \psi_i, \psi'_i \rangle|^2 \geqslant 1 - \delta$$

Now let $\{\psi_i\}$ be linearly dependent (this is the first premiss in a reductio proof); i.e. for some i,

(iii) $$-\psi_i = \sum_{\substack{i' \\ i' \neq i}} C_{i'} \psi_{i'}$$

Taking scalar product of both sides of (iii) with ψ'_i, and the taking modulus squared, gives

(iv) $$|\langle \psi_i, \psi'_i \rangle|^2 = |\sum_{\substack{i' \\ i' \neq i}} C_{i'} \langle \psi_{i'}, \psi'_i \rangle|^2$$

† Note also we that we agree with Bunge that the Stern-Gerlach experiment is not really a description of a measurement—not, at least, until the interaction between the target screen and the incident electron is spelled out (Bunge, 1967, p. 281). The answer to the perennial question 'which component is the electron *really* in' is contained in the answer to the Schrödinger cat paradox—see later.

where, by (ii), the left-hand side of (iv) $\geqslant 1 - \delta$, and, by Hölder's inequality, the right-hand side of (iv)

$$\leqslant \sum_{\substack{i' \\ i' \neq i}} |C_{i'}|^2 \sum_{\substack{i' \\ i' \neq i}} |\langle \psi_{i'}, \psi_i' \rangle|^2$$

which, by (i),

$$\leqslant \sum_{\substack{i' \\ i' \neq i}} |C_{i'}|^2 \delta$$

Hence

(v) $$(1 - \delta) \leqslant \delta \sum_{\substack{i' \\ i' \neq i}} |C_{i'}|^2$$

Hence, if we let K be the maximum value of $|C_{i'}|$, then, from (v), we get tha

(vi) $$K^2 \geqslant \frac{(1 - \delta)}{(N - 1)\delta}$$

Now let $\{\psi'_{i\alpha}\}$ be the complete orthonormal set for which

$$\psi'_{i1} \approx \psi_i \text{ to order } \delta.$$

Hence, from Theorem 1$'$

(vii) $$\psi_i = \sum_{i'\alpha'} C^i_{i'\alpha'} \psi'_{i'\alpha'}$$

where

(viii) $$\sum_{\substack{i'\alpha' \\ (i'\alpha') \neq (i1)}} |C^i_{i'\alpha'}|^2 < \delta$$

Also, from (vii),

(ix) $$|C^i_{i1}|^2 = |\langle \psi_{i1}, \psi_i \rangle|^2 \leqslant 1$$

Hence, from (vii),

$$|\langle \psi_{i'}, \psi_i \rangle| = |\sum_{i''\alpha} \bar{C}^{i'}_{i''\alpha} C^i_{i''\alpha}|$$

$$\leqslant |\sum_{\substack{i''\alpha \\ (i''\alpha) \neq (i'1) \text{ and } (i''\alpha) \neq (i1)}} \bar{C}^{i'}_{i''\alpha} C^i_{i''\alpha}| + |\bar{C}^{i'}_{i'1}||C^i_{i'1}| + |\bar{C}^{i'}_{i1}||C^i_{i1}|$$

which by Hölder's inequality, (vii), (viii) and (ix),

$$\leqslant | \sum_{\substack{i''\alpha \\ (i''\alpha) \neq (i'1)}} |\bar{C}^{i'}_{i''\alpha}|^2 \sum_{\substack{i''\alpha \\ (i''\alpha) \neq (i1)}} |C^i_{i''\alpha}|^2 |^{1/2} + \delta^{1/2} + \delta^{1/2}$$

By (viii), we therefore get that

(x) $$|\langle \psi_{i'}, \psi_i \rangle| \leqslant \delta + 2\delta^{1/2} \leqslant 3\delta^{1/2}$$

Now, from (iii) for any $i'' \neq i$,

(xi) $$C_{i''} = \langle \psi_{i''}, \psi_i \rangle - \sum_{\substack{i' \\ i' \neq i \text{ or } i''}} C_{i'} \langle \psi_{i''}, \psi_i \rangle$$

In particular if we let $|C_{i''}| = K$, then we get, from (x) and (xi) (where K is maximum of $|C_{i'}|$)

$$K \leqslant 3\delta^{1/2} + (N-2)3\delta^{1/2}K$$

Hence

$$K(1-(N-2)3\delta^{1/2}) \leqslant 3\delta^{1/2}$$

Hence, if $(1-(N-2)3\delta^{1/2}) \neq 0$, we have

$$K^2 \leqslant \frac{9\delta}{(1-(N-2)3\delta^{1/2})^2}$$

Hence

(xii) $$K^2 \leqslant \frac{9\delta}{(1-(N-2)3\delta^{1/2})^2} \qquad \text{if } 3\delta^{1/2}(N-2) < 1$$

Now comparing (xii) and (vi), we see that we have a contradiction if both $3\delta^{1/2}(N-2) < 1$ and

(xiii) $$\frac{9\delta}{(1-(N-2)3\delta^{1/2})^2} < \frac{(1-\delta)}{(N-1)\delta}$$

But for $3\delta^{1/2}(N-2) \ll 1$ we can put $(1-(N-2)3\delta^{1/2}) \approx 1$ and $(1-\delta) \approx 1$ and hence (xiii) comes to

(xiv) $$\delta^2 \leqslant \frac{1}{9(N-1)}$$

which is obviously implied by $3\delta^{1/2}(N-2) \ll 1$.

Hence for δ very small, we see that we have a contradiction, and hence, for δ very small, (iii) is false. Hence, for δ small enough, $\{\psi_i\}$ is linearly independent, and since the $\{\psi_i\}$ was arbitrarily chosen–one vector from each C_i–it follows, from Definition 3′ that $\{C_i\}$ is linearly independent. Q.E.D.

We now use the previous theorems. It is apparent that we would like to be able to come to conclusions like 'there is a probability p_i that $\bar{A}$ has value a_i in S at t' on the basis of information about the state of S at t. To do this, we make use of the full degree of generality in (ii), which only required the $\{P[\psi_\alpha]\}$ to be linearly independent (and *not* orthogonal). We also extend (iii) to:

Axiom II′. If S at t is associated with $\{p_{i\beta}, \psi_{i\beta}\}$, where $\psi_{i\beta}$ is in the cluster C_i of $\bar{A}$, for each i, β, then $\bar{A}$ has the value a_i in S at t if and only if S at t is in one of the $\{\psi_{i\beta}\}_i$ out of the set $\{\psi_{i\beta}\}$.

We can then see that if $\mathbf{W}(S, t) = \Sigma\, p_i \mathbf{P}[\psi_i]$, where ψ_i is in C_i for each i, *and if* the C_i are linearly independent (Definition 3′), then there is a probability p_i that $\bar{A}$ has value a_i.† (Proof follows from (ii), II′, and probability theory.)

The problem then is how to guarantee the linear independence of the $\{C_i\}$, so that we can quite generally infer from '$\mathbf{W}(S, t) = \Sigma\, p_i \mathbf{P}[\psi_i]$, where ψ_i is in C_i a cluster of $\bar{A}$' to 'there is probability p_i that $\bar{A}$ has value a_i'. Now we know from Theorem 3′, that linear independence of $\{C_i\}$ is guaranteed by there being a finite number of $\{C_i\}$, and by their being approximately orthogonal to degree δ, for δ small. This suggests that we postulate the $\{C_i\}$ for any m-variable $\bar{A}$ in S, to be approximately orthogonal to degree δ, where δ is very small; and that we either assume observed systems to be restricted to states which are superpositions of vectors from a finite number of the $\{C_i\}$, or that there *are* only a finite number of members of $\{C_i\}$. Of the latter two alternatives, we prefer the second because it is less arbitrary, although it does have the controversial consequences that macroscopic variables necessarily have finite (albeit arbitrarily large) ranges, and hence must be infinitely degenerate (whenever the system's Hilbert space is infinite dimensional). Formally, then we suggest,

Axiom III′. The number of clusters associated with any m-variable is finite.

Axiom IV′. For any m-variable $\bar{A}$, the degree δ to which its vectors approximate the vectors of the associated variable A, is so small that $3\delta^{1/2}(N-2) \ll 1$, where N is the number of clusters of $\bar{A}$ associated with A.

Theorem 3′b. $\{C_i\}$ is linearly independent.

Proof. Trivially from III′, IV′ and Theorem 3′.

Comment. N may depend on $\bar{A}$, although we do not explicitly display this dependence. Note also that we could alter the condition that $3\delta^{1/2}(N-2) \ll 1$ in IV′, and even forgo IV′ entirely, if some significantly weaker condition can be found to guarantee $\{C_i\}$ linearly independent.

We also note that in practice III′ is usually obeyed any way. For example, in the Schrödinger cat paradox (to be discussed later), the m-variable has *two* clusters associated with it—one for the 'cat-dead states' and the other for the 'cat-alive states'.

Finally, we note that the approximate orthogonality of the $\{C_i\}$ is just the condition which Araki & Yanase (1960) found was dictated by independent considerations, based on the study of 'ideal measuring processes'.

3. *Evolution of m-Variables*

There is one crucial question which must be resolved, if we are to maintain the model for macroscopic variables which we have suggested in the preceding part. That question is why there are no interference effects observed between

† The probability here may be non-unique. This is a problem which will be dealt with in another paper, where we show the non-uniqueness is restricted to within a small range—see Part 3 of Krips (1974).

macroscopically distinct states. One answer would be that superpositions of macroscopically dinstinct states do not occur. But this restriction seems just too *ad hoc*. Furthermore, we shall now show this restriction to be inconsistent with quantum dynamics (via a certain lemma).

Let S at t be in the pure state ψ_i in the cluster C_i—where C_i is the cluster of vectors corresponding to the value a_i of some macro-variable $\bar{A}$. It is a fact that even in isolated systems the values taken by some macro-variables do, on occasions, change. Therefore, for some S, $\bar{A}$, and t', S will be isolated from t to t' and will be in the pure state $\psi_{i'}$ at t', where $\psi_{i'}$ is in the cluster $C_{i'}$, and $i \neq i'$. We can assume that the state-vector of S is not in any of the $\{C_i\}$, other than C_i or $C_{i'}$ during t, t' with no loss of generality;† because, if S at t'' were in $C_{i''}$, where $i'' \neq i$ or i' and $t < t'' < t'$, then we could change t' to t'' and consider the new interval $[t, t'']$ instead.

Now, since S is isolated between t, and t', it follows that for any t_1, t_2 in the interval t, t', we have that S at t_1 is in the pure state $\psi_{(t_1)}$, and $\psi_{(t_2)} = \mathbf{U}_{(t_2, t_1)}\psi_{(t_1)}$; and in particular $\psi_{i'} = \mathbf{U}_{(t', t)}\psi_1$. ($\mathbf{U}_{(t_2, t_1)}$ is the Schrödinger propagator from t_1 to t_2; and the set of $\mathbf{U}(t_2, t_1)$, for constant t_1, forms a one-parameter continuous group of unitary transformations,‡ so that, for any ψ or ψ' in H, if $t'_2 \to t_2$, then $\langle \psi, \mathbf{U}_{(t'_2, t_1)}\psi' \rangle \to \langle \psi, \mathbf{U}_{(t_2, t_1)}\psi' \rangle$.)

We can now prove the following:

Lemma. If for any two members $\psi_{(t_1)}$, $\psi_{(t_2)}$ of the continuous sequence of vectors $\psi_{(t)}$, we have that $\psi_{(t_2)} = \mathbf{U}(t_2, t_1)\,\psi_{(t_1)}$; and if $\psi_{(t_1)}$ is in $C_{i'''}$ and $\psi_{(t_2)}$ is in $C_{i''}$ then $i'' = i'''$ if $|t_1 - t_2|$ is sufficiently small.

Proof. There is a set $\{\psi'_i\}$, which is a complete orthonormal set of vectors of the variable A which corresponds to the m-variable $\bar{A}$ (to which the $\{C_i\}$ belong). Let $\psi_{(t_1)}$ be in $C_{i'''}$ and $\psi_{(t_2)}$ be in $C_{i''}$ where $i'' \neq i'''$. By Theorem 1′, if $\psi_{(t_1)} = \Sigma\, C_i^1 \psi'_i$ and $\psi_{(t_2)} = \Sigma\, C_i^2 \psi'_i$ then

$$\text{(i)} \qquad |C_{i''}^2|^2 > 1 - \delta \qquad \text{and} \qquad |C_{i''}^1|^2 < \delta$$

Now let

$$\psi_{(t_2)} = \mathbf{U}_{(t_2, t_1)}\psi_{(t_1)}$$

Taking scalar product of both sides with $\psi'_{i''}$, and the modulus squared, gives

$$\text{(ii)} \qquad |\langle \psi'_{i''}, \mathbf{U}_{(t_2, t_1)}\psi'_i \rangle|^2 = |C_{i''}^2|^2$$

But, as $t_1 \to t_2$,

$$|\langle \psi'_{i''}, \mathbf{U}_{(t_2, t_1)}\psi_{(t_1)} \rangle|^2 \to |\langle \psi'_{i''}, \psi_{(t_1)} \rangle|^2 = |C_{i''}^1|^2$$

† The state-vector of S may of course be in a linear-superposition of vectors from various $\{C_i\}$ during $[t, t']$, without being in one of the $\{C_i\}$.

‡ We shall not be discussing the axioms of quantum dynamics. An excellent attempt at doing just this, starting from very elementary principles, is to be found in Eckstein (1967, 1969). Details of one-parameter groups are given in Riesz & Nagy (1965).

(since the group of transformations is continuous, and $\mathbf{U}_{(t_1, t_1)} = \mathbf{I}$). Hence from (i), we see that, as $t_1 \to t_2$, the inequality (ii) is not possible. Hence, the supposition $i'' \neq i'''$, as $t_1 \to t_2$, is not possible. Q.E.D.

Now introduce the function $i(t'')$, which takes the value i in S at t'' if and only if S at t has a state-vector in C_i. *Ex hypothesi* $i(t'')$ takes the value i or i' for t'' in $[t, t']$. If and only if S at t'' has a state-vector which is in none of the $\{C_i\}$, then $i(t'')$ is undefined. We shall now show, using the preceding lemma, that $i(t'')$ must be undefined at some t'' in $[t, t']$.

Proof. Suppose that $i(t'')$ is defined for all t'' in $[t, t']$. Then the preceding lemma entails that $i(t'')$ is a continuous function of t'' on the domain $[t, t']$. Hence, from the fundamental property of continuous functions, and since $i(t) = i$ and $i(t') = i'$, we see that $i(t'')$ must take any value between i and i' as t'' takes various values in $[t, t']$. This is impossible however, since $i(t'')$, *ex hypothesi*, only has values i or i', where $i \neq i'$. Therefore, $i(t'')$ must be undefined for at least one t'' in $[t, t']$.

From the preceding proof it immediately follows that the state-vector of S is not in one of the $\{C_i\}$, at some t'' in $[t, t']$. But the set $\{C_i\}$ is complete in H (since it includes the complete set of eigenvectors of the variable A corresponding to the m-variable $\bar{A}$); and therefore, at some t'' in $[t, t']$, $\psi_{(t'')}$ is a non-trivial linear combination of vectors from various $\{C_i\}$ *and* is not in one of the $\{C_i\}$.†

This completes the proof that to restrict the occurrence of superpositions of macroscopically distinct states to joint systems is inconsistent with quantum dynamics. It follows that some other way must be found to explain the failure to observe interference effects between macroscopically distinct states.

The way which we shall suggest now is a slight variant of the traditional 'phase wash-out theories'—see Margenau (1967), Van Kampen (1962).

The first point is the macro-variables change their values slowly; i.e., there is, by and large, stability of their values over small time-intervals. This in turn suggests an approximate correlation between the various $\{C_i\}$, for an m-variable $\bar{A}$ of the isolated system S, and the 'energy shells' (eigenspaces of the Hamiltonian $\mathbf{H}$) of S. The correlation must not of course be an exact one—otherwise we contradict the fact mentioned earlier, that macro-variables do change their values in isolated systems. (We are, of course, assuming conservation of energy for isolated systems.) This approximate correlation may be expressed as follows. Let $\psi_{m\beta}$ be the eigenvector of $\mathbf{H}$ for energy value $E_{m\beta}$, where $E_{m\beta} = E_m$ for all m, β. Then, if ψ_i' is a vector of A to which vectors in C_i (a cluster of $\bar{A}$) are approximately orthogonal, and if

$$\psi_i' = \sum_m \sum_\beta C^i_{m\beta} \psi_{m\beta}$$

† The second conjunct in the latter conclusion is *not* redundant, because a non-trivial linear combination of vectors from various $\{C_i\}$ *may* be in one of the $\{C_i\}$—since the $\{C_i\}$ are *not* closed linear manifolds, but are unions of hyperspheres (see Section 1 of this paper)

we have that

$$\sum_m \sum_\beta |C^i_{m\beta}|^2 < \delta \qquad \text{for } E_i > E_m \geqslant E_i + \Delta E_i$$

where δ is small, and all the semi-closed intervals $(E_i, E_i + \Delta E_i]$ are disjoint.

Our second point is the one mentioned earlier that at the macro-level we do not, in fact, observe the probability of measuring $\bar{A}$ to have a_i in S at t; but instead observe the time-average quantity

$$\frac{1}{T} \int_{t-T_0}^{t+T_1} P[A, a_i; S, t']\, dt'$$

where $T_1 + T_0 = T > 0$, and T is the error accepted in locating times at the macro-level. Furthermore, we shall assume that $T\,\Delta E \geqslant 1/\delta$, where ΔE is the minimum energy difference between the energy shells correlated with the $\{C_i\}$. We use units of $\hbar = 1$.

It is tempting to try to justify the latter assumption by referring to the Heisenberg relation '$\Delta E \Delta t \geqslant 1$'; but Allcock (1969), in a penetrating and thorough series of articles, warned against doing this lightly. According to Allcock, we need to examine the details of the measuring apparatus used, before interpreting 'Δt' in the above relation as the indeterminacy in the time variable. We therefore leave this assumption as unjustified; but one which will, it is to be hoped, be vindicated by a complete theory of macroscopic phenomena.

Now suppose that

$$\mathbf{W}(S, t) = \sum p_y \mathbf{P}[\psi_y]$$

We also suppose that S is isolated; and for simplicity we assume that, for each i, any vector in C_i is approximately orthogonal to the same ψ_i' (i.e. A is non-degenerate). Hence, for any y,

(vii) $$\psi_y = \sum C_i^y \psi_i'$$

As above, for any i, if

(viii) $$\psi_i' = \sum_m \sum_\beta C^i_{m\beta} \psi_{m\beta}$$

then

(ix) $$\sum_m \sum_\beta |C^i_{m\beta}|^2 < \delta \qquad \text{for } E_i > E_m \geqslant E_i + \Delta E_i$$

Also, from (i) and (iv) (in Section 1)

$$P[A, i; S, t] = \sum_y p_y |C_i^y|^2$$

One way for there to be no observed interference effects between the various $\{C_i\}$ of $\bar{A}$, would be for there to be no interference effects between the various $\{\psi_i'\}$ of A (since $\bar{A}$ is measured via A); i.e. for the density operator of S at t to be

$$\text{(x)} \qquad \mathbf{W} = \sum \sum p_y \, |C_i^y|^2 \mathbf{P}[\psi_i']$$

We shall now show that time-averaging $\mathbf{W}(S, t)$ does, in fact, smooth out the interference terms; so that, after a macroscopic time-averaging, $\mathbf{W}(S, t) \approx \mathbf{W}$. This is sufficient to show that interference effects are not significant at the macro-level, because of the second point we raised above; viz. that only time-averaged quantities are significant at the macro-level.

From the rules of quantum dynamics, the time-averaged density operator for S at t is

$$\begin{aligned}\bar{\mathbf{W}}(S, t) &= \frac{1}{T} \int_{t-T_0}^{t+T_1} \mathbf{W}(S, t')\, dt' \\ &= \frac{1}{T} \int_{t-T_0}^{t+T_1} [\exp -i\mathbf{H}(t'-t)]\mathbf{W}(S, t)[\exp i\mathbf{H}(t'-t)]\end{aligned}$$

where $\mathbf{H}$ is the time-dependent Hamiltonian for the isolated system S. Hence,

$$\begin{aligned}\bar{\mathbf{W}}(S, t) &= \frac{1}{T} \int_{t-T_0}^{t+T_1} [\exp - i\mathbf{H}(t'-t)] \sum_y p_y \sum_{ii'} C_i^y |\psi_i'\rangle\langle\psi_{i'}'| \bar{C}_{i'}^y \\ &\quad \times [\exp i\mathbf{H}(t'-t)]\, dt' \\ &= \frac{1}{T} \int_{t-T_0}^{t+T_1} [\exp - i\mathbf{H}(t'-t) \sum_y p_y \sum_{ii'} \sum_{m\beta} \sum_{m'\beta'} C_i^y C_{m\beta}^i |\psi_{m\beta}\rangle \\ &\quad \langle\psi_{m'\beta'}| \bar{C}_{m'\beta'}^{i'} \bar{C}_{i'}^y [\exp i\mathbf{H}(t'-t)]\, dt' \\ &= \frac{1}{T} \sum_y \sum_{ii'} \sum_{m\beta} \sum_{m'\beta'} p_y C_i^y \bar{C}_{i'}^y C_{m\beta}^i \bar{C}_{m'\beta'}^{i'} \int_{t-T_0}^{t+T_1} dt' \\ &\quad \times [\exp - i(E_m - E_{m'})(t'-t)] \, |\psi_{m\beta}\rangle\langle\psi_{m'\beta'}| \\ &= \sum_y \sum_{ii'} \sum_{m\beta} \sum_{\substack{m'\beta' \\ E_m \neq E_m}} p_y C_i^y \bar{C}_{i'}^y C_{m\beta}^i \bar{C}_{m'\beta'}^{i'} |\psi_{m\beta}\rangle\langle\psi_{m'\beta'}| \\ &\quad \times [\exp - \tfrac{1}{2} i(E_m - E_{m'})(T_1 - T_0)] \, \frac{[\sin \frac{1}{2}(E_m - E_{m'})T]}{\frac{1}{2}(E_m - E_{m'})T}\end{aligned}$$

provided $[\sin \frac{1}{2}(E_m - E_{m'})T]/\frac{1}{2}(E_m - E_{m'})T$ is understood as 1 if $E_m = E_{m'}$.

We now evaluate Tr $\mathbf{R}^*\mathbf{R}$, where $\mathbf{R}$ is got from the above expression of $\overline{\mathbf{W}}(S, t)$ by putting $i \neq i'$. This will be a measure of the degree to which $\mathbf{W}(S, t)$ approximates $\mathbf{W}$, after time-averaging;† and it is this quantity which we require to be small.

Now

$$\operatorname{Tr} \mathbf{R}^*\mathbf{R} = \sum_{m\beta} \sum_{m'\beta'} \langle \psi_{m\beta}\mathbf{R}, \psi_{m'\beta'} \rangle \langle \psi_{m'\beta'}, \mathbf{R}\psi_{m\beta} \rangle$$

$$= \sum_{m\beta} \sum_{m'\beta'} \left| \sum_{y} \sum_{\substack{ii' \\ i \neq i'}} p_y C_i^y \bar{C}_{i'}^y C_{m\beta}^{i'} \bar{C}_{m'\beta'}^{i'} [\exp - \tfrac{1}{2} i(E_m - E_{m'})(T_1 - T_0)] \times \frac{[\sin \frac{1}{2}(E_m - E_{m'})T]}{\frac{1}{2}(E_m - E_{m'})T} \right|^2$$

Informally, one can see that this expression can be made small, because, from (v), the only large modulus terms will be those for which

$$E_i > E_m \geqslant E_i + \Delta E_i \qquad \text{and} \qquad E_{i'} > E_{m'} \geqslant E_{i'} + \Delta E_i$$

But, for any term for which the latter inequalities hold, $|E_m - E_{m'}| > \Delta E$, since $i \neq i'$, and therefore $|\frac{1}{2}(E_m - E_{m'})T| > 1/\delta$ (since we decided above that $T\Delta E > 1/\delta$); and hence, for small δ, the term has small modulus. Assuming the appropriate convergence conditions, the required result then follows.

Another problem which the m-variable model for macro-variables has to face is the following. It is a fact that if a macro-variable starts out with a certain value in S at t, then its value does not spread (although it may change) at later times in S. To capture this fact in the m-variable model, the Hamiltonians for macro-systems will have to be so constructed that if S at t has a small spread in the values of A (necessary for it to be in one of the C_i), then that spread will remain small. Van Kampen (1962) discusses this requirement. This restriction does not, however, guarantee that S will always be in one of the $\{C_i\}$. Indeed it is clear from the first considerations in Section 3, that S must be in none of the $\{C_i\}$, at some time, if it is to change to some other $\{C_i\}$. This discrepancy between m-variables and macroscopic variables can, however, be explained by realising that, at the macro-level, the accepted errors are so large that they mask the transitions between the discrete $\{C_i\}$. (This is also the reason why classical theory gives satisfactory answers at the macro-level, despite containing the false assumption that the macro-variables have continuous ranges of values.)

4. *Schrödinger Paradox*

We shall here present a generalised version of Schrödinger's paradox (Schrödinger, 1935). We have already discussed the paradox in Krips (1969);

† Obviously Tr $(\overline{\mathbf{W}}(S, t) - \overline{\mathbf{W}})\mathbf{A}$ = Tr $\mathbf{RA}$ (where $\overline{\mathbf{W}}$ is the time-average of W—see (x)); and hence $\overline{\mathbf{W}}(S, t)$ is effectively $\overline{\mathbf{W}}$—at least as far as the expectation values of any variables go—if $|\operatorname{Tr} \mathbf{RA}|$ is small. But $|\operatorname{Tr} \mathbf{RA}| \leqslant \operatorname{Tr} \mathbf{R}^*\mathbf{R} \operatorname{Tr} \mathbf{A}^*\mathbf{A}$; and hence Tr $\mathbf{RA}$ is small if Tr $\mathbf{R}^*\mathbf{R}$ is—at least if $\mathbf{A}$ is Hilbert-Schmidt (and hence if $\mathbf{A}$ is a projection operator).

but what we have to say here is somewhat different in the light of the preceding theory of m-variables.

Suppose there is a conservative measurement interaction from time t to t', between S, which is in the pure state $\Sigma\, C_i\psi_i$ at t, and the measuring apparatus M which is in the pure state ϕ at t. The measured variable A is non-degenerate, with vectors $\{\psi_i\}$. Then at t', $S+M$ is in the pure state $\Sigma\, C_i\phi_i \times \psi_i$, where $\{\phi_i\}$ is a set of macroscopically distinct states–M at t' in ϕ_i (out of $\{\phi_i\}$) is to be interpreted as M registering the value a_i for A.

Note that the latter condition, from our theory of m-variables, implies that ϕ_i is in a cluster C_i of some $\bar{A}$ of M, for each i–and hence the $\{\phi_i\}$ need *not* be orthogonal (indeed, according to Araki & Yanase (1960), they *cannot* be).

The Schrödinger paradox points out that $S+M$ at t' is in a pure *superposition* of vectors corresponding to macroscopically distinct states–which is at variance with the 'observed fact' that, after measurement, M does register one of the values for A.

Where this paradox breaks down is in making the implicit assumption that if $S+M$ at t' is in a pure superposition of vectors corresponding to macroscopically distinct states, then M is not in one of a set of possible macroscopically distinct states. In fact this assumption is wrong. From the density operator for $S+M$ at t', we can deduce (via (iv) in section 1) the density operator for M at t' to be $\Sigma\, |C_i|^2\, \mathbf{P}[\phi_i]$. Then, from the theorems in Section 2 of this paper, we get that, in M at t', $\bar{A}$ does in fact have one of the values $\bar{a}_i$. ($\bar{A}$ has $\bar{a}_i$ with probability $|C_i|^2$.) We note that in Schrödinger's example, the variable 'i' has a finite range–in fact it is two-valued (C_1 corresponds to a cat being dead, and C_2 to the cat being alive). This fits in with III$'$.

References

Allcock, G. (1969). *Annals of Physics,* **53**, 253, 286, 311.

Araki, H. and Yanase, M. (1960). *Physical Review,* **120**, 622.

Bunge, M. (1967). *Foundations of Physics.* Springer.

Eckstein, H. (1967). *Physical Review,* **153**, 1397; (1969). **184**, 1315.

Huang, K. (1963). *Statistical Mechanics.* Wiley.

Van Kampen, N. (1962). In *Fundamental Problems in Statistical Mechanics* (Ed. Cohen). Amsterdam.

Krips, H. (1969). *Philosophy of Science,* **86**, 145.

Krips, H. (1974). Foundations of quantum theory, Parts 1, 2, and 3, Part 1 in *Foundations of Physics,* 4, 181; other parts in subsequent issues.

Ludwig, G. (1954). *Die Grundlagen der Quantemechanik.* Springer.

Margenau, H. (1967). In *Studies in the Foundations, Methodology and Philosophy of Science,* Vol. 1 (Ed. Bunge). Springer.

Riesz, F. and Nagy, B. (1965). *Functional Analysis.* Unger.

Schiff, L. (1955). *Quantum Mechanics.* New York.

Schrödinger, E. (1935). *Naturwissenschaften,* **48**, 52.