Aust. J. Phys., 1995, **48**, 613–33

Contiguity and the Quantum Theory of Measurement

H. S. Green

Department of Physics and Mathematical Physics,
University of Adelaide, Adelaide, SA 5005, Australia.

Abstract

This paper presents a comprehensive treatement of the problem of measurement in microscopic physics, consistent with the indeterministic Copenhagen interpretation of quantum mechanics and information theory. It is pointed out that there are serious difficulties in reconciling the deterministic interpretations of quantum mechanics, based on the concepts of a universal wave function or hidden variables, with the principle of contiguity. Quantum mechanics is reformulated entirely in terms of observables, represented by matrices, including the statistical matrix, and the utility of information theory is illustrated by a discussion of the EPR paradox. The principle of contiguity is satisfied by all conserved quantities. A theory of the operation of macroscopic measuring devices is given in the interaction repesentation, and the attentuation of the indeterminacy of a microscopic observable in the process of measurement is related to observable changes of entropy.

1. Introduction

In the introduction to his Waynflete Lectures, Born (1949) included among the more important concepts of physics the *principle of contiguity*. This may be regarded as denying the necessity, if not the possibility, of action at a distance, and requiring a spatially continuous relation between any cause and effect. In a relativistic context, it implies also a time-like relation between cause and effect, communicated with a velocity not exceeding the velocity of light. Born contrasted Newton's theory of gravitation, formulated originally as a theory of action at a distance, with Einstein's general theory of relativity, a theory satisfying the principle of contiguity. The principle is satisfied also by Maxwell's formulation of electromagnetic theory, and indeed by any field theory, subject to the condition that the Lagrangian density depends only on local values of the field variables and their derivatives.

Since wave mechanics can be formulated as a field theory, and there is an established procedure for the quantization of field theories (Pauli 1941), it would appear that there should be no difficulty in reconciling quantum mechanics with the principle of contiguity. However, as Green and Triffet (1991) have pointed out, quantum mechanics is the synthesis of a generally accepted dynamical theory first formulated by Born and Jordan (1925), and a theory of the measurement of microscopic quantities that is, even today, in a very rudimentary state. Most accounts of the process of measurement appear to have been influenced by von

0004-9506/95/040613$10.00

Neumann (1955), by whom it was represented as the discontinuous change of a state vector not necessarily localized in space. When translated into the terms of wave mechanics, this 'collapse' affects the entire wave function, which may extend over an arbitrarily large region of space. This violation of the principle of contiguity obviously influenced Einstein's objections to the quantum theory, ultimately formulated in what is known as the EPR paradox (Einstein *et al.* 1935).

On the other hand, some authors, from Janossy (1952, 1953) to Stapp (1992, 1993) have been prepared to accept the 'collapse of the wave function' as a physical event. As these authors readily acknowledge, this requires the acceptance not only of action at a distance, but of the transmission of observable effects with velocities greater than the velocity of light.

The validity of the concept of contiguity in quantum mechanics is closely related to the interpretation given to the theory, and both are determined by the theory of measurement. It is possible to distinguish between two major interpretations that regard the future as predetermined, and an indeterministic interpretation that is realistic in the sense that it assumes no more than our present knowledge can justify.

The deterministic interpretations may be summarized as follows:

(1) The concept of a *universal wave function* (or state vector) that determines past, present and future was essentially due to Schrödinger (1926). It was Schrödinger's view that not only light but atoms and indeed all other forms of matter consist of waves, and the fact that they appear in discrete quanta was a consequence of the mathematical properties of his wave equation. The existence of a universal wave function has been assumed in many subsequent contributions to quantum mechanics (such as Gottfried 1966, Zurek 1982 and Bell 1990). Though, as a field theory, wave mechanics is clearly consistent with the principle of contiguity, it suffers from the lack of a coherent theory of measurement. If one rejects attempts, like Janossy's, to account for the 'collapse of the wave packet' as a physical event, it seems necessary to adopt the *many worlds* interpretation of quantum mechanics, introduced explicitly by Everett III (1957), and further developed in a volume edited by De Witt and Graham (1973), in which Everett III advocates a theory of measurement similar to von Neumann's. The many worlds theory is completely deterministic. It does not explicitly violate the principle of contiguity, but assumes that every possible result of a measurement that is allowed by quantum mechanics is actually realized, and that a separate universe is evolved to accommodate each possibility. It concedes that an observer in one of these universes can have no experience of his or her counterpart in the others.

It could be said that the many worlds interpretation requires the creation of innumerable unobservable systems to account for our inability to predict the behaviour of the one that is observed. It is by no means obvious that the principle of contiguity is satisfied in the process of multiplication of worlds, because there is no adequate description of this process.

(2) The interpretation of quantum mechanics in terms of *hidden variables* is generally attributed to Bohm (1952). The wave function is supposed to represent an ensemble of point-like particles, moving in accordance with deterministic laws, and subject to forces that are partly classical and partly quantum mechanical.

Each particle has its own position and momentum, the hidden variables, both of which can in principle be observed. By supposing that hidden variables can also be associated with macroscopic measuring devices, apparent violations of Heisenberg's uncertainty principle and von Neumann's proof that hidden variables were incompatible with quantum mechanics are mitigated. However, the quantum mechanical forces needed to accommodate this interpretation are highly artificial, with singularities at all zeros of the wave function, and are incompatible with the principle of contiguity.

These deterministic ideas are in sharp contrast with those advanced by Born (1926), in his original papers on the theory of measurement. This theory accepted that the result of any measurement made on a microscopic system could not be predicted, in general, and that it was possible to determine only the probability that any particular result will be obtained. It was therefore considered necessary to abandon the determinism of the classical theories of Newton and his followers. This view became the basis of what is generally known as the *Copenhagen interpretation* of quantum mechanics. In the form advocated by Bohr (1928, 1933), it incorporated a subtle dualistic principle, according to which, at the microscopic level, matter has both wave-like and particle-like properties either of which may predominate, depending on the way it is observed. This principle should not be understood as meaning that wave-like and particle-like properties are mutually exclusive (Ghose *et al.* 1992; Mizobuchi and Ohtaké 1992). However, it does exclude both of the deterministic interpretations outlined above.

Here we shall adopt the Copenhagen interpretation of quantum theory in essence, but shall develop a more satisfactory theory of measurement in terms applicable to both microscopic and macroscopic systems, the latter being distinguished only by an extremely large number of independent degrees of freedom. Since this approach is new, it will be developed in sufficient detail to show that it can also provide a self-sufficient basis for the accepted results of quantum mechanics. The theory may be regarded simply as an application of quantum mechanics to the interaction of a microscopic system with a measuring device. In part, it may also be regarded as the generalization of a theme developed previously by the author (Green 1958; Green and Triffet 1991).

However, it also develops an idea, suggested by various authors from Brillouin (1956) to Busch *et al.* (1991), that quantum mechanics should be regarded as an extension of information theory. We shall make a point, moreover, of never assuming more information than can be made available experimentally. This of course forbids the introduction of hidden variables or a universal wave function, and *a fortiori* the use of a single wave function or state vector to represent a system of any kind. We are then committed to a version of quantum mechanics similar to that inaugurated by Born and Jordan, and motivated by Heisenberg's (1925) proposal that quantum theory should not contain any numerical quantity that cannot, in principle, be measured.

Incidentally, we shall show for the first time that the resulting indeterministic theory is consistent with the principle of contiguity. In the process of measurement of a microscopic observable, there are large changes of information on a time scale characteristic of the measuring device, but there is no discontinuity in the information or any other fundamental observable in the process of measurement. Contiguity is secured by the definition of a density and a flux density for all such observables. It is an inevitable feature of the theory that, although it becomes extremely tenuous, the thread connecting reality to 'other worlds' in the acquisition of information is never completely broken. However, as Bohr insisted, the theory has no result that is not consistent with common sense derived from macroscopic experience, and paradoxes such as the EPR paradox can be understood without invoking either action at a distance or superluminal velocities.

2. The Statistical Matrix in Quantum Theory

We wish to consider the measurement of some property of a system S, usually of atomic dimensions; then it can only be detected with the help of a macroscopic device, such as a counter or cloud chamber. From the present point of view, the essential result of the detection of the system is an exchange of information. The device must have some component that is initially in a metastable state, but makes the transition to a stable state in the process of detection; this transition will be characterized as an increase of entropy that is equivalent to a loss of thermodynamic information. The loss of information concerning the macrosopic device is compensated in part by the gain of information concerning the microscopic system.

The detection of a microscopic system does not necessarily provide an observer with more than evidence of its existence, and a very rough idea of its position at the time. However, by a suitable experimental arrangement, usually involving the placement of apertures and electromagnetic guides, it can be ensured that a suitably chosen detector will interact with systems having only a certain value or values of some observable or set of observables, such as momentum, energy and helicity. By the use of a composite system of detection with several independently functioning components, it may be possible to distinguish between different values of the selected observables. The experimental arrangement that secures this possibility also determines the type of information obtained by detection of the microscopic system, and in the following discussion of quantum mechanics we wish to provide the means to quantify this type of information.

In most existing theories of measurement, it is assumed that, before it interacts with the detector, the microscopic system is in a pure state that can be represented by a single wave function (or state vector). It is, however, a feature of wave mechanics that if a system S consists of two sub-systems S′ and S″ in interaction, the wave function ψ of the joint system, supposing that there is one, cannot be factorized in the form $\psi'\psi''$, but is always a sum of such products. There is, therefore, no unique wave function representing a system that is or has been in interaction with any other system: there are no pure states in nature, and it is not possible to create one. Of course there can be no objection whatsoever to the use of wave functions for computational purposes, in particular for the determination of the eigenvalues of physical observables and transition probabilities. They may also have a role in the discussion of certain ideal, as distinct from actual measurements. But in the theory of measurement, and of macroscopic phenomena in general, it is important to recognize that it is not possible to attribute a physical significance to any particular wave function.

On the other hand, there is always a statistical matrix for any system, and in the theory of measurement it appears reasonable to suppose that this can be expressed as a sum of two terms, one representing the possibility that a given system does activate a particular macroscopic component of a detector, and the other representing the possibility that it does not. Moreover, as we have noted above, it may be legitimate to conclude that, if it does activate the detector, one or more observables of the system had values that are commonly attributed to a pure state. An adequate theory of measurement therefore requires an explanation of how it is possible for a statistical matrix to be reduced effectively to a sum of two or more terms.

In this Section, we shall begin with a brief and uncontroversial account of how quantum mechanics can be formulated without reference to wave functions or state vectors. According to Heisenberg's uncertainty principle, and Bohr's principle of complementarity, there are pairs of complementary observables associated with any system, such that an experimental arrangement to measure one of the observables accurately precludes the possibility of measuring the other. In Born and Jordan's (1925) quantum mechanics, this was provided for by representing any microscopic observable by a matrix operator O, whose eigenvalues o_r are the possible values that may result from the measurement of the observable. The subscript r in general denotes a vector $(r_1, \ldots r_R)$, where R is the number of degrees of freedom of the system, and if the eigenvalues are discrete, the matrix O can be expressed in terms of a discrete set of matrices g_r, with numerical coefficients o_r, thus:

$$O = \sum_r o_r g_r. \tag{1}$$

The g_r are minimal hermitean idempotents or projections, satisfying

$$g_r g_s = \delta_{rs} g_s, \qquad \mathrm{tr}(g_r) = 1, \qquad \sum_r g_r = 1, \tag{2}$$

so that $o_r = \mathrm{tr}(O g_r)$. Here and in the following, tr denotes the *trace*, or the sum of the diagonal elements, of the matrix which follows, and, in a relation between matrices, 1 denotes the unit matrix. If O and $\widehat{O}$ represent two observables that are complementary, in Bohr's sense, or cannot be measured by the same detector, then they do not commute: if $\widehat{O} = \sum_s \widehat{o}_s \widehat{g}_s$, then $O\widehat{O} \neq \widehat{O}O$ and $g_r \widehat{g}_s - \widehat{g}_s g_r \neq 0$ in general.

Any matrix C of the system S can be expressed in terms of products $g_r \widehat{g}_s$ of the idempotents of complementary observables:

$$C = \sum_{r,s} c_{rs} g_r \widehat{g}_s, \qquad c_{rs} = \mathrm{tr}(C \widehat{g}_s g_r). \tag{3}$$

If C is hermitean, it is also possible to write $C = \sum_{s,r} c^\dagger_{sr} \widehat{g}_s g_r$, where $c^\dagger_{sr}$ is the complex conjugate of c_{rs}. The eigenvalues of C are then real, but do not necessarily form a discrete set, so that observables with continuous spectra can be represented in this way. However, there is also a spectral decomposition $C = \int c_r dE_r$ of such observables, where formally $E_r = \int^r g_r dr$, and the summation in (1) is interpreted as an integration.

For any system there is a special observable P, called the statistical matrix, or the statistical operator in von Neumann's (1955) terminology, that plays a central part in quantum-mechanical information theory. In many applications, even in the absence of any measurement, there is a relation between the statistical matrix P of a microscopic system and what will be called *selected* observables. There is a selected observable O^*, whose eigenvalues p_v^* $(v = 1, 2...)$ determine the probability p_v^* that the measurement of O^*, with a suitably chosen detector D*, will yield the one of the values o_s^* $(s = 1, 2...)$. We note that these values are not necessarily all different from one another. If

$$O^* = \sum_v o_v^* g_v^*, \tag{4}$$

then the statistical matrix may be defined by

$$P = \sum_v p_v^* g_v^*. \tag{5}$$

Since $\sum_v p_v^* = 1$, then $\mathrm{tr}(P) = 1$. It will be noticed that P commutes with the selected observable O^*, but not, in general, with another observable O. As already suggested, the nature of the selected observable may be determined by an experimental arrangement designed to ensure that only the value o_v^* of the selected observable O^* will be detected by a particular component of the detector.

According to Born's theory of measurement, the probability that the measurement of *any* observable $O = \sum_r o_r g_r$ by a suitable detector D will yield the value o_s is

$$p_r = \mathrm{tr}(g_r P) = \sum_v p_{rv} p_v^*, \qquad p_{rv} = \mathrm{tr}(g_r g_v^*). \tag{6}$$

In view of the interpretation of p_r and p_v^* as probabilities, p_{rv} can be interpreted as the probability of a 'transition', or the conditional probability that the measurement of O will yield the value o_r provided that it is certain that the selected observable O^* has the value o_v^*. It is obvious that the above formula is correct if O and O^* are the same, or commuting observables, but it is also true if O and O^* do not commute. In an adequate theory of measurement, this proposition should not simply be assumed, but should be established on the basis of the quantum mechanics of the interaction of the microscopic system with the macroscopic detector.

In macroscopic physics, there is usually no experimental arrangement to measure any observable, and the only information is of statistical, or thermodynamical nature. The selected observables in such applications include the energy, or, more generally, the derivative of the action with respect to time, which may depend on several fundamental observables.

We shall next discuss briefly the definition of the matrices representing the fundamental observables, such as the position, momentum, angular momentum and energy of a system S, possibly but not necessarily a particle or set of particles in the microscopic domain. This is a simple application of the theory of Lie groups. We shall regard the observables as independent of the conditions under which they are measured, or indeed whether they are measured at all.

However, in general the probabilities p_s^*, and therefore the statistical matrix P, obviously depend on the time, the spatial configuration and state of motion.of the macroscopic detector D. Assuming that the latter is an rigid unaccelerated body, a sufficient specification includes the position $\mathbf{x}$ of the centre of mass of the detector at any time t, its orientation $\mathbf{u}$, and its velocity $\mathbf{v}$ relative to some other rigid unaccelerated body D_0. If the detector of the microscopic system is D, we denote the statistical matrix of the microscopic system by $P = P(t, \mathbf{x}, \mathbf{u}, \mathbf{v})$. But, if the detector is D_0, the statistical matrix is $P_0 = P(0, \mathbf{0}, \mathbf{0}, \mathbf{0})$. Because of the interpretation of the eigenvalues of this matrix as probabilities, these two statistical matrices must be related by a *similarity* transformation:

$$P(t, \mathbf{x}, \mathbf{u}, \mathbf{v}) = U(t, \mathbf{x}, \mathbf{u}, \mathbf{v})P_0\overline{U}(t, \mathbf{x}, \mathbf{u}, \mathbf{v}), \quad \overline{U}(t, \mathbf{x}, \mathbf{u}, \mathbf{v})U(t, \mathbf{x}, \mathbf{u}, \mathbf{v}) = 1, \tag{7}$$

where 1 again stands for the unit matrix. In fact, as the statistical matrices are hermitean, this is also a *unitary* transformation: $\overline{U}$ is not merely the inverse of U but its hermitean conjugate.

If we consider a small change δt in the time t at which the measurement is made, we may write

$$U(t + \delta t, \mathbf{x}, \mathbf{u}, \mathbf{v}) = (1 - iH\delta t/\hbar)U(t, \mathbf{x}, \mathbf{u}, \mathbf{v}), \tag{8}$$

where $\hbar$ is Planck's constant and H is an hermitean matrix. It then follows that

$$i\hbar\frac{\partial U}{\partial t} = HU, \qquad i\hbar\frac{\partial P}{\partial t} = HP - PH. \tag{9}$$

Thus the matrix U satisfies the same equation as the wave function in Schrödinger's wave mechanics. Adopting Heisenberg's definition of the energy as the agent of change with time (in suitable units with Planck's constant equal to 1), we conclude that the observable H is the energy of the microscopic system. So long as a microscopic system is free of external interactions, all probabilities, and therefore the statistical matrix of the system, are independent of the time; it then follows that $HP = PH$, and the energy is a selected observable, as defined above.

There are similar definitions, analogous to Heisenberg's definition of the energy, for the momentum $\mathbf{P}$, the angular momentum $\mathbf{J}$, and the central vector $\mathbf{K}$ ($=M\mathbf{Q}$, where M is the mass and $\mathbf{Q}$ the position of the mass centre) of a microscopic system, as the agents of change with position, orientation and velocity, respectively. In consequence of these definitions, we have

$$-i\hbar\frac{\partial U}{\partial \mathbf{x}} = \mathbf{P}U, \qquad -i\hbar\frac{\partial U}{\partial \mathbf{u}} = \mathbf{J}U, \qquad i\hbar\frac{\partial U}{\partial \mathbf{v}} = \mathbf{K}U = M\mathbf{Q}U\,, \tag{10}$$

and we may also write

$$U = \exp(-iA), \qquad A = Ht - \mathbf{P}\,.\,\mathbf{x} - \mathbf{J}\,.\,\mathbf{u} + \mathbf{K}\,.\,\mathbf{v}\,. \tag{11}$$

The last equation, together with (7), shows how the statistical matrix, and all the fundamental observables of a microscopic system, are completely determined by the *action* A, regarded as a function $A(t, \mathbf{x}, \mathbf{u}, \mathbf{v})$ of the set of parameters introduced above. In the Lagrangian formulation of Newtonian mechanics, the parameters are regarded specifying the system adopted by a particular observer O relative to that of some other observer O_0, and the interpretation given above in terms of detectors is quite consistent with this idea. In quantum mechanics and quantized field theory the action associated with a microscopic system is treated as an observable and, for reasons already discussed, represented by a matrix, but the derivation of the observables is unchanged. It is clear from (7) that the statistical matrix P is completely determined by its value P_0 for $t=0$ and $\mathbf{x}=\mathbf{u}=\mathbf{v}=\mathbf{0}$, and the action A.

The derivation of the commutation rules satisfied by the fundamental observables, which determine their matrix representations, will be discussed briefly in the next Section. We conclude the present Section with an outline of the extension of the above considerations to *composite* systems. It is sufficient to consider a system S consisting of two subsystems S$'$ and S$''$. We assume that these systems are disjoint, without any parts in common, and denote observables of the two systems separately by O' and O''. When these are regarded as matrices of the composite system, however, they are expressed as direct (outer) products of the form $O^{(1)}=O'\otimes 1$ and $O^{(2)}=1\otimes O''$ with idempotents $g_r^{(1)}=g_r\otimes 1$ and $g_v^{(2)}=1\otimes g_v$. By a generalization of (3), an arbitrary matrix C of the joint system can be expressed in terms of direct products

$$C=\sum_{r,s.v,w} c_{rs,vw}g_r\widehat{g}_s\otimes g_v\widehat{g}_w, \qquad c_{rs,vw}=\mathrm{tr}[C(\widehat{g}_sg_r\otimes\widehat{g}_wg_v)] \tag{12}$$

of matrices $g_r\widehat{g}_s$ and $g_v\widehat{g}_w$ of the subsystems.

Of course there is a statistical matrix P for the composite system, but it is also always possible to define statistical matrices P' and P'' for the separate subsystems as partial traces of P, thus:

$$P'=\sum_{r,s} g_r\widehat{g}_s\mathrm{tr}[P(\widehat{g}_sg_r\otimes 1)], \qquad P''=\sum_{v,w} g_v\widehat{g}_w\mathrm{tr}[P(1\otimes\widehat{e}_wg_v)]\,. \tag{13}$$

If $O'=\sum_u o'_ug_u$ is an observable of the first subsystem, the probability p'_u that a measurement of O' will yield the value o'_u, according to (6), is $\mathrm{tr}'(g_uP')$, and this is the same as $\mathrm{tr}(g_u^{(1)}P)$, as required, since $\mathrm{tr}[g_u(g_r\widehat{g}_s)]=\delta_{ur}$.

If the two subsystems of the composite system are not interacting and have not interacted, directly or indirectly, in the recent past, they are statistically independent of one another, and the joint probability that the measurement of selected observables $O'=\sum_s o'_sg'_s$ and $O''=\sum_s o''_sg''_s$ of the subsystems will yield the values o'_s and o''_s must be the product, $p_s=p'_sp''_s$, of the separate probabilities. From (5) it can be seen that this condition for independence implies that the statistical matrix P of the composite systems is the direct product of the statistical matrices $P^{(1)}$ and $P^{(2)}$ of the subsystems:

$$P=P'\otimes P''\,. \tag{14}$$

It can also be seen from (5) that a unitary transformation U of the type considered that there must, subject to the same restriction, also be a direct product: $U = U' \otimes U''$. It follows from (11) that the action observable A is additive in the sense that $A = A^{(1)} + A^{(2)}$, and the same is true for all the fundamental observables:

$$H = H^{(1)} + H^{(2)}, \qquad \mathbf{P} = \mathbf{P}^{(1)} + \mathbf{P}^{(2)}, \qquad \mathbf{J} = \mathbf{J}^{(1)} + \mathbf{J}^{(2)},$$
$$\mathbf{K} = \mathbf{K}^{(1)} + \mathbf{K}^{(2)}, \qquad M = M^{(1)} + M^{(2)}. \tag{15}$$

However, (15) and (16) apply only as long as the systems have had no interaction, and with interaction we must write

$$H = H^{(1)} + H^{(2)} + V, \tag{16}$$

where V is the interaction energy. Often, V is a scalar function of the relative position $\mathbf{Q}^{(2)} - \mathbf{Q}^{(1)}$, and the observables $\mathbf{P}$, $\mathbf{J}$, $\mathbf{K}$ and M are still simply additive. To generalize (15), it is necessary to make use of the theory of the interaction representation, and this will be done in Section 5.

3. Contiguity in Quantum Mechanics

The differential equations (9) and (10) determine the way in which the outcome of any measurement made with the device D will depend on the a set of ten parameters (t, x_1, x_2, x_3, u_1, u_2, u_3, v_1, v_2, v_3), where the subscripts denote cartesian components of the vectors. The parameters are those of the Galilei group of non-relativistic physics or the inhomogeneous Lorentz group of relativistic physics. The matrix U is an element of the group, and the observables H, P_1, P_2, P_3, J_1, J_2, J_3, K_1, K_2 and K_3 are the elements of the corresponding Lie algebra, for which we have adopted a matrix representation. The mass M is an invariant, and in the non-relativistic formulation, it can be regarded as an additional fundamental observable and an element of the Lie group, associated with the parameter $\mathbf{x} \cdot \mathbf{v}$; however, as it commutes with all the other matrices it may be treated as a numerical multiple of the unit matrix.

The *commutation relations* satisfied by the observables are easily derived from the geometrical and kinematical relations between the parameters of the group (as in Cornwell 1984) and in the Galilean approximation include the well known results

$$M(Q_\lambda H - HQ_\lambda) = i\hbar P_\lambda = iM\dot{Q}_\lambda\,,$$
$$Q_\lambda P_\mu - P_\mu Q_\lambda = i\hbar\delta_{\lambda\mu} \tag{17}$$

of Born and Jordan. The complete set of commutation relations can be used to determine the form of the matrices representing the observables unambiguously, apart from a similarity transformation. We note that $2MH - \mathbf{P}^2$ is an invariant of the enveloping algebra of the Lie algebra.

Let P be the statistical matrix of any system S, and C be any conserved observable, so that, according to (9),

$$i\hbar\frac{\partial(CP)}{\partial t} = H(CP) - (CP)H. \tag{18}$$

We define the density $\rho(\mathbf{x}, t; C)$ and the flux density $\sigma(\mathbf{x}, t; C)$ of C by

$$\rho(\mathbf{x}, t, C) = \mathrm{tr}[CP\delta(\mathbf{x} - \mathbf{Q})] \equiv \frac{1}{(2\pi)^3} \int \mathrm{tr}\{CP \exp[i\mathbf{p} \cdot (\mathbf{x} - \mathbf{Q})]\} d^3 p,$$

$$\sigma(\mathbf{x}, t; C) = \mathrm{tr}[\tfrac{1}{2}(\dot{\mathbf{Q}}CP + CP\dot{\mathbf{Q}})\delta(\mathbf{x} - \mathbf{Q})]\,. \tag{19}$$

Then, from (9), (10) and the commutation relations, we verify the macroscopic conservation law

$$\frac{\partial \rho(\mathbf{x}, t; C)}{\partial t} + \nabla \cdot \sigma(\mathbf{x}, t; C) = 0 \tag{20}$$

(where $\nabla = \partial/\partial \mathbf{x}$). The validity of the principle of contiguity for any conserved observable C is thus established.

Where the system S consists of two disjoint subsystems S$'$ and S$''$, we write $C = C^{(1)} + C^{(2)}$, where $C^{(1)} = C' \otimes 1$ and $C^{(2)} = 1 \otimes C''$, and $M\mathbf{Q} = M^{(1)}\mathbf{Q}^{(1)} + M^{(2)}\mathbf{Q}^{(2)}$; then a similar construction can be given for the densities $\rho(\mathbf{x}, t, C')$ and $\rho(\mathbf{x}, t, C'')$ of the conserved quantities associated with the subsystems. The existence of the corresponding flux densities $\sigma(\mathbf{x}, t, C')$ and $\sigma(\mathbf{x}, t, C'')$ demonstrates the validity of the principle of contiguity also for the subsystems.

If $C = 1$ (the unit matrix), the density $\rho(\mathbf{x}, t, C)$ is of course the probability, per unit volume, that with a suitable detector, the system will be found in the neighbourhood of the point $\mathbf{x}$ at time t, and $M\sigma(\mathbf{x}, t, C)$ is the expectation value of the momentum per unit volume associated with the system. Similar considerations apply if C is the energy, momentum, or angular momentum, but also if C is any function of the statistical matrix P itself. From the definition of information in (22) below, it will be evident that when $C = -\log(P)$, $\rho(\mathbf{x}, t, C)$ is the density and $\sigma(\mathbf{x}, t, C)$ the flux density of information. Since the principle of contiguity is satisfied, it is completely unnecessary to suppose that the transfer of any conserved observable is accomplished by action at a distance.

4. Information Theory in Quantum Mechanics

The theory of measurement in quantum mechanics may be regarded as an application of information theory. In the classical theory, due to originally to Shannon (1949), the information gained by observation of any event is defined as the expectation value of $-\log(P)$, if P is the probability of a particular outcome of the event. But in statistical mechanics, according to Boltzmann, the macroscopic entropy associated with any system is the expectation value of $-k\log(P)$, where k is Boltzmann's constant and P is the probability that the system will be found in a particular state. When an extremely small unit is chosen for the absolute temperature, then $k = 1$, and in a thermodynamical context there is therefore no difference between information and entropy.

In quantum mechanics, it is easy to generalize the definition of information in the following way. Again we consider the measurement of a microscopic observable with the help of a detector. The information I to be gained from the measurement of any observable $O = \sum_r o_r g_r$ of a system S by a suitably chosen detector D is

$$I = -\sum_r p_r \log(p_r)\,, \tag{21}$$

where p_r is the probability that the measurement will yield the value o_r; this is given by $p_r = \mathrm{tr}(g_r P)$, according to Born's hypothesis, as formulated in (6) above.

It is evident that the information depends on the observable. If $O^* = \sum_r o_r^* g_r^*$ is a selected observable, then it commutes with P, so that $P = \sum_r p_r^* g_r^*$ and the information to be gained by measuring the selected observables, including O^*, is

$$I^* = -\sum_r p_r^* \log(p_r^*) = -\mathrm{tr}[P\log(P)] = -\langle \log(P)\rangle\,, \tag{22}$$

where the logarithm is defined by $\log(P) = \sum_r g_r^* \log(p_r^*)$. In the last Section, it was shown how to define a density and a flux density for conserved macroscopic quantities of this kind, and there is no doubt that it satisfies the principle of contiguity.

But I^* is in general different from and in fact less than the information I to be gained from the measurement of the observable O, and it is therefore meaningless to speak of the information to be gained by observation of a microscopic system without reference to the detector or what is being measured. In this respect, there is an important difference between the classical and quantum mechanical theory of information. Moreover, it is by no means evident that the information defined in (21) always satisfies the principle of contiguity. That will be made evident in the following Sections of this paper, but for the present we shall state a view consistent with the Copenhagen interpretation of quantum mechanics, but based essentially on macroscopic experience.

When information is first gained concerning a microscopic system, it is localized at the detector or detectors, where it may possibly be recorded, coded, and transmitted elsewhere by electromagnetic or by other physical means. However, the detection is generally distributed over an extended region of space. Let us consider, for instance, a pair of complementary observables, such as the momentum $\mathbf{P}$ and the position $\mathbf{Q}$ of the centre of mass of a system S. The momentum of the system can only be measured, even approximately, by the detection of wave-like properties of the system over an extended region, and the information gained in this way is extensively non-local. The position may be measured approximately by the observation of particle-like properties, for instance by placing a detector immediately behind a small aperture through which the system must pass; the information gained in this way is localized at the detector. Bohr's discussion of complementary measurements made the point that one of these measurements excludes the other. But, as already pointed out, Bohr's 'principle of complementarity' should not be interpreted as meaning that the wave-like and particle-like properties of matter are mutually exclusive (Ghose *et al.* 1992; Mizobuchi and Ohtaké 1992). In such experiments, information concerning particle and wave-like properties is gained by detecting coincidences or anti-coincidences and is at least partly non-local. We shall show below that information theory is useful in resolving paradoxes, such as the EPR paradox, that were originally formulated with the intention of raising doubts concerning the indeterministic (Copenhagen) interpretation of quantum mechanics.

We suppose first that the measurement of a selected observable O^* by a detector D*, possibly a component of a larger system, can yield only the eigenvalue o_u^*. There are three different ways in which this can be reconciled with our earlier observation that there are no pure states in nature:

[1] One possibility is that, in an actual measurement with the detector D*, a system for which the eigenvalue of O^* is different from o_u^* cannot interact with D*. The state of any system interacting with the detector is not pure, and there is information, not exceeding I^*, to be gained from the measurement of O^*. This is one of the possibilities taken into account in the considerations of the following Sections.

[2] It is also possible that, in an actual measurement with the detector D*, the eigenvalue of the selected observable O^* (e.g. the spin of the system) is certain to be o_u^*, because, for very many values of r, $o_r^* = o_u^*$, while for other values $p_r^* = 0$. Then the state of any system interacting with the detector is not pure, but the information gained by a measurement of O^* alone is zero.

[3] In an ideal as distinct from an actual measurement, a system interacting with the detector may not have interacted previously with any other system, and it is then possible for the system to be in a pure state in which the eigenvalue of the observable O^* is o_u^*. It follows that $p_r^* = \delta_{ru}$, that the statistical matrix P for the system reduces to a single idempotent g_u^*, and that the information gained from a measurement of O^* is zero. But if O does not commute with P, the information I to be gained from the measurement of O is given by (1) and has a value limited only by the accuracy of the measurement.

The third possibility listed above has often been discussed in the literature, and is not excluded in the following discussion, which however is also consistent with the more realistic options. A typical application is to the EPR paradox (Einstein *et al.* 1935), which concerns a composite system S consisting of two microscopic subsystems S$'$ and S$''$ that are not in interaction but have interacted previously. The statistical matrix P for the composite system is not simply a direct product of the statistical matrices P' and P'' of the separate subsystems, and these statistical matrices are certainly not idempotents. If $O' = \sum_r o'_r g_r$ is any observable of the first subsystem, and $\widehat{O}' = \sum_s \widehat{o}'_s \widehat{g}_s$ is a complementary observable of the same subsystem, the statistical matrix P can always be written in the form

$$P = \sum_{r,s} g_r \widehat{g}_s \otimes p_{rs}\,, \tag{23}$$

where the p_{rs} are matrices of the second subsystem. Because P is not a single direct product, the p_{rs} cannot have the same value for all values of r and s. According to (13), the statistical matrices of the subsystems are

$$P' = \sum_{r,s} g_r \widehat{g}_s \mathrm{tr}''(p_{rs}), \qquad P'' = \sum_{v,w} \mathrm{tr}'(g_v \widehat{g}_w) p_{vw}\,. \tag{24}$$

Let us now suppose that $O' = \sum_r o'_r g_r$ is a selected observable of the first subsystem; then $P' = \sum_r p'_r g_r$, where p'_r is the probability that a measurement of O' will yield the value o'_r. This is only possible if p_{rs} has the same value $p'_r g_r$ for every value of s, so that if the measurement of O' yields the value o'_u, then the measurement of $O'' = \sum_v o''_v g_v$ must yield the value o''_u. The information gained from the measurement is given by (22), with $p^*_r = p'_r$. On the other hand, if $\widehat{O}'$ is the selected observable, $P' = \sum_s \widehat{p}'_s \widehat{g}_s$, so that p_{rs} must have the same value $\widehat{p}'_s \widehat{g}_s$ for every value of r, and if the measured value of $\widehat{O}'$ is $\widehat{o}'_u$, then the measured value of $\widehat{O}''$ is $\widehat{o}''_u$. The information gained from the measurement is still given by (22), but with $p^*_r = \widehat{p}'_r$. It is not possible to measure complementary observables in the same experiment, and the probabilities $\widehat{p}'_r$ are in general quite different from the p'_r .

All actual experiments have so far confirmed the validity of these predictions of quantum mechanics. But if the second subsystem, in an experiment of the type just considered, is far removed from the first, it may not be easy to see how the choice of a selected observable for the first subsystem can affect the form of the statistical matrix, and the result of a measurement performed on the second subsystem. It was this paradox that led Einstein to conclude that quantum mechanics was an incomplete theory. We have referred to proposals that there is some form of action at a distance that allows a measurement made in one place to affect a distant system. This in fact seems inevitable if quantum mechanics is interpreted either in terms of a universal wave function or in terms of hidden variables. But when the EPR and similar quantum mechanical paradoxes are interpreted in the light of indeterministic information theory, the problematic concept of action at a distance will be found unnecessary.

The first point to be made is that, as we have shown above, it is no surprise that measurements made in different places should be correlated. Consider, in the interests of clarity, an idealized composite system for which the commuting observables $\mathbf{P}^{(1)} + \mathbf{P}^{(2)}$ (the resultant momentum) and $\mathbf{Q}^{(2)} - \mathbf{Q}^{(1)}$ (the relative position) of two subsystems have definite eigenvalues $\mathbf{p}$ and $\mathbf{r}$. It is not surprising that if the measurement of $\mathbf{P}^{(1)}$ yields the value $\mathbf{p}^{(1)}$, a distant measurement of $\mathbf{P}^{(2)}$ should yield the value $\mathbf{p} - \mathbf{p}^{(1)}$, or that if, alternatively, the measurement of $\mathbf{Q}^{(1)}$ yields the value $\mathbf{q}^{(1)}$, a distant measurement of $\mathbf{Q}^{(2)}$ should yield the value $\mathbf{r} + \mathbf{q}^{(1)}$. As we have pointed out above, this is a valid prediction of the theory. But measurements of the complementary observables $\mathbf{P}^{(1)}$ and $\mathbf{Q}^{(2)}$ are mutually exclusive and yield different information, and it is important to avoid the assumption that a suitable experiment could yield precise information concerning the values of both $\mathbf{p} - \mathbf{p}^{(1)}$ and $\mathbf{r} + \mathbf{q}^{(1)}$, or that either was already determinate, in some sense, before any measurement was made.

The second and more essential point is that the information gained from the measurement of momentum is associated with an extended region of space; on the other hand, the information gained from the idealized measurement of relative position is also non-local. Whatever its nature, it is not necessary to suppose that information is transmitted by action at a distance. The laws of quantum mechanics, and especially the conservation laws and the principle of contiguity, satisfied by the fundamental observables are sufficient to ensure that no surprising discrepancies should be found between information concerning the same system, acquired in different ways. There are certainly serious problems if the information obtained from a measurement on a microscopic system is supposed to be predetermined, but, viewed from the point of view of indeterministic information theory, quantum mechanics offers no affront to common sense.

In the following, we shall analyse the requirements for the realistic detection of a microscopic system in more detail, and so obtain a better insight into the mechanism by which, in the process of measurement, an indeterminate observable acquires a definite value.

5. The Interaction Representation

When two systems interact, there is in general an exchange of both energy and information, and while the total energy is unchanged, there may be a loss of total information over a period of time. This applies in particular in the interaction between a microscopic system and a macroscopic detector, which will be considered later. In quantum mechanics the technique of the *interaction representation* is very useful in dealing with such problems

Let us consider a composite system S consisting of two subsystems S$'$ and S$''$, and suppose that the statistical matrix of the composite system at some initial time $t = 0$ is P_0. As shown in (13), the statistical matrices of the subsystems at the initial time are $P_0' = \mathrm{tr}''(P_0)$ and $P_0'' = \mathrm{tr}'(P_0)$, but P_0 can only be expressed as the direct product of P_0' and P_0'' when the subsystems have not been in interaction up to that time. It follows that the information to be gained concerning the composite system is the same as the joint information to be gained from the two subsystems only if they have not interacted previously.

We choose an observational frame of reference so that the centre of mass of the composite system is at rest ($\mathbf{v} = 0$) and is appropriately oriented ($\mathbf{u} = 0$) at the origin of coordinates ($\mathbf{x} = 0$). Then, according to (7) and (11) and (16), the statistical matrix P of the system S at time t is given in terms of its value P_0 at the initial time by

$$P = UP_0\overline{U}, \qquad U = \exp(-iHt/\hbar), \qquad H = H^{(1)} + H^{(2)} + V\,. \tag{25}$$

Here H, $H^{(1)}$ and $H^{(2)}$ are the energies of the composite system and the two subsystems, respectively, and V is the interaction energy; also, $\overline{U}$ is the hermitean conjugate of U, which satisfies $U\overline{U} = 1$. It is obvious that the unitary matrix U satisfies

$$i\hbar\frac{\partial U}{\partial t} = (H^{(1)} + H^{(2)} + V)U\,, \tag{26}$$

and if we write

$$U_0 = \exp[i(H^{(1)} + H^{(2)})t/\hbar]\,,$$

$$U = \exp[-i(H^{(1)} + H^{(2)})t/\hbar]T = \overline{U}_0 T\,, \tag{27}$$

it follows from (26) that the unitary matrix $T = T(t)$ (the S-matrix for large t) satisfies the equations

$$i\hbar\frac{\partial T}{\partial t} = \widetilde{V}T, \qquad \widetilde{V} = U_0 V \overline{U}_0 ,$$

$$T(t) = 1 - i\int_0^t \widetilde{V}(\tau)T(\tau)d\tau/\hbar , \tag{28}$$

of which the last can be solved by iteration or otherwise. The expression given in (25) for the statistical matrix can then be rewritten in the form

$$P = \overline{U}_0 \widetilde{P} U_0, \qquad \widetilde{P} = T P_0 \overline{T} , \tag{29}$$

where $\widetilde{P}$ is the statistical matrix in the interaction representation and satisfies

$$i\hbar\frac{\partial \widetilde{P}}{\partial t} = \widetilde{V}\widetilde{P} - \widetilde{P}\widetilde{V} . \tag{30}$$

The expectation value $\langle O\rangle$ of an observable O at time t is $\mathrm{tr}(PO)$, where P obviously depends on the time, so that although, in the Heisenberg representation, O is independent of the time, its expectation value $\langle O\rangle$ varies with the time, as it should. However, it follows from (29) that

$$\langle O\rangle = \mathrm{tr}(PO) = \mathrm{tr}(\widetilde{P}\widetilde{O}), \qquad \widetilde{O} = U_0 O \overline{U}_0 , \tag{31}$$

where $\widetilde{O}$ is the observable in the interaction representation. It is evident that, in principle, the last result, together with (28), provides a method for the calculation of the expectation value of any observable of the composite system, as a function of time. The interaction representation is in fact used extensively in quantum field theoretical calculations.

The matrix U_0 in (31) is the direct product $U'\otimes U''$, where $U' = \exp(iH't/\hbar)$ and $U'' = \exp(iH''t/\hbar)$ are unitary matrices of the separate subsystems. The explicit evaluation of these matrices requires only the determination of the eigenvalues of H' and H''. This can be done by Schrödinger's method, which involves the discussion of the solutions of differential equations with special boundary conditions, but, as Green and Triffet (1969) have shown, matrix methods are generally much simpler.

The theory of the interaction between two systems is somewhat simpler if they have not experienced any interaction at the initial time $t = 0$. Then the statistical matrix P_0 of the composite system at that time can be expressed as a direct product $P_0'\otimes P_0''$ of the statistical matrices of the subsystems, so that the result (29) can be written

$$\widetilde{P} = T(P_0'\otimes P_0'')\overline{T}, \tag{32}$$

where, for instance, we may write $P_0' = \sum_s p_s^* g_s^*$, if p_s^* is the probability that any selected observable $O^* = \sum_s o_s^* g_s^*$ of the first subsystem has the value o_s^* at the initial time.

There is an obvious generalization of all these considerations to a system consisiting of any number of subsystems in interaction. In the following, we do not in fact exclude the possibility that one of the systems (S′) is microscopic and the other is a macroscopic system, such as a measuring device, constructed from any number of disjoint macroscopic components S_1, ... S_d, each consisting of a very large number of microscopic subsystems. Very special instances of this type with $d = 1$ and $d = 2$, respectively, were considered previously by Green (1958) and Green and Triffet (1991).

It is in fact sufficient to consider the *possibility* of the interaction of S' with a single component. For if $d > 1$, the components must be chosen so that they are well separated. The statistical matrix of the measuring device is then a direct product of the statistical matrices of its components. The interaction energy in (30) will then be a sum of terms: $\widetilde{V} = \sum_j \widetilde{V}_j$ where $\widetilde{V}_j$ is the interaction with the jth component of the measuring device, and has no significant effect on factors of the statistical matrix corresponding to other components.

The question of interest is then whether S′ either interacts or does not interact, with a particular component S″. In the many worlds scenario, both possibilities are realized, though they exist in different worlds. An alternative view sometimes expressed is that neither possibility is realized in the absence of a conscious observer. Here, however, we shall develop a reasoned argument that *macroscopic events, irrespective of their cause, cannot remain indeterminate to any significant extent.* It will follow that, in an efficient measuring device, the unique component of the measuring system that interacts with S′, though itself indeterminate until the measurement is made, determines the measured value. We still have to show that this conclusion, consistent with observation, is also consistent with a theory of measurement based on application of information theory and quantum mechanics to the interaction of a microscopic system with a measuring device.

We shall express the T-matrix of a particular microscopic system (S′) which may interact with the detector (S″) in the form

$$T = \sum_{r+} g_{r+} \otimes T_{r+} + \sum_{r-} g_{r-} \otimes 1 \tag{33}$$

in which the g_{r+} and g_{r-} are minimal idempotents of an observable O' of S′, corresponding to the eigenvalue o'_r (usually degenerate), and other eigenvalues, respectively. Then it follows from (32) that

$$\widetilde{P} = P_{++} + P_{+-} + P_{-+} + P_{--}, \qquad P_{++} = \sum_{r+,s+} g_{r+} P'_0 g_{s+} \otimes T_{r+} P''_0 \overline{T}_{s+} ,$$

$$P_{+-} = \sum_{r+,s-} g_{r+} P'_0 g_{s-} \otimes T_{r+} P''_0, \qquad P_{--} = \sum_{r-,s-} g_{r-} P'_0 g_{s-} \otimes P''_0 , \tag{34}$$

and P_{-+} is the hermitean conjugate of P_{+-}. The statistical matrices of the separate subsystems S′ and S″ are $\widetilde{P}'$ and $\widetilde{P}''$, respectively, given by

$$\widetilde{P}' = P'_{++} + P'_{+-} + P'_{-+} + P'_{--}, \qquad P'_{++} = \sum_{r+,s+} \mathrm{tr}''(T_{r+} P''_0 \overline{T}_{s+}) g_{r+} P'_0 g_{s+} ,$$

$$P'_{+-} = \sum_{r+,s-} \mathrm{tr}''(T_{r+} P''_0) g_{r+} P'_0 g_{s-}, \qquad P'_{--} = \sum_{r-,s-} g_{r-} P'_0 g_{s-} \tag{35}$$

and

$$\widetilde{P}'' = P''_+ + P''_-, \qquad P''_+ = \sum_{r+} p'_{r+}(T_{r+}P''_0\overline{T}_{r+}), \qquad p'_{r+} = \mathrm{tr}'(g_{r+}P'_0)\,,$$
$$P''_- = (1 - p_r)P''_0, \qquad p_r = \sum_{r+} p'_{r+}\,. \tag{36}$$

Here p_r is identified as the total probability that measurement of the observable O' at time t should yield the value o'_r, with contributions given by

$$p'_r = \sum_v p'_{r,v}p^*_v, \qquad p'_{r,v} = \mathrm{tr}'(\sum_{r+} g_{r+}g^*_v)\,, \tag{37}$$

where p^*_v is the probability that a selected observable O^* of S$'$ has the value o^*_v and, as in (6), $p'_{r,v}$ is the probability for a transition between states represented by the idempotent g^*_v, and any of the g_{r+}. Of course $1 - p_r$ is then the total probability that the measurement yields some other value, or that the microscopic system is not detected. These results are entirely consistent with Born's statistical interpretation of quantum mechanics. The fact that the statistical matrix (36) is expressed as a sum of two terms (non-trivially, when $T_{r+} \neq 1$) implies not the existence of more than one world, but that, following the interaction, additional information may be gained from examination of the component S$''$ of the measuring system.

However, the nature of the information is coded in the statistical matrix $\widetilde{P}'$ of the system S$'$, and it should be expected that, with an efficient measuring device, this also reduces effectively for large t to the sum $P_{++} + P_{--}$ of two terms, since the component P'_{+-} of $\widetilde{P}'$ is a measure of the residual indeterminacy of the measurement of the value o'_{r+} of O'. From (35) it is evident that this component is given by

$$P_{+-} = \sum_{r+,s-} C_{r+}(g_{r+}P'_0g_{s-}), \qquad C_{r+} = \mathrm{tr}''(T_{r+}P''_0)\,, \tag{38}$$

where the coefficients C_{r+} are determined by the interaction of the microscopic system with the measuring device. The evaluation of these traces is carried out in the following Section.

6. Measurement and Observation

We are concerned with the quantum mechanics of the interaction of a microscopic system S$'$ with another system S$''$, designed to measure the observable O' of S$'$. The second system must then be a measuring device, or a component of a such a device, characterized as a detector, from the operation of which it should be possible to determine whether the observable has one or more of the values o'_s, and so gain information that can be acquired by a conscious observer. We shall first enquire into those properties of the detector that are essential for this purpose.

We again consider an experiment designed so that a particular component of the measuring system will function when and only when one particular value o'_r of the observable O' of the microscopic system is realized. This can be accomplished by an arrangement of guides and channels such that only systems distinguished by the value o'_r can enter a particular channel and interact with the rth subsystem of the detector. A composite detector may detect any of a number of eigenvalues, and it is also possible that some eigenvalues may escape detection. There are, however, two important requirements:

[1] Since only macroscopic phenomena can be observed directly, some macroscopic feature of a detector must change rapidly and significantly as a result of the interaction. It is obviously necessary that the detector should itself be a macroscopic object. It is also necessary that the interaction with the microscopic system S$'$ should result in an appreciable change in the expectation value of a selected observable O'' of some component S$''$ of the measuring device, corresponding to the eigenvalue o'_r, affecting very many (R) of the disjoint microscopic components from which S$''$ is made. If $O'' = -P'' \log(P'')$, where P'' is the statistical matrix of S$''$, the expectation value is the information I'' associated with S$''$, or the entropy in suitable units. A significant change in the entropy can result, for instance, if S$''$ is in a thermodynamically unstable or metastable state.

[2] It is also necessary to ensure that, if the observable O' is measured, the statistical matrix $\widetilde{P}'$ appearing in (35) should reduce to the form $P_{++} + P_{--}$ after the interaction with the detector has occurred. This condition is rigorously satisfied if and only if the traces C_{r+} in (38) are zero whenever there is some significant change in the information I'' associated with the component S$''$ of the measuring device. But to meet this requirement in fact poses an essential difficulty, since, as already mentioned, the C_{r+} are coefficients in the expression P_{+-}, which can be interpreted as a measure of the indeterminacy of the microscopic observable O' following the interaction with the detector; moreover, in quantum mechanics the indeterminacy associated with any finite system is never zero. We must therefore be content with the demonstration that the C_{r+} are immeasurably small, and thus effectively zero, under the conditions stated.

To formulate these requirements, we first express the statistical matrix P''_0 of S$''$ at the initial time in terms of the idempotents g^*_{m,s_m} ($m = 1, \ldots R$) of its microscopic subsystems, thus:

$$P''_0 = \sum_{s_1} p^*_{s_1} g^*_{1,s_1} \otimes P_{s_1}, \qquad P_{s_1 \ldots s_{m-1}} = \sum_{s_m} p^*_{s_1 \ldots s_m} g^*_{m,s_m} \otimes P_{s_1 \ldots s_m}. \tag{39}$$

Here the coefficient $p^*_{s_1 \ldots s_m}$ is a conditional probability, and $P_{s_1 \ldots s_m}$ is the reduced statistical matrix for the part of the detector excluding the first m of its microscopic subsystems. As foreshadowed in Section 2, the subscript $s+$ that labels the states of the detector is being treated as an R-component vector $(s_1, \ldots s_r)$. In the absence of interaction with S$'$, the statistical matrix $P^{(2)}$ of S$''$ at time t and also any selected observable $O^{(2)}$ of S$''$ can be expressed in the same way. However, at time t the statistical matrix of S$''$ in the interaction representation is given by (36), in which the unitary matrix T_{r+} can be written

$$T_{r+} = \sum_{r_1} t_{r_1} g_{1,r_1} \otimes T_{r_1}, \qquad T_{r_1 \ldots r_{m-1}} = \sum_{r_m} t_{r_1 \ldots r_m} g_{m,r_m} \otimes T_{r_1 \ldots r_m}, \tag{40}$$

where the coefficient $t_{r_1 \ldots r_m}$ is a complex number of modulus 1, and $T_{r_1 \ldots r_m}$ is unitary. All of these vary with the time.

Assuming that the detector functions, the change in entropy, or information can be determined from the components $T_{r+}P_0''\overline{T}_{r+}$ of the statistical matrix, which may be developed with the help of (39) and (40), thus:

$$T_{r+}P_0''\overline{T}_{r+} = \sum_{s_1} p_{s_1}^* g_{1,r_1} g_{1,s_1}^* g_{1,r_1} \otimes T_{r_1} P_{s_1} \overline{T}_{r_1} , \tag{41}$$

etc. The trace of any function of this or similar matrices can be evaluated in terms of the traces

$$c_{r_m s_m} = \mathrm{tr}(g_{m,r_m} g_{m,s_m}^*) , \tag{42}$$

and these are the only time-dependent factors. Thus a macroscopically significant change in the entropy requires a significant change in a large number of these transition probabilities. In fact, at time $t = 0$, $T_{r+} = 1$, and $g_{m,r_m} = g_{m,r_m}^*$, so that the $c_{r_m s_m}$ all reduce to 1; but for $t > 0$, when the above requirements [1] and [2] are met, very many of them must rapidly deviate from their initial values. A particular example of this, in which the detector was modelled by a set of coupled oscillators in a metastable state, was given in the author's original paper (Green 1958), but the present discussion is very much more general.

It is now easy to obtain an explicit expression for the indeterminacy coefficient C_{r+}, as defined in (38), by a similar method. Again with the help of (39) and (40), we have

$$C_{r+} = \prod_{m=1}^{R} t_{r_1 \ldots r_m} \sum_{s_m} p_{s_1 \ldots s_m}^* c_{r_m s_m} . \tag{43}$$

The transition probabilities $c_{r_m s_m}$ are always non-negative real numbers not greater than 1, and for $t \gg 0$, under the conditions stated, many of them will have values significantly less than 1 (typically $\frac{1}{2}$). This is therefore true also of the statistical averages $\sum_{s_m} p_{s_1 \ldots s_m} c_{r_m s_m}$. The coefficients $t_{r_1 \ldots r_m}$ under the product of (43) are of modulus 1. If R is sufficiently large, therefore, the products under the summation are all immeasurably small in absolute magnitude. Thus, the effect of the measurement is to reduce the indeterminacy coefficients of the observable O' in (38) to a completely negligible level.

The statistical matrix $\widetilde{P}'$ of S$'$ in (34) is thus reduced effectively to the form $P'_{++} + P$,

$$\widetilde{P}' \simeq P'_{++} + P'_{--} , \tag{44}$$

and as $\mathrm{tr}'(P'_{++}) = p'_r$, the information to be gained concerning the system from the operation of the detector $-\sum_r p'_r \log(p'_r)$. This has a maximum value, related to the accuracy of the measurement, when nothing is known initially concerning the microsopic observable O', and is only zero when the value of this observable is already known. In summary, the essential features of the detector were its macroscopic character, to ensure that the amplitudes C_{r+} were completely negligible, and the metastability of the functional components of the detector, to ensure a sufficiently rapid and readily observable change in entropy following the detection of the microscopic system.

The fact that the C_{r+} are not rigorously zero, however, is of some interest, in showing that at the microscopic level there is always a tenuous link with 'other worlds' of the many worlds interpretation of quantum mechanics. This link, however, is negligibly weak with a macroscopic measuring system, due to the amount of information transacted during the detection of the microscopic system.

When quantum mechanics is interpreted in the light of information theory there is no reason to accept the many worlds interpretation. For the information gained from a measurement of the type described is objective in exactly the same sense as the macroscopic change of entropy in any other irreversible process in nature is objective, and has nothing to do with the presence of a conscious observer. But the phenomenon of consciousness would be unintelligible in any deterministic theory, such as the many worlds theory, since it could have no effect on the course of events. In an indeterministic theory, it is left to the conscious observer to make the information gained by the measurement, in a certain sense, his or her own. As Green and Triffet (1991) have noted, the conscious process of noticing the result of the measurement is quite comparable with the action of a detector in making the measurement. However, the microscopic system involved is then not inanimate but, very probably, a component of the potential in the cortex of the observer. The information gained by conscious observation is usually in exact correspondence with some external macroscopic event, and may correspond to the information gained from the measurement of a microscopic observable.

References

Bell, J. S. (1990). *In* 'Sixty-two years of Uncertainty' (Ed. A. I. Miller), pp. 17–31 (Plenum: New York).

Bohm, D. (1952). *Phys. Rev.* **85**, 166–79; 180–93.

Bohr, N. (1928). *Nature* **121**, 580–90.

Bohr, N. (1933). *Nature* **131**, 421–3; 457–9.

Born, M. (1926). *Z. Phys.* **37**, 863–7; *ibid.* **38**, 803–27.

Born, M. (1949). 'Natural Philosophy of Cause and Chance' (Clarendon: Oxford).

Born, M., and Jordan, P. (1925). *Z. Phys.* **34**, 858–88.

Brillouin, L. (1956). 'Science and Information Theory' (Academic: New York).

Busch, P., Lahti, P. J., and Mittelstaedt, P. (1991) 'The Quantum Theory of Measurement' (Springer: Berlin).

Cornwell, J. F. (1984). 'Group Theory In Physics' (Academic: London).

De Witt, B. S., and Graham, N. (Eds) (1973). 'The Many-Worlds Interpretation of Quantum Mechanics' (Princeton University Press).

Einstein, A., Podolsky, B., and Rosen, N. (1935). *Phys. Rev.* **90**, 777–80.

Everett, III, H. (1957). *Rev. Mod. Phys.* **29**, 454–62.

Ghose, P., Home, D., and Agarwal, G. S. (1992). *Phys. Lett.* A **168**, 95–9.

Green, H. S. (1958). *Nuovo Cimento* **9**, 880–9.

Green, H. S., and Triffet, T. (1969). *J. Math. Phys.* **10**, 1069–89.

Green, H. S., and Triffet, T. (1991). *Aust. J. Phys.* **44**, 323–34.

Gottfried, K. (1966). 'Quantum Mechanics' (Benjamin: New York).

Heisenberg, W. (1925). *Z. Phys.* **33**, 879–93.

Janossy, L. (1952). *Acta Phys. Hungar.* 1, 391–422; 423–68, also (1953) as *Ann. Phys. (Leipzig)*, **11**, 293–323; 323–61.

Mizobuchi, Y., and Ohtaké, Y. (1992). *Phys. Lett.* A **168**, 1–5.

Neumann, J. von (1955). 'Mathematical Foundations of Quantum Mechanics' (Princeton University Press).

Pauli, W. (1941). *Rev. Mod. Phys.* **13**, 203–32.
Schrödinger, E. (1926). *Ann. Phys. (Leipzig)* **79**, 734–56.
Shannon, C. E. (1949). 'The Mathematical Theory of Communiication' (Univ. of Illinois: Urbana).
Stapp, H. P. (1992). *Phys. Rev.* A **46**, 6860–8.
Stapp, H. P. (1993). 'Mind, Matter and Quantum Mechanics' (Springer: Berlin).
Zurek, W. H. (1982). *Phys. Rev.* D **26**, 1862–80.

Manuscript received 21 October 1994, accepted 24 February 1995